COMPOSITIO MATHEMATICA

Common graphs with arbitrary chromatic number

Daniel Kráľ, Jan Volec and Fan Wei



Compositio Math. **161** (2025), 594–634.

doi: 10.1112/S0010437X24007681

FOUNDATION
COMPOSITIO
MATHEMATICA



LONDON
MATHEMATICAL
SOCIETY
EST. 1865



https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Compositio Math. 161 (2025) 594–634
doi:10.1112/S0010437X24007681

Common graphs with arbitrary chromatic number

Daniel Kráľ, Jan Volec and Fan Wei

Abstract

Ramsey's theorem guarantees for every graph H that any 2-edge-coloring of a sufficiently large complete graph contains a monochromatic copy of H. In 1962, Erdős conjectured that the random 2-edge-coloring minimizes the number of monochromatic copies of K_k, and the conjecture was extended by Burr and Rosta to all graphs. In the late 1980s, the conjectures were disproved by Thomason and Sidorenko, respectively. A classification of graphs whose number of monochromatic copies is minimized by the random 2-edge-coloring, which are referred to as common graphs, remains a challenging open problem. If Sidorenko's conjecture, one of the most significant open problems in extremal graph theory, is true, then every 2-chromatic graph is common and, in fact, no 2-chromatic common graph unsettled for Sidorenko's conjecture is known. While examples of 3-chromatic common graphs were known for a long time, the existence of a 4-chromatic common graph was open until 2012, and no common graph with a larger chromatic number is known.

We construct connected k-chromatic common graphs for every k. This answers a question posed by Hatami *et al.* [*Non-three-colourable common graphs exist*, Combin. Probab. Comput. **21** (2012), 734–742], and a problem listed by Conlon *et al.* [*Recent developments in graph Ramsey theory*, in *Surveys in combinatorics 2015*, London Mathematical Society Lecture Note Series, vol. 424 (Cambridge University Press, Cambridge, 2015), 49–118, Problem 2.28]. This also answers in a stronger form the question raised by Jagger *et al.* [*Multiplicities of subgraphs*, Combinatorica **16** (1996), 123–131] whether there exists a common graph with chromatic number at least four.

Contents

1 Introduction **595**
1.1 Proof motivation . . . 597
1.2 Proof overview . . . 598
2 Preliminaries **600**
2.1 Graph limits . . . 600
2.2 Rooted graphs . . . 602
2.3 Spectral properties of graphons and kernels . . . 603
2.4 Estimates on densities . . . 604
2.5 Entropy-based estimates . . . 606

Received 18 August 2022, accepted in final form 30 September 2024, published online 3 July 2025.
2020 Mathematics Subject Classification 05D10, 05D40, 05C15, 05C35 (primary).
Keywords: Ramsey theory; graph limits; chromatic number; spectral method; probabilistic method.
© The Author(s), 2025. The publishing rights in this article are licensed to Foundation Compositio Mathematica under an exclusive licence.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

3 Locally Sidorenko component **608**
4 Local regime **618**
5 Non-local regime **626**
6 Main result **627**
7 Conclusion **630**
Appendix A **631**
Acknowledgements **631**
References **632**

1. Introduction

Ramsey's theorem [Ram30] led to a large body of results on the existence of well-behaved substructures in large structures, known as Ramsey theory, with links to many other areas of mathematics; see, e.g., [GRS13, CFS15]. Prominent examples of such links are Hindman's theorem [Hin74] and Szemerédi's theorem [Sze75] in number theory. In one of its simplest forms, the Ramsey's theorem asserts that for every graph H, there exists an integer N such that any 2-edge-coloring of the complete graph with N vertices contains a monochromatic copy of H. Determining the smallest such N, known as the Ramsey number $r(H)$ of a graph H, is a famous open problem, even in the supposedly simplest case when H is a complete graph. The best lower bound on the Ramsey number of a complete graph is by a random construction, which pioneered the development of the probabilistic method in combinatorics [AS16]; however, whether this random construction is (asymptotically) optimal is still widely open despite recent major progress [Con09, Sah23] and, in particular, the recent major breakthrough [CGMS23] on the upper bound. This problem is an example of a classical theme in extremal combinatorics: when is a random graph construction (close to) optimal? In this paper, we are concerned with the quantitative version of this problem, known as Ramsey multiplicity, which asks how many monochromatic copies of a graph H necessarily exist in any 2-edge-coloring of the complete graph with N vertices, and when the bound coming from the random construction is optimal.

Goodman's theorem [Goo59] implies that the number of monochromatic copies of the triangle K_3 is asymptotically minimized by the random 2-edge-coloring, i.e. when each edge of a complete graph is colored randomly with one of two colors with probability $1/2$. We say that a graph H is *common* if the number of monochromatic copies of H is asymptotically minimized by the random 2-edge-coloring of a complete graph. In particular, K_3 is common and, more generally, every cycle is common [Sid89].

In 1962, Erdős [Erd62] conjectured that every complete graph is common, and later Burr and Rosta [BR80] conjectured that every graph is common. Both conjectures turned out to be false: in the late 1980s, Sidorenko [Sid86, Sid89] showed that a triangle with a pendant edge is not common and Thomason [Tho89] showed that K_4 is not common. More generally, any graph containing K_4 is not common [JST96] (and so almost every graph is not common), and there exist graphs H and 2-edge-colorings with the number of monochromatic copies of H sublinear in the number of monochromatic copies of H in the random 2-edge-coloring [Cla92, Fox08].

A characterization of the class of common graphs is an intriguing open problem and there is even no conjectured description of the class. This problem is closely related to the famous conjecture of Sidorenko [Sid93] and of Erdős and Simonovits [ES83], which asserts that every bipartite graph H has the Sidorenko property, i.e. the number of copies of H in any graph is asymptotically at least the number of its copies in the random graph with the same density. Since every graph H with the Sidorenko property is common, as the number of copies of H in each color class is at least the expected number of its copies in the random edge-coloring, the conjecture,

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

if true, would imply that all bipartite are common. Hence, families of bipartite graphs known to have the Sidorenko property [BR65, Sid89, Sid91, CFS10, CL17, CKLL18, CL21] provide examples of bipartite graphs that are common.

Common graphs that are not bipartite, i.e. their chromatic number is larger than two, are scarce. In particular, Jagger, Štovíček and Thomason [JST96] asked whether there exists a common graph with chromatic number at least four. While odd cycles [Sid89] and even wheels [JST96, Sid96] are examples of 3-chromatic common graphs, also see [GLLV22], the existence of a common graph with chromatic number at least four was open until 2012 when the 5-wheel was shown to be common [HHK$^+$12] as one of the first application of the flag algebra method of Razborov [Raz07]. The question whether there exist common graphs with arbitrarily large chromatic number has been reiterated in [HHK$^+$12], and also by Conlon, Fox and Sudakov in the survey paper 'Recent developments in graph Ramsey theory', which is now a classical reference for results and directions related to Ramsey theory.

Problem (Conlon *et al.* [CFS15, Problem 2.28]). Do there exist common graphs of all chromatic numbers?

In this paper, we solve this problem by establishing the following.

Theorem 1.1. *For every $\ell \in \mathbb{N}$, there exists a connected common graph with chromatic number ℓ.*

We treat common graphs using methods from the theory of graph limits while employing spectral tools from the operator theory; graph limit methods permit lower-order terms to be avoided when analyzing large graphs, which makes our arguments simpler to present and in some cases also more natural. To prove Theorem 1.1, we show that every graph of sufficiently large girth can be embedded in a graph H such that any 2-edge-coloring of a complete graph has asymptotically at least as many monochromatic copies of H as the random 2-edge-coloring. The proof of Theorem 1.1 is split into two cases based on whether the considered 2-edge-coloring is close to the random coloring or not; we refer to the two cases as the local regime and the non-local regime (we provide an overview of the proof in § 1.2). The core of the proof is formed by the arguments related to the local regime, which is described by the existence of a dominant eigenvalue of the operator associated with the 2-edge-coloring. While both the Sidorenko property and common graphs have been studied in the local regime [Lov11, FW17, CHL23, HKKV23], the proof of Theorem 1.1 requires new spectral arguments to control (in)dependence of monochromatic embeddings of different parts of H in the host edge-colored complete graph. Our techniques could be used in other settings, e.g. the setting of a conjecture of Kohayakawa *et al.* [KNRS10] as discussed in § 7.

Our techniques also apply in the setting of k-common graphs introduced in [JST96]: for a positive integer k, a graph H is k-common if the random k-edge-coloring of a complete graph asymptotically minimizes the number of monochromatic copies of H among all k-edge-colorings. This notion provides another link to the Sidorenko property: a graph H has the Sidorenko property if and only if the graph H is k-common for all $k \geq 2$ (see [KNN$^+$22]). If H is k-common, then H is k'-common for all $k' = 2, \ldots, k$, and thus k-common graphs for $k \geq 3$ are even more rare than common graphs (as such graphs are necessarily also common). In fact, the question of Jagger *et al.* [JST96] about the existence of a non-bipartite k-common graph for $k \geq 3$ has been resolved only recently in [KNN$^+$22]. Using the techniques developed to prove Theorem 1.1, we also prove the following.

Theorem 1.2. *For every integer $k \geq 2$ and positive integer ℓ, there exists a connected k-common graph with chromatic number ℓ.*

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

1.1 Proof motivation

To motivate our proof, we sketch a possible strategy for proving the existence of a *disconnected* common graph with high chromatic number. We present the strategy in the language of graph limits, which is reviewed in § 2. Fix a graph H with girth four and chromatic number $k \geq 3$ and consider the disjoint union of the graph H and a complete bipartite graph, denoted as $H \cup K_{n,n}$, for a large n. Showing that $H \cup K_{n,n}$ is common amounts to showing that

$$t(H \cup K_{n,n}, W) + t(H \cup K_{n,n}, 1-W) \geq 2^{-e(H)-n^2+1} \tag{1}$$

for every graphon W, where $e(H)$ denotes the number of edges in H. Note that the right-hand side of (1) is equal to the left-hand side when W is the constant graphon equal $1/2$, i.e. W is the limit graphon for the sequence of the Erdős–Rényi random graphs $G_{m,1/2}$. Since the inequality (1) is symmetric in W and $1-W$, it is enough to establish (1) for graphons W with density at least $1/2$, i.e. when $\int_{[0,1]^2} W \geq 1/2$.

The graph $H \cup K_{n,n}$ is locally common in the following sense as shown by Fox and the last author in a forthcoming paper [FW17].

THEOREM 1.3 (Fox and Wei [FW17]). *If a graph G has even girth, then for any $p \in [0,1]$, any graphon W with density p such that $\|W-p\|_\square \leq p^2 2^{-48e(G)-2}$, and $\|W\|_\infty \leq 2p$ satisfies that $t(G,W) \geq p^{e(G)}$.*

Hence, if a graphon W is close (in the cut distance $\|\cdot\|_\square$) to the $1/2$-graphon, then Theorem 1.3 almost implies (1); the issue lies in the fact that if W has density $p > 1/2$, then the graphon $1-W$, whose density is $1-p$, may fail to satisfy $\|1-W\|_\infty \leq 2(1-p)$ needed to apply Theorem 1.3. The proof of Theorem 1.3 in [FW17] can be extended, at the expense of having a stronger bound on $\|W-p\|_\square$, to the setting with the condition $\|W\|_\infty \leq 2p$ weakened to $\|W\|_\infty \leq Cp$ for any fixed $C \in \mathbb{N}$; however, the bound on $\|W-p\|_\square$ would depend on H (in addition to C). Such an extension of Theorem 1.3 would be sufficient for the case of the disconnected graph $H \cup K_{n,n}$; however, the proofs of Theorems 6.1 and 6.2 in § 6 require a bound on $\|W-p\|_\square$ that is independent of the actual graph H for a specific family of graphs H. Indeed, the value of ε_0 in Theorem 3.1, which gives a variant of Theorem 1.3 tailored for our setting, does not depend on ℓ. Without removing the dependance on H (or at least proving a significantly better dependance than in Theorem 1.3), which is the crucial contribution of the arguments presented in § 3, it would not be possible to set up the parameters to prove Theorems 6.1 and 6.2 and, thus, Theorems 1.1 and 1.2.

We now continue a sketch of the argument that $H \cup K_{n,n}$ is common. As discussed, if W is close to the $1/2$-graphon, then the inequality (1) holds by a suitable extension of Theorem 1.3. If a graphon W is not close (in the cut distance) to the $1/2$-graphon, we can use that the graph $K_{n,n}$ has the Sidorenko property in the strong quantitative sense as given by Lemmas 2.1 and 2.4:

$$t(K_{n,n}, W) \geq \left(p^4 + \frac{\|W-p\|_\square^4}{8}\right)^{n^2/4}.$$

Hence, the quantity $t(K_{n,n}, W)$ is significantly larger than p^{n^2}, which can balance out a possible drop of the density of H in W and so implies that

$$t(H \cup K_{n,n}, W) \geq 2^{-e(H)-n^2+1}.$$

However, if a graphon W contains a sparse part, it may even happen that $t(H,W) = 0$, so $t(H \cup K_{n,n}, W) = 0$, and the argument fails (in fact, it can be shown that the presence of a

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

sparse part is essentially the only obstacle for the argument to proceed). However, if W contains a sparse part of measure $\alpha_0 > 0$ with density q close to 0, then this part has density $1-q$ in $1-W$ and its contribution to $t(H \cup K_{n,n}, 1-W)$ (as formalized in Lemma 5.1) is roughly $\alpha_0^{v(H)+2n}(1-q)^{e(H)+n^2}$; the value of α_0 depends on H only and so if n is sufficiently large, then

$$\alpha_0^{v(H)+2n}(1-q)^{e(H)+n^2} \geq 2^{-e(H)-n^2+1}.$$

Here, we use that the number of edges of $H \cup K_{n,n}$ grows superlinearly in the number of vertices of $H \cup K_{n,n}$ with n tending to infinity.

1.2 Proof overview

On a high level, the proofs of Theorems 6.1 and 6.2, which imply Theorems 1.1 and 1.2, respectively, follow the lines described in § 1.1; the main steps of the two proofs are visualized in Figure 1.

In the proof of Theorem 6.1, we show that a graph H_0 obtained by joining a suitable high-girth high-chromatic chromatic graph H by a suitably long path to a complete bipartite graph $K_{m,n}$ is common. The length ℓ of the path joining H and the complete bipartite graph needs to be carefully controlled. On the one hand, the path needs to be long enough to make copies of H and $K_{m,n}$ in a graphon W sufficiently independent in the sense that $t(H_0, W)$ is approximately $t(H, W)t(K_{m,n}, W)p^\ell$ for any graphon W with density p; note that $t(H \cup K_{m,n}, W) = t(H, W)t(K_{m,n}, W)$. At the same time, the number of edges of the graph H_0 needs to stay superlinear in the number of the vertices of H_0, i.e. the length ℓ needs to be $o(v(H) + mn)$, so that we are able to handle the case when the graphon W contains a sparse part of measure $\alpha_0 > 0$ with density q close to 0. In such case, the density of H_0 in W can be zero but the density of H_0 in $1-W$ is approximately $\alpha_0^{v(H_0)}(1-q)^{e(H_0)}$, which needs to be at least $2^{-e(H_0)+1}$ in order to yield that H_0 is common.

As the first step towards proving Theorems 6.1 and 6.2, we deal with the *local regime* and establish a counterpart of Theorem 1.3, which is given as Theorem 3.1 in § 3. While Theorem 3.1 applies to graphs of a particular form only, the bound on the cut distance given in Theorem 1.3 is independent of the length ℓ of the path attached to the graph H (in addition to relaxing the technical assumption that $\|W\|_\infty \leq 2p$ as discussed in § 1.1). The independence on ℓ is essential for proving Theorem 4.1 in § 4, where the obtained bound on the cut distance has to be independent of the parameters ℓ, m and n so that it is sufficiently strong to eventually yield proofs of Theorems 6.1 and 6.2. The proof of Theorem 4.1, which is the technically most challenging part of our argument, is obtained using spectral arguments, which allow us to control how densities of rooted subgraphs change when walking along a path in a graphon. In particular, we argue that the density of H_0 in a graphon W is at least the density of H_0 in the constant graphon with the same density as W (cases (i) and (ii) in the proof) unless the graphon W has a sparse part (case (iii)), which, in turn, implies that it also has a denser part where sufficiently many copies of the graph H_0 can be found.

The *non-local regime* is addressed in § 5. First, Lemma 5.1 yields that the density of a fixed graph H in any graphon W is positive unless W contains a sparse part; however, the constants in the statement of the lemma depend on H. Lemma 5.1 is then used to prove Theorem 5.2 that asserts that if a graphon W is far in the cut distance from the constant (quasirandom) graphon, then the density of the graph H_0 in W is large enough unless W contains a sparse part whose size is independent of the parameters ℓ, m and n (from a suitably chosen range).

