Sidorenko's Conjecture: Elementary Cases and Graphon Methods

Rohan Singh

June 2026

Abstract

Sidorenko's conjecture predicts that $t_H(G) \geq t_{K_2}(G)^{|E(H)|}$ for every bipartite graph H and every finite graph G. We introduce the conjecture through homomorphism densities, prove the finite-host inequality for P_3, C_4, and the complete bipartite graphs $K_{m,n}$, and characterize the bipartite hosts attaining each bound. We then formulate the conjecture for graphons, prove the conjecture for all trees, and apply three analytic methods: reflection positivity, which together with the tree bound proves the conjecture for every even cycle; graph norms; and entropy. A final section examines the limitations of these methods on the open case $K_{5,5} \setminus C_{10}$ and surveys the Forcing Conjecture and the Common Graph Conjecture.

1 Introduction

1.1 Background and History

Random graphs serve as natural baselines for subgraph counts in extremal graph theory. Given a fixed edge density p, Sidorenko's conjecture posits that the value attained in $G(n, p)$ is a lower bound for the homomorphism density of every bipartite graph. The conjecture is known for trees, even cycles, and complete bipartite graphs [1], for hypercubes [16], and for bipartite graphs with a vertex complete to the opposite part [10], but remains open for general bipartite H. Blakley and Roy proved the case of paths through an inequality for symmetric matrices with nonnegative entries [9]. Related questions arose in the Turán-type problems of Erdős and Simonovits [6], who later gave an equivalent formulation in terms of supersaturation, asking how many copies of H are forced once the number of edges exceeds the Turán threshold [7]. Sidorenko stated and popularized the modern conjecture for bipartite graphs in 1991 [1, 2].

1.2 Motivation and Applications

Homomorphism counts also occur outside extremal graph theory. Analogous sums appear in the partition functions of statistical mechanics, and in database theory counting homomorphisms is closely related to estimating the output size of a multi-relation join [8], where bounds on such counts inform the design of query algorithms. Below, we review homomorphism densities and establish the conjecture for P_3, C_4, and $K_{m,n}$. We then introduce graphons together with reflection positivity, graph norms, and entropy, proving in particular the conjecture for all trees and all even cycles.

1.3 Definitions and Statement of the Conjecture

Definition 1 (Graph homomorphism)**.** Let H and G be graphs. A *homomorphism* from H to G is a map

$$\phi : V(H) \to V(G)$$

such that $\phi(u)\phi(v) \in E(G)$ whenever $uv \in E(H)$. Homomorphisms need not be injective, so nonadjacent vertices of H may share an image. For instance, if $H = P_3$ has vertices a, b, c and edges ab, bc, and $xy \in E(G)$, then both $(\phi(a), \phi(b), \phi(c)) = (x, y, z)$ for any $z \in N(y)$ and $(\phi(a), \phi(b), \phi(c)) = (x, y, x)$ are homomorphisms.

Definition 2 (Homomorphism density)**.** Let H be a pattern graph and G a host graph. The *homomorphism density* $t_H(G)$ is the probability that a uniformly random map $\phi : V(H) \to V(G)$ is a homomorphism:

$$t_H(G) \coloneqq \frac{\hom(H, G)}{|V(G)|^{|V(H)|}},$$

where $\hom(H, G)$ is the number of homomorphisms from H to G.

Definition 3 (Edge density)**.** Let G be a finite simple graph. Its *homomorphism edge density* is

$$p \coloneqq t_{K_2}(G) = \frac{\hom(K_2, G)}{|V(G)|^2} = \frac{2|E(G)|}{|V(G)|^2}.$$

Conjecture 1 (Sidorenko, 1991)**.** *Let H be a bipartite graph and let G be any finite graph. Then*

$$t_H(G) \geq t_{K_2}(G)^{|E(H)|}.$$

The exponent is explained by the Erdős–Rényi random graph $G(n,p)$, in which each possible edge is included independently with probability p. Any injective $\phi : V(H) \to V(G)$ is a homomorphism with probability exactly $p^{|E(H)|}$, and the proportion of noninjective maps tends to zero, so

$$\mathbb{E}\big[t_H(G(n,p))\big] \longrightarrow p^{|E(H)|}, \qquad \mathbb{E}\big[t_{K_2}(G(n,p))\big] \longrightarrow p.$$

The probabilistic method therefore supplies hosts G_n whose H-density is no greater than the expectation above. Moreover, almost every graph in the model has edge density close to p once n is large. Hence $p^{|E(H)|}$ is the best lower bound one can hope for, and Sidorenko's conjecture asserts that no host of edge density p falls below it. The bipartite hypothesis on H is necessary. If $H = K_3$ and G is any bipartite graph with at least one edge, then $t_{K_3}(G) = 0$ while $t_{K_2}(G) > 0$. When $H = K_2$ the inequality holds with equality by Definition 3, since $|E(K_2)| = 1$. The conjecture posits that this lower bound persists for all bipartite H.

2 Elementary Cases of Sidorenko's Conjecture

We first establish part-respecting bounds for bipartite hosts $G = (X, Y; E)$ with fixed parts X and Y, and define the *bipartite edge density*

$$p_{\mathrm{bi}} \coloneqq \frac{|E|}{|X||Y|}.$$

If H has specified parts A and B, we count only homomorphisms sending A to X and B to Y, and write

$$t_H^{X,Y}(G) \coloneqq \frac{\hom_{X,Y}(H, G)}{|X|^{|A|}|Y|^{|B|}}$$

for the resulting part-respecting density. The results below give $t_H^{X,Y}(G) \geq p_{\mathrm{bi}}^{|E(H)|}$. Although the proofs are phrased for bipartite hosts, they already imply the ordinary finite-host statement: Lemma 3 replaces an arbitrary host by a bipartite graph having exactly the same relevant homomorphism and edge densities. Let $d(v)$ denote the degree of v, and for $x_1, \ldots, x_m \in X$ let

$$N_\cap(x_1, \ldots, x_m) := \left| \bigcap_{i=1}^m N(x_i) \right|$$

denote the size of their common neighborhood in Y, so that $N_\cap(x) = d(x)$.

2.1 Preliminary Inequalities

We record two standard moment inequalities:

Lemma 1 (Cauchy–Schwarz inequality). *For any real sequences $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$,*

$$\left(\sum_{i=1}^n a_i b_i \right)^2 \leq \left(\sum_{i=1}^n a_i^2 \right) \left(\sum_{i=1}^n b_i^2 \right).$$

Proof. For every real λ we have $\sum_i (a_i - \lambda b_i)^2 \geq 0$, so this quadratic in λ has nonpositive discriminant. □

Corollary 1 (Sum-of-squares form). *For any reals $a_1, \ldots, a_n$,*

$$\sum_{i=1}^n a_i^2 \geq \frac{1}{n} \left(\sum_{i=1}^n a_i \right)^2,$$

with equality if and only if $a_1 = \cdots = a_n$.

