

An approximate version of Sidorenko's conjecture

David Conlon*

Jacob Fox[†]

Benny Sudakov[‡]

Abstract

A beautiful conjecture of Erdős-Simonovits and Sidorenko states that if H is a bipartite graph, then the random graph with edge density p has in expectation asymptotically the minimum number of copies of H over all graphs of the same order and edge density. This conjecture also has an equivalent analytic form and has connections to a broad range of topics, such as matrix theory, Markov chains, graph limits, and quasirandomness. Here we prove the conjecture if H has a vertex complete to the other part, and deduce an approximate version of the conjecture for all H . Furthermore, for a large class of bipartite graphs, we prove a stronger stability result which answers a question of Chung, Graham, and Wilson on quasirandomness for these graphs.

1 Introduction

A fundamental problem in extremal graph theory (see [5] and its references) is to determine or estimate the minimum number of copies of a graph H which must be contained in another graph G of a certain order and size. The special case where one wishes to determine the minimum number of edges in a graph on N vertices which guarantee a single copy of H has received particular attention. The case where H is a triangle was solved by Mantel more than a century ago. This was generalized to cliques by Turán and the Erdős-Stone-Simonovits theorem determines the answer asymptotically if H is not bipartite. For bipartite graphs H , a classical result of Kővári, Sós, and Turán implies that $O(N^{2-\epsilon_H})$ edges are sufficient for some $\epsilon_H > 0$, but, despite much effort by researchers, the asymptotics, and even good estimates for the largest possible ϵ_H , are understood for relatively few bipartite graphs.

The general problem can be naturally stated in terms of subgraph densities. The *edge density* of a graph G with N vertices and M edges is $M/\binom{N}{2}$. More generally, the *H -density* of a graph G is the fraction of all one-to-one mappings from the vertices of H to the vertices of G which map edges of H to edges of G . The general extremal problem asks for the minimum possible H -density over all graphs on N vertices with edge density p . For fixed H , the asymptotic answer as $N \rightarrow \infty$ is a function of p . Determining this function is a classical problem and notoriously difficult even in the case where H is the complete graph of order r . Early results in this case were obtained by Erdős, Goodman, Lovász, Simonovits, Bollobás, and Fisher. Recently, Razborov [28] using flag algebras and Nikiforov [26] using

*St John's College, Cambridge CB2 1TP, United Kingdom. E-mail: D.Conlon@dpmms.cam.ac.uk. Research supported by a Junior Research Fellowship at St John's College.

[†]Department of Mathematics, Princeton University, Princeton, NJ 08544. Email: jacobfox@math.princeton.edu. Research supported by an NSF Graduate Research Fellowship and a Princeton Centennial Fellowship.

[‡]Department of Mathematics, UCLA, Los Angeles, CA 90095. Email: bsudakov@math.ucla.edu. Research supported in part by NSF CAREER award DMS-0812005 and by a USA-Israeli BSF grant.

a combination of combinatorial and analytic arguments gave an asymptotic answer in the cases $r = 3$ and $r = 4$, respectively.

There is a simple upper bound on the minimum H -density in terms of the edge density. Suppose that H has m edges. By taking G to be a random graph with edge density p , it is easy to see that the minimum possible H -density is at most p^m . The beautiful conjectures of Erdős and Simonovits [32] and Sidorenko [31] suggest that this bound is sharp for bipartite graphs. That is, for any bipartite H there is a $\gamma(H) > 0$ such that the number of copies of H in any graph G on N vertices with edge density $p > N^{-\gamma(H)}$ is asymptotically at least the same as in the N -vertex random graph with edge density p . This is known to be true in a few very special cases, e.g., for complete bipartite graphs, trees, even cycles (see [31]) and, recently, for cubes [20].

The original formulation of the conjecture by Sidorenko is in terms of graph homomorphisms. A *homomorphism* from a graph H to a graph G is a mapping $f : V(H) \rightarrow V(G)$ such that, for each edge (u, v) of H , $(f(u), f(v))$ is an edge of G . Let $h_H(G)$ denote the number of homomorphisms from H to G . We also consider the normalized function $t_H(G) = h_H(G)/|G|^{|H|}$, which is the fraction of mappings $f : V(H) \rightarrow V(G)$ which are homomorphisms.

Conjecture 1 (*Sidorenko*) *For every bipartite graph H with m edges and every graph G ,*

$$t_H(G) \geq t_{K_2}(G)^m.$$

Sidorenko's conjecture also has the following nice analytical form. Let μ be the Lebesgue measure on $[0, 1]$ and let $h(x, y)$ be a bounded, non-negative, symmetric and measurable function on $[0, 1]^2$. Let H be a bipartite graph with vertices u_1, \dots, u_t in the first part and vertices v_1, \dots, v_s in the second part. Denote by E the set of edges of H , i.e., all the pairs (i, j) such that u_i and v_j are adjacent, and let $m = |E|$. The analytic formulation of Sidorenko's conjecture states that

$$\int \prod_{(i,j) \in E} h(x_i, y_j) d\mu^{s+t} \geq \left(\int h d\mu^2 \right)^m. \quad (1)$$

The expression on the left hand side of this inequality is quite common. For example, Feynman integrals in quantum field theory, Mayer integrals in statistical mechanics, and multicenter integrals in quantum chemistry are of this form (see Section 6 of [30] and its references). Unsurprisingly then, Sidorenko's conjecture has connections to a broad range of topics, such as matrix theory [2, 4], Markov chains [3, 27], graph limits [24], and quasirandomness.