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Theorem 18 (local regime). *W with density $p \geq p_0$ and $\|W - p\|_\square \leq \varepsilon \Rightarrow$ $t(H_0, W) \geq p^{e(H_0)}$*

Theorem 22 (non-local regime). *W with density $p \geq p_0$, no sparse part and $\|W - p\|_\square \geq \varepsilon \Rightarrow t(H_0, W) \geq p^{e(H_0)}$*

Theorem 23. *H_0 is common, i.e., $t(H_0, W) + t(H_0, 1 - W) \geq 2 \cdot 2^{-e(H_0)}$*

- *W has a part with measure α_0 and density q close to zero*
  $\Rightarrow t(H_0, 1 - W) \geq \alpha_0^{v(H_0)}(1 - q)^{e(H_0)} \geq 2 \cdot 2^{-e(H_0)}$
- *$1 - W$ has a part with measure α_0 and density q close to zero*
  $\Rightarrow t(H_0, W) \geq \alpha_0^{v(H_0)}(1 - q)^{e(H_0)} \geq 2 \cdot 2^{-e(H_0)}$
- *W has density $\geq 1 - p_0$*
  Theorem 18 or Theorem 22 $\Rightarrow t(H_0, W) \geq (1 - p_0)^{e(H_0)} \geq 2 \cdot 2^{-e(H_0)}$
- *$1 - W$ has density $\geq 1 - p_0$*
  Theorem 18 or Theorem 22 $\Rightarrow t(H_0, 1 - W) \geq (1 - p_0)^{e(H_0)} \geq 2 \cdot 2^{-e(H_0)}$
- *W has density $p \in [p_0, 1 - p_0]$ and no sparse part*
  Theorem 18 or Theorem 22 $\Rightarrow t(H_0, W) \geq p^{e(H_0)}$
  Theorem 18 or Theorem 22 $\Rightarrow t(H_0, 1 - W) \geq (1 - p)^{e(H_0)}$
  $t(H_0, W) + t(H_0, 1 - W) \geq p^{e(H_0)} + (1 - p)^{e(H_0)} \geq 2 \cdot 2^{-e(H_0)}$

Theorem 24. *H_0 is k-common, i.e., $t(H_0, W_1) + \cdots t(H_0, W_k) \geq k \cdot k^{-e(H_0)}$*

- *$\exists i$ W_i has a part with measure α_0 and density close to zero*
  induction $\Rightarrow t(H_0, W_1) + \cdots + t(H_0, W_k) \gtrapprox \alpha_0^{v(H_0)}(k - 1)^{-e(H_0)} \geq k \cdot k^{-e(H_0)}$
- *$\exists i$ W_i has density smaller than $p_0 \Rightarrow \exists j$ density of W_j is at least $(k - 1)^{-1}$*
  $\Rightarrow t(H_0, W_j) \gtrapprox (k - 1)^{-e(H_0)} \geq k \cdot k^{-e(H_0)}$
- *Each W_i has density $p_i \geq p_0$ bounded away from zero and no sparse part*
  Theorem 18 or Theorem 22 $\Rightarrow t(H_0, W_i) \geq p_i^{-e(H_0)}$
  $t(H_0, W_1) + \cdots + t(H_0, W_k) \geq p_1^{e(H_0)} + \cdots + p_k^{e(H_0)} \geq k \cdot k^{-e(H_0)}$

Figure 1. Informal statements of Theorems 3.1 and 5.2, and the main steps in the proofs of Theorems 6.1 and 6.2, which assert that a suitable graph H_0 is common and k-common, respectively.

Finally, § 6 presents our main results: Theorems 6.1 and 6.2; the main steps of their proofs are visualized in Figure 1. If a graphon W or $1 - W$ has a sparse part, then the density of H_0 in the complement of the sparse part is sufficiently large to establish that H_0 is common; here, we need that ℓ is sublinear in $v(H) + mn$. Otherwise, if a graphon W is close to a constant graphon, then the density of H_0 in both W and $1 - W$ is at least the density of H_0 in the constant graphons with the same density by Theorem 4.1, and if W is not close to a constant graphon (and does not have a sparse part), then the density of H_0 in both W and $1 - W$ is at least the density of H_0 in the constant graphons with the same density by Theorem 5.2. A simple convexity argument now yields that $t(H_0, W) + t(H_0, 1 - W) \geq 2 \cdot 2^{-e(H_0)}$.

The proof of the more general Theorem 6.2, which concern k-common graphs, proceeds by induction on k, which is the number of colors. Individual color classes are represented by

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

graphons $W_1, \ldots, W_k$ such that $W_1 + \cdots + W_k$ is a graphon with density close to one (rather than exactly equal to one); this makes the induction argument significantly easier. If one of graphons $W_1, \ldots, W_k$ has a very sparse part, we apply induction to the remaining $k-1$ colors restricted to the sparse part. If one of the color classes is sparse, we argue that we find sufficiently many copies of H_0 in the densest color class (here, we use that ℓ is sublinear in $v(H) + mn$). If neither of these two cases apply, we use for each graphon W_i, $i = 1, \ldots, k$, Theorem 4.1 or Theorem 5.2, depending whether W_i is close to a constant graphon or not, i.e. whether the corresponding color class falls into the local or non-local regime, to derive that the density of H_0 in each W_i is at least as in the constant graphon with the same density, and we eventually derive using convexity of the function $p^{e(H_0)}$ that the sum of densities of H_0 in $W_1, \ldots, W_k$ is sufficiently large.

2. Preliminaries

In this section, we fix notation used throughout the paper and present auxiliary results needed in our arguments. We start with basic notation. The set of the first k positive integers is denoted by $[k]$, and the set of all k-element subsets of A is denoted by $\binom{A}{k}$. If A is a measurable subset of $\mathbb{R}^k$, we write $\mu(A)$ for the measure of A; throughout the paper, we always consider the standard Borel measure on $\mathbb{R}^k$.

All graphs considered in this paper are simple and loopless. If G is a graph, then the vertex set of G is denoted by $V(G)$ and the edge set by $E(G)$; the numbers of vertices and edges of G are denoted by $v(G)$ and $e(G)$, respectively. If A is a subset of vertices of G, then $G[A]$ is the subgraph induced by A, i.e. the graph with the vertex set A and the edge set $E(G) \cap \binom{A}{2}$. Finally, if F is a subset of edges of G, then $\langle F \rangle$ is the spanning subgraph of G with the edge set F; the host graph G will always be clear from the context. A *homomorphism* from a graph H to G is mapping f from $V(H)$ to $V(G)$ such that if uv is an edge of H, then $f(u)f(v)$ is an edge of G, and the *homomorphism density* of H in G, denoted by $t(H, G)$, is the probability that a random mapping from $V(H)$ to $V(G)$ is a homomorphism.

The n-vertex path is denoted by P_n, the n-vertex cycle by C_n, the complete graph with n vertices by K_n, and the complete bipartite graph with parts consisting of a and b vertices by $K_{a,b}$. Finally, $K_{a|\ell,b}$ is the graph obtained from $K_{a,b}$ by adding an ℓ-edge path to one of the vertices in the a-vertex part. In particular, $K_{a|0,b}$ is just the graph $K_{a,b}$.

2.1 Graph limits

We treat the notion of common graphs using tools from the theory of graph limits. In this subsection, we provide an introduction to the most important concepts; we refer the reader to the monograph by Lovász [Lov12] for a more complete treatment. In the theory of graph limits, large graphs are represented by an analytic object called graphon. A *graphon* is a measurable function $W : [0,1]^2 \to [0,1]$ that is symmetric, i.e. $W(x,y) = W(y,x)$ for all $x, y \in [0,1]$. When no confusion can arise, we use p to denote the graphon equal to $p \in [0,1]$ everywhere. More generally, a *kernel* is a symmetric measurable function $U : [0,1]^2 \to \mathbb{R}$. In particular, a graphon is a kernel with range in $[0,1]$. A graphon can be thought of as a continuous version of the adjacency matrix of a graph. Because of this analogy, we refer to the elements of the domain $[0,1]$ as to vertices of W. The *degree* of a vertex $x \in [0,1]$ of a graphon W is defined as

$$\deg_W(x) = \int_{[0,1]} W(x,y)\, dy.$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

The *homomorphism density* of a graph H in a graphon W is defined as

$$t(H,W)=\int_{[0,1]^{V(H)}}\prod_{uv\in E(H)}W(x_u,x_v)\,dx_{V(H)}. \tag{2}$$

For brevity, we often speak about the density of H in W instead of the homomorphism density of H in W. The *density* of a graphon W is the density of K_2 in W, i.e. $t(K_2,W)$. Analogously, we define the homomorphism density of a graph H in a kernel U using (2), and we define the density of a kernel U to be $t(K_2,U)$.

We now cast the definition of graphs with the Sidorenko property and commons graphs in the language of graph limits; see, e.g., [KNN+22] for further details. A graph H has the *Sidorenko property* if the following holds for every graphon W:

$$p^{e(H)}\le t(H,W)$$

where p is the density of W. A graph H is *common* if it holds that

$$2^{1-e(H)}\le t(H,W)+t(H,1-W)$$

for every graphon W. Finally, for an integer $k\ge 2$, we say that a graph H is *k-common* if

$$k^{1-e(H)}\le t(H,W_1)+\cdots+t(H,W_k)$$

holds for all graphons $W_1,\dots,W_k$ such that $W_1+\cdots+W_k=1$.

We next define the notion of a norm for kernels, which gives a metric on the space of kernels and so on graphons. The *cut norm* of a kernel U is defined as

$$\|U\|_\square=\sup_{S,T\subseteq[0,1]}\left|\int_{S\times T}U(x,y)\,dx\,dy\right|,$$

where the supremum is taken over all measurable subsets S and T of $[0,1]$. If U is a kernel with $\|U\|_\infty\le 1$, then it holds [Lov12, Lemma 8.12] that

$$\|U\|_\square^4\le t(C_4,U)\le 4\|U\|_\square. \tag{3}$$

Similarly, it holds that $t(P_2,U)\le 2\|U\|_\square$. Since C_4 has the Sidorenko property, it holds that $t(C_4,W)\ge p^4$ for every graphon W with density p. If a graphon W with density p is actually far from the p-constant graphon, then the density of C_4 is much larger than p^4 as given in the lemma.

LEMMA 2.1 (Cooper *et al.* [CKM18, Lemma 11]). *The following holds for every graphon W with density p:*

$$p^4+\frac{\|W-p\|_\square^4}{8}\le t(C_4,W).$$

On the other hand, the counting lemma, which we now state, implies that $t(C_4,W)\le p^4+4\|W-p\|_\square$ for every graphon W with density p.

LEMMA 2.2 (Counting lemma, Lovász [Lov12, Lemma 10.23]). *The following holds for every graph H and all graphons W_1 and W_2:*

$$|t(H,W_1)-t(H,W_2)|\le e(H)\cdot\|W_1-W_2\|_\square.$$

We next define a notion used in [KNN+22] that can be viewed as an analogue of a subgraph in a graphon. Consider a measurable function $h:[0,1]\to[0,1]$ such that $\|h\|_1>0$. Let $f:[0,\|h\|_1]\to[0,1]$ be the measurable function defined as

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$$f(z) := \inf\left\{t \in [0,1] \text{ such that } \int_{[0,t]} h(x)\,dx \geq z\right\}.$$

Note that it holds that $\int_A h(x)\,dx = \mu(f^{-1}(A))$ for every measurable subset $A \subseteq [0,1]$. We define the graphon $W[h]$ by setting

$$W[h](x,y) = W(f(x \cdot \|h\|_1),\ f(y \cdot \|h\|_1)) \quad \text{for } (x,y) \in [0,1]^2.$$

Note that if h is the indicator function of a measurable subset A, then $W[h]$ is obtained by restricting W to A and rescaling; if this is the case, we write $W[A]$ for the graphon $W[h]$. It can be shown that

$$t(H, W[h]) = \frac{1}{\|h\|_1^{v(H)}} \int_{[0,1]^{V(H)}} \prod_{u \in V(H)} h(u) \prod_{uv \in E(H)} W(x_u, x_v)\,dx_{V(H)}$$

holds for every graph H. In particular, it holds that

$$t(H, W) \geq \|h\|_1^{v(H)} \cdot t(H, W[h]). \tag{4}$$

We conclude this subsection by defining a notion of an almost independent set in a graphon. For a graphon W and $\delta > 0$, $\mathcal{H}_\delta(W)$ is the set of all measurable functions $h : [0,1] \to [0,1]$ with $\|h\|_1 > 0$ such that the density of $W[h]$ is at most δ. The *δ-independence ratio* of W is defined as

$$\alpha_\delta(W) = \sup_{h \in \mathcal{H}_\delta(W)} \|h\|_1;$$

if the set $\mathcal{H}_\delta(W)$ is empty, then we set $\alpha_\delta(W) = 0$. We remark that this notion is closely related to the notion of (ρ, d)-dense graphs widely used in Ramsey theory: a graph is *(ρ, d)-dense* if any subset of at least $\rho \cdot v(G)$ vertices of G induces a subgraph with density at least d. Indeed, if a graphon W is a limit graphon of a sequence of (ρ, d)-dense graphs, then $\alpha_\delta(W) < \rho$ for every $\delta < d$.

2.2 Rooted graphs

A *rooted graph* is a graph with one or more vertices distinguished as *roots*. We use superscripts to emphasize that a graph is rooted and indicate the number of roots. In particular, $H^\bullet$ will denote a rooted graph with a single root, $H^{\bullet\bullet}$ will denote a rooted graph with two roots, and $H^{\bullet\cdots\bullet}$ will denote a rooted graph with three or more roots. In all rooted graphs considered in this paper, the set of roots will always form an *independent set*. To simplify our notation, if $H^\bullet$ is a rooted graph, $v(H)$ and $e(H)$ denote the number of vertices and edges of $H^\bullet$, respectively, as these quantities are the same regardless of the presence or the absence of roots. Examples of rooted graphs considered further include the following. The rooted graph $P_n^\bullet$ is the graph obtained from the n-vertex path by choosing one of its end vertices as a root. The rooted graph $K_{a,b}^\bullet$ is the graph obtained from the complete bipartite graph $K_{a,b}$ by choosing one of the vertices of the a-vertex part as a root. Finally, the rooted graph $K_{a|\ell,b}^\bullet$ is the graph obtained from $K_{a|\ell,b}$ by choosing the end vertex of the appended ℓ-edge path as a root.

If $G^\bullet$ and $H^\bullet$ are two rooted graphs, then $G^\bullet \oplus H^\bullet$ is the (unrooted) graph obtained by identifying their roots. Note that $K_{a|\ell,b} = P_{\ell+1}^\bullet \oplus K_{a,b}^\bullet$. Similarly, if $G^{\bullet\bullet}$ and $H^{\bullet\bullet}$ are two rooted graphs with two roots, then $G^{\bullet\bullet} \oplus H^{\bullet\bullet}$ is the graph obtained by identifying the corresponding roots, and if $G^{\bullet\cdots\bullet}$ and $H^{\bullet\cdots\bullet}$ are two rooted graphs with the same number of roots, then $G^{\bullet\cdots\bullet} \oplus H^{\bullet\cdots\bullet}$ is the graph obtained by identifying the corresponding roots; the correspondence will always be clear from the context.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

We next extend the notion of homomorphism density to rooted graphs. Let $H^{\bullet\cdots\bullet}$ be a rooted graph with k roots, and let $v_1,\dots,v_n$ be the vertices of $H^{\bullet\cdots\bullet}$ listed in a way that the vertices $v_1,\dots,v_k$ are the roots. If U is a kernel and $x_1,\dots,x_k\in[0,1]$, we define $t_{x_1,\dots,x_k}(H^{\bullet\cdots\bullet},U)$ as

$$t_{x_1,\dots,x_k}(H^{\bullet\cdots\bullet},U)=\int_{[0,1]^{n-k}}\prod_{v_iv_j\in E(H^{\bullet\cdots\bullet})}U(x_i,x_j)\,dx_{k+1}\cdots dx_n. \tag{5}$$

Observe that if W is a graphon, then $\deg_W(x)=t_x(P_2^{\bullet},W)$ for $x\in[0,1]$. Finally, if $G^{\bullet\cdots\bullet}$ and $H^{\bullet\cdots\bullet}$ are two rooted graphs such that each has k roots and the roots induce an independent set, it holds that

$$t(G^{\bullet\cdots\bullet}\oplus H^{\bullet\cdots\bullet},U)=\int_{[0,1]^k}t_{x_1,\dots,x_k}(G^{\bullet\cdots\bullet},U)t_{x_1,\dots,x_k}(H^{\bullet\cdots\bullet},U)\,dx_1\cdots dx_k \tag{6}$$

for every kernel U; note that the assumption that the roots of $G^{\bullet\cdots\bullet}$ and $H^{\bullet\cdots\bullet}$ form independent sets is needed for (6) to hold.

2.3 Spectral properties of graphons and kernels

We now review spectral properties of graphons and kernels; we refer to [Lov12, § 7.5] for further details. Fix a kernel U with $\|U\|_\infty\le 1$. We can think of U as a Hilbert–Schmidt integral operator from $L_2[0,1]$ to $L_2[0,1]$ defined as

$$(Uf)(x)=\int_0^1 U(x,y)f(y)\,dy.$$

Let $\lambda_1,\lambda_2,\dots$ be the non-zero eigenvalues of U listed in the non-increasing order of the absolute value (with multiplicities), and let $f_1,f_2,\dots$ be the corresponding orthonormal eigenfunctions, i.e. $\|f_i\|_2=1$ for every i and the functions are orthogonal to each other in $L_2[0,1]$. Note that

$$\sum_i \lambda_i f_i(x)f_i(y)$$

converges to U in the L_2-norm. We remark that if U is a graphon with density p, then λ_1 is at least p. Since we have assumed that $\|U\|_\infty\le 1$, it holds that $(Uf_i)(x)\le\|f_i\|_2=1$ for every $x\in[0,1]$, i.e. $\|Uf_i\|_\infty\le 1$, which implies that $\|f_i\|_\infty\le|\lambda_i|^{-1}$.