Proof. Take $b_i = 1$ in Lemma 1, so that $\sum_i b_i^2 = n$. Equality there requires (a_i) and (b_i) to be proportional. □

Lemma 2 (Power mean inequality). *For any nonnegative reals $a_1, \ldots, a_N$ and any integer $k \geq 1$,*

$$\sum_{i=1}^N a_i^k \geq \frac{1}{N^{k-1}} \left(\sum_{i=1}^N a_i \right)^k.$$

For $k \geq 2$, equality holds if and only if $a_1 = \cdots = a_N$; for $k = 1$ the two sides always agree.

Proof. The map $t \mapsto t^k$ is convex on $[0, \infty)$, so Jensen's inequality gives

$$\frac{1}{N}\sum_{i=1}^{N} a_i^k \geq \left(\frac{1}{N}\sum_{i=1}^{N} a_i\right)^k,$$

and multiplying by N gives the statement. For $k \geq 2$ the map is strictly convex, so Jensen's inequality is strict unless the a_i are equal. $\square$

2.2 The Path P_3

For P_3, the homomorphism count reduces to the second moment of the vertex degrees, so Corollary 1 applies directly.

Definition 4 (The path P_3)**.** Let P_3 be the path on three vertices, with center c joined to leaves ℓ_1 and ℓ_2 (Figure 1), so $|E(P_3)| = 2$. Set $A = \{c\}$ and $B = \{\ell_1, \ell_2\}$.

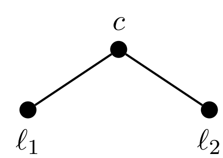

Figure 1: The path P_3: a center vertex joined to two leaves.

Theorem 1. *Let G be a bipartite graph with parts X and Y. Then $t_{P_3}^{X,Y}(G) \geq p_{\mathrm{bi}}^2$.*

Proof. Let $T = \hom_{X,Y}(P_3, G)$, so that $t_{P_3}^{X,Y}(G) = T/(|X||Y|^2)$, the denominator counting the $|X|$ choices for c and $|Y|^2$ choices for the leaves. Each $x \in X$ admits $d(x)$ choices for each leaf, so

$$T = \sum_{x \in X} d(x)^2.$$

Every edge has exactly one endpoint in X, whence $\sum_{x \in X} d(x) = |E|$. Corollary 1 with $n = |X|$ and $a_x = d(x)$ therefore gives

$$T \;\geq\; \frac{|E|^2}{|X|} \;=\; \frac{\big(p_{\mathrm{bi}}|X||Y|\big)^2}{|X|} \;=\; p_{\mathrm{bi}}^2|X||Y|^2,$$

and dividing by $|X||Y|^2$ completes the proof. $\square$

2.3 A Worked Example

To illustrate, let $X = \{x_1, x_2, x_3, x_4\}$ and $Y = \{y_1, y_2, y_3, y_4\}$, and let G be the graph of Figure 2, with edges

$$x_1 \sim y_1, y_2, y_3, y_4 \qquad x_2 \sim y_1, y_2 \qquad x_3 \sim y_1 \qquad x_4 \sim y_2.$$

Here $|E| = 8$ and $p_{\mathrm{bi}} = 8/16 = 1/2$, so the bound is $p_{\mathrm{bi}}^2 = 1/4$. The degree

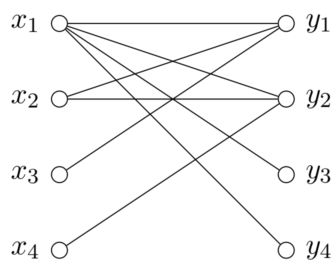

Figure 2: The bipartite graph used in the P_3 worked example.

sequence on X is $(4, 2, 1, 1)$, giving

$$T = \sum_{x \in X} d(x)^2 = 4^2 + 2^2 + 1^2 + 1^2 = 22, \qquad t_{P_3}^{X,Y}(G) = \frac{22}{64} \approx 0.344 \;\geq\; \frac{1}{4},$$

where $|X||Y|^2 = 64$ is the number of part-respecting maps. Because the degree sequence is nonconstant, Corollary 1 is strict. If instead $|E| = 8$ and every vertex of X has degree 2, then $T = 16$ and $t_{P_3}^{X,Y}(G) = 1/4$, attaining equality.

2.4 The Four-Cycle (C_4)

Next we consider the four-cycle, with bipartition $\{x_1, x_2\}$ and $\{y_1, y_2\}$ as in Figure 3.

Theorem 2. *Let G be a bipartite graph with bipartite edge density p_{bi}. Then $t_{C_4}^{X,Y}(G) \geq p_{\mathrm{bi}}^4$.*

Proof. For $u, v \in X$ the quantity $N_\cap(u, v)$ is the codegree of u and v. The two vertices may coincide, and $N_\cap(u, u) = d(u)$. A part-respecting homomorphism is a tuple (x_1, y_1, x_2, y_2) with cycle edges x_1y_1, y_1x_2, x_2y_2, y_2x_1.

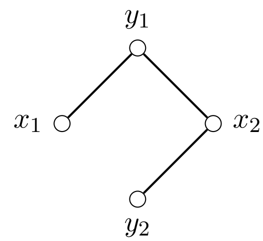

Figure 3: The four-cycle C_4, with x_1, x_2 in one bipartition class and y_1, y_2 in the other.

For fixed $x_1, x_2 \in X$ the vertices y_1 and y_2 are chosen independently from the $N_\cap(x_1, x_2)$ common neighbors, so

$$T = \sum_{x_1, x_2 \in X} N_\cap(x_1, x_2)^2,$$

and Corollary 1 applied over X^2 gives

$$T \geq \frac{1}{|X|^2} \left(\sum_{x_1, x_2 \in X} N_\cap(x_1, x_2) \right)^2.$$

Let

$$\mathcal{S} \coloneqq \{(x_1, x_2, y) \in X \times X \times Y : x_1 y, x_2 y \in E\}.$$

Counting $\mathcal{S}$ by the pair (x_1, x_2) gives $|\mathcal{S}| = \sum_{x_1, x_2 \in X} N_\cap(x_1, x_2)$, while counting by $y \in Y$ gives $|\mathcal{S}| = \sum_{y \in Y} d(y)^2$, since x_1 and x_2 range independently over the $d(y)$ neighbors of y. A second application of Corollary 1, with $\sum_{y \in Y} d(y) = |E|$, gives $\sum_{y \in Y} d(y)^2 \geq |E|^2/|Y|$. Hence

$$T \;\geq\; \frac{1}{|X|^2} \left(\frac{|E|^2}{|Y|} \right)^2 = \frac{|E|^4}{|X|^2 |Y|^2} = p_{\mathrm{bi}}^4 |X|^2 |Y|^2,$$

and dividing by the number $|X|^2|Y|^2$ of part-respecting maps completes the proof. $\square$

2.5 Complete Bipartite Graphs ($K_{m,n}$)

We now generalize these bounds to $K_{m,n}$. The family includes both earlier cases, since $P_3 = K_{1,2}$ and $C_4 = K_{2,2}$.