The study of quasirandom graphs was introduced by Thomason [35] and Chung, Graham, and Wilson [10]. They showed that a large number of interesting graph properties satisfied by random graphs are all equivalent. This idea has been quite influential, leading to the study of quasirandomness in other structures such as hypergraphs [8, 18], groups [19], tournaments, permutations, sequences and sparse graphs (see [9] and its references), and progress on problems in different areas (see, e.g., [11, 18, 19]). It is closely related to Szemerédi's regularity lemma and its recent hypergraph generalization and all proofs of Szemerédi's theorem on long arithmetic progressions in dense subsets of the integers use some notion of quasirandomness. Finally, there is also the fast-growing study of properties of quasirandom graphs, which has recently attracted lots of attention both in combinatorics and theoretical computer science (see, e.g., [23]).

A sequence $(G_n : n = 1, 2, \dots)$ of graphs is called *quasirandom with density p* (where $0 < p < 1$) if, for every graph H ,

$$t_H(G_n) = (1 + o(1))p^{|E(H)|}. \quad (2)$$

Note that (2) is equivalent to saying that the H -density of G_n is $(1 + o(1))p^{|E(H)|}$, since the proportion of mappings from $V(H)$ to $V(G_n)$ which are not one-to-one tends to 0 as $|V(G_n)| \rightarrow \infty$. This property is equivalent to many other properties shared by random graphs. One such property is that the edge density between any two vertex subsets of G_n of linear cardinality is $(1 + o(1))p$. A surprising fact, proved in [10], is that it is enough that (2) holds for $H = K_2$ and $H = C_4$ for a graph to be quasirandom. That is, a graph with edge density p is quasirandom with density p if the C_4 -density is approximately p^4 . A question of Chung, Graham, and Wilson [10] which has received considerable attention (see, e.g., [6]) asks for which graphs H is it true that if (2) holds for K_2 and H , then the sequence is quasi-random with density p . Such a graph H is called *p -forcing*. We call H *forcing* if it is p -forcing for all p . Chung, Graham, and Wilson prove that even cycles C_{2t} and complete bipartite graphs $K_{2,t}$ with $t \geq 2$ are forcing. Skokan and Thoma [33] generalize this result to all complete bipartite graphs $K_{a,b}$ with $a, b \geq 2$.

There are two simple obstacles to a graph being forcing. It is easy to show that a forcing graph must be bipartite. Further, for any forest H , (2) is satisfied for any sequence of nearly regular graphs of edge density tending to p . The property of being nearly regular is not as strong as being quasirandom. Hence, a forcing graph must be bipartite and have at least one cycle. Skokan and Thoma [33] ask whether these properties characterize the forcing graphs. We conjecture the answer is yes and refer to it as the *forcing conjecture*.

Conjecture 2 *A graph H is forcing if and only if it is bipartite and contains a cycle.*

It is not hard to see that the forcing conjecture is stronger than Sidorenko's conjecture, and it further gives a stability result for Sidorenko's conjecture. A *stability* result not only characterizes the extremal graphs for an extremal problem, but also shows that if a graph is close to being optimal for the extremal problem, then it is close in a certain appropriate metric to an extremal graph. In recent years, there has been a great amount of research done toward proving stability results in extremal combinatorics. The forcing conjecture implies that if H is bipartite with m edges and contains a cycle, then G satisfies $t_H(G)$ is close to $t_{K_2}(G)^m$ if and only if it is quasirandom with density $t_{K_2}(G)$.

As consequences of the following theorem, we prove Sidorenko's conjecture and the forcing conjecture for a large class of bipartite graphs.

Theorem 1.1 *Let H be a bipartite graph with m edges which has $r \geq 1$ vertices in the first part complete to the second part, and the minimum degree in the first part is at least d . Then*

$$t_H(G) \geq t_{K_{r,d}}(G)^{\frac{m}{rd}}.$$

From Theorem 1.1, by taking $r = 1$ and $d = 1$, we have the following corollary.

Theorem 1.2 *Sidorenko's conjecture holds for every bipartite graph H which has a vertex complete to the other part.*

From Theorem 1.2, we may easily deduce an approximate version of Sidorenko's conjecture for all graphs. For a connected bipartite graph H with parts V_1, V_2 , define the bipartite graph \bar{H} with parts V_1, V_2 such that $(v_1, v_2) \in V_1 \times V_2$ is an edge of \bar{H} if and only if it is not an edge of H . Define the *width* of H to be the minimum degree of \bar{H} . If H is not connected, the width of H is the sum of the widths of the connected components of H . Note that the width of a connected bipartite graph is 0 if and only if it has a vertex that is complete to the other part. Also, the width of a bipartite graph with n vertices is at most $n/2$.

Corollary 1.1 *If H is a bipartite graph with m edges and width w , then $t_H(G) \geq t_{K_2}(G)^{m+w}$ holds for every graph G .*

We also obtain the following result from Theorem 1.1.