The spectrum of U can be used to express the density of cycles and paths. Expressing the density of cycles is easier: it holds for every $n\ge 3$ that

$$t(C_n,U)=\sum_i\lambda_i^n. \tag{7}$$

To express the density of paths, we need to introduce additional notation. Let $\mathfrak{j}:[0,1]\to[0,1]$ be the constant function equal to 1, and set $\alpha_i\in[0,\pi/2]$ to be the real such that $\langle\mathfrak{j},f_i\rangle=\cos\alpha_i$ (we may assume without loss of generality by replacing f_i with $-f_i$ that the product is non-negative). Further set $\delta=1-\cos\alpha_1$. Informally speaking, for a graphon W, δ measures how close W is to having all the degrees the same, in particular, $\delta=0$ if and only if $\deg_W(x)=p$ for almost every $x\in[0,1]$ where p is the density of W. The following holds for every $n\ge 2$ (for $n=2$, the equality is a particular case of (7.21) in [Lov12]):

$$t(P_n,U)=\langle\mathfrak{j},U^{n-1}\mathfrak{j}\rangle=\sum_i\lambda_i^{n-1}\cos^2\alpha_i. \tag{8}$$

In the rest of this subsection, we deal with graphons only and estimate some of the introduced parameters. These estimates will be used repeatedly in the proofs of Theorems 3.1 and 4.1. Fix

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

a graphon W with density $p \in [0,1]$, and let $\gamma = t(C_4, W) - p^4$. Note that $\gamma \geq 0$ since the cycle C_4 has the Sidorenko property; on the other hand, $\gamma \leq 4\|W-p\|_\square$ by the counting lemma (Lemma 2.2). Since the density of C_4 in W is at least λ_1^4, it follows that

$$\lambda_1 \leq p + \frac{\gamma}{4p^3} \leq p + \frac{\|W-p\|_\square}{p^3}. \tag{9}$$

We derive from $\lambda_1 \geq p$ and (7) that

$$|\lambda_i| \leq \gamma^{1/4} \tag{10}$$

holds for every $i \geq 2$. Hence, we derive from (10) that

$$\sum_{i\geq 2} \lambda_i^m \leq \gamma^{\frac{m-4}{4}} \sum_{i\geq 2} \lambda_i^4 \leq \gamma^{m/4} \tag{11}$$

holds for every $m \geq 4$. In addition, (10) yields that it holds for every $m \geq 0$ that

$$\left| \sum_{i\geq 2} \lambda_i^m \cos^2 \alpha_i \right| \leq \sum_{i\geq 2} |\lambda_i|^m \cos^2 \alpha_i \leq \max_{i\geq 2} |\lambda_i|^m \cdot \sum_{i\geq 2} \cos^2 \alpha_i \leq \max_{i\geq 2} |\lambda_i|^m \leq \gamma^{m/4},$$

which implies by (8) that

$$\lambda_1^m (1-\delta)^2 - \gamma^{m/4} \leq t(P_{m+1}, W) \leq \lambda_1^m (1-\delta)^2 + \gamma^{m/4}. \tag{12}$$

We next estimate δ in terms of γ and subsequently in terms of $\|W-p\|_\square$. Since the path P_5 has the Sidorenko property and (8) holds, we obtain that

$$\begin{aligned} p^4 \leq t(P_5, W) &= \sum_i \lambda_i^4 \cos^2 \alpha_i \leq \lambda_1^4 (1-\delta)^2 + \sum_{i\geq 2} \lambda_i^4 = \lambda_1^4 (1-\delta)^2 - \lambda_1^4 + \sum_i \lambda_i^4 \\ &\leq p^4 (1-\delta)^2 - p^4 + \sum_i \lambda_i^4 = p^4 (1-\delta)^2 + \gamma \leq p^4 (1-\delta) + \gamma. \end{aligned}$$

It follows that

$$\delta \leq \frac{\gamma}{p^4} \leq \frac{4\|W-p\|_\square}{p^4}. \tag{13}$$

Since the eigenfunctions f_i are orthonormal, we obtain that

$$\sum_i \cos^2 \alpha_i = \sum_i \langle \mathrm{j}, f_i \rangle^2 \leq \|\mathrm{j}\|_2^2 = 1. \tag{14}$$

The inequality (14) implies that

$$\sum_{i\geq 2} \cos^2 \alpha_i \leq 1 - (1-\delta)^2 = 2\delta - \delta^2 \leq 2\delta. \tag{15}$$

In particular, $\cos \alpha_i \leq 2\delta^{1/2}$ for every $i \geq 2$. Using (8) for $n = 2$, (10) and (15), we obtain that

$$p = \sum_i \lambda_i \cos^2 \alpha_i \leq \lambda_1 (1-\delta)^2 + \gamma^{1/4} \sum_{i\geq 2} \cos^2 \alpha_i \leq \lambda_1 (1-\delta)^2 + 2\delta\gamma^{1/4}.$$

This implies for $\delta \in [0,1]$ that

$$\lambda_1 \geq \frac{p - 2\delta\gamma^{1/4}}{(1-\delta)^2} \geq (p - 2\delta\gamma^{1/4})(1+\delta)^2 \geq p(1+2\delta) - 2\delta\gamma^{1/4}(1+\delta)^2 \geq p(1+2\delta) - 8\delta\gamma^{1/4}. \tag{16}$$

2.4 Estimates on densities

We now review several estimates on densities of graphs in graphons and kernels. A large number of the estimates that we present is standard and we state them explicitly just for referencing in

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

the rest of the paper. We start with one of such estimates, which follows by the standard use of Jensen's inequality; see, e.g., [KNN+22, Proposition 1.10]. The two lemmas are also implied by the weak Hölder property that all complete bipartite graphs have since they are weakly norming [Hat10]; also see [Lov12, § 14.1] for in-depth discussion.

Lemma 2.3. *The following holds for every graphon W and all integers m, m', n such that $m \geq m'$:*

$$t(K_{m,n}, W) \geq t(K_{m',n}, W)^{m/m'}.$$

Two applications of Lemma 2.3 yield the following.

Lemma 2.4. *The following holds for every graphon W and all integers $m, n \geq 2$:*

$$t(K_{m,n}, W) \geq t(C_4, W)^{mn/4}.$$

The estimates in the rest of the subsection concern kernels. We start with a well-known estimate, which follows by a straightforward application of the Cauchy–Schwarz inequality.

Lemma 2.5. *The following holds for any kernel U:*

$$0 \leq t(K_{1,2}, U) \leq t(C_4, U)^{1/2}. \tag{17}$$

The next lemma follows by using Jensen's inequality and the following two identities:

$$t(K_{1,k}, U) = \int_{[0,1]} t_x(P_2^{\bullet}, U)^k \, dx \quad \text{and} \quad t(K_{2,k}, U) = \int_{[0,1]^2} t_{x,y}(K_{2,1}^{\bullet\bullet}, U)^k \, dx \, dy.$$

Lemma 2.6. *The following holds for any kernel U with $\|U\|_\infty \leq 1$ and every integer $k \geq 2$:*

$$|t(K_{1,k}, U)| \leq t(K_{1,2}, U)^{k/2} \leq t(K_{1,2}, U) \quad \text{and} \quad |t(K_{2,k}, U)| \leq t(C_4, U)^{k/2} \leq t(C_4, U).$$

The remaining estimates that we present were proven by Lovász in [Lov11]; we state them with corresponding references to the statements in the paper [Lov11], which contain them or imply them. We remark that the second inequality in Lemma 2.9 follows using Lemma 2.5.

Lemma 2.7 (Lovász [Lov11, Lemma 3.3]). *Let $f_1, \dots, f_n : [0,1]^k \to \mathbb{R}$ be bounded measurable functions such that for each variable there are at most two functions f_i that depend on that variable. It holds that*

$$\int_{[0,1]^k} \prod_{i \in [n]} f_i(x_1, \dots, x_n) \, dx_1 \cdots dx_k \leq \prod_{i \in [n]} \|f_i\|_2.$$

Lemma 2.8 (Lovász [Lov11, Corollary 3.12]). *The following holds for any kernel U with $\|U\|_\infty \leq 1$ and every integer $k \geq 2$:*

$$t(C_{2k}, U) \leq t(C_4, U)^{k/2}.$$

Lemma 2.9 (Lovász [Lov11, Lemma 3.14]). *The following holds for any kernel U with $\|U\|_\infty \leq 1$ and every integer $k \geq 1$:*

$$t(P_{k+3}, U)^4 \leq t(P_3, U)^4 t(C_4, U)^k \leq t(C_4, U)^{k+2}.$$

Lemma 2.10 (Lovász [Lov11, Lemma 3.19]). *The following holds for any kernel U with $\|U\|_\infty \leq 1$ and every bipartite graph G with minimum degree two and girth at least four that is not a single cycle or a complete bipartite graph:*

$$|t(G, U)| \leq t(C_4, U)^{5/4}.$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Lemma 2.11 (Lovász [Lov11, Lemma 3.21]). *The following holds for any kernel U with $\|U\|_\infty \le 1$ and any tree T that is not a star:*

$$|t(T,U)| \le t(P_3,U)t(C_4,U)^{1/4}.$$

Lemma 2.12 (Lovász [Lov11, Lemma 3.22]). *The following holds for any kernel U with $\|U\|_\infty \le 1$ and every bipartite graph G with girth at least four that has exactly one vertex of degree one:*

$$|t(G,U)| \le \frac{(t(C_4,U)+t(P_3,U))\cdot t(C_4,U)^{1/8}}{2}.$$

2.5 Entropy-based estimates

We next estimate the density of a graph $K_{a|\ell,b}$. The next lemma asserts that a graph $K_{a|\ell,b}$ has the Sidorenko property in a strong sense as $t(K_{1,2},G) \ge t(K_2,G)^2$. Our proof is based on the entropy argument, which has been pioneered in this context by Szegedy [Sze14].

Lemma 2.13. *Let a, b and ℓ be any positive even integers. It holds that for any graph G,*

$$t(K_{a|\ell,b},G) \ge t(K_{1,2},G)^{(ab+\ell)/2}.$$

Proof. Fix integers a, b and ℓ, and a graph G, and let n be the number of vertices of G. We say that a triple (x_1,x_2,y) of vertices of G is *feasible* if x_1y and x_2y are edges of G. We define several probability distributions on vertices of G and their tuples. The first distribution is the uniform distribution on $t(K_{1,2},G)n^3$ feasible triples (x_1,x_2,y); note that the entropy of this distribution is

$$H(x_1,x_2,y) = 3\log n + \log t(K_{1,2},G). \tag{18}$$

The second distribution is the distribution on $(a+1)$-tuples $(x_1,\dots,x_a,y)$ that we first choose y according to the marginal distribution of the uniform distribution on feasible triples and we then choose $x_1,\dots,x_a$ to be neighbors of y chosen uniformly and independently of each other. Observe that the marginal distribution on y and any two of the vertices $x_1,\dots,x_a$ is the uniform distribution on feasible triples and so the following holds:

$$\begin{aligned} H(x_1,\dots,x_a,y) &= \frac{a}{2}H(x_1,x_2|y) + H(y) = \frac{a}{2}H(x_1,x_2,y) - \frac{a-2}{2}H(y) \\ &\ge \frac{a}{2}H(x_1,x_2,y) - \frac{a-2}{2}\log n. \end{aligned}$$

The third distribution is the distribution on $(a+b)$-tuples $(x_1,\dots,x_a,y_1,\dots,y_b)$ that we first choose $x_1,\dots,x_a$ according to the marginal distribution derived from the second distribution and we then choose each of $y_1,\dots,y_b$ according to the second distribution conditioned on the choice of $x_1,\dots,x_a$. Observe that

$$\begin{aligned} H(x_1,\dots,x_a,y_1,\dots,y_b) &= b\cdot H(y|x_1,\dots,x_a) + H(x_1,\dots,x_a) \\ &= b\cdot H(x_1,\dots,x_a,y) - (b-1)H(x_1,\dots,x_a) \\ &\ge b\cdot H(x_1,\dots,x_a,y) - a(b-1)\log n \\ &\ge \frac{ab}{2}H(x_1,x_2,y) - \frac{(a-2)b}{2}\log n - a(b-1)\log n \\ &= \frac{ab}{2}H(x_1,x_2,y) - \frac{3ab-2a-2b}{2}\log n. \end{aligned}$$

Finally, the last distribution, which is a distribution on $(a+b+\ell)$-tuples of vertices, is obtained as follows. We first choose $(x_1,\dots,x_a,y_1,\dots,y_b)$ according to the third distribution; note that

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

the marginal distribution of x_1 is the same as its marginal distribution of the uniform distribution on feasible triples. Next choose (x_1, y_1', x_1') uniformly among all feasible triples with the first coordinate equal to x_1, then (x_1', y_2', x_2') uniformly among all feasible triples with the first coordinate equal to x_1', then (x_2', y_3', x_3') uniformly among all feasible triples with the first coordinate equal to x_2' and so on until the triple $(x'_{\ell/2-1}, y'_{\ell/2}, x'_{\ell/2})$ has been chosen. Observe that

$$H(x_1, y_1', x_1', \dots, y'_{\ell/2}, x'_{\ell/2}) = \frac{\ell}{2} H(y, x_2|x_1) + H(x_1) = \frac{\ell}{2} H(x_1, y, x_2) - \left(\frac{\ell}{2} - 1\right) H(x_1).$$

It follows that $H(x_1, \dots, x_a, y_1, \dots, y_b, y_1', x_1', \dots, y'_{\ell/2}, x'_{\ell/2})$ is equal to

$$\begin{aligned}
&H(x_1, y_1', x_1', \dots, y'_{\ell/2}, x'_{\ell/2}) + H(x_1, \dots, x_a, y_1, \dots, y_b) - H(x_1) \\
&\quad \geq \frac{ab+\ell}{2} H(x_1, x_2, y) - \frac{3ab - 2a - 2b}{2} \log n - \frac{\ell}{2} H(x_1) \\
&\quad \geq \frac{ab+\ell}{2} H(x_1, x_2, y) - \frac{3ab - 2a - 2b + \ell}{2} \log n \\
&\quad = (a + b + \ell) \log n + \frac{ab+\ell}{2} \log t(K_{1,2}, G),
\end{aligned}$$

where the last inequality is implied by (18). Since $H(x_1, \dots, x_a, y_1, \dots, y_b, y_1', x_1', \dots, y'_{\ell/2}, x'_{\ell/2})$ is at most $\log\left(t(K_{a|\ell,b}, G) n^{a+b+\ell}\right)$, we conclude that $\log t(K_{a|\ell,b}, G) \geq \frac{ab+\ell}{2} \log t(K_{1,2}, G)$, i.e. $t(K_{a|\ell,b}, G) \geq t(K_{1,2}, G)^{(ab+\ell)/2}$ as desired. □

Our second estimate on the density of a graph $K_{a|\ell,b}$ is a quantitative bound on the density of $K_{a|\ell,b}$ in a graphon W depending on its distance from the constant graphon with the same density.

LEMMA 2.14. *For all positive even integers a, b and ℓ such that $\ell \leq ab/4$, it holds that*

$$t(K_{a|\ell,b}, W) \geq p^{ab+\ell}(1 + 10^{-9}\|W - p\|_\square^{16}))^{ab/4+\ell/4}$$

for every graphon W with density p.

Proof. Fix integers a, b and ℓ, and a graphon W with density p; set $\varepsilon = \|W - p\|_\square$. Let A be the set of $x \in [0, 1]$ such that $\deg_W(x) \leq p(1 - 0.001\varepsilon^5)$. We distinguish two cases depending on the measure of A.

The first case is the case when $\mu(A) < 0.001p\varepsilon^5$, i.e. the degree of most of the vertices of W is close to p. Consider the graphon W_1 defined as

$$W_1(x, y) = \begin{cases} 0 & \text{if } x \in A \text{ or } y \in A, \text{ and} \\ W(x, y) & \text{otherwise.} \end{cases}$$

Observe that

$$\deg_{W_1}(x) \geq p(1 - 0.001\varepsilon^5) - \mu(A) \geq p(1 - 0.002\varepsilon^5) \tag{19}$$

for every $x \notin A$. It follows that the density p_1 of W_1 is at least

$$p_1 \geq (1 - \mu(A))(1 - 0.002p\varepsilon^5) \geq (1 - 0.001p\varepsilon^5)p(1 - 0.002\varepsilon^5) \geq p(1 - 0.003\varepsilon^5).$$

Note that the cut distance between W and W_1 is at most $2\mu(A) + \mu(A)^2 \leq 3\mu(A) \leq 0.003p\varepsilon^5$. Hence, the triangle inequality implies that

$$\|W_1 - p_1\|_\square \geq \|W - p\|_\square - \|W - W_1\|_\square - |p - p_1| \geq \varepsilon - 0.006p\varepsilon^5 \geq 0.994\varepsilon.$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Lemma 2.1 now yields that

$$t(C_4, W_1) \geq p_1^4 + \frac{0.994^4\varepsilon^4}{8} \geq p_1^4(1 + 0.122\varepsilon^4) \geq p^4(1 - 0.003\varepsilon^5)^4(1 + 0.122\varepsilon^4) \geq p^4(1 + 0.1\varepsilon^4).$$

Hence, Lemma 2.4 implies (the last inequality follows from $\ell \leq ab/4$) that

$$t(K_{a,b}, W_1) \geq t(C_4, W_1)^{ab/4} \geq p^{ab}(1 + 0.1\varepsilon^4)^{ab/4} \geq p^{ab}(1 + 0.03\varepsilon^4)^{ab/4+2\ell}.$$

Therefore, we obtain using (19) that

$$t(K_{a|\ell,b}, W_1) \geq p^{ab+\ell}(1 + 0.03\varepsilon^4)^{ab/4+2\ell}(1 - 0.002\varepsilon^5)^{\ell} \geq p^{ab+\ell}(1 + 0.03\varepsilon^4)^{ab/4+\ell}.$$

As $t(K_{a|\ell,b}, W) \geq t(K_{a|\ell,b}, W_1)$, the estimate in the statement of the lemma follows.