Theorem 3. *Let $m, n \geq 1$. Let G be a bipartite graph with parts X and Y and bipartite edge density p_{bi}. Then $t^{X,Y}_{K_{m,n}}(G) \geq p_{\mathrm{bi}}^{mn}$.*

Proof. A part-respecting homomorphism is a tuple $(x_1, \ldots, x_m, y_1, \ldots, y_n)$ with $x_i y_j \in E(G)$ for all i, j. For fixed $(x_1, \ldots, x_m)$ each y_j is chosen independently from the common neighborhood, so

$$T = \sum_{x_1,\ldots,x_m \in X} N_\cap(x_1, \ldots, x_m)^n,$$

the sum running over all $|X|^m$ ordered tuples. Lemma 2 with $N = |X|^m$ and $k = n$ gives

$$T \;\geq\; \frac{1}{|X|^{m(n-1)}} \left(\sum_{x_1,\ldots,x_m \in X} N_\cap(x_1, \ldots, x_m) \right)^n .$$

Now count the tuples $(x_1, \ldots, x_m, y)$ with y adjacent to every x_i. Counting by the x-tuple gives $N_\cap(x_1, \ldots, x_m)$ choices of y; counting by y gives $d(y)^m$ tuples, since each x_i ranges independently over the neighbors of y. Hence

$$\sum_{x_1,\ldots,x_m \in X} N_\cap(x_1, \ldots, x_m) = \sum_{y \in Y} d(y)^m,$$

and Lemma 2 with $N = |Y|$ and $k = m$ gives

$$\sum_{y \in Y} d(y)^m \;\geq\; \frac{1}{|Y|^{m-1}} \left(\sum_{y \in Y} d(y) \right)^m = \frac{|E|^m}{|Y|^{m-1}}.$$

Combining,

$$T \;\geq\; \frac{1}{|X|^{m(n-1)}} \left(\frac{|E|^m}{|Y|^{m-1}} \right)^n = \frac{|E|^{mn}}{|X|^{m(n-1)}\,|Y|^{n(m-1)}},$$

and dividing by the $|X|^m |Y|^n$ unrestricted maps gives

$$t^{X,Y}_{K_{m,n}}(G) \;\geq\; \frac{|E|^{mn}}{|X|^{mn} |Y|^{mn}} = p_{\mathrm{bi}}^{mn}. \qquad \square$$

The bipartite double cover transfers these part-respecting bounds to arbitrary hosts.

Lemma 3 (Bipartite double cover)**.** *Let H be a bipartite graph with fixed bipartition $A \cup B$. Given a finite graph G, let B_G be the bipartite graph with parts $X = Y = V(G)$ in which $x_u \in X$ is adjacent to $y_v \in Y$ exactly when $uv \in E(G)$. Then*

$$p_{\mathrm{bi}}(B_G) = t_{K_2}(G) \qquad \textit{and} \qquad t_H^{X,Y}(B_G) = t_H(G).$$

Proof. Every edge of G produces the two cross-edges $x_u y_v$ and $x_v y_u$ in B_G, so $p_{\mathrm{bi}}(B_G) = 2|E(G)|/|V(G)|^2 = t_{K_2}(G)$. Sending each vertex of A to its copy in X and each vertex of B to its copy in Y is a bijection between homomorphisms $H \to G$ and part-respecting homomorphisms $H \to B_G$, and the normalizing denominators agree. □

Theorems 1, 2, and 3 therefore yield the finite-host Sidorenko inequality for their respective patterns.

2.6 Equality Cases

Each theorem above applies Lemma 2, and the resulting equality conditions separate P_3 sharply from the other cases.

Proposition 4. *Let $G = (X, Y; E)$ be a bipartite graph with $|E| > 0$, and let $m, n \geq 1$.*

(i) Equality holds in Theorem 3 if and only if $N_\cap(x_1, \ldots, x_m)$ is constant on X^m and, when $m \geq 2$, $d(y)$ is constant on Y.

(ii) For $m = 1$, condition (i) states that G is regular on the X side. In particular, equality holds in Theorem 1 exactly when d is constant on X.

(iii) For $m, n \geq 2$, condition (i) forces G to be the complete bipartite graph on $X \cup Y$, so $p_{\mathrm{bi}} = 1$. In particular, the bound of Theorem 2 is attained only by the complete bipartite host.

Proof. The proof of Theorem 3 applies Lemma 2 twice, the second time composed with an increasing function of its output, so equality requires equality at both steps. By Lemma 2 this means $N_\cap$ is constant on X^m and, when the second exponent m is at least 2, that d is constant on Y. This is (i), and (ii) follows since $N_\cap(x) = d(x)$ and the second application has exponent

$k = 1$. For (iii), let $m \geq 2$ and suppose $N_\cap$ is constant on X^m with value c. Evaluating on $(x, x, \ldots, x)$ gives $d(x) = c$ for all $x \in X$. Evaluating on a tuple with first entry x_1 gives $|N(x_1) \cap \cdots \cap N(x_m)| = c = |N(x_1)|$, so the intersection is all of $N(x_1)$ and hence $N(x_1) \subseteq N(x_i)$ for each i. As the tuple was arbitrary, all vertices of X share a neighborhood $N \subseteq Y$ with $|N| = c$. Then $d(y) = |X|$ for $y \in N$ and $d(y) = 0$ otherwise, so constancy of d on Y forces $N = Y$. $\square$

Equality for P_3 thus requires only regularity on one side, which the comparison graph of Section 2.3 satisfies, whereas for C_4 and the larger complete bipartite patterns it determines the host completely. The Forcing Conjecture of Section 4.2.1 is the approximate form of this dichotomy. This exact statement is finite and part-respecting: for C_4, equality in the bipartite inequality at positive density forces $p_{\mathrm{bi}} = 1$. It does not conflict with the asymptotic statement in Section 4.2.1. There, quasirandom graph sequences of any limiting density $p \in (0, 1)$ have $t_{C_4}(G_n) \to p^4$ without any individual finite host attaining equality. Thus exact equality is substantially more rigid than asymptotic near-equality.