Theorem 1.3 *The forcing conjecture holds for every bipartite graph H which has two vertices in one part complete to the other part, which has at least two vertices.*

Sidorenko further conjectured that the assumption that h is symmetric in (1) can be dropped. This has the following equivalent discrete version. For bipartite graphs $H = (V_1, V_2, E)$ and $G = (U_1, U_2, F)$ where H has $m = |E|$ edges and G has edge density $p = \frac{|F|}{|U_1||U_2|}$ between its parts, the density of mappings $f : V(H) \rightarrow V(G)$ with $f(V_i) \subset U_i$ for $i = 1, 2$ that are homomorphisms is at least p^m . It is not hard to check that the proofs of Theorems 1.1, 1.2, and 1.3 and Corollary 1.1 can be extended to prove stronger asymmetric versions. We leave the details to the interested reader.

Although several authors (e.g., Sidorenko [30] and Lovász [24]) suggested that one might need an analytic approach to attack Conjecture 1, our proof of Theorem 1.1 given in the next section uses simple combinatorial tools. We conclude with a discussion of some related problems.

2 Proof of Theorem 1.1

We begin with some simple observations. For a vertex v in a graph G , the *neighborhood* $N(v)$ is the set of vertices adjacent to v . For a sequence S of vertices of a graph G , the *common neighborhood* $N(S)$ is the set of vertices adjacent to every vertex in S . The identity

$$h_{K_{a,b}}(G) = \sum_T |N(T)|^a, \tag{3}$$

where the sum is over all sequences T of b vertices of G , follows by counting homomorphisms of $K_{a,b}$ which fix the second part. Here we allow sequences of vertices to include repeated vertices.

The previous approaches toward proving Sidorenko's conjecture mainly used clever applications of Hölder's inequality. We propose a new approach using a probabilistic technique known as dependent random choice (for more details, see, e.g., the survey [16]). The first attempt to use this technique to estimate subgraph densities was made in [15]. Roughly, the idea is that most small subsets of the neighborhood of a random subset of vertices have large common neighborhood. Our proof uses an equivalent counting version.

Before going into the details of the proof of Theorem 1.1, we first give a brief outline of the proof idea in the case $r = 1$ and $d = 1$. Suppose u is a vertex in the bipartite graph $H = (V_1, V_2, E)$ on n vertices which is complete to V_2 , and G is a graph on N vertices with edge density p . The bulk of the proof is geared toward proving a seemingly weaker result, Lemma 2.3, which shows that the bound in Theorem 1.1 is tight apart from a positive constant factor which only depends on $n = |H|$. To obtain this result, we use dependent random choice to show that (see Lemma 2.1) an average vertex v of G (weighted by its degree) has the property that almost all small subsets S of $N(v)$ satisfy $|N(S)| \geq c_n p^{|S|} N$, which, apart from the factor c_n , is the expected size of the common neighborhood of a subset of vertices of size $|S|$ in the random graph $G(N, p)$. We will give a lower bound on the number of homomorphisms $f : V(H) \rightarrow V(G)$ from H to G as follows. We first pick $f(u) = v$ so that almost all small subsets of vertices in $N(v)$ have large common neighborhood. Having picked $f(u) = v$, we then randomly pick a sequence of $|V_2|$ vertices from $N(v)$ to be $f(V_2)$. With large probability, for all subsets $S \subset f(V_2)$, we have $|N(S)| \geq c_n p^{|S|} N$. For any vertex $u' \in V_1 \setminus \{u\}$, we can pick $f(u')$ to be any vertex in the common neighborhood of $S = f(N(u')) \subset f(V_2)$. To summarize, we get a lower bound on the number of homomorphisms from H to G by first choosing the image of u , then the image of V_2 , and finally the image of the remaining vertices in V_1 . The homomorphism count is within a positive constant factor, depending only on n , of $p^m N^n$, which is asymptotically the expected homomorphism count in $G(N, p)$. We complete the proof of Theorem 1.1 by using a tensor power trick to get rid of the factor depending on n .

Lemma 2.1 *Let G be a graph with N vertices and d, n , and r be positive integers with $d \leq n$. Call a sequence S of k vertices of G rare if $|N(S)| \leq (2n)^{-2n} t_{K_{r,d}}(G)^{\frac{k}{rd}} N$. Call a sequence $T = (v_1, \dots, v_r)$ of r vertices bad with respect to k if the number of rare sequences of k vertices in $N(T)$ is at least $\frac{1}{2n} |N(T)|^k$. Call T good if, for all $d \leq k \leq n$, it is not bad with respect to k . Then the sum of $|N(T)|^d$ over all good sequences T is at least $h_{K_{r,d}}(G)/2$.*

Proof: We write $T \sim k$ to denote that T is bad with respect to k . Let X_k denote the number of pairs (T, S) with S a rare sequence of k vertices and T a sequence of r vertices which are adjacent to every vertex in S . For a given sequence S , the number of pairs (T, S) with T a sequence of r vertices adjacent to every vertex in S is $|N(S)|^r$. As the number of sequences S of k vertices is N^k , we have, by summing over rare S ,

$$X_k = \sum_{S \text{ rare}} |N(S)|^r \leq N^k \cdot \left((2n)^{-2n} t_{K_{r,d}}(G)^{\frac{k}{rd}} N \right)^r = (2n)^{-2nr} t_{K_{r,d}}(G)^{\frac{k}{d}} N^{k+r}. \quad (4)$$