We next analyze the complementary case, i.e. the case when the measure of A is at least $0.001p\varepsilon^5$. We obtain the following estimate on $t(K_{1,2}, W)$ using the Cauchy–Schwarz inequality:

$$\begin{aligned}
t(K_{1,2}, W) &= \int_{[0,1]} \deg_W(x)^2\,dx = \int_A \deg_W(x)^2\,dx + \int_{[0,1]\setminus A} \deg_W(x)^2\,dx \\
&\geq \frac{1}{\mu(A)}\left(\int_A \deg_W(x)\,dx\right)^2 + \frac{1}{1-\mu(A)}\left(\int_{[0,1]\setminus A} \deg_W(x)\,dx\right)^2 \\
&= \frac{1}{\mu(A)}\left(\int_A \deg_W(x)\,dx\right)^2 + \frac{1}{1-\mu(A)}\left(p - \int_A \deg_W(x)\,dx\right)^2 \\
&= p^2 + \frac{\mu(A)}{1-\mu(A)}\left(p - \frac{\int_A \deg_W(x)\,dx}{\mu(A)}\right)^2 \\
&\geq p^2 + \frac{0.001p\varepsilon^5}{1 - 0.001p\varepsilon^5}(0.001p\varepsilon^5)^2 \geq p^2 + 10^{-9}\varepsilon^{15}p^3.
\end{aligned}$$

Observe that $\varepsilon = \|W - p\|_\square \leq \|W\|_\square + \|p\|_\square = 2p$ and so $t(K_{1,2}, W) \geq p^2(1 + 10^{-9}\varepsilon^{16}/2)$. Therefore, Lemma 2.13 now yields that

$$t(K_{a|\ell,b}, W) \geq t(K_{1,2}, W)^{(ab+\ell)/2} \geq p^{ab+\ell}(1 + 10^{-9}\varepsilon^{16}/2)^{ab/2+\ell/2} \geq p^{ab+\ell}(1 + 10^{-9}\varepsilon^{16})^{ab/4+\ell/4}.$$

The estimate given in the statement of the lemma is now established. □

3. Locally Sidorenko component

In this section, we establish that a certain family of graphs that contain any fixed high-girth graph as an induced subgraph has the Sidorenko property in a local sense. As discussed in §§ 1.1 and 1.2, the main contribution of this section is Theorem 3.1 asserting that the value of ε_0 is independent of ℓ and so, when $H^\bullet$ is fixed, of the graph $H^\bullet \oplus P_\ell^\bullet$. The proof is inspired by the proof of an extension of a result by Lovász [Lov11] due to Fox and the last author [FW17]; however, it does not follow from either of these results even when the bound is not required to be uniform. We remark that the lower bound of 50 on the girth of G given in Theorem 3.1 can certainly be decreased, possibly all the way to 4, at the expense of making the proof of Theorem 3.1 significantly more technical; we decided to prove the theorem with the lower bound of 50 as proving it with an even larger lower bound on the girth would not simplify the proof substantially.

THEOREM 3.1. *For every graph G with girth at least 50 and every real number $p_0 \in (0, 1)$, there exist a real $\varepsilon_0 > 0$ and a connected rooted graph $H^\bullet$ containing G as an induced subgraph such*

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

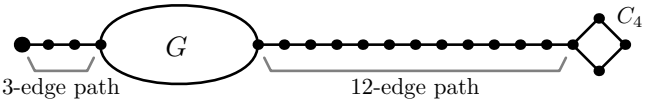


FIGURE 2. The graph $H^\bullet$ from the statement of Theorem 3.1.

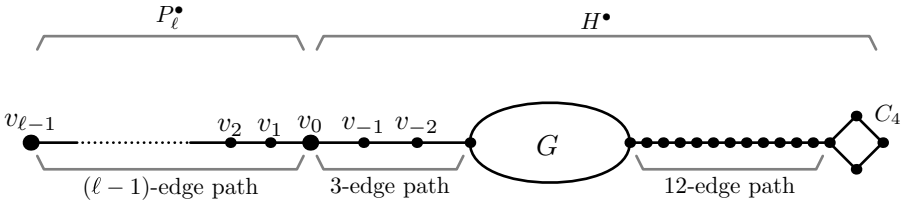


FIGURE 3. The graph $H_\ell = H^\bullet \oplus P_\ell^\bullet$ from the statement of Theorem 3.1 and notation concerning the vertices $v_{-2}, v_{-1}, \ldots, v_{\ell-1}$ used in its proof.

that the following holds for every graphon W with density $p \geq p_0$ and $\|W - p\|_\square \leq \varepsilon_0$:

$$t(H^\bullet \oplus P_\ell^\bullet, W) \geq p^{e(H)+\ell-1}$$

for every integer $\ell \geq 9$.

Proof. Fix a graph G and a real $p_0 \in (0,1)$. The graph $H^\bullet$ is obtained from the graph G by attaching a 3-edge path and a 12-edge path to two different vertices of G, attaching a cycle of length four to the end vertex of the 12-edge path not contained in G, and choosing the end vertex of the 3-edge path not contained in G as the root. The construction of $H^\bullet$ is illustrated in Figure 2. In what follows, we use H for the non-rooted graph with the same vertices and edges as $H^\bullet$, i.e. $H^\bullet$ can be obtained from H be distinguishing the vertex v_0 as the root.

We show that the statement of the theorem holds for ε_0 chosen as

$$\varepsilon_0 = \min\left\{\frac{p_0^4}{10^{12}}, \frac{p_0^{64}}{2^{8e(H)+90}}\right\}.$$

In particular, ε_0 satisfies that

$$\frac{3\varepsilon_0^{1/4}}{p_0} \leq \frac{1}{100}\left(1 - \frac{4\varepsilon_0}{p_0^4}\right)^2. \tag{20}$$

We remark that the estimate (20) will be used later to show that the middle inequality in (31) holds for all $m \geq 2$ and the first inequality in (33) holds for all $m = 2, \ldots, \ell$ whenever $\lambda_1 \in [p_0, 1]$, $\delta \in [0, 4\varepsilon_0/p_0^4]$ and $\gamma \in [0, 4\varepsilon_0]$.

Consider a graph $H^\bullet \oplus P_\ell^\bullet$ for $\ell \geq 9$ and a graphon W with density $p \geq p_0$ such that $\|W - p\|_\square \leq \varepsilon_0$. To simplify our notation, we write H_ℓ for $H^\bullet \oplus P_\ell^\bullet$. The vertices of the path $P_\ell^\bullet$ are denoted by $v_0, \ldots, v_{\ell-1}$ in a way that v_0 is the root; the two internal vertices of the 3-edge path attached to G when constructing $H^\bullet$ will be denoted by v_{-2} and v_{-1} in a way that v_{-1} is the neighbor of v_0. The notation is illustrated in Figure 3.

Let $U = W - p$ and observe that $t(K_2, U) = 0$ (as the density of W is p) and $\|U\|_\infty \leq 1$. Since a path P_m has the Sidorenko property for every $m \geq 1$, we obtain that

$$t(P_m, p + U) = t(P_m, W) \geq p^{m-1} \tag{21}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

for every $m \geq 1$. The density of H_ℓ can be expressed (see [Lov11] and [Lov12, Proof of Proposition 16.27]) as

$$t(H_\ell, W) = t(H_\ell, p+U) = p^{e(H)+\ell-1} + \sum_{\substack{F_0 \subseteq E(H_\ell) \\ F_0 \neq \emptyset}} p^{e(H)+\ell-1-|F_0|} t(\langle F_0 \rangle, U). \tag{22}$$

Hence, in order to establish the statement of the theorem, it is enough to show the following (note that the sum from (22) is divided by $p^{e(H)}$):

$$\sum_{\substack{F_0 \subseteq E(H_\ell) \\ F_0 \neq \emptyset}} p^{\ell-1-|F_0|} t(\langle F_0 \rangle, U) \geq 0. \tag{23}$$

We split the non-empty $2^{e(H)+\ell-1}-1$ subsets F_0 of $E(H_\ell)$ into eight groups given below, which we analyze separately. We say that a component of $\langle F_0 \rangle$ is a *core* component if it contains an edge of $H^\bullet$.

(a) The graph $\langle F_0 \rangle$ has no core component.

(b) The graph $\langle F_0 \rangle$ has a single core component and this component is C_4.

(c) The graph $\langle F_0 \rangle$ has a single core component and this component is a star.

(d) The graph $\langle F_0 \rangle$ has a core component containing the vertex v_8.

(e) The graph $\langle F_0 \rangle$ has no core component containing v_8 and there is a core component containing a cycle of G.

(f) The graph $\langle F_0 \rangle$ has no core component containing v_8 and the only cycle in $\langle F_0 \rangle$ is C_4 but the core component containing C_4 is not just C_4.

(g) The graph $\langle F_0 \rangle$ has no core component containing v_8, all core components are acyclic or C_4, and at least one of the acyclic core components is not a star.

(h) The graph $\langle F_0 \rangle$ has no core component containing v_8, all core components are stars or C_4, and there are at least two core components.

We briefly verify that each subset F_0 belongs to exactly one of the groups (a)–(h). Since any core component containing the vertex v_8 can be neither C_4 nor a star, group (d) is disjoint from the remaining seven groups. If F_0 is not covered by group (a) or (d), then $\langle F_0 \rangle$ has at least one core component and no core component of $\langle F_0 \rangle$ contains v_8. We restrict our attention to such subsets F_0 in the rest of this paragraph. If $\langle F_0 \rangle$ has a core component containing a cycle of G (note that the length of this cycle is at least 50), then F_0 belongs to group (e) and it cannot belong to any other group. If $\langle F_0 \rangle$ has no core component containing a cycle of G but it has a core component that contains the cycle C_4 and this core component is not solely this cycle, then F_0 belongs to group (f) and it cannot belong to any other group. Hence, we can further restrict our attention to only those subsets F_0 such that no core component of $\langle F_0 \rangle$ contains v_8 and each core component of $\langle F_0 \rangle$ is either acyclic or solely the cycle C_4. If one of core components is not a star, then F_0 belongs to group (g) and it cannot belong to any other group. Otherwise, all core components of $\langle F_0 \rangle$ are stars or solely the cycle C_4. If F_0 has a single core component, then it belongs to group (b) or group (c); otherwise, i.e. when F_0 has at least two core components, it belongs group (h). We conclude that each subset F_0 belongs to exactly one of the groups (a)–(h).

We next analyze the contribution of the terms for F_0 based on the groups (a)–(h) in Claims 3.2–3.8. Claims 3.2 and 3.8 will readily yield that the left-hand side of (23) is at

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

least

$$t(P_\ell, p+U)(t(P_3,U)+t(C_4,U)) - t(P_\ell,p+U)\left(\frac{t(P_3,U)}{2}+\frac{3t(C_4,U)}{8}\right)$$
$$\geq t(P_\ell,p+U)\left(\frac{t(P_3,U)}{2}+\frac{5t(C_4,U)}{8}\right) \geq 0,$$

which establishes that (23) holds. In particular, when Claims 3.2 and 3.8 are established, the proof of the theorem will be completed. We remark that Claims 3.3–3.7 describe auxiliary results used to prove Claim 3.8.

CLAIM 3.2. *The sum of the terms in the left-hand side of* (23) *that involve F_0 from one of the groups (a)–(c) is at least*

$$t(P_\ell, p+U)(t(P_3,U)+t(C_4,U)).$$

Proof of Claim 3.2. The sum of the terms that involve a set F_0 from group (a) is equal to

$$\sum_{\substack{F\subseteq E(P_\ell)\\ F\neq\emptyset}} p^{\ell-1-|F|}t(\langle F\rangle,U) = \sum_{F\subseteq E(P_\ell)} p^{\ell-1-|F|}t(\langle F\rangle,U) - p^{\ell-1} = t(P_\ell,p+U)-p^{\ell-1}\geq 0, \quad (24)$$

where the last inequality follows from the fact that P_ℓ has the Sidorenko property and $p+U=W$.

Along the same lines, we derive that the sum of the terms that involve a set F_0 from group (b) (note that such F_0 contain all the edges of C_4 and some edges of $P_\ell^\bullet$, and F_0 can consist of the four edges of C_4 only) is equal to

$$\sum_{F\subseteq E(P_\ell)} p^{\ell-1-|F|-4}t(\langle F\rangle,U)t(C_4,U) = t(P_\ell,p+U)\frac{t(C_4,U)}{p^4} \geq t(P_\ell,p+U)t(C_4,U). \quad (25)$$

It remains to analyze the sum of terms that involve a set F_0 from group (c). For any edge e of $H^\bullet$, the sum of the terms that involve a set F_0 from group (c) such that the core component of $\langle F_0\rangle$ is formed by the edge e only is equal to zero as $t(K_2,U)=0$. Hence, the sum of the terms that involve a set F_0 from group (c) is equal to:

- the sum of the terms that involve such F_0 with $\{v_{-2}v_{-1}, v_{-1}v_0\}\subseteq F_0$;
- the sum of the terms that involve such F_0 with $\{v_{-1}v_0, v_0v_1\}\subseteq F_0$; and
- the sum of the terms that involve such F_0 that the core component of F_0 has at least two edges but does not contain v_0.

The sum of the terms that involve such F_0 with $\{v_{-2}v_{-1}, v_{-1}v_0\}\subseteq F_0$ is equal to (note that $t(K_{1,2},U)\geq 0$ by Lemma 2.5)

$$\sum_{F\subseteq E(P_{\ell-1})} p^{\ell-2-|F|-1}t(\langle F\rangle,U)t(K_{1,2},U) = t(P_{\ell-1},p+U)\frac{t(K_{1,2},U)}{p} \geq 0. \quad (26)$$

Similarly, the sum of the terms that involve such F_0 with $\{v_{-1}v_0, v_0v_1\}\subseteq F_0$ is equal to

$$\sum_{F\subseteq E(P_{\ell-2})} p^{\ell-3-|F|}t(\langle F\rangle,U)t(K_{1,2},U) = t(P_{\ell-2},p+U)t(K_{1,2},U) \geq 0. \quad (27)$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Finally, sum of the terms that involve such F_0 from group (c) such that the core component of F_0 has at least two edges but does not contain v_0 is equal to the following:

$$\begin{aligned}
&\sum_{F\subseteq E(P_\ell)}\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\sum_{d=2}^{\deg_H(v)}\binom{\deg_H(v)}{d}p^{\ell-1-|F|-d}t(\langle F\rangle,U)t(K_{1,d},U)\\
&\quad=\Bigg(\sum_{F\subseteq E(P_\ell)}p^{\ell-1-|F|}t(\langle F\rangle,U)\Bigg)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\sum_{d=2}^{\deg_H(v)}\binom{\deg_H(v)}{d}p^{-d}t(K_{1,d},U)\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\sum_{d=2}^{\deg_H(v)}\binom{\deg_H(v)}{d}p^{-d}t(K_{1,d},U). \qquad (28)
\end{aligned}$$

Since it holds that $t(K_{1,1},U)=t(K_2,U)=0$, we obtain that (28) is equal to

$$\begin{aligned}
&t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\sum_{d=1}^{\deg_H(v)}\binom{\deg_H(v)}{d}p^{-d}t(K_{1,d},U)\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\sum_{d=1}^{\deg_H(v)}\binom{\deg_H(v)}{d}\int_{[0,1]}\left(\frac{t_x(K_2^\bullet,U)}{p}\right)^d dx\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\int_{[0,1]}\left(1+\frac{t_x(K_2^\bullet,U)}{p}\right)^{\deg_H v}-1\,dx. \qquad (29)
\end{aligned}$$

Recall that Pascal's inequality asserts $(1+x)^n\geq 1+nx$ for all integers $n\geq 0$ and reals $x\geq -1$. As $U\geq -p$, it holds that $t_x(K_2^\bullet,U)\geq -p\geq -1$ and so (29) is at least

$$\begin{aligned}
&t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\int_{[0,1]}\left(1+\frac{(\deg_H v-1)t_x(K_2^\bullet,U)}{p}\right)\left(1+\frac{t_x(K_2^\bullet,U)}{p}\right)-1\,dx\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\int_{[0,1]}\frac{\deg_H v\cdot t_x(K_2^\bullet,U)}{p}+\frac{(\deg_H v-1)t_x(K_2^\bullet,U)^2}{p^2}\,dx\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}\deg_H v\,\frac{t(K_2,U)}{p}+(\deg_H v-1)\frac{t(K_{1,2},U)}{p^2}\\
&\quad=t(P_\ell,p+U)\sum_{v\in V(H)\setminus\{v_0,v_{-1}\}}(\deg_H v-1)\frac{t(K_{1,2},U)}{p^2}\geq t(P_\ell,p+U)t(K_{1,2},U), \qquad (30)
\end{aligned}$$

where the last inequality follows since $p\leq 1$ and the degree of at least one vertex of H different from v_0 and v_{-1} is at least two.