3 Analytic Methods

We now formulate the conjecture for graphons and develop three analytic methods: reflection positivity, graph norms, and entropy. Besides recovering cases from Section 2, the graphon framework proves the conjecture for all trees and lets reflection positivity handle every even cycle. Section 4.1 compares the directions in which the methods extend.

3.1 Graphons and their Continuous Formulation

In graphon language homomorphism densities become integrals, and inequalities such as Cauchy–Schwarz apply without discrete approximation. A graphon is a symmetric Lebesgue-measurable function $W : [0, 1]^2 \to [0, 1]$ [4]; the standard reference is Lovász [5]. A graphon is a measurable analogue of an adjacency matrix, with $W(x, y)$ playing for points $x, y \in [0, 1]$ the role that the (i, j) entry plays for vertices i and j, interpreted as the probability of an edge between the sampled vertices. For finite graphs the correspondence is exact: if G has vertices $v_1, \ldots, v_n$, partition $[0, 1]$ into intervals $I_1, \ldots, I_n$ of equal length and define the *step graphon*

$$W_G(x,y) \coloneqq \begin{cases} 1, & \text{if } x \in I_i,\ y \in I_j,\ \text{and } v_iv_j \in E(G), \\ 0, & \text{otherwise.} \end{cases}$$

Choosing a uniform point of I_i is equivalent to choosing v_i, so $t_H(W_G) = t_H(G)$ for every finite H. Graphons also describe limits of sequences of increasingly large dense graphs; Figure 4 shows adjacency matrices sampled from a single graphon at three sizes, with rows and columns ordered by the underlying point of $[0,1]$, and the block pattern stabilizes as n grows even though the individual entries remain random. Choosing independent uni-

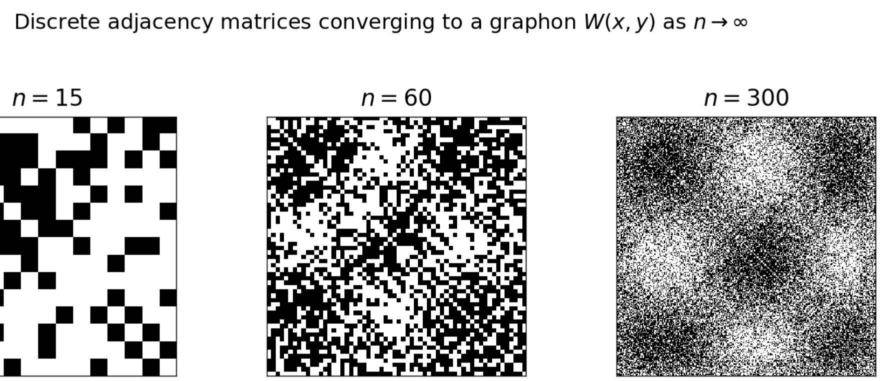

Figure 4: Three adjacency matrices sampled from a fixed graphon, at $n = 15$, 60, and 300.

form points $x, y \in [0,1]$, the edge is present with probability $W(x,y)$, so integrating gives

$$p \coloneqq t_{K_2}(W) = \mathbb{E}[W(x,y)] = \int_0^1 \int_0^1 W(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

in agreement with Definition 3. More generally, let $V(H) = \{1, \ldots, k\}$ and choose independent points $x_1, \ldots, x_k \in [0,1]$. The edge $ij \in E(H)$ then occurs with probability $W(x_i, x_j)$, and the edge events are conditionally independent, so every edge of H occurs with probability $\prod_{ij \in E(H)} W(x_i, x_j)$. Integrating defines

$$t_H(W) \coloneqq \int_{[0,1]^k} \prod_{ij \in E(H)} W(x_i, x_j)\,\mathrm{d}x_1 \cdots \mathrm{d}x_k,$$

which generalizes the finite density $t_H(G)$.

Remark 1 (Multigraph convention)**.** If H is a multigraph, the product is taken with multiplicity, so an edge appearing r times contributes $W(x_i, x_j)^r$. This is needed in Section 3.3, where gluing labeled graphs can create parallel edges.

In graphon terms, Sidorenko's conjecture states that for every bipartite H and every graphon W,

$$\int_{[0,1]^k} \prod_{ij \in E(H)} W(x_i, x_j)\, \mathrm{d}x_1 \cdots \mathrm{d}x_k \geq \left(\int_{[0,1]^2} W(x,y)\, \mathrm{d}x\, \mathrm{d}y \right)^{|E(H)|}.$$

This statement is already in arbitrary-host form, so the bipartite bookkeeping of Section 2 is not needed below.

3.2 Trees

The tree case needs a little more information than the total density of a smaller tree: one must retain how rooted copies are distributed. The following change of measure does exactly that.

Theorem 5 (Trees)**.** *If T is a tree with e edges and W is a graphon of edge density p, then*

$$t_T(W) \geq p^e.$$

Proof. Write $d(x) = \int_0^1 W(x,y)\, \mathrm{d}y$. The claim is immediate when $p = 0$, so assume $p > 0$ and root T at a vertex r. Define a random map $(X_v)_{v \in V(T)}$ into $[0,1]$ as follows. Choose X_r with density $d(x)/p$. Once $X_u = x$ has been chosen, choose each child v of u, independently given its parent, with conditional density

$$\frac{W(x,y)}{d(x)}.$$

Points with $d(x) = 0$ occur with probability zero, so the construction is well defined. Symmetry of W shows inductively that every X_v has the same marginal density $d(x)/p$: if this holds for a parent, then the density of its child at y is

$$\int_0^1 \frac{d(x)}{p} \frac{W(x,y)}{d(x)}\, \mathrm{d}x = \frac{d(y)}{p}.$$

Let q be the joint density of this random map and put

$$F(\mathbf{x}) = \prod_{uv\in E(T)} W(x_u, x_v), \qquad \text{so that} \qquad t_T(W) = \int F(\mathbf{x})\,\mathrm{d}\mathbf{x}.$$

Then

$$q(\mathbf{x}) = \frac{d(x_r)}{p} \prod_{u\to v} \frac{W(x_u, x_v)}{d(x_u)},$$

and hence, on the support of q,

$$\frac{F(\mathbf{x})}{q(\mathbf{x})} = p \prod_{v\in V(T)} d(x_v)^{\deg_T(v)-1}.$$

Jensen's inequality for the concave logarithm now gives

$$\begin{aligned}
\log t_T(W) &= \log \mathbb{E}_q\left[\frac{F}{q}\right] \\
&\geq \mathbb{E}_q\left[\log \frac{F}{q}\right] \\
&= \log p + \sum_{v\in V(T)} (\deg_T(v) - 1)\frac{1}{p}\int_0^1 d(x)\log d(x)\,\mathrm{d}x.
\end{aligned}$$

The convexity of $z \mapsto z\log z$, with $0\log 0 = 0$, yields

$$\frac{1}{p}\int_0^1 d(x)\log d(x)\,\mathrm{d}x \geq \log p.$$

Finally, because T is a tree,

$$\sum_{v\in V(T)} (\deg_T(v) - 1) = 2e - |V(T)| = e - 1.$$

Therefore $\log t_T(W) \geq e\log p$, which is equivalent to the desired inequality.