Of course, X_k is at least the number of such pairs (T, S) with T having the additional property that it is bad with respect to k . Hence, by summing over such T , we have

$$X_k \geq \sum_{T, T \sim k} \frac{1}{2n} |N(T)|^k \geq \frac{1}{2n} N^r \left(\sum_{T, T \sim k} |N(T)|^d / N^r \right)^{k/d} = \frac{1}{2n} N^{r-rk/d} \left(\sum_{T, T \sim k} |N(T)|^d \right)^{k/d}. \quad (5)$$

The second inequality follows from convexity of the function $f(x) = x^{k/d}$ together with the fact that there are at most N^r sequences T . From (4) and (5) and simplifying, we get

$$\begin{aligned}
\sum_{T, T \sim k} |N(T)|^d &\leq \left(2nN^{\frac{rk}{d}-r}X_k\right)^{d/k} \leq \left(2nN^{\frac{rk}{d}-r}(2n)^{-2nr}t_{K_{r,d}}(G)^{\frac{k}{d}}N^{k+r}\right)^{d/k} \\
&= (2n)^{(1-2nr)d/k}t_{K_{r,d}}(G)N^{d+r} \leq \frac{1}{2n}t_{K_{r,d}}(G)N^{d+r} = \frac{1}{2n}h_{K_{r,d}}(G).
\end{aligned}$$

Hence, using (3), we have

$$\sum_{T \text{ good}} |N(T)|^d \geq \sum_T |N(T)|^d - \sum_{k=1}^n \sum_{T, T \sim k} |N(T)|^d \geq h_{K_{r,d}}(G) - n \cdot \frac{1}{2n}h_{K_{r,d}}(G) = h_{K_{r,d}}(G)/2.$$

□

The bound on $|N(S)|$ in the definition of a rare sequence S in the above lemma is quite natural. Indeed, in the case $G = G(N, p)$, $t_{K_{r,d}}(G) \approx p^{rd}$ and a sequence S of order k is rare if $|N(S)| \leq c_n t_{K_{r,d}}(G)^{\frac{k}{rd}} N \approx c_n p^k N$ with $c_n = (2n)^{-2n}$, which, apart from the constant factor c_n , is roughly the size of the common neighborhood of every subset of order k .

A *hypergraph* $\mathcal{H} = (V, E)$ consists of a set V of vertices and a set E of edges, which are subsets of V . A *strongly directed hypergraph* $\mathcal{H} = (V, E)$ consists of a set V of vertices and a set E of edges, which are sequences of vertices in V .

Lemma 2.2 *Suppose \mathcal{H} is a hypergraph with h vertices and at most e edges such that each edge has size at least d , and \mathcal{G} is a strongly directed hypergraph on N vertices with the property that for each k , $d \leq k \leq h$, the number of sequences of k vertices of \mathcal{G} that do not form an edge of \mathcal{G} is at most $\frac{1}{2e}N^k$. Then the number of homomorphisms from \mathcal{H} to \mathcal{G} is at least $\frac{1}{2}N^h$.*

Proof: Consider a random mapping from the vertices of \mathcal{H} to the vertices of \mathcal{G} . The probability that a given edge of \mathcal{H} does not map to an edge of \mathcal{G} is at most $\frac{1}{2e}$. By the union bound, the probability that there is an edge of \mathcal{H} that does not map to an edge of \mathcal{G} is at most $e \cdot \frac{1}{2e} = 1/2$. Hence, with probability at least $1/2$, a random mapping gives a homomorphism, so there are at least $\frac{1}{2}N^h$ homomorphisms from \mathcal{H} to \mathcal{G} . □

Lemma 2.3 *Let $H = (V_1, V_2, E)$ be a bipartite graph with n vertices and m edges such that there are r vertices $u_1, \dots, u_r \in V_1$ which are adjacent to all vertices in V_2 , and the minimum degree in V_1 is at least d . Then, for every graph G ,*

$$t_H(G) \geq (2n)^{-2n^2} t_{K_{r,d}}(G)^{\frac{m}{rd}}.$$

Proof: Let N denote the number of vertices of G . Let $n_i = |V_i|$ for $i \in \{1, 2\}$. We will give a lower bound on the number of homomorphisms $f : V(H) \rightarrow V(G)$ that map u_1, \dots, u_r to a good sequence $T = (v_1, \dots, v_r)$ of r vertices of G . Suppose we have already picked $f(u_i) = v_i$ for $1 \leq i \leq r$. Let \mathcal{H} be the hypergraph with vertex set V_2 , where $S \subset V_2$ is an edge of \mathcal{H} if there is a vertex $w \in V_1 \setminus \{u_1, \dots, u_r\}$