Since the quantities estimated in (24), (26) and (27) are non-negative, the sum of the terms that involve a set F_0 from one of the groups (a)–(c) is at least the sum of the final estimates given in (25) and (30), i.e. at least $t(P_\ell,p+U)(t(K_{1,2},U)+t(C_4,U))$. This yields the bound stated in the claim. □

Before we proceed with proving Claims 3.3–3.7, we recall an estimate on the density of P_ℓ in $W=p+U$ from § 2.3 and provide more refined estimates that are needed in the proofs of some of the claims. Let λ_1 be the largest eigenvalue of W viewed as a Hilbert–Schmidt operator on $L_2[0,1]$, let $\gamma=t(C_4,W)-p^4$ and let δ be defined as in § 2.3. Recall that $\gamma\leq 4\varepsilon_0$ by (3),

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$p \le \lambda_1 \le p + \varepsilon_0/p^3$ by (9), and $\delta \le 4\varepsilon_0/p^4$ by (13). By (12), it holds that

$$\lambda_1^m(1-\delta)^2 - \gamma^{m/4} \le t(P_{m+1}, p+U) \le \lambda_1^m(1-\delta)^2 + \gamma^{m/4}$$

for every integer $m \ge 0$. Next, the choice of ε_0 implies that $\gamma^{1/4} \le (4p_0^4/10^{12})^{1/4} \le p_0/700$ and $\delta \le 4\varepsilon_0/p^4 \le 10^{-11}$, and so we obtain that

$$\begin{aligned}
\frac{t(P_m, p+U)}{t(P_{m+1}, p+U)} &\le \frac{\lambda_1^{m-1}(1-\delta)^2 + \gamma^{(m-1)/4}}{\lambda_1^m(1-\delta)^2 - \gamma^{m/4}} \\
&\le \frac{\lambda_1^{m-1}(1-10^{-11})^2 + (p_0/700)^{m-1}}{\lambda_1^m(1-10^{-11})^2 - (p_0/700)^m} \\
&\le \frac{\lambda_1^{m-1}(1-10^{-11})^2 + (\lambda_1/700)^{m-1}}{\lambda_1^m(1-10^{-11})^2 - (\lambda_1/700)^m} \\
&\le \frac{(1-10^{-11})^2 + 1/700}{(1-10^{-11})^2 - 1/700^2} \cdot \frac{1}{\lambda_1} \le \frac{101}{100\lambda_1} \le \frac{101}{100p}
\end{aligned} \tag{31}$$

for every integer $m \ge 2$; since $t(P_1, p+U) = 1$ and $t(P_2, p+U) = p$, the estimate (31) also holds for $m = 1$. Hence, the following holds for every $i = 0, \dots, 8$:

$$t(P_{\ell-i}, p+U) \le \left(\frac{101}{100p}\right)^i t(P_\ell, p+U) \le \frac{2}{p^8} t(P_\ell, p+U) \le \frac{2}{p_0^8} t(P_\ell, p+U). \tag{32}$$

Similarly to (31), the choice of ε_0 implies that the following holds for every integer $m = 2, \dots, \ell$:

$$\frac{t(P_m, p+U)}{t(P_\ell, p+U)^{(m-1)/(\ell-1)}} \le \frac{\lambda_1^{m-1}(1-\delta)^2 + \gamma^{(m-1)/4}}{(\lambda_1^{\ell-1}(1-\delta)^2 - \gamma^{(\ell-1)/4})^{\frac{m-1}{\ell-1}}} \le \frac{101}{100}. \tag{33}$$

If $m = 1$, the left-hand side of (33) is 1, and so the estimate (33) also holds for $m = 1$.

CLAIM 3.3. *The absolute value of the sum of the terms on the left-hand side of* (23) *that involve* F_0 *from group* (d) *is at most*

$$\frac{1}{8} t(P_\ell, p+U) t(C_4, U).$$

Proof of Claim 3.3. Consider a set F_0 from group (d). Let F_c be the edges of F_0 contained in the core components of $\langle F_0 \rangle$, and let k be the largest index such that v_k is contained in a core component of F_0. Note $k \ge 8$ by the definition of group (d). Let $S^\bullet$ be the rooted graph obtained from $\langle F' \rangle$ by removing the path $v_0 \cdots v_k$ and choosing the vertex v_0 to be the root. Observe that

$$\begin{aligned}
|t(\langle F_c \rangle, U)| &= \left| \int_{[0,1]} t_x(P_{k+1}^\bullet, U) t_x(S^\bullet, U)\, dx \right| \\
&\le \left(\int_{[0,1]} t_x(P_{k+1}^\bullet, U)^2\, dx \right)^{1/2} \left(\int_{[0,1]} t_x(S^\bullet, U)^2\, dx \right)^{1/2} \\
&\le \left(\int_{[0,1]} t_x(P_{k+1}^\bullet, U)^2\, dx \right)^{1/2} dx = t(P_{2k+1}, U)^{1/2} \le t(C_4, U)^{k/4},
\end{aligned}$$

where the last inequality follows from Lemma 2.9. As there are at most $2^{e(H)}$ choices of a set F_c such that a core component of F_0 contains v_k but not v_{k+1}, the absolute value of the sum of the terms that involve F_0 from group (d) is at most

$$2^{e(H)} \left(t(C_4, U)^{\ell/4} + \sum_{k=8}^{\ell-1} t(P_{\ell-k}, p+U) t(C_4, U)^{k/4} \right). \tag{34}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Using (33), we estimate the sum in (34) as follows:

$$\begin{aligned} &t(C_4,U)^{\ell/4}+\sum_{k=8}^{\ell-1} t(P_{\ell-k},p+U)t(C_4,U)^{k/4}\\ &\quad\le t(C_4,U)^{\ell/4}+\frac{101}{100}\sum_{k=8}^{\ell-1} t(P_\ell,p+U)^{\frac{\ell-k-1}{\ell-1}}t(C_4,U)^{k/4}\\ &\quad\le \frac{101}{100}\sum_{k=8}^{\ell} t(P_\ell,p+U)^{\frac{\ell-k-1}{\ell-1}}t(C_4,U)^{k/4}. \end{aligned}$$

We next estimate the last sum using that $t(P_\ell,p+U)\ge p^{\ell-1}$:

$$\begin{aligned} \frac{101}{100}\sum_{k=8}^{\ell} t(P_\ell,p+U)^{\frac{\ell-k-1}{\ell-1}}t(C_4,U)^{k/4} &\le \frac{101}{100}t(P_\ell,p+U)\sum_{k=8}^{\ell}\left(\frac{t(C_4,U)^{1/4}}{t(P_\ell,p+U)^{\frac{1}{\ell-1}}}\right)^k\\ &\le \frac{101}{100}t(P_\ell,p+U)\sum_{k=8}^{\ell}\left(\frac{t(C_4,U)^{1/4}}{p}\right)^k\\ &= \frac{101}{100}t(P_\ell,p+U)\frac{t(C_4,U)^2}{p^8}\sum_{m=0}^{\ell-8}\left(\frac{t(C_4,U)^{1/4}}{p}\right)^m\\ &\le \frac{101}{100}t(P_\ell,p+U)\frac{t(C_4,U)^2}{p^8}\sum_{m=0}^{\ell-8}\left(\frac{(4\varepsilon_0)^{1/4}}{p}\right)^m, \end{aligned}$$

which we further estimate using $\varepsilon_0\le p_0^4/10^{12}$ to be at most

$$\begin{aligned} \frac{101}{100}t(P_\ell,p+U)\frac{t(C_4,U)^2}{p^8}\sum_{m=0}^{\ell-8}\left(\frac{p_0}{100p}\right)^m &\le \frac{101}{100}t(P_\ell,p+U)\frac{t(C_4,U)^2}{p^8}\sum_{m=0}^{\infty}\left(\frac{1}{100}\right)^m\\ &\le 2t(P_\ell,p+U)\frac{t(C_4,U)^2}{p^8}. \end{aligned}$$

We now plug in this estimate to (34) and derive using (3) and the choice of ε_0 that the absolute value of the sum of the terms that involve F_0 from group (d) is at most

$$2^{e(H)+1}\cdot\frac{t(P_\ell,p+U)t(C_4,U)^2}{p^8}\le 2^{e(H)+1}\cdot\frac{4\varepsilon_0 t(P_\ell,p+U)t(C_4,U)}{p^8}\le\frac{1}{8}t(P_\ell,p+U)t(C_4,U).$$

The proof of the claim is now completed. □

Claim 3.4. *The absolute value of the sum of the terms on the left-hand side of (23) that involve F_0 from group (e) is at most*

$$\tfrac{1}{8}t(P_\ell,p+U)(t(P_3,U)+t(C_4,U)).$$

Proof of Claim 3.4. Consider a set F_0 such that a core component of $\langle F_0\rangle$ contains a cycle of G. Let F_c be the edges of F_0 contained in the core components of $\langle F_0\rangle$, let C be the shortest cycle of G contained in $\langle F_c\rangle$, let $L\ge 50$ be the length of C, and let q be the number of vertices of degree at least three in $\langle F_c\rangle$ that are contained in C. If $q\le 7$, a part of the cycle C is an 8-edge path whose all internal vertices have degree two in $\langle F_c\rangle$. Let $S^{\bullet\bullet}$ be the rooted graph obtained from $\langle F_c\rangle$ by removing the edges of such an 8-edge path and choosing the end vertices of the removed path to be two roots of $S^{\bullet\bullet}$. We obtain using the Cauchy–Schwarz inequality and Lemma 2.8

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

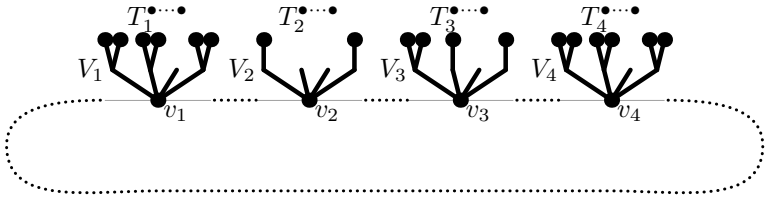


FIGURE 4. Illustration of the definition of v_i, V_i and $T_i^{\bullet\cdots\bullet}$ in the proof of Claim 3.4; roots are the vertices depicted in the figure.

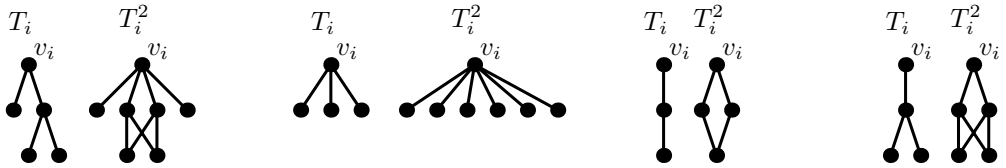


FIGURE 5. Examples of rooted trees $T_i^{\bullet\cdots\bullet}$ and the associated trees T_i^2 in the proof of Claim 3.4.

that

$$\begin{aligned}|t(\langle F_c\rangle, U)| &= \left| \int_{[0,1]^2} t_{x,x'}(P_9^{\bullet\bullet}, U) t_{x,x'}(S^{\bullet\bullet}, U)\, dx\, dx' \right| \\ &\leq \left(\int_{[0,1]} t_{x,x'}(P_9^{\bullet\bullet}, U)^2\, dx\, dx' \right)^{1/2} \left(\int_{[0,1]} t_{x,x'}(S^{\bullet\bullet}, U)^2\, dx\, dx' \right)^{1/2} \\ &\leq t(C_{16}, U)^{1/2} \leq t(C_4, U)^2. \end{aligned} \tag{35}$$

We next assume that $q \geq 8$. Let $v_1, \ldots, v_4$ be any four vertices of C with degree at least three in $\langle F_c\rangle$ such that no two of them are consecutive on the cycle C; the choice of C as the shortest cycle contained in G implies that the vertices $v_1, \ldots, v_4$ form an independent set. Let V_i for $i \in [4]$ be the set consisting of the vertex v_i and the vertices at distance at most two from v_i in $\langle F_c\rangle$ that are not contained in the cycle C; see Figure 4. Observe that the sets $V_1, \ldots, V_4$ are disjoint and there is no edge joining a vertex of V_i and V_j for $i \neq j$; otherwise, $\langle F_c\rangle$ would contain a cycle shorter than C, which would be formed by a path from v_i through V_i and then V_j back to v_j and the shorter of the two parts of C delimited by v_i and v_j. For $i = 1, \ldots, 4$, let $T_i^{\bullet\cdots\bullet}$ be the rooted graph obtained from the subgraph of $\langle F_c\rangle$ induced by V_i by choosing the vertex v_i and all vertices at distance exactly two from v_i to be the roots. Furthermore, let $S^{\bullet\cdots\bullet}$ be the graph obtained from $\langle F_c\rangle$ by deleting the non-root vertices of $T_1^{\bullet\cdots\bullet}, \ldots, T_4^{\bullet\cdots\bullet}$ and choosing the other vertices shared with $T_1^{\bullet\cdots\bullet}, \ldots, T_4^{\bullet\cdots\bullet}$ as roots. Let m be the number of roots of $S^{\bullet\cdots\bullet}$. Observe that it holds that (we omit the exact subscripts at the functions t for clarity)

$$|t(\langle F_c\rangle, U)| = \left| \int_{[0,1]^m} t_{\ldots}(S^{\bullet\cdots\bullet}, U) \prod_{i=1}^{4} t_{\ldots}(T_i^{\bullet\cdots\bullet}, U)\, dx_1 \cdots dx_m \right|.$$

For $i = 1, \ldots, 4$, we write T_i^2 for the graph obtained by taking two copies of $T_i^{\bullet\cdots\bullet}$ and identifying the pairs of corresponding roots; see Figure 5. By Lemma 2.7, we obtain using that $\|U\|_\infty \leq 1$ that

$$|t(\langle F_c\rangle, U)| \leq \left(\int_{[0,1]^m} t_{\ldots}(S^{\bullet\cdots\bullet}, U)^2\, dx_1 \cdots dx_m \right)^{1/2} \prod_{i=1}^{4} \left(\int_{[0,1]^m} t_{\ldots}(T_i^{\bullet\cdots\bullet}, U)^2\, dx_1 \cdots dx_m \right)^{1/2}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$$\leq \prod_{i=1}^{4} \left(\int_{[0,1]^m} t_{\dots}(T_i^{\bullet\cdots\bullet}, U)^2 \, dx_1 \cdots dx_m \right)^{1/2} = \prod_{i=1}^{4} t(T_i^2, U)^{1/2}. \tag{36}$$

Consider $i \in [4]$. If the vertex v_i has degree one in $T_i^{\bullet\cdots\bullet}$, then T_i^2 is the complete bipartite graph $K_{2,d}$ where d is the number of roots of $T_i^{\bullet\cdots\bullet}$. If $d = 1$, then $t(T_i^2, U) = t(P_3, U)$. If $d \geq 2$, we obtain using Lemma 2.6 that $t(T_i^2, U) \leq t(C_4, U)$. If the vertex v_i has a neighbor of degree one in $T_i^{\bullet\cdots\bullet}$, it holds that $t_{x_1,\dots,x_r}(T_i^{\bullet\cdots\bullet}, U)^2 \leq t_{x_1}(P_2^{\bullet}, U)^2$ (assuming that x_1 corresponds to the root vertex v_i) and so $t(T_i^2, U) \leq t(P_3, U)$. Otherwise, T_i^2 satisfies the assumptions of Lemma 2.10 and we obtain that

$$t(T_i^2, U) \leq t(C_4, U)^{5/4} \leq t(C_4, U).$$

In all three cases, it holds that

$$t(T_i^2, U) \leq \max\{t(P_3, U), t(C_4, U)\} = t(P_3, U) + t(C_4, U).$$

Hence, we obtain from (36) that

$$|t(\langle F_c \rangle, U)| \leq (t(P_3, U) + t(C_4, U))^2. \tag{37}$$

As there are at most $2^{e(H)+7}$ possible sets F_c such that an associated set F_0 can be contained in group (e), we conclude using (32), (35) and (37) that the absolute value of the sum of the terms that involve F_0 from group (e) is at most

$$\frac{2}{p_0^8} t(P_\ell, p + U) 2^{e(H)+7} (t(P_3, U) + t(C_4, U))^2. \tag{38}$$

The choice of ε_0 implies using (3) that

$$t(C_4, U) \leq 4p_0^{64}/2^{8e(H)+90} \tag{39}$$

and so

$$t(P_3, U) \leq t(C_4, U)^{1/2} \leq 2p_0^{32}/2^{4e(H)+45}, \tag{40}$$

which yields that (38) does not exceed $\frac{1}{8} t(P_\ell, p + U)(t(P_3, U) + t(C_4, U))$. □

CLAIM 3.5. *The absolute value of the sum of the terms on the left-hand side of (23) that involve F_0 from group (f) is at most*

$$\tfrac{1}{8} t(P_\ell, p + U)(t(P_3, U) + t(C_4, U)).$$

Proof of Claim 3.5. Consider a set F_0 from group (f) and let F_c be the edges of F_0 contained in the core components of $\langle F_0 \rangle$. Let m be the number of edges of the 12-edge path from C_4 to the graph G in H that are contained in the component of $\langle F_0 \rangle$ that contains C_4. Note that $m \geq 1$ by the definition of group (f). If $m \leq 7$, we obtain using $\|U\|_\infty \leq 1$ and Lemma 2.12 that

$$t(\langle F_c \rangle, U) \leq \tfrac{1}{2}(t(C_4, U) + t(P_3, U)) t(C_4, U)^{1/8}. \tag{41}$$