□

3.3 Reflection Positivity

Reflection positivity generates Cauchy–Schwarz inequalities for graphon densities from positive semidefinite matrices. Let F be a graph with some vertices labeled $1,\ldots,k$, and for fixed $x_1,\ldots,x_k \in [0,1]$ let

$$f_F(x_1,\ldots,x_k)$$

denote the integral defining $t_F(W)$ with the variables at labeled vertices held fixed and only those at unlabeled vertices integrated. If F_1 and F_2 carry the same labels, let $F_1 \star F_2$ be obtained from their disjoint union by identifying vertices with the same label, retaining edge multiplicities and reading densities through Remark 1. Fubini's theorem gives

$$t_{F_1 \star F_2}(W) = \int_{[0,1]^k} f_{F_1}(\mathbf{x}) f_{F_2}(\mathbf{x}) \, \mathrm{d}\mathbf{x}.$$

Lemma 4 (Reflection positivity)**.** *Let W be a graphon and let $F_1, \ldots, F_r$ be graphs with the same k labeled vertices. Then the $r \times r$ matrix*

$$M_{ij} := t_{F_i \star F_j}(W)$$

is positive semidefinite. In particular every 2×2 principal minor is nonnegative, giving

$$t_{F_1 \star F_2}(W)^2 \leq t_{F_1 \star F_1}(W) \, t_{F_2 \star F_2}(W).$$

Proof. For $c_1, \ldots, c_r \in \mathbb{R}$ the Fubini identity gives

$$\sum_{i,j=1}^{r} c_i c_j M_{ij} = \int_{[0,1]^k} \left(\sum_{i=1}^{r} c_i f_{F_i}(\mathbf{x}) \right)^2 \mathrm{d}\mathbf{x} \geq 0,$$

so M is positive semidefinite and its principal minors are nonnegative. $\square$

A single 2×2 minor now recovers Theorem 2.

Theorem 6. *For every graphon W with edge density p,*

$$t_{C_4}(W) \;\geq\; t_{P_3}(W)^2 \;\geq\; p^4.$$

Proof. Let F be a path of length two with labeled endpoints, so that

$$f_F(x, z) := \int_0^1 W(x, y) W(y, z) \, \mathrm{d}y.$$

Gluing two copies of F identifies the endpoints and leaves the middle vertices distinct, so $F \star F = C_4$ with no parallel edges, and

$$t_{C_4}(W) = \int_{[0,1]^2} f_F(x, z)^2 \, \mathrm{d}x \, \mathrm{d}z.$$

Let F_0 carry the same two labels and no edges, so $f_{F_0} \equiv 1$. Then $t_{F \star F_0}(W) = \int f_F = t_{P_3}(W)$ and $t_{F_0 \star F_0}(W) = 1$, so the minor of Lemma 4 for F and F_0 is

$$\begin{pmatrix} t_{C_4}(W) & t_{P_3}(W) \\ t_{P_3}(W) & 1 \end{pmatrix},$$

whose nonnegative determinant gives $t_{C_4}(W) \geq t_{P_3}(W)^2$. Writing

$$\deg_W(y) \coloneqq \int_0^1 W(x, y)\, \mathrm{d}x$$

for the graphon degree function, Jensen's inequality gives

$$t_{P_3}(W) = \int_0^1 \deg_W(y)^2\, \mathrm{d}y \geq \left(\int_0^1 \deg_W(y)\, \mathrm{d}y \right)^2 = p^2,$$

and combining the two bounds gives $t_{C_4}(W) \geq p^4$. □

The same reflection step proves the whole family of even cycles.

Theorem 7 (Even cycles)**.** *For every integer $k \geq 2$ and every graphon W of edge density p,*

$$t_{C_{2k}}(W) \geq p^{2k}.$$

Proof. Let P_k be the path of length k with its two endpoints labeled, and let $f_{P_k}(x, z)$ be its rooted density function. Two copies of P_k glued at both labeled endpoints form C_{2k}, so the Fubini identity gives

$$t_{C_{2k}}(W) = \int_{[0,1]^2} f_{P_k}(x, z)^2\, \mathrm{d}x\, \mathrm{d}z.$$

Applying the reflection-positivity inequality against the edgeless graph with the same two labels, or equivalently applying Cauchy–Schwarz over the pair (x, z), gives

$$t_{C_{2k}}(W) \geq \left(\int_{[0,1]^2} f_{P_k}(x, z)\, \mathrm{d}x\, \mathrm{d}z \right)^2 = t_{P_k}(W)^2.$$

Since P_k is a tree with k edges, Theorem 5 gives $t_{P_k}(W) \geq p^k$. Therefore $t_{C_{2k}}(W) \geq p^{2k}$. □

The case $k = 2$ recovers Theorem 6. The proof illustrates why one should reflect across both endpoints at once: attempting to remove the path one edge at a time would require normalizing the weight $W(x, y)\,\mathrm{d}y$, which is not generally a probability measure.

Gluing labeled graphs thus corresponds to an inner product on rooted density functions, with Lemma 4 encoding the associated Cauchy–Schwarz inequalities. For C_4 the matrix is 2×2 and the analytic content is the step already used in Section 2.4; for C_{2k} the same construction uses a longer path as the reflected piece.

3.4 A Graphon Proof for Complete Bipartite Graphs

Theorem 8. *For every graphon W with edge density p and all positive integers m, n,*

$$t_{K_{m,n}}(W) \geq p^{mn}.$$

Proof. For $\mathbf{x} = (x_1, \dots, x_m) \in [0, 1]^m$ set

$$A(\mathbf{x}) \coloneqq \int_0^1 \prod_{i=1}^m W(x_i, y)\,\mathrm{d}y,$$

the common-neighborhood measure of $x_1, \dots, x_m$. The same two Jensen steps used in Theorem 3 now run directly over $[0, 1]^m$. Indeed, Fubini's theorem gives the middle identity below, and hence

$$\begin{aligned} t_{K_{m,n}}(W) &= \int_{[0,1]^m} A(\mathbf{x})^n\,\mathrm{d}\mathbf{x} \\ &\geq \left(\int_{[0,1]^m} A(\mathbf{x})\,\mathrm{d}\mathbf{x} \right)^n \\ &= \left(\int_0^1 \deg_W(y)^m\,\mathrm{d}y \right)^n \\ &\geq \left(\int_0^1 \deg_W(y)\,\mathrm{d}y \right)^{mn} = p^{mn}. \end{aligned}$$

□

The case $m = n = 2$ agrees with Theorem 6. Evaluating on a step graphon recovers the discrete result.