such that $N(w) = S$. The number of vertices of \mathcal{H} is n_2 , which is at most n , and the number of edges of \mathcal{H} is at most $n_1 - r$, which is at most n . Let \mathcal{G} be the strongly directed hypergraph on $N(T)$, where a sequence R of k vertices in $N(T)$ is an edge of \mathcal{G} if $|N(R)| \geq (2n)^{-2n} t_{K_{r,d}}(G) \frac{k}{r^d} N$. Since T is good, for each k , $d \leq k \leq n$, the number of sequences of k vertices of \mathcal{G} that are not edges of \mathcal{G} is at most $\frac{1}{2n} |N(T)|^k$. Hence, by Lemma 2.2, there are at least $\frac{1}{2} |N(T)|^{n_2}$ homomorphisms g from \mathcal{H} to \mathcal{G} . Pick one such homomorphism g , and let $f(x) = g(x)$ for $x \in V_2$. By construction, once we have picked T and $f(V_2)$, there are at least $(2n)^{-2n} t_{K_{r,d}}(G) \frac{|N(w)|}{r^d} N$ possible choices for $f(w)$ for each vertex $w \in V_1 \setminus \{u_1, \dots, u_r\}$. Hence, the number of homomorphisms from H to G satisfies

$$\begin{aligned}
t_H(G)N^n &\geq \sum_{T \text{ good}} \frac{1}{2} |N(T)|^{n_2} \prod_{w \in V_1 \setminus \{u_1, \dots, u_r\}} (2n)^{-2n} t_{K_{r,d}}(G) \frac{|N(w)|}{r^d} N \\
&= \frac{1}{2} (2n)^{-2n(n_1-r)} t_{K_{r,d}}(G) \frac{m-rn_2}{r^d} N^{n_1-r} \sum_{T \text{ good}} |N(T)|^{n_2} \\
&\geq \frac{1}{2} (2n)^{-2n(n_1-r)} t_{K_{r,d}}(G) \frac{m-rn_2}{r^d} N^{n_1-r} N^r \left(\sum_{T \text{ good}} |N(T)|^d / N^r \right)^{n_2/d} \\
&= \frac{1}{2} (2n)^{-2n(n_1-r)} t_{K_{r,d}}(G) \frac{m-rn_2}{r^d} N^{n_1-rn_2/d} \left(\sum_{T \text{ good}} |N(T)|^d \right)^{n_2/d} \\
&\geq \frac{1}{2} (2n)^{-2n(n_1-r)} t_{K_{r,d}}(G) \frac{m-rn_2}{r^d} N^{n_1-rn_2/d} (h_{K_{r,d}}(G)/2)^{n_2/d} \\
&= 2^{-1-n_2/d} (2n)^{-2n(n_1-r)} t_{K_{r,d}}(G) \frac{m}{r^d} N^{n_1+n_2} \\
&\geq (2n)^{-2n^2} t_{K_{r,d}}(G) \frac{m}{r^d} N^n.
\end{aligned}$$

In the first equality, we use $\sum_{w \in V_1 \setminus \{u_1, \dots, u_r\}} |N(w)| = m - rn_2$, which follows from the fact that the vertices u_1, \dots, u_r each have degree n_2 . The second inequality uses the convexity of the function $q(x) = x^{n_2/d}$ together with the fact that there are at most N^r sequences T . The third inequality follows by Lemma 2.1, and the third equality follows from substituting $h_{K_{r,d}}(G) = t_{K_{r,d}}(G)N^{d+r}$. Dividing by N^n , we get the desired inequality. \square

We next complete the proof of Theorem 1.1 by improving the inequality in the previous lemma using a tensor power trick. This technique was used by Alon and Ruzsa [1] to give an elementary proof of Sidorenko's conjecture for trees, which implies the Blakley-Roy matrix inequality [4]. This technique has also been used in many other areas, and Tao [34] has collected a number of these applications. The tensor product $F \times G$ of two graphs F and G has vertex set $V(F) \times V(G)$ and any two vertices (u_1, u_2) and (v_1, v_2) are adjacent in $F \times G$ if and only if u_i is adjacent with v_i for $i \in \{1, 2\}$. Let $G^1 = G$ and $G^s = G^{s-1} \times G$. Note that $t_H(F \times G) = t_H(F) \times t_H(G)$ for all F, G, H .

Proof of Theorem 1.1: Suppose, for contradiction, that there is a graph G such that $t_H(G) < t_{K_{r,d}}(G) \frac{m}{r^d}$. Let $c = t_H(G) / t_{K_{r,d}}(G) \frac{m}{r^d} < 1$. Let s be such that $c^s < (2n)^{-2n^2}$. Then

$$t_H(G^s) = t_H(G)^s = c^s t_{K_{r,d}}(G) \frac{ms}{r^d} = c^s t_{K_{r,d}}(G^s) \frac{m}{r^d} < (2n)^{-2n^2} t_{K_{r,d}}(G^s) \frac{m}{r^d}.$$

However, this contradicts Lemma 2.3 applied to H and G^s . This completes the proof. \square

Proof of Corollary 1.1: First assume H is connected. Let H' be obtained from H by making the vertex u of minimum degree in \bar{H} complete to the other part. Note that H is a subgraph of H' and H' has exactly w more edges than H . Hence, by Theorem 1.2, $t_H(G) \geq t_{H'}(G) \geq t_{K_2}(G)^{m+w}$. If H is not connected, letting H_1, \dots, H_r denote the connected components of H , we have $t_H(G) = \prod_{i=1}^r t_{H_i}(G)$, and the corollary easily follows from the case where H is connected. \square