If $m \geq 8$, let $S^{\bullet\bullet}$ be the rooted graph obtained from $\langle F_c \rangle$ by removing the edges of an 8-edge subpath of the m-edge path starting at C_4 and choosing the end vertices of the path to be two roots of $S^{\bullet\bullet}$. As in (35), we obtain that

$$|t(\langle F_c \rangle, U)| \leq t(C_4, U)^2. \tag{42}$$

The estimates (41) and (42) imply that the following holds in either of the cases:

$$|t(\langle F_c \rangle, U)| \leq (t(C_4, U) + t(P_3, U)) t(C_4, U)^{1/8}.$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

As there are at most $2^{e(H)+7}$ possible sets F_c such that an associated F_0 can be contained in group (f), we conclude using (32) and the choice of ε_0, which implies that (39) holds, that the absolute value of the sum of the terms that involve F_0 from group (f) is at most

$$\frac{2}{p_0^8}t(P_\ell, p+U)2^{e(H)+7}(t(P_3,U)+t(C_4,U))t(C_4,U)^{1/8} \le \frac{1}{8}t(P_\ell,p+U)(t(P_3,U)+t(C_4,U)).$$

This completes the proof of the claim. □

CLAIM 3.6. *The absolute value of the sum of the terms on the left-hand side of* (23) *that involve* F_0 *from group (g) is at most*

$$\tfrac{1}{8}t(P_\ell,p+U)t(P_3,U).$$

Proof of Claim 3.6. Consider a set F_0 from group (g) and let F_c be the edges of F_0 contained in the core components of $\langle F_0\rangle$. We obtain using Lemma 2.11 that

$$|t(\langle F_c\rangle,U)| \le t(P_3,U)t(C_4,U)^{1/4}.$$

As there are at most $2^{e(H)+7}$ possible sets F_c such that an associated set F_0 can be contained in group (g), we conclude using (32) and the choice of ε_0, which implies that (39) holds, that the absolute value of the sum of the terms that involve F_0 from group (g) is at most

$$\frac{2}{p_0^8}t(P_\ell,p+U)2^{e(H)+7}t(P_3,U)t(C_4,U)^{1/4} \le \frac{1}{8}t(P_\ell,p+U)t(P_3,U).$$

The proof of the claim is now completed. □

CLAIM 3.7. *The absolute value of the sum of the terms on the left-hand side of* (23) *that involve* F_0 *from group (h) is at most*

$$\tfrac{1}{8}t(P_\ell,p+U)t(P_3,U).$$

Proof of Claim 3.7. Consider a set F_0 from group (h) and let F_c be the edges of F_0 contained in the core components of $\langle F_0\rangle$, and let S_1 and S_2 be any two core components of $\langle F\rangle$. If S_i is a single edge, then $t(S_i,U)=0$, and if S_i is C_4, then $t(S_i,U)=t(C_4,U)$. Finally, if S_i is a star with at least two leaves, we obtain using Lemma 2.6 that

$$|t(S_i,U)| \le t(K_1,U)=t(P_3,U).$$

Hence, we obtain that

$$|t(\langle F_c\rangle,U)| \le |t(S_1,U)t(S_2,U)| \le t(P_3,U)(t(P_3,U)+t(C_4,U)).$$

As there are at most $2^{e(H)+7}$ possible sets F_c such that an associated set F_0 can be contained in group (h), we conclude using (32) and the choice of ε_0, which implies that (40) holds, that the absolute value of the sum of the terms that involve F_0 from group (h) is at most

$$\frac{2}{p_0^8}t(P_\ell,p+U)2^{e(H)+7}t(P_3,U)(t(P_3,U)+t(C_4,U)) \le \frac{1}{8}t(P_\ell,p+U)t(P_3,U),$$

which completes the proof of the claim. □

Summing the bounds obtained in Claims 3.3–3.7 yields the following.

CLAIM 3.8. *The absolute value of the sum of the terms on the left-hand side of* (23) *that involve a set* F_0 *from one of the groups (d)–(h) is at most*

$$t(P_\ell,p+U)\left(\frac{t(P_3,U)}{2}+\frac{3t(C_4,U)}{8}\right).$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Since we have proven both Claims 3.2 and 3.8, the proof of the theorem is now complete. □

4. Local regime

In this section, we prove that the graph $H^\bullet \oplus K_{m|\ell,n}$ is locally Sidorenko (assuming W has density bounded away from zero), where the bound on locality does not depend on the parameters ℓ, m and n, under the assumption that all three parameters are sufficiently large. Similarly to the proof of Theorem 3.1, the proof of Theorem 4.1 is split into several separate claims, which are stated and proven inside the proof of the theorem since they rely on the notation introduced at the beginning of the proof of the theorem.

THEOREM 4.1. *For every graph G with girth at least 50 and every real number $p_0 \in (0,1)$, there exist a real $\gamma_0 > 0$, a positive integer n_0 and a connected rooted graph $H^\bullet$ that contains G as an induced subgraph such that the following holds for graphon W with density $p \geq p_0$ such that $t(C_4, W) - p^4 \leq \gamma_0$:*

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) \geq p^{e(H)+mn+\ell} \tag{43}$$

for all even integers $m, n, \ell \geq n_0$ such that m is divisible by 5 and $\ell \geq n + e(H)$.

Proof. We first apply Theorem 3.1 with G and p_0 to get a connected rooted graph $H^\bullet$ that contains G as an induced subgraph and a real $\varepsilon_H > 0$. We prove that the statement of the theorem holds with

$$\gamma_0 = \min\left\{\frac{\varepsilon_H^4}{8}, \frac{p_0^{160}}{2^{200}}\right\} \quad \text{and} \quad n_0 = 64.$$

Note that Theorem 3.1 implies that

$$t(H^\bullet \oplus P^\bullet_{\ell+1}, W) \geq p^{e(H)+\ell} \tag{44}$$

for every graphon W with density $p \geq p_0$ such that $t(C_4, W) - p^4 \leq \gamma_0$, and for every integer $\ell \geq 8$; note that $\|W - p\|_\square \leq \varepsilon_H$ indeed holds for any such graphon by Lemma 2.1.

Fix even integers $m, n, \ell \geq n_0$ such that m is divisible by 5 and $\ell \geq n + e(H) \geq n_0 \geq 64$. Assume that (43) fails for $H^\bullet$, m, n, ℓ and γ_0, and let q be the supremum over all $p \in [p_0, 1]$ such that there exists a graphon W with density p and $t(C_4, W) - p^4 \leq \gamma_0$ such that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) < p^{e(H)+mn+\ell}. \tag{45}$$

Fix such a graphon W with density $p > q/(1 + p_0^{mn}/4)$ for the rest of the proof. Observe that the choices of q and p implies that every graphon W' with density $p' \geq p(1 + p_0^{mn}/4)$ and $t(C_4, W) - (p')^4 \leq \gamma_0$ satisfies (43).

We next analyze spectral properties of the operator associated with the graphon W using results presented in § 2.3. Let λ_i be the non-zero eigenvalues of W listed in the non-increasing order of the absolute value (with multiplicities), and let f_i be the corresponding orthonormal eigenfunctions. Furthermore, let $\alpha_i \in [0, \pi/2]$ be the real such that $\langle \mathrm{j}, f_i \rangle = \cos \alpha_i$, and set $\delta = 1 - \cos \alpha_1$. Finally, let $\gamma = t(C_4, W) - p^4$. Note that $\gamma \leq \gamma_0$.

We start with giving a lower bound on λ_1 in terms of p and δ. We obtain using (16) and $\gamma \leq p_0^4/2^{20}$ that

$$\lambda_1 \geq p(1 + 2\delta) - 8\delta\gamma^{1/4} \geq p(1 + 2\delta) - \delta p_0/4 \geq p(1 + 7\delta/4). \tag{46}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

On the other hand, it holds that $\lambda_1 \leq p + \gamma/(4p^3)$ by (9). Hence, we obtain using (46) and $\gamma \leq p_0^5/2^{20}$ that

$$\delta \leq \frac{4}{7} \cdot \frac{\lambda_1 - p}{p} \leq \frac{\gamma}{7p^3} \leq \frac{p_0^2}{2^{20}} \leq \frac{1}{2^{20}}. \tag{47}$$

Let $g_H : [0,1] \to [0,1]$ be defined as $g_H(x) = t_x(H^\bullet, W)$ and let $g_K : [0,1] \to [0,1]$ be defined as $g_K(x) = t_x(K^\bullet_{m,n}, W)$. Note that it holds that $g_K(x) = t_x(K^\bullet_{m,n}, W) \leq t(K_{m-1,n}, W)$ for every $x \in [0,1]$, which implies that

$$\|g_K\|_2 \leq \|g_K\|_\infty \leq t(K_{m-1,n}, W). \tag{48}$$

Set $\sigma_i = \langle g_H, f_i \rangle$ and $\kappa_i = \langle g_K, f_i \rangle$. Since $\|g_H\|_2 \leq \|g_H\|_\infty \leq 1$ and $\|g_K\|_2 \leq \|g_K\|_\infty \leq 1$, we obtain that $|\sigma_i| \leq 1$ and $|\kappa_i| \leq 1$ for every i. Next observe that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) = \langle g_H, W^\ell g_K \rangle = \sigma_1 \kappa_1 \lambda_1^\ell + \sum_{i \geq 2} \sigma_i \kappa_i \lambda_i^\ell. \tag{49}$$

In the next two claims, we give lower bounds on the product $\sigma_1 \lambda_1^\ell$ and κ_1.

CLAIM 4.2. *For every $\ell' \geq 8$, it holds that*

$$\sigma_1 \lambda_1^{\ell'} \geq p^{e(H)+\ell'} - 2\gamma^{\ell'/4}\delta^{1/2}.$$

Proof of Claim 4.2. Consider a positive integer $\ell' \geq 8$. The estimate (44) implies that

$$p^{e(H)+\ell'} \leq t(H^\bullet \oplus P^\bullet_{\ell'+1}, W) = \sigma_1 \lambda_1^{\ell'}(1-\delta) + \sum_{i \geq 2} \sigma_i \lambda_i^{\ell'} \cos \alpha_i. \tag{50}$$

We combine (50) with (11) and (15) to obtain that

$$\begin{aligned} p^{e(H)+\ell'} \leq \sigma_1 \lambda_1^{\ell'}(1-\delta) + \sum_{i \geq 2} \sigma_i \lambda_i^{\ell'} \cos \alpha_i &\leq \sigma_1 \lambda_1^{\ell'} + \sum_{i \geq 2} |\lambda_i|^{\ell'} \cdot |\cos \alpha_i| \\ &\leq \sigma_1 \lambda_1^{\ell'} + \left(\sum_{i \geq 2} \lambda_i^{2\ell'} \right)^{1/2} \left(\sum_{i \geq 2} \cos^2 \alpha_i \right)^{1/2} \\ &\leq \sigma_1 \lambda_1^{\ell'} + 2\gamma^{\ell'/4}\delta^{1/2}. \end{aligned}$$

Hence, we can conclude that

$$\sigma_1 \lambda_1^{\ell'} \geq p^{e(H)+\ell'} - 2\gamma^{\ell'/4}\delta^{1/2} \tag{51}$$

holds for every $\ell' \geq 8$. The estimate given in the statement of the claim now follows. □

CLAIM 4.3. *It holds that*

$$\kappa_1 \geq t(K_{m-1,n}, W)((p^4 + \gamma)^{n/4} - 2\delta^{1/2}).$$

Proof of Claim 4.3. We first obtain using the Cauchy–Schwarz inequality that

$$t(K_{m,n}, W) = \langle g_K, \mathbb{j} \rangle = \sum_{i \geq 1} \kappa_i \cos \alpha_i \leq \kappa_1 \cos \alpha_1 + (1 - \cos^2 \alpha_1)^{1/2} \|g_K\|_2,$$

which implies using the definition of δ and (48) that

$$t(K_{m,n}, W) \leq \kappa_1 \cos \alpha_1 + 2\delta^{1/2} t(K_{m-1,n}, W).$$

Hence, it follows using Lemma 2.3 that

$$\kappa_1 \geq t(K_{m,n}, W) - 2\delta^{1/2} t(K_{m-1,n}, W) \geq t(K_{m-1,n}, W)(t(K_{m-1,n}, W)^{1/(m-1)} - 2\delta^{1/2}).$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Since Lemma 2.4 implies that

$$t(K_{m-1,n},W)^{1/(m-1)} \geq t(C_4,W)^{n/4} = (p^4+\gamma)^{n/4},$$

the statement of the claim now follows. □

We next distinguish the following three cases, which are covered in Claims 4.4, 4.6 and 4.7, respectively. In each of the cases, we show that the fixed graphon W satisfies (43), which would contradict the choice of W and so complete the proof of the theorem.

(i) It holds that $\delta \leq \frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$.

(ii) It holds that $\delta > \frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$ and $\|g_K\|_2 \leq \frac{\kappa_1}{\gamma^{\ell/8}}$.

(iii) It holds that $\delta > \frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$ and $\|g_K\|_2 > \frac{\kappa_1}{\gamma^{\ell/8}}$.

Intuitively speaking, case (i) corresponds to the situation when the degrees of almost all vertices of W are the same, which is equivalent to that δ is small. Case (ii) corresponds to the situation when the sum in (49) is small compared with the leading term, i.e. the term $\sigma_1\kappa_1\lambda_1^\ell$ controls the expression in (49). In case (iii), which covers the cases not covered in cases (i) and (ii), we show that the reason for the sum in (49) being large comparatively to the leading term is that the graphon W contains a part denser than q where we identify sufficiently many copies of $H^\bullet \oplus K^\bullet_{m|\ell,n}$; here, we use the choice of q as the supremum among the densities of graphons violating (43) for fixed m, n and ℓ.

CLAIM 4.4. *If δ is at most $\frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$, then (43) holds for m, n, ℓ and the graphon W.*

Proof of Claim 4.4. We start with refining the bound given in Claim 4.3 using the assumption that $\delta \leq \frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$, which is equivalent to $\frac{40\delta^{1/2}p^4}{np^n} \leq \gamma$, and Pascal's inequality:

$$\begin{aligned}
\kappa_1 &\geq t(K_{m-1,n},W)((p^4+\gamma)^{n/4} - 2\delta^{1/2}) \\
&\geq t(K_{m-1,n},W)\left(p^n\left(1+\frac{n\gamma}{4p^4}\right) - 2\delta^{1/2}\right) \\
&\geq t(K_{m-1,n},W)\left(p^n\left(1+\frac{10\delta^{1/2}}{p^n}\right) - 2\delta^{1/2}\right) \\
&= t(K_{m-1,n},W)(p^n + 8\delta^{1/2}) \geq t(K_{m-1,n},W)p^n(1+8\delta^{1/2}).
\end{aligned} \tag{52}$$

Hence, we conclude using Claim 4.2, the bound $\gamma \leq p^8/2^{20}$, (52) and $\ell \geq e(H)$ that the first term in (49) can be bounded from below as follows:

$$\begin{aligned}
\sigma_1\kappa_1\lambda_1^\ell &\geq (p^{e(H)+\ell} - 2\gamma^{\ell/4}\delta^{1/2})t(K_{m-1,n},W)p^n(1+8\delta^{1/2}) \\
&\geq (p^{e(H)+\ell} - p^{2\ell}\delta^{1/2}/16)t(K_{m-1,n},W)p^n(1+8\delta^{1/2}) \\
&= t(K_{m-1,n},W)p^n(p^{e(H)+\ell} + 8p^{e(H)+\ell}\delta^{1/2} - p^{2\ell}\delta^{1/2}/16 - p^{2\ell}\delta/2) \\
&\geq t(K_{m-1,n},W)p^n(p^{e(H)+\ell} + 8p^{e(H)+\ell}\delta^{1/2} - p^{2\ell}\delta^{1/2}) \\
&\geq t(K_{m-1,n},W)p^n(p^{e(H)+\ell} + 8p^{2\ell}\delta^{1/2} - p^{2\ell}\delta^{1/2}) \geq t(K_{m-1,n},W)p^{e(H)+\ell+n}.
\end{aligned} \tag{53}$$

On the other hand, the absolute value of the sum in (49) can be bounded from above using Cauchy–Schwarz inequality, (11) and (48) as follows:

$$\left|\sum_{i\geq 2}\sigma_i\kappa_i\lambda_i^\ell\right| \leq \left(\sum_{i\geq 2}\kappa_i^2\sigma_i^2\right)^{1/2}\left(\sum_{i\geq 2}\lambda_i^{2\ell}\right)^{1/2}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$$\leq \left(\sum_{i\geq 1}\kappa_i^2\sigma_i^2\right)^{1/2}\left(\sum_{i\geq 1}\lambda_i^{2\ell}\right)^{1/2}$$
$$= \|g_K\|_2\cdot\gamma^{\ell/4} \leq t(K_{m-1,n},W)\gamma^{\ell/4}. \tag{54}$$

Hence, we conclude using (49), (53), (54), Lemma 2.4, Pascal's inequality, $\gamma\leq p^{16}/2^{32}$ and $\ell\geq n+e(H)\geq 8$ that

$$\begin{aligned}
t(H^\bullet\oplus K^\bullet_{m|\ell,n},W) &\geq t(K_{m-1,n},W)(p^{e(H)+\ell+n}-\gamma^{\ell/4})\\
&\geq (p^4+\gamma)^{(m-1)n/4}(p^{e(H)+\ell+n}-\gamma^{1+\ell/8})\\
&\geq p^{(m-1)n}\left(1+\frac{\gamma}{p^4}\right)^{(m-1)n/4}(p^{e(H)+\ell+n}-\gamma^{1+\frac{\ell+n+e(H)}{16}})\\
&\geq p^{(m-1)n}\left(1+\frac{(m-1)n\gamma}{4p^4}\right)\left(p^{e(H)+\ell+n}-\frac{\gamma p^{e(H)+\ell+n}}{4}\right)\\
&\geq p^{e(H)+\ell+mn}\left(1+\frac{\gamma}{2}\right)\left(1-\frac{\gamma}{4}\right)\geq p^{e(H)+\ell+mn},
\end{aligned}$$

which completes the proof of the claim. □

Before stating and proving Claims 4.6 and 4.7, we show that if δ is large, then κ_1 is also large as stated in the next claim.