Corollary 2. *For every finite graph G and all positive integers m, n,*

$$t_{K_{m,n}}(G) \geq t_{K_2}(G)^{mn}.$$

Proof. Apply Theorem 8 to W_G and use $t_H(W_G) = t_H(G)$. $\square$

The graphon route reaches this without the intermediate bipartite host of Lemma 3, since $\deg_W$ already plays that role. The equality conditions mirror the discrete case. Both applications of Jensen's inequality are strict unless the function being averaged is constant, so equality in Theorem 8 forces $\deg_W$ to be constant almost everywhere, hence equal to p, and forces $A(\mathbf{x})$ to be constant almost everywhere. These are the continuous forms of the two conditions in Proposition 4(i): the degree is the same at every point, and the common neighborhood of m points has the same measure wherever those points lie. The constant graphon with value p satisfies both.

3.5 Graph Norms

A second method replaces the counting identity with a norm inequality. Throughout this subsection U denotes a symmetric measurable function $[0,1]^2 \to \mathbb{R}$, not necessarily valued in $[0,1]$, with $t_H(U)$ given by the same integral as before.

Definition 5 (Norming and weakly norming graphs)**.** For a graph H with $e(H)$ edges, set

$$\|U\|_H := \left|t_H(U)\right|^{1/e(H)}.$$

Then H is *norming* if $\|\cdot\|_H$ is a norm on the space of all such U, and *weakly norming* if it is a norm on the cone of nonnegative U.

Proposition 9. *Let H be weakly norming. Then $t_H(G) \geq t_{K_2}(G)^{e(H)}$ for every finite graph G.*

Proof. The statement is due to Hatami [16]. Two properties of $\|\cdot\|_H$ are used. The first is homogeneity: replacing U by cU for $c \geq 0$ multiplies each of the $e(H)$ factors in the integrand by c, so $t_H(cU) = c^{e(H)} t_H(U)$ and $\|cU\|_H = c\|U\|_H$. The second is the triangle inequality, which holds on nonnegative kernels precisely because H is weakly norming. Let G have N vertices $v_1, \ldots, v_N$ and at least one edge, let $p = t_{K_2}(G)$, and let $I_1, \ldots, I_N$ be the intervals defining W_G. For a permutation π of $\{1, \ldots, N\}$ let W^π take the value 1 on $I_a \times I_b$ when $v_{\pi(a)}v_{\pi(b)} \in E(G)$ and 0 otherwise. Translating each I_a onto $I_{\pi(a)}$ carries W_G to W^π, so the two differ only by a renaming of the integration variables and

$$t_F(W^\pi) = t_F(W_G) = t_F(G)$$

for every graph F; in particular $\|W^\pi\|_H = \|W_G\|_H$. Set $\overline{W} \coloneqq \frac{1}{N!}\sum_\pi W^\pi$ and fix $x \in I_a$, $y \in I_b$. If $a \neq b$ then $(\pi(a), \pi(b))$ is a uniformly random ordered pair of distinct indices, so

$$\overline{W}(x, y) = \frac{2|E(G)|}{N(N-1)} \eqqcolon q,$$

while if $a = b$ then $v_{\pi(a)}v_{\pi(a)}$ is never an edge and $\overline{W}(x, y) = 0$. Hence $\overline{W} = q\, W_{K_N}$, where W_{K_N} is the step graphon of the complete graph on N vertices. The triangle inequality followed by homogeneity gives

$$\|W_G\|_H = \frac{1}{N!}\sum_\pi \|W^\pi\|_H \;\geq\; \Big\|\tfrac{1}{N!}\sum_\pi W^\pi\Big\|_H = \|\overline{W}\|_H = q\, t_H(K_N)^{1/e(H)},$$

Since $q = 2|E(G)|/(N(N-1)) \geq p$, raising this inequality to the power $e(H)$ yields

$$t_H(G) \;\geq\; q^{e(H)}\, t_H(K_N) \;\geq\; p^{e(H)}\, t_H(K_N). \tag{1}$$

To remove the factor $t_H(K_N)$, let $G[m]$ be obtained from G by replacing each vertex with m copies and joining two copies exactly when the original vertices are adjacent. Distinct copies of one vertex are left nonadjacent, consistent with G having no loops, so $G[m]$ has the same step graphon as G; hence $t_F(G[m]) = t_F(G)$ for every F, and $G[m]$ has edge density p. Applying (1) to $G[m]$, which has mN vertices, gives

$$t_H(G) = t_H(G[m]) \;\geq\; p^{e(H)}\, t_H(K_{mN})$$

for every $m \geq 1$. Every injective map $V(H) \to V(K_M)$ is a homomorphism, so

$$t_H(K_M) \;\geq\; \frac{M(M-1)\cdots(M-|V(H)|+1)}{M^{|V(H)|}} = \prod_{i=0}^{|V(H)|-1}\Big(1 - \frac{i}{M}\Big) \;\longrightarrow\; 1$$

as $M \to \infty$, and letting $m \to \infty$ gives $t_H(G) \geq p^{e(H)}$. $\square$

Hatami applied this framework to the hypercube case cited in Section 1.1, showing that hypercubes are norming [16]. The argument yields no equality information, since it compares W_G only with a large complete graph.

3.6 Entropy Methods

Entropy methods bound homomorphism counts through the information content of a random homomorphism. All logarithms are base 2, and for a random variable Z with finitely many values the *Shannon entropy* is

$$H(Z) \coloneqq -\sum_{z} \Pr[Z = z] \log \Pr[Z = z].$$

Lemma 5 (Uniform bound and chain rule)**.** *Let Z, Z_1, Z_2 take finitely many values.*

(E1) If Z takes values in a finite set S, then $H(Z) \le \log |S|$, with equality exactly when Z is uniform on S.

(E2) $H(Z_1, Z_2) = H(Z_1) + H(Z_2 \mid Z_1)$.

Property (E1) is used below in both directions: to bound the size of a set from the entropy of a variable supported on it, and to cap the entropy of a single choice by the number of options available.