Proof of Theorem 1.3: Let $H = (V_1, V_2, E)$ be a bipartite graph with n vertices, m edges such that there are 2 vertices in V_1 which are adjacent to all vertices in V_2 and $|V_2| \geq 2$. We may suppose H has no isolated vertices as they do not affect homomorphism density counts. Fix $0 < p < 1$. Suppose G has N vertices and $(1 + o(1))pN^2/2$ edges so that $t_{K_2}(G) = (1 + o(1))p$, and $t_H(G) = (1 + o(1))p^m$. We first prove the case of Theorem 1.3 where the minimum degree of H is at least 2. By Theorem 1.1 with $r = d = 2$, we have $t_H(G) \geq t_{K_{2,2}}(G)^{m/4}$. From the above bounds on $t_H(G)$, we get $t_{K_{2,2}}(G) \leq (1 + o(1))p^4$. Also, since Sidorenko's conjecture holds for $K_{2,2}$, we have $t_{K_{2,2}}(G) \geq (1 + o(1))p^4$. Thus, $t_{K_{2,2}}(G) = (1 + o(1))p^4$. Since $K_{2,2}$ is forcing, this implies G is quasirandom with density p , and hence H is forcing.

Suppose now that H has $s \geq 1$ vertices of degree 1. Then all of these vertices are in V_1 . Let H' be the induced subgraph of H with $n - s$ vertices and $m - s$ edges obtained by deleting all degree 1 vertices. Note that H' has minimum degree at least 2 and also has at least two vertices in one part complete to the other part. By the above argument, H' is forcing. We can apply Theorem 1.1 with $r = 2$ and $d = 1$ to get $t_H(G) \geq t_{K_{1,2}}(G)^{m/2}$. From the bounds on $t_H(G)$, we get $t_{K_{1,2}}(G) \leq (1 + o(1))p^2$. This inequality implies the following claim:

Claim: All but $o(N)$ vertices of G have degree at least $(1 + o(1))pN$.

Proof: Suppose, instead, that there are ϵN vertices with degree less than $(1 - \epsilon)pN$. For each vertex v , let $\delta(v) = |N(v)| - pN$. Since $t_{K_2}(G) = (1 + o(1))p$, we have $\sum_v \delta(v) = o(pN^2)$. Moreover, since $\delta(v) \leq -\epsilon pN$ for ϵN vertices, $\sum_v \delta^2(v) \geq \epsilon^3 p^2 N^3$. Therefore,

$$\begin{aligned} h_{K_{1,2}}(G) &= \sum_v |N(v)|^2 = \sum_v (pN + \delta(v))^2 \\ &= p^2 N^3 + 2pN \sum_v \delta(v) + \sum_v \delta^2(v) \\ &\geq (1 + o(1) + \epsilon^3) p^2 N^3. \end{aligned}$$

Thus

$$t_{K_{1,2}}(G) = N^{-3} h_{K_{1,2}}(G) \geq (1 + o(1) + \epsilon^3) p^2,$$

which contradicts $t_{K_{1,2}}(G) \leq (1 + o(1))p^2$ and verifies the claim. \square

Any set X of $o(N)$ vertices of G , and in particular the $o(N)$ vertices of least degree, can only change the subgraph density count by $o(1)$ for any fixed subgraph. Indeed, suppose F has k vertices, then the number of mappings from $V(F)$ to $V(G)$ whose image $f(F)$ has nonempty intersection with X is $N^k - (N - |X|)^k = o(N^k)$, and hence the fraction of mappings whose image has nonempty intersection with X is $o(1)$. Every homomorphism from H' to G whose image does not intersect the $o(N)$ vertices of least degree can be extended to at least $((1 + o(1))pN)^s$ homomorphisms from H to G by picking the

images of the s vertices of H of degree 1. By normalizing, we get $t_H(G) \geq (1 + o(1))p^s t_{H'}(G) - o(1)$. Comparing the bounds on $t_H(G)$, we get $t_{H'}(G) \leq (1 + o(1))p^{m-s}$. Since H' satisfies Sidorenko's conjecture, $t_{H'}(G) \geq (1 + o(1))p^{m-s}$, and hence $t_{H'}(G) = (1 + o(1))p^{m-s}$. As H' is forcing, this implies G is quasirandom with density p , and hence H is also forcing. \square

3 Concluding Remarks

- The classical theorem of Ramsey states that every 2-edge-coloring of a sufficiently large complete graph K_N contains at least one monochromatic copy of a given graph H . Let $c_{H,N}$ denote the fraction of copies of H in K_N that must be monochromatic in any 2-edge-coloring. By an averaging argument, $c_{H,N}$ is a monotone increasing function in N , and therefore has a limit c_H as $N \rightarrow \infty$. The constant c_H is known as the *Ramsey multiplicity constant* for the graph H . For H with m edges, the uniform random 2-edge-coloring of K_N shows that $c_H \leq 2^{1-m}$. Erdős [13] and, in a more general form, Burr and Rosta [7] conjectured that this bound is tight. However, Thomason [36] proved that these conjectures are already false for K_4 . In fact, it is false for almost all graphs [22], and there are graphs H with m edges for which $c_H \leq m^{-\frac{m}{2} + o(m)}$, showing that the bound can be far from tight [14]. The situation for bipartite graphs is very different as Sidorenko's conjecture implies the Ramsey multiplicity conjecture holds for bipartite graphs. Thus, for a large class of bipartite graphs, this conjecture is implied by Theorem 1.2 and further a stability result follows from Theorem 1.3. For all bipartite graphs, an approximate version of this conjecture follows from Corollary 1.1.