CLAIM 4.5. *If δ is larger than $\frac{\gamma^2n^2p^{2n}}{1600p^8}$, then $\kappa_1\geq p^{mn}$.*

Proof of Claim 4.5. We start with a bound on the density of $K_{1,2}$:

$$t(K_{1,2},W)=\langle \mathrm{j},W^2\mathrm{j}\rangle=\sum_{i\geq 1}\lambda_i^2\cos^2\alpha_i\geq\lambda_1^2\cos^2\alpha_1=\lambda_1^2(1-\delta)^2.$$

Since $mn/10$ is an even integer, which follows from the facts that m and n are even integers and m is divisible by 5, Lemma 2.13 implies that

$$t(K_{m|mn/10,n},W)\geq\lambda_1^{11mn/10}(1-\delta)^{11mn/10}.$$

On the other hand, we obtain using (11) that

$$\begin{aligned}
t(K_{m|mn/10,n},W)=\langle \mathrm{j},W^{mn/10}g_K\rangle &= \kappa_1(1-\delta)\lambda_1^{mn/10}+\sum_{i\geq 2}\kappa_i\cos\alpha_i\lambda_i^{mn/10}\\
&\leq \kappa_1\lambda_1^{mn/10}+\sum_{i\geq 2}|\lambda_i|^{mn/10}\\
&\leq \kappa_1\lambda_1^{mn/10}+\gamma^{mn/40}.
\end{aligned}$$

We obtain by combining the just obtained lower and upper bounds on $t(K_{m|mn/10,n},W)$ that

$$\kappa_1\geq\lambda_1^{mn}(1-\delta)^{11mn/10}-\gamma^{mn/40}/\lambda_1^{mn/10}. \tag{55}$$

We next plug the estimate (46) to (55) and we derive using $\gamma\leq p^{160}$, Pascal's inequality, and $n_0\geq 20$ that

$$\begin{aligned}
\kappa_1 &\geq p^{mn}(1+7\delta/4)^{mn}(1-\delta)^{11mn/10}-\gamma^{mn/40}/p^{mn/10}\\
&\geq p^{mn}((1+7\delta/4)^{10}(1-\delta)^{11})^{mn/10}-\gamma^{mn/160}p^{2mn}\\
&\geq p^{mn}((1+70\delta/4)(1-11\delta))^{mn/10}-\gamma^2p^{2n+mn}\\
&\geq p^{mn}(1+70\delta/4-11\delta-770\delta^2/4)^{mn/10}-\gamma^2p^{2n+mn},
\end{aligned}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

which yields using $\delta \le 2^{-20}$ given by (47) that

$$\kappa_1 \ge p^{mn}(1+4\delta)^{mn/10} - \gamma^2 p^{2n+mn} \ge p^{mn}\left(1 + \frac{2\delta mn}{5} - \gamma^2 p^{2n}\right).$$

We now use the assumption from the statement of the claim that $\delta > \gamma^2 n^2 p^{2n}/1600$, and the fact that $n_0 \ge 20$ to derive that

$$\kappa_1 \ge p^{mn}\left(1 + \frac{\gamma^2 mn^3 p^{2n}}{4000} - \gamma^2 p^{2n}\right) \ge p^{mn}, \tag{56}$$

which completes the proof of the claim. □

We are now ready to analyze cases (ii) and (iii), which are covered in the next two claims.

CLAIM 4.6. *If δ is larger than $\frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$ and $\|g_K\|_2$ is at most $\frac{\kappa_1}{\gamma^{\ell/8}}$, then (43) holds for m, n, ℓ and the graphon W.*

Proof of Claim 4.6. We first plug the assumption $\|g_K\|_2 \le \frac{\kappa_1}{\gamma^{\ell/8}}$ to (49) and obtain using (11) and Cauchy–Schwarz inequality that

$$\begin{aligned} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) &= \sigma_1\kappa_1\lambda_1^\ell + \sum_{i\ge 2}\sigma_i\kappa_i\lambda_i^\ell \\ &\ge \sigma_1\kappa_1\lambda_1^\ell - \sum_{i\ge 2}|\kappa_i|\cdot|\lambda_i|^\ell \\ &\ge \sigma_1\kappa_1\lambda_1^\ell - \|g_K\|_2\gamma^{\ell/4} \ge \kappa_1(\sigma_1\lambda_1^\ell - \gamma^{\ell/8}). \end{aligned} \tag{57}$$

We next estimate $\sigma_1\lambda_1^\ell$ using Claim 4.2 with $\ell' = \ell - 1$, (46) and $\ell \ge 10$, as follows:

$$\begin{aligned} \sigma_1\lambda_1^\ell = \sigma_1\lambda_1^{\ell-1}\lambda_1 &\ge (p^{e(H)+\ell-1} - 2\gamma^{(\ell-1)/4}\delta^{1/2})p(1+7\delta/4) \\ &\ge p^{e(H)+\ell}\left(1 - \frac{2\gamma^{(\ell-1)/4}\delta^{1/2}}{p^{e(H)+\ell-1}}\right)(1+7\delta/4) \\ &\ge p^{e(H)+\ell}(1-2\gamma^{\ell/8}\delta^{1/2})(1+7\delta/4). \end{aligned} \tag{58}$$

We obtain using Claim 4.5, (57), (58), $\gamma \le p^{64}$ and $\ell \ge n + e(H)$ that

$$\begin{aligned} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) &\ge p^{mn}(\sigma_1\lambda_1^\ell - \gamma^{\ell/8}) \\ &\ge p^{mn}(p^{e(H)+\ell}(1-2\gamma^{\ell/8}\delta^{1/2})(1+7\delta/4) - \gamma^{\ell/8}) \\ &\ge p^{e(H)+\ell+mn}((1-\gamma^{\ell/16}\delta^{1/2}/2^{10})(1+7\delta/4) - \gamma^{\ell/16}). \end{aligned} \tag{59}$$

We next estimate $\gamma^{\ell/16}$ using that the assumption that $\frac{\gamma^2 n^2 p^{2n}}{1600 p^8} < \delta$, $n \ge 64$ and $\ell \ge n$ (and $\ell \ge 64$) as follows:

$$\gamma^{\ell/16} \le \gamma^2\gamma^{\ell/32} \le \frac{1600\delta}{n^2 p^{2n}} p^{160\ell/32} \le \frac{\delta}{p^{2n}} p^{160n/32} \le \delta.$$

Hence, we obtain from (59) that

$$\begin{aligned} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) &\ge p^{e(H)+\ell+mn}((1-\delta^{3/2}/2^{10})(1+7\delta/4) - \delta) \\ &\ge p^{e(H)+\ell+mn}(1+7\delta/4 - \delta^{3/2}/2^{10} - 7\delta^{5/2}/2^{12} - \delta) \ge p^{e(H)+\ell+mn}, \end{aligned}$$

which yields the statement of the claim. □

CLAIM 4.7. *If δ is larger than $\frac{\gamma^2 n^2 p^{2n}}{1600 p^8}$ and $\|g_K\|_2$ is larger than $\frac{\kappa_1}{\gamma^{\ell/8}}$, then (43) holds for m, n, ℓ and the graphon W.*

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Proof of Claim 4.7. Our aim is to show that W contains a significantly denser part and use this part to obtain a contradiction with the choice of q at the beginning of the proof. Observe that we can assume the eigenfunction f_1 is non-negative since the graphon W is non-negative (if needed, change f_1 on a set of measure zero). We define subsets S and S' of $[0,1]$ (also see Figure 6) as follows:

$$S = \{x \in [0,1] \text{ such that } f_1(x) \leq \gamma^{\ell/100}\},$$
$$S' = \{x \in [0,1] \text{ such that } f_1(x) \leq 1 - \delta^{1/4}\}.$$

Note that $\gamma^{\ell/100} \leq \gamma_0^{n_0/100} \leq 2^{-128}$ by the choice of n_0 and γ_0 and $\delta^{1/4} \leq 1/32$ by (47), and so it holds that $S \subseteq S'$.

We start with bounding the measures of the sets S and S' from above. Since $\|f_1\|_1 = \langle f_1, \mathbb{j}\rangle = 1-\delta$ and $0 \leq f_1(x) \leq 1$ for all $x \in [0,1]$, we obtain that $\mu(S)\gamma^{\ell/100} + 1 - \mu(S) \geq 1 - \delta$, which implies using (47) that

$$\mu(S) \leq \frac{\delta}{1-\gamma^{\ell/100}} \leq \frac{p_0}{2^{20}(1-\gamma^{\ell/100})} \leq \frac{p}{2^{16}}. \tag{60}$$

Similarly, it holds that $\mu(S')(1-\delta^{1/4}) + 1 - \mu(S') \geq 1 - \delta$, which implies that

$$\mu(S') \leq \delta^{3/4}. \tag{61}$$

Our next step is proving the following lower bound on the measure of the set S:

$$\mu(S) \geq (1-\gamma^{\ell/20})p^{mn} \geq \frac{p^{mn}}{2}. \tag{62}$$

Suppose that (62) does not hold, i.e. $\mu(S) < (1-\gamma^{\ell/20})p^{mn}$. We first bound κ_1 as follows (we use the Sidorenko property of $K_{m,n}$, i.e. $t(K_{m,n}, W) \geq p^{mn}$, in the penultimate inequality):

$$\begin{aligned}
\kappa_1 &= \int_{[0,1]} f_1(x) g_K(x)\, dx \\
&\geq \int_{[0,1]} \gamma^{\ell/100} g_K(x)\, dx + \int_S (f_1(x) - \gamma^{\ell/100}) g_K(x)\, dx \\
&\geq \gamma^{\ell/100} \int_{[0,1]} g_K(x)\, dx - \gamma^{\ell/100}\mu(S) \\
&> \gamma^{\ell/100}(t(K_{m,n}, W) - (1-\gamma^{\ell/20})p^{mn}) \\
&\geq \gamma^{\ell/100} t(K_{m,n}, W)(1 - 1 + \gamma^{\ell/20}) \geq \gamma^{\ell/16} t(K_{m,n}, W).
\end{aligned} \tag{63}$$

On the other hand, we obtain using the assumption $\|g_K\|_2 > \frac{\kappa_1}{\gamma^{\ell/8}}$ that

$$\begin{aligned}
\kappa_1 &< \gamma^{\ell/8}\|g_K\|_2 = \gamma^{\ell/8}\left(\int_{[0,1]} g_K(x)^2\, dx\right)^{1/2} \\
&\leq \gamma^{\ell/8}\|g_K\|_1^{1/2}\|g_K\|_\infty^{1/2} \\
&\leq \gamma^{\ell/8} t(K_{m,n}, W)^{1/2} t(K_{m-1,n}, W)^{1/2},
\end{aligned}$$

which implies using Lemma 2.3, the Sidorenko property of $K_{m,n}$, $\gamma \leq p^{16}$ and $\ell \geq n + e(H)$ that

$$\begin{aligned}
\kappa_1 &< \gamma^{\ell/8} t(K_{m,n}, W)^{1/2} t(K_{m,n}, W)^{(m-1)/2m} \\
&\leq \gamma^{\ell/16} p^n t(K_{m,n}, W)^{(2m-1)/2m} \leq \gamma^{\ell/16} t(K_{m,n}, W).
\end{aligned} \tag{64}$$

However, the estimates (63) and (64) contradict each other. Hence, we conclude that (62) holds.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

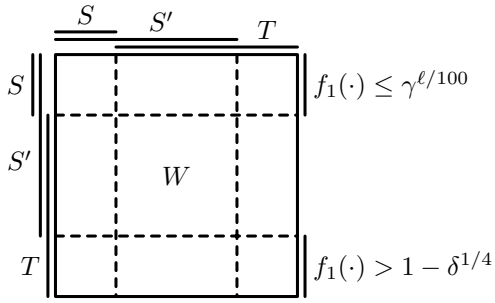


FIGURE 6. Illustration of the definition of the sets S, S' and T from the proof of Claim 4.7.

We next analyze the graphon $W[T]$ where $T = [0,1] \setminus S$ and start with estimating the density of $W[T]$; we refer to Figure 6 for the illustration of the relation of the sets S, S' and T. As the first step to estimate the density of $W[T]$, we bound the density of W between the sets S and $[0,1] \setminus S'$ (we refer to Figure 6 and recall that $\lambda_1 \in [p_0, 1]$):

$$\begin{aligned}\mu(S)\gamma^{\ell/100} &\ge \int_S f_1(x)\,dx = \int_S \frac{1}{\lambda_1}\left(\int_{[0,1]} W(x,y)f_1(y)\,dy\right)dx \ge \int_{S\times[0,1]} W(x,y)f_1(y)\,dx\,dy\\ &\ge \int_{S\times([0,1]\setminus S')} W(x,y)f_1(y)\,dx\,dy \ge (1-\delta^{1/4})\int_{S\times([0,1]\setminus S')} W(x,y)\,dx\,dy,\end{aligned}$$

which yields that

$$\int_{S\times([0,1]\setminus S')} W(x,y)\,dx\,dy \le \frac{\gamma^{\ell/100}\mu(S)}{1-\delta^{1/4}} \le 2\gamma^{\ell/100}\mu(S). \tag{65}$$

Hence, we can estimate the density of W on T^2 using (65) as follows (we again refer to Figure 6):

$$\begin{aligned}\int_{T^2} W(x,y)\,dx\,dy &= \int_{[0,1]^2} W(x,y)\,dx\,dy - \int_{S^2} W(x,y)\,dx\,dy - 2\int_{S\times([0,1]\setminus S)} W(x,y)\,dx\,dy\\ &\ge p - \mu(S)^2 - 2\mu(S)\cdot\mu(S'\setminus S) - \int_{S\times([0,1]\setminus S')} W(x,y)\,dx\,dy\\ &\ge p - \mu(S)^2 - 2\mu(S)\cdot\mu(S'\setminus S) - 4\gamma^{\ell/100}\lambda_1\mu(S)\\ &\ge p - \mu(S)^2 - 2\mu(S)\cdot\mu(S') - 4\gamma^{\ell/100}\lambda_1\mu(S),\end{aligned}$$

which combines with estimates (60), (61), $\delta \le p^2/2^{20}$ by (47), and $\gamma^{\ell/100} \le \gamma_0^{n_0/100} \le p^{100}/2^{128}$ by the choice of n_0 and γ_0 to

$$\begin{aligned}\int_{T^2} W(x,y)\,dx\,dy &\ge p - \frac{p\mu(S)}{2^{16}} - 2\mu(S)\delta^{3/4} - 4\gamma^{\ell/100}\mu(S)\\ &\ge p - \frac{p\mu(S)}{2^{16}} - \frac{p\mu(S)}{2^{14}} - \frac{p\mu(S)}{2^{126}} \ge p\left(1 - \frac{\mu(S)}{2^{10}}\right).\end{aligned}$$

Hence, the density p_T of the graphon $W[T]$ is at least

$$p_T \ge \frac{\int_{T^2} W(x,y)\,dx\,dy}{(1-\mu(S))^2} \ge p\left(1-\frac{\mu(S)}{2^{10}}\right)(1+2\mu(S)) \ge p\left(1+\frac{511\mu(S)}{256}\right), \tag{66}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

which combines with (62) to

$$p_T \geq p\left(1+\frac{511\mu(S)}{512}\right) > p(1+p^{mn}/4). \tag{67}$$

Note that if q were equal to 1, then (67) would imply that $p_T > 1$ (recall that $q \leq p(1+p_0^{mn}/4)$), which is impossible, and so it holds that $q < 1$. Finally, we estimate $\gamma_T = t(C_4, W[T]) - p_T^4$ as follows:

$$\begin{aligned}
\gamma_T = t(C_4, W[T]) - p_T^4 &\leq \frac{t(C_4, W)}{(1-\mu(S))^4} - p^4\left(1+\frac{511\mu(S)}{256}\right)^4 \\
&\leq (p^4+\gamma)(1+5\mu(S)) - p^4\left(1+\frac{511\mu(S)}{256}\right)^4 \\
&\leq \gamma + 5\gamma\mu(S) + 5p^4\mu(S) - \frac{511p^4\mu(S)}{64} \\
&\leq \gamma + \frac{5p^4\mu(S)}{2^{200}} + 5p^4\mu(S) - \frac{511p^4\mu(S)}{64} \leq \gamma.
\end{aligned} \tag{68}$$

Since the density p_T of $W[T]$ is larger than q by (67) (again recall that $q \leq p(1+p_0^{mn}/4)$) and $\gamma_T \leq \gamma_0$ by (68), the choice of q implies that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W[T]) \geq p_T^{e(H)+\ell+mn}.$$

Hence, we obtain using (4), (60) and (66) that

$$\begin{aligned}
t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) &\geq (1-\mu(S))^{v(H)+\ell+m+n-2} p_T^{e(H)+\ell+mn} \\
&\geq p^{e(H)+\ell+mn}(1-\mu(S))^{v(H)+\ell+m+n-2}\left(1+\frac{511\mu(S)}{256}\right)^{e(H)+\ell+mn} \\
&\geq p^{e(H)+\ell+mn}\left((1-\mu(S))\left(1+\frac{511\mu(S)}{256}\right)\right)^{v(H)+\ell+m+n-2} \\
&= p^{e(H)+\ell+mn}\left(1-\mu(S)+\frac{511\mu(S)}{256}-\frac{511\mu(S)^2}{256}\right)^{v(H)+\ell+m+n-2} \\
&\geq p^{e(H)+\ell+mn}\left(1-\mu(S)+\frac{511\mu(S)}{256}-\frac{511\mu(S)}{2^{24}}\right)^{v(H)+\ell+m+n-2} \\
&\geq p^{e(H)+\ell+mn},
\end{aligned}$$

which implies that W satisfies (45) under the assumptions made in the statement of the claim. □

Claims 4.4, 4.6 and 4.7 imply that W satisfies (45), which contradicts the choice of W as violating (45), and so completes the proof of the theorem. □

Theorem 4.1 together with the counting lemma (Lemma 2.2) applied for C_4 yields the following.