Theorem 10 (Entropy form of Theorem 1)**.** *Let $G = (X, Y; E)$ be bipartite with $|E| > 0$. Then $t^{X,Y}_{P_3}(G) \ge p^2_{\mathrm{bi}}$.*

Proof. Let S be the set of part-respecting homomorphisms $P_3 \to G$ with center in X, and $T = |S|$. Choose an edge of G uniformly at random with endpoints $A \in X$ and $B_1 \in Y$, then choose B_2 uniformly from $N(A)$, independently of B_1 given A. Every outcome lies in S, so (E1) gives

$$\log T \;\ge\; H(A, B_1, B_2),$$

and (E2) applied twice gives

$$H(A, B_1, B_2) = H(A) + H(B_1 \mid A) + H(B_2 \mid A, B_1).$$

Given $A = x$, each of B_1 and B_2 is uniform on the $d(x)$ neighbors of x, and B_2 is conditionally independent of B_1. Writing $q_x \coloneqq d(x)/|E|$ for the distribution of A,

$$H(B_1 \mid A) = H(B_2 \mid A, B_1) = \sum_{x \in X} q_x \log d(x),$$

$$H(A) = \log |E| - \sum_{x \in X} q_x \log d(x).$$

Indeed, expanding the definition of entropy and using $q_x = d(x)/|E|$ gives

$$\begin{aligned} H(A) &= -\sum_{x\in X} q_x \log q_x \\ &= -\sum_{x\in X} q_x \log\left(\frac{d(x)}{|E|}\right) \\ &= -\sum_{x\in X} q_x \log d(x) + \left(\sum_{x\in X} q_x\right)\log|E| \\ &= \log|E| - \sum_{x\in X} q_x \log d(x). \end{aligned}$$

Adding the three entropy terms leaves

$$H(A, B_1, B_2) = \log|E| + \sum_{x\in X} q_x \log d(x).$$

Applying (E1) to A, which takes values in X, gives $H(A) \leq \log|X|$, and substituting the expression for $H(A)$ yields

$$\sum_{x\in X} q_x \log d(x) \;\geq\; \log|E| - \log|X|.$$

Hence $H(A, B_1, B_2) \geq 2\log|E| - \log|X|$, so $T \geq |E|^2/|X|$, the bound obtained from Corollary 1 in Theorem 1. Dividing by $|X||Y|^2$ and substituting $|E| = p_{\mathrm{bi}}|X||Y|$ gives the result. $\square$

This proof has the shape of every entropy argument for the conjecture: the homomorphism is built one vertex at a time along a chosen order, the chain rule splits its entropy into one term per vertex, and each term is bounded by the options available at that step. Unlike the moment inequalities of Section 2, the scheme requires no moment structure in the count, only a vertex whose conditional distribution can be controlled. Conlon, Fox, and Sudakov used that freedom to prove Sidorenko's inequality for every bipartite graph with a vertex adjacent to all vertices of the opposite part, sampling that vertex first and bounding the conditional entropies of the remainder [10].

4 Scope, Open Problems, and Further Directions

4.1 Scope and Limitations of the Methods

Sections 2 and 3 supply four methods: the moment inequalities of Cauchy–Schwarz and the power mean, reflection positivity, graph norms, and entropy. Comparing them on a single graph shows how different their structural requirements are. The underlying difficulty is that edges in a host graph need not behave independently. In a random graph one edge gives no information about its neighbors, while in a structured graph the conditional edge probabilities vary across regions, so dense regions carry many homomorphisms of H and sparse regions few. Each method is a way of showing that the surplus outweighs the deficit. The moment inequalities prove the examples of Section 2 because those counts reduce to moments: $\sum_x d(x)^2$ for P_3, $\sum_{x_1,x_2} N_\cap(x_1,x_2)^2$ for C_4, and a higher moment of common-neighborhood sizes for $K_{m,n}$. A general bipartite pattern need not have a count of this form. Reflection positivity replaces a direct count with the matrix of Lemma 4. Proving a new case requires labeled pieces whose gluing produces the target and whose matrix inequalities reduce its density to quantities already bounded. Symmetry often suggests useful pieces but does not supply the inequalities. Graph norms require t_H to define a norm, which by Proposition 9 suffices but is a strong hypothesis on H and can fail. Entropy uses concentration instead. A vertex adjacent to a large portion of the opposite part can be sampled first, as in Theorem 10, and used to bound the conditional entropy of the remaining choices. The graph $K_{5,5} \setminus C_{10}$ exhibits all four limitations. It is obtained by deleting a Hamiltonian cycle from $K_{5,5}$, has ten vertices, and is 3-regular; it is neither tree-like nor complete bipartite, and has no vertex complete to the opposite part. Conlon and Lee identify it as the smallest open case and prove the conjecture for its square [11]. Rossman gives a conditional spectral approach to bipartite Möbius ladders: a conjectured monotonicity property for an associated matrix spectrum, in a particular parameter case, would imply Sidorenko's conjecture for the family [19]. Each method meets an obstruction here. Its count is not a single common-neighborhood moment, so the argument for $K_{m,n}$ does not transfer. It has no vertex complete to the opposite part, so the Conlon–Fox–Sudakov scheme has no distinguished first vertex. Its symmetry does not by itself produce labeled combinations reducing its density to controlled quantities. Finally, Lee and Schülke showed that $K_{5,5} \setminus C_{10}$ is not weakly norming [17], ruling out Proposition 9. Only this last obstruction is a theorem about the graph; the first three concern the particular arguments given here rather than the methods in general.

4.2 Related Conjectures

Several related problems concern the structure of graphs that nearly attain the Sidorenko bound.

4.2.1 The Forcing Conjecture

First posed as a question by Skokan and Thoma [13] and later formalized by Conlon, Fox, and Sudakov [10], the Forcing Conjecture asks which hosts can come close to the bound, a question Sidorenko's conjecture leaves open. Quasirandomness formalizes asymptotic similarity to a random graph.

Definition 6 (Quasirandomness)**.** A sequence of graphs G_n with $|V(G_n)| \to \infty$ and $t_{K_2}(G_n) \to p$ is *quasirandom of density* p if $t_{C_4}(G_n) \to p^4$.

Chung, Graham, and Wilson showed this single condition equivalent to $t_H(G_n) \to p^{|E(H)|}$ for every H, and to several further properties of random graphs [12]. Quasirandomness is therefore the statement that a sequence meets the Sidorenko benchmark for all patterns simultaneously.

Trees show why any forcing statement must impose a cyclic hypothesis. Let T be a tree and G a d-regular graph on n vertices, with $p = d/n$. Ordering the vertices of T from a root so that each has its parent already placed gives n choices for the root and d for each of the remaining $|V(T)| - 1 = |E(T)|$ vertices, so $\hom(T, G) = n\, d^{|E(T)|}$ and

$$t_T(G) = \frac{n\, d^{|E(T)|}}{n^{|V(T)|}} = p^{|E(T)|}.$$

Every regular graph therefore sits exactly at the tree benchmark. For a concrete failure, let G_n be a disjoint union of k cliques of size n/k. Its edge density tends to $p = 1/k$ and it is regular, so $t_T(G_n) \to p^{|E(T)|}$ for every tree T; but a homomorphic image of C_4 must lie inside a single clique, giving $t_{C_4}(G_n) \to k^{-3} = p^3$ rather than p^4.