Despite the fact that the conjectures of Erdős and Burr-Rosta are false for K_4 , Franek and Rödl [17] proved that a local version is true in this case. They showed that a small perturbation of a quasirandom coloring will not give a counterexample to these conjectures. In terms of graphons, in which we consider the limit of the graphs of one color for a sequence of colorings, the uniform graphon, which is the limit of a sequence of quasirandom graphs with density $1/2$, is a local minimum for the density of monochromatic copies of the graph K_4 , but not always the global minimum. Very recently, Lovász [25] proved the analogous local version of Sidorenko's conjecture.

- It is tempting to conjecture that the products of integrals in the left hand side of (1) are always non-negatively correlated. Equivalently, if H is a bipartite graph and H_1, \dots, H_r are subgraphs which edge-partition H , then for every graph G we have

$$t_H(G) \geq \prod_{i=1}^r t_{H_i}(G). \tag{6}$$

Sidorenko's conjecture is the case $H_i = K_2$ for each i , and the forcing conjecture would follow by taking H_1 to be a cycle in H and all other $H_i = K_2$. Unfortunately, (6) does not hold in general. As noted in [30, 31], a counterexample to this inequality with $H = P_3$, $H_1 = P_2$, and $H_2 = P_1$, where P_i denotes the path with i edges, was found in 1966 by London (see [30]).

- Let $\|h\|_H$ be the integral on the left hand side of (1) raised to the power $1/|E(H)|$. Lovász asked for which graphs H is $\|\cdot\|_H$ a norm. When H is an even cycle, for example, they are the classical Schatten-von Neumann norms. Hatami [20] showed that if $\|\cdot\|_H$ is a norm then H has the following

property. For all subgraphs H' of H and all h ,

$$\|h\|_H \geq \|h\|_{H'}. \quad (7)$$

Therefore, if $\|\cdot\|_H$ is a norm, then, taking $H' = K_2$, we see that Sidorenko's conjecture holds for H , and if H also has a cycle, then, taking H' to be that cycle, we see that H is forcing. Property (7) implies that the non-negative correlation inequality (6) holds for H . Hence, property (7) does not hold for $H = P_3$. Hatami proved that if $\|\cdot\|_H$ is a norm, then the degrees of any two vertices in the same part of H are equal. This shows, in combination with Theorem 1.2, that the class of graphs for which $\|\cdot\|_H$ is a norm is considerably smaller than that for which Sidorenko's conjecture holds. In the positive direction, Hatami proves that if H is a cube or a complete bipartite graph, then $\|\cdot\|_H$ is a norm.

This discussion leads to the following general question: for which pairs of graphs H, H' do we have $\|\cdot\|_H \geq \|\cdot\|_{H'}$. Sidorenko's conjecture is that if $H' = K_2$, then this inequality holds if and only if H is bipartite. Theorem 1.1 shows that this inequality holds if $H' = K_{r,d}$ and H has r vertices in the first part complete to the second part and the minimum degree in the first part is at least d . The case where H and H' are trees of the same size is studied in [30]. From the fact the inequality holds if H is a complete bipartite graph and $H' = K_2$, one may naturally guess that the inequality holds if H is a complete tripartite graph and $H' = K_3$, but a counterexample is given in [29]. It may be the case that the general problem is undecidable. There are other similar questions involving counting graph homomorphism densities which are undecidable by a reduction to Hilbert's 10th Problem (see, e.g., [21]).

- Call a family \mathcal{F} of graphs p -forcing if, whenever $t_F(G_n) = (1+o(1))p^{|E(F)|}$ for all $F \in \mathcal{F}$, the sequence $(G_n : n = 1, 2, \dots)$ is p -quasirandom. If \mathcal{F} is p -forcing for all p , we simply say that the family \mathcal{F} is forcing. Note that the statement that a graph H is forcing is equivalent to the family $\mathcal{F} = \{K_2, H\}$ being forcing. The first examples of forcing families not involving any bipartite graphs were given in [12]. For example, let L_3 be the line graph of the cube, that is, it has 12 vertices corresponding to the edges of the cube and two vertices are connected if and only if the corresponding edges meet. Then the pair consisting of L_3 and the triangle K_3 is forcing. It would be interesting to extend this result and determine what other graphs may be coupled with the triangle to give a forcing pair.
- We conclude by mentioning that counterexamples are given in [31] to the natural generalization of inequality (1) to functions of more than two variables. That is, the hypergraph analogue of Sidorenko's conjecture is false.

References

- [1] N. Alon and I. Z. Ruzsa, Non-averaging subsets and non-vanishing transversals, *J. Combin. Theory Ser. A* **86** (1999), 1–13.
- [2] F. V Atkinson, G. A. Watterson, P. A. D. Moran, A matrix inequality, *Quart. J. Math. Oxford II* **11** (1960), 137–140.

- [3] I. Benjamini and Y. Peres, A correlation inequality for tree-indexed Markov chains, in *Seminar on Stochastic Processes*, Proc. Semin., Los Angeles/CA (USA) 1991, Prog. Probab. 29, 1992, 7–14.
- [4] G.R. Blakley and P.A. Roy, A Hölder type inequality for symmetric matrices with nonnegative entries, *Proc. Amer. Math. Soc.* **16** (1965) 1244–1245.
- [5] B. Bollobás, **Extremal Graph Theory** Academic Press, London, 1978.
- [6] B. Bollobás, **Random graphs** (2nd edition), Cambridge Studies in Advanced Mathematics 73, Cambridge Univ. Press, 2001.
- [7] S. A. Burr and V. Rosta, On the Ramsey multiplicity of graphs - problems and recent results, *J. Graph Theory* **4** (1980), 347–361.
- [8] F. R. K. Chung and R. L. Graham, Quasi-random set systems, *J. Amer. Math. Soc.* **4** (1991), 151–196.
- [9] F. R. K. Chung and R. L. Graham, Quasi-random graphs with given degree sequences, *Random Structures Algorithms* **12** (2008), 1–19.
- [10] F. R. K. Chung, R. L. Graham, R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362.
- [11] D. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. of Math.* **170** (2009), 941–960.
- [12] D. Conlon, H. Hàn, Y. Person and M. Schacht, Weak quasi-randomness for uniform hypergraphs, submitted.
- [13] P. Erdős, On the number of complete subgraphs contained in certain graphs, *Publ. Math. Inst. Hung. Acad. Sci.*, VII, ser. A, **3** (1962), 459–462.
- [14] J. Fox, There exist graphs with super-exponential Ramsey multiplicity constant, *J. Graph Theory* **57** (2008), 89–98.
- [15] J. Fox and B. Sudakov, Density theorems for bipartite graphs and related Ramsey-type results, *Combinatorica* **29** (2009), 153–196.
- [16] J. Fox and B. Sudakov, Dependent random choice, *Random Structures Algorithms*, to appear.
- [17] F. Franek and V. Rödl, Ramsey problem on multiplicities of complete subgraphs in nearly quasirandom graphs, *Graphs Combin.* **8** (1992), 299–308.
- [18] W. T. Gowers, Hypergraph Regularity and the Multidimensional Szemerédi Theorem, *Ann. of Math.* **166** (2007) 897–946.
- [19] W. T. Gowers, Quasirandom groups, *Combin. Probab. Comput.* **17** (2008), 363–387.
- [20] H. Hatami, Graph norms and Sidorenko’s conjecture, *Israel J. Math.*, to appear.
- [21] H. Hatami and S. Norine, A solution to Lovász’s Seventeenth problem, preprint.

- [22] C. Jagger, P. Štoviček, and A. Thomason, Multiplicities of subgraphs, *Combinatorica* **16** (1996), 123–141.
- [23] M. Krivelevich and B. Sudakov, Pseudorandom graphs, in *More Sets, Graphs and Numbers*, Bolyai Society Mathematical Studies 15, Springer, 2006, 199–262.
- [24] L. Lovász, Very large graphs, *Current Developments in Mathematics Volume 2008* (2009), 67–128.
- [25] L. Lovász, Subgraph densities in signed graphons and the local Sidorenko conjecture, preprint.
- [26] V. Nikiforov, The number of cliques in graphs of given order and size, *Transactions of AMS*, to appear.
- [27] R. Pemantle and Y. Peres, Domination between trees and application to an explosion problem, *Ann. Probab.* **22** (1994), 180–194.
- [28] A. Razborov, On the minimal density of triangles in graphs, *Combin. Probab. Comput.* **17** (2008), 603–618.
- [29] A. Shapira and R. Yuster, On the density of a graph and its blowup, submitted.
- [30] A. F. Sidorenko, An analytic approach to extremal problems for graphs and hypergraphs, in *Extremal problems for finite sets* (Visegrád, 1991), 423–455, Bolyai Soc. Math. Stud., 3, János Bolyai Math. Soc., Budapest, 1994.
- [31] A. F. Sidorenko, A correlation inequality for bipartite graphs, *Graphs Combin.* **9** (1993), 201–204.
- [32] M. Simonovits, Extremal graph problems, degenerate extremal problems and super-saturated graphs, in *Progress in graph theory* (Waterloo, Ont., 1982), Academic Press, Toronto, ON, 1984, 419–437.
- [33] J. Skokan and L. Thoma, Bipartite subgraphs and quasi-randomness, *Graphs Combin.* **20** (2004), 255–262.
- [34] T. Tao, **Structure and randomness: pages from year one of a mathematical blog**, AMS, 2008.
- [35] A. G. Thomason, Pseudorandom graphs, in *Random graphs '85* (Poznań, 1985), North-Holland Math. Stud., vol. 144, North-Holland, Amsterdam, 1987, 307–331.
- [36] A. G. Thomason, A disproof of a conjecture of Erdős in Ramsey theory, *J. London Math. Soc.* **39** (1989), 246–255.