Motivated by this obstruction, a bipartite graph H containing a cycle is *forcing* if, for every $p \in (0,1)$, every sequence G_n with

$$t_{K_2}(G_n) \to p \qquad \text{and} \qquad t_H(G_n) \to p^{|E(H)|}$$

is quasirandom. The Forcing Conjecture asserts that every bipartite graph containing a cycle is forcing.

4.2.2 The Common Graph Conjecture

A graph H is *common* if the number of monochromatic copies of H in a 2-coloring of the edges of a large complete graph is asymptotically minimized by a uniformly random coloring. Identifying a coloring with a graphon W and its complement $1-W$, this says

$$t_H(W) + t_H(1-W) \;\geq\; 2^{1-|E(H)|}$$

for every graphon W, the right side being the value at $W \equiv \frac{1}{2}$. The notion goes back to a 1962 conjecture of Erdős that all complete graphs are common, extended to all graphs by Burr and Rosta [14]. Both were disproven, the first by Thomason, who showed K_4 is not common [15], and the second by Sidorenko, who showed a triangle with a pendant edge is not common [3]. Every Sidorenko graph is common: with $p = t_{K_2}(W)$, applying the inequality to W and to $1-W$, whose edge density is $1-p$, gives

$$t_H(W) + t_H(1-W) \geq p^{|E(H)|} + (1-p)^{|E(H)|} \geq 2^{1-|E(H)|},$$

the second step being Corollary 1 in power-mean form, minimized at $p = \frac{1}{2}$. Whether every bipartite common graph satisfies Sidorenko's conjecture is open [20], and the discrepancy hinges on the number of colors. Call H k-*common* if the number of monochromatic copies of H in a k-coloring is asymptotically minimized by a uniformly random coloring. Králˇ, Noel, Norin, Volec, and Wei proved that a bipartite graph satisfies Sidorenko's conjecture if and only if it is k-common for every k [18], so a bipartite counterexample would be common for two colors but not for all.

4.3 Algorithmic Considerations

In database theory, homomorphism counts govern relational join sizes, and exact subgraph counting in large graphs is typically intractable, so algorithms rely on density approximations. Atserias, Grohe, and Marx bound the output size of a multi-relation join from above in terms of the fractional edge cover number of the query hypergraph, and their bound is tight in the worst case [8]. Sidorenko's conjecture supplies the complementary estimate, a lower bound on the number of homomorphisms of a bipartite pattern in terms of edge density alone. Upper bounds rule out large outputs, while lower bounds rule out small ones and help a planner decide whether a subquery is worth materializing.

4.4 Concluding Perspective

Sidorenko's conjecture characterizes the local subgraph counts forced by a global edge-density condition. Degree moments and common-neighborhood moments already establish the conjectured bound for several basic families, and Proposition 4 shows that for all but the smallest of them the bound is attained only by the complete bipartite host. The graphon formulation places these arguments in an analytic setting. Reflection positivity, graph norms, and entropy apply under different structural hypotheses, but none of the arguments presented here resolves the same ten-vertex graph. Extending these ideas to all bipartite graphs remains a central open problem.

Acknowledgements

I am grateful to Simon Rubinstein-Salzedo and the Euler Circle community for their guidance and support during the development of this paper. I also thank Shahzar Rizvi for his detailed feedback on earlier drafts.

References

[1] A. F. Sidorenko, *Inequalities for functionals generated by bipartite graphs* (in Russian), Diskret. Mat. 3 (1991), no. 3, 50–65.

[2] A. Sidorenko, *A correlation inequality for bipartite graphs*, Graphs and Combinatorics 9 (1993), 201–204.

[3] A. F. Sidorenko, *Cycles in graphs and functional inequalities*, Matematicheskie Zametki 46 (1989), no. 5, 72–79; English translation in Mathematical Notes 46 (1989), 877–882.

[4] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, Journal of Combinatorial Theory, Series B, 96 (2006), no. 6, 933–957.

[5] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society Colloquium Publications 60, American Mathematical Society, Providence, RI, 2012.

[6] P. Erdős and M. Simonovits, *A limit theorem in graph theory*, Studia Scientiarum Mathematicarum Hungarica 1 (1966), 51–57.

[7] P. Erdős and M. Simonovits, *Cube-supersaturated graphs and related problems*, in Progress in Graph Theory (Waterloo, Ont., 1982), Academic Press, Toronto, ON, 1984, 203–218.

[8] A. Atserias, M. Grohe, and D. Marx, *Size bounds and query plans for relational joins*, SIAM Journal on Computing 42 (2013), no. 4, 1737–1767.

[9] G. R. Blakley and P. Roy, *A Hölder type inequality for symmetric matrices with nonnegative entries*, Proceedings of the American Mathematical Society 16 (1965), 1244–1245.

[10] D. Conlon, J. Fox, and B. Sudakov, *An approximate version of Sidorenko's conjecture*, Geometric and Functional Analysis 20 (2010), no. 6, 1354–1366.

[11] D. Conlon and J. Lee, *Sidorenko's conjecture for blow-ups*, Discrete Analysis 2021:2, 13 pp.

[12] F. R. K. Chung, R. L. Graham, and R. M. Wilson, *Quasi-random graphs*, Combinatorica 9 (1989), no. 4, 345–362.

[13] J. Skokan and L. Thoma, *Bipartite subgraphs and quasi-randomness*, Graphs and Combinatorics 20 (2004), 255–262.

[14] S. A. Burr and V. Rosta, *On the Ramsey multiplicities of graphs—problems and recent results*, Journal of Graph Theory 4 (1980), no. 4, 347–361.

[15] A. Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, Journal of the London Mathematical Society (2) 39 (1989), no. 2, 246–255.

[16] H. Hatami, *Graph norms and Sidorenko's conjecture*, Israel Journal of Mathematics 175 (2010), 125–150.

[17] J. Lee and B. Schülke, *Convex graphon parameters and graph norms*, Israel Journal of Mathematics 242 (2021), no. 2, 549–563.

[18] D. Kráľ, J. A. Noel, S. Norin, J. Volec, and F. Wei, *Non-bipartite k-common graphs*, Combinatorica 42 (2022), no. 1, 87–114.

[19] B. Rossman, *On Sidorenko's Conjecture for Bipartite Möbius Ladders*, preprint, 2025. Available at `https://users.cs.duke.edu/~br148/sidorenko-mobius.pdf` (accessed August 6, 2026).

[20] A. Grzesik, J. Lee, B. Lidický, and J. Volec, *On tripartite common graphs*, Combinatorics, Probability and Computing 31 (2022), no. 5, 907–923.