COROLLARY 4.8. *For every graph G with girth at least 50 and every real number $p_0 \in (0,1)$, there exist a real $\varepsilon_0 > 0$, a positive integer n_0 and a connected rooted graph $H^\bullet$ that contains G as an induced subgraph such that the following holds for every graphon W with density $p \geq p_0$ such that $\|W-p\|_\square \leq \varepsilon_0$:*

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) \geq p^{e(H)+mn+\ell}$$

for all even integers $m, n, \ell \geq n_0$ such that m is divisible by 5 and $\ell \geq n + e(H)$.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

5. Non-local regime

In this section, we present the result covering the non-local regime, i.e. the regime when one of the k color classes is far from being quasirandom with density $1/k$. We start with the lemma that informally says that if the density of a graph H in a given graphon is small, then the graphon contains a large sparse part. Since it is folklore that complete graphs satisfy the Kohayakawa–Nagle–Rödl–Schacht conjecture (Conjecture 7.1 in § 7), a simple proof of this fact can be found in [Lee21, Theorem 4.2], the next lemma holds for H being a complete graph and so it holds for all graphs H. We include a self-contained proof in Appendix A for completeness.

Lemma 5.1. *For every graph H and real $\delta \in (0,1)$ there exist reals $\omega_0 > 0$ and $\alpha_0 > 0$ such that every graphon W satisfies at least one of the following:*

- $t(H,W) \geq \omega_0$; *or*
- $\alpha_\delta(W) \geq \alpha_0$.

The next theorem is the main result of this section. The theorem asserts that if a graphon W is far from the quasirandom graphon and does not contain a sparse part, then the density of $H^\bullet \oplus K^\bullet_{m|\ell,n}$ in W is at least the density of $H^\bullet \oplus K^\bullet_{m|\ell,n}$ in the quasirandom graphon with the same density.

Theorem 5.2. *For all positive reals $p_0 \in (0,1)$, $\delta \in (0,1)$ and $\varepsilon_0 \in (0,1)$, and rooted graph $H^\bullet$, there exist a positive real α_0 and a positive integer n_0 such that every graphon W with density $p \geq p_0$ and with $\|W - p\|_\square \geq \varepsilon_0$ satisfies at least one of the following:*

- $\alpha_\delta(W) \geq \alpha_0$; *or*
- *it holds for all integers $m, n \geq n_0$ and $\ell \leq mn/4$ that*

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) \geq p^{e(H)+mn+\ell}.$$

Proof. Fix the reals p_0, δ and ε_0 and the rooted graph $H^\bullet$. Let ω_0 and α_0' be the reals obtained by applying Lemma 5.1 with δ and H, and let $d_0 = \omega_0(p_0\varepsilon_0/10)^{16v(H)}$. Choose $n_0 \geq 2$ such that

$$d_0(1 + 10^{-11}\varepsilon_0^{16})^{n_0^2/4} \geq 1.$$

We show that the statement of the theorem holds for this choice of n_0 and for $\alpha_0 = \alpha_0' p_0(\varepsilon_0/10)^{16}$.

Fix a graphon W and let A be the set of $x \in [0,1]$ such that $t_x(H^\bullet, W) \leq d_0$. First suppose that $\mu(A) > p_0(\varepsilon_0/10)^{16}$ and note that $t(H, W[A]) \leq d_0/\mu(A)^{v(H)} < \omega_0$. Hence, the graphon $W[A]$ does not satisfy the first conclusion of Lemma 5.1 and so it satisfies the second conclusion, i.e. $\alpha_\delta(W[A]) \geq \alpha_0'$. It follows that $\alpha_\delta(W) \geq \alpha_0'\mu(A) \geq \alpha_0$. Therefore, we can assume in the rest of the proof of the theorem that $\mu(A) \leq p_0(\varepsilon_0/10)^{16}$.

Let W' be the graphon obtained from the graphon W by setting

$$W'(x,y) = \begin{cases} 0 & \text{if } x \in A \text{ or } y \in A, \text{ and} \\ W(x,y) & \text{otherwise.} \end{cases}$$

Note that the density p' of W' is at least $p - 3\mu(A) \geq p(1 - 3(\varepsilon_0/10)^{16})$ and

$$\|W - p'\|_\square \geq \|W - p\|_\square - |p - p'| \geq \|W - p\|_\square - 3(\varepsilon_0/10)^{16}.$$

As $\|W - p\|_\square \geq \varepsilon_0$, we obtain using Lemma 2.14 (note that we need that $\ell \leq mn/4$ to apply the lemma) that

$$\begin{aligned} t(K_{m|\ell,n}, W') &\geq p^{mn+\ell}(1 - 3(\varepsilon_0/10)^{16})^{mn+\ell}(1 + 10^{-9}(1 - 3\varepsilon_0^{15}/10^{16})^{16}\varepsilon_0^{16})^{mn/4+\ell/4} \\ &\geq p^{mn+\ell}(1 - 12(\varepsilon_0/10)^{16})^{mn/4+\ell/4}(1 + 10^{-10}\varepsilon_0^{16})^{mn/4+\ell/4} \end{aligned}$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$$\geq p^{mn+\ell}((1-10^{-14}\varepsilon_0^{16})(1+10^{-10}\varepsilon_0^{16}))^{mn/4+\ell/4}$$
$$\geq p^{mn+\ell}(1+10^{-11}\varepsilon_0^{16})^{mn/4+\ell/4}. \tag{69}$$

We next estimate the density of $H^\bullet \oplus K^\bullet_{m|\ell,n}$ in the graphon W as follows:

$$\begin{aligned} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) &= \int_{[0,1]} t_x(H^\bullet, W) t_x(K^\bullet_{m|\ell,n}, W)\, dx \\ &\geq d_0 \int_{[0,1]\setminus A} t_x(K^\bullet_{m|\ell,n}, W)\, dx \geq d_0 \int_{[0,1]\setminus A} t_x(K^\bullet_{m|\ell,n}, W')\, dx \\ &= d_0 t(K_{m|\ell,n}, W'). \end{aligned}$$

Hence, we obtain using (69) and the choice of n_0 that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) \geq d_0 p^{mn+\ell}(1+10^{-11}\varepsilon_0^{16})^{mn/4} \geq p^{mn+\ell} \geq p^{e(H)+mn+\ell}.$$

The statement of the theorem now follows. □

6. Main result

We are now ready to prove our main results: Theorem 1.1 and the more general Theorem 1.2. Since there are graphs with arbitrarily large chromatic number and girth at least 50 by a classical result of Erdős [Erd59] (see also [AS16]), the former is an immediate corollary of the next theorem, which is a special case of Theorem 6.2 proven later in this section. However, we include a short proof of the next theorem to highlight the main steps of the argument in the case of two colors.

THEOREM 6.1. *For every graph G with girth at least 50, there exist a positive integer N_0 and a connected rooted graph $H^\bullet$ that contains G as an induced subgraph such that the following holds for any graphon W:*

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) + t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, 1-W) \geq 2^{1-e(H^\bullet \oplus K^\bullet_{m|n+e(H),n})} = 2^{1-2e(H)-(m+1)n}$$

for all even integers $m, n \geq N_0$ such that m is divisible by 5.

Proof. Fix a graph G and an integer k as in the statement. We apply Corollary 4.8 with $p_0 = 1/4$ to obtain a positive real ε_0, a positive integer n_0 and a connected rooted graph $H^\bullet$ containing G as an induced subgraph such that every graphon W with density $p \geq 1/4$ and with $\|W-p\|_\square \leq \varepsilon_0$ satisfies that

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) \geq p^{2e(H)+(m+1)n}$$

for all even integers $m, n \geq n_0$ such that m is divisible by 5. We next apply Theorem 5.2 with $p_0 = 1/4$, $\delta = \min\{\varepsilon_0/2, 1/16\}$, ε_0 and $H^\bullet$ to obtain a positive real α_0 and an integer n_0 with the properties given in Theorem 5.2. We now set N_0 in a way that N_0 is at least $\max\{8, n_0, e(H)\}$ and the following holds for all $m, n \geq N_0$:

$$\alpha_0^{v(H)+e(H)+m+2n-1} \geq 1.5^{1-2e(H)-(m+1)n}. \tag{70}$$

We now show that the statement of theorem holds for any graphon W and all even integers $m, n \geq N_0$ such that m is divisible by 5. To do so, fix a graphon W. By symmetry, we can assume that the density p of W is at least $1/2$ (otherwise, we consider the graphon $1-W$ instead of W).

Suppose that $\alpha_\delta(1-W) \geq \alpha_0$, and let $h : [0,1] \to [0,1]$ be such that $|h|_1 \geq \alpha_0$ and the density of $(1-W)[h] = 1 - W[h]$ is at most δ. Let q be the density of $W[h]$; note that $q \geq 1-\delta$.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

The triangle inequality implies that $\|W[h]-q\|_\square \le \|1-W[h]\|_\square + (1-q) \le 2\delta \le \varepsilon_0$. Hence, Corollary 4.8, (4) and (70) yield that

$$\begin{aligned} t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) &\ge \alpha_0^{v(H)+e(H)+m+2n-1} t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W[h]) \\ &\ge \alpha_0^{v(H)+e(H)+m+2n-1} q^{2e(H)+(m+1)n} \\ &\ge 1.5^{1-2e(H)-(m+1)n}(1-\delta)^{2e(H)+(m+1)n} \ge 2^{1-2e(H)-(m+1)n}. \end{aligned}$$

We conclude that if $\alpha_\delta(1-W) \ge \alpha_0$ or (symmetrically) $\alpha_\delta(W) \ge \alpha_0$, then the statement of the theorem holds. In the rest of the proof, we assume that both $\alpha_\delta(W)$ and $\alpha_\delta(1-W)$ are at most α_0.

We now derive from Corollary 4.8 if $\|W-p\|_\square \le \varepsilon_0$ or from Theorem 5.2 if $\|W-p\|_\square > \varepsilon_0$ that

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) \ge p^{2e(H)+(m+1)n}. \tag{71}$$

In particular, if $p \ge 3/4$, it follows that

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) \ge 1.5^{2e(H)+(m+1)n} \cdot 2^{-2e(H)-(m+1)n} \ge 2^{1-2e(H)-(m+1)n},$$

and so the statement of the theorem holds. On the other hand, if $p < 3/4$, then the density of $1-W$ is at least $p_0 = 1/4$, and we can derive from Corollary 4.8 or from Theorem 5.2, depending whether the cut distance between $1-W$ and $1-p$ is at most ε_0 or larger, that

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, 1-W) \ge (1-p)^{2e(H)+(m+1)n}. \tag{72}$$

We combine (71) and (72) to obtain that

$$t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, W) + t(H^\bullet \oplus K^\bullet_{m|n+e(H),n}, 1-W) \ge p^{2e(H)+(m+1)n} + (1-p)^{2e(H)+(m+1)n},$$

which is at least $2^{1-2e(H)-(m+1)n}$ by convexity of the function $x^{2e(H)+(m+1)n}$. This completes the proof of the theorem. □

We are now ready to prove the more general result, which implies Theorem 1.2. To display the dependance on the length of the path between $H^\bullet$ and $K^\bullet_{m,n}$, we decided to use ℓ throughout the proof to denote the length of this path.

THEOREM 6.2. *For every graph G with girth at least 50 and every integer $k \ge 2$, there exist a positive integer N_k and a connected rooted graph $H^\bullet$ that contains G as an induced subgraph such that the following holds for any k graphons $W_1, \dots, W_k$ such that $W_1 + \cdots + W_k = 1$:*

$$\sum_{i=1}^{k} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i) \ge k^{1-e(H)-mn-\ell}$$

for all even integers $m, n \ge N_k$ and $\ell = n + e(H)$ such that m is divisible by 5.

Proof. Fix the graph G and the integer k. Apply Corollary 4.8 with $p_0 = 1/2k$ to obtain a positive real ε_0, a positive integer n_0 and a connected rooted graph $H^\bullet$ containing G as an induced subgraph such that every graphon W with density $p \ge 1/2k$ and with $\|W-p\|_\square \le \varepsilon_0$ satisfies that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W) \ge p^{e(H)+mn+\ell} \tag{73}$$

for all even integers $m, n, \ell \ge n_0$ such that m is divisible by 5 and $\ell \ge n + e(H)$. Without loss of generality, we may assume that $\varepsilon_0 \le 1/(4k)$.

We next repeatedly apply Theorem 5.2. Set $\delta_1 = \varepsilon_0/2 \le 1/8$. For $k' = 2, \dots, k$, apply Theorem 5.2 with $p_0 = 1/2k$, $\delta = \delta_{k'-1}/2$, ε_0 and $H^\bullet$ to obtain a positive real $\alpha_{k'}$ and an integer

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

$n_{k'}$, and set $\delta_{k'} = \min\{\delta_{k'-1}\alpha_{k'}^2/2, 1/4k'\}$. Finally, set N_k in a way that N_k is at least n_0 and the following holds for all $m, n \geq N_k$:

$$n + e(H) \leq mn/4, \tag{74}$$

$$(1+1/4k)^{2e(H)+(m+1)n} \geq k, \tag{75}$$

and

$$\alpha_{k'}^{v(H)+e(H)+m+2n-1}\left(\frac{1-\delta_{k'-1}}{k'-1}\right)^{2e(H)+(m+1)n} \geq \left(\frac{1}{k'}\right)^{2e(H)+(m+1)n-1} \tag{76}$$

for $k' = 2, \ldots, k$. Note that such N_k exists since it holds $\frac{1-\delta_{k'-1}}{k'-1} \geq \frac{4k'-1}{4(k'-1)k'} > \frac{1}{k'}$ for $k' = 2, \ldots, k$.

We now prove the following claim, which implies that the statement of the theorem.

CLAIM 6.3. *Let $k' \in [k]$. Any k' graphons $W_1, \ldots, W_{k'}$ such that $\|W_1 + \cdots + W_{k'}\|_\infty \leq 1$ and $t(K_2, W_1 + \cdots + W_{k'}) \geq 1 - \delta_{k'}$ satisfy that*

$$\sum_{i=1}^{k'} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i) \geq k' \cdot \left(\frac{t(K_2, W_1 + \cdots + W_{k'})}{k'}\right)^{e(H)+mn+\ell} \tag{77}$$

for all even integers $m, n \geq N_k$ and $\ell = n + e(H)$ such that m is divisible by 5.

Proof of Claim 6.3. The proof proceed by induction on $k' = 1, \ldots, k$. The base of the induction is the case $k' = 1$. Consider a graphon W_1 such that $t(K_2, W_1) \geq 1 - \delta_1 = 1 - \varepsilon_0/2$, i.e. the density p of W_1 is at least $1 - \varepsilon_0/2$. Note that $\|W_1 - p\|_\square \leq \|W_1 - 1\|_\square + |1 - p| \leq \varepsilon_0$. Hence, Corollary 4.8 implies that (73) holds for $W = W_1$, which is equivalent to (77) for $k' = 1$.

We now present the induction step. Consider $k' \in \{2, \ldots, k\}$ and assume that the claim holds for $k' - 1$. Consider graphons $W_1, \ldots, W_{k'}$ and let $p_1, \ldots, p_{k'}$ be their respective densities; note that $p_1 + \cdots + p_{k'} \geq 1 - \delta_{k'}$. First suppose that there exists i such that $\alpha_{\delta_{k'-1}/2}(W_i) \geq \alpha_{k'}$. By symmetry, we may assume that $i = k'$. Hence, there exists $h : [0,1] \to [0,1]$ with $\|h\|_1 \geq \alpha_{k'}$ such that the density of $W_{k'}[h]$ is at most $\delta_{k'-1}/2$. It follows that the density of $(W_1 + \cdots + W_{k'-1})[h]$ is at least

$$1 - \frac{\delta_{k'}}{\|h\|_1^2} - \frac{\delta_{k'-1}}{2} \geq 1 - \frac{\delta_{k'}}{\alpha_{k'}^2} - \frac{\delta_{k'-1}}{2} \geq 1 - \frac{\delta_{k'-1}}{2} - \frac{\delta_{k'-1}}{2} = 1 - \delta_{k'-1}.$$

By the induction hypothesis, it holds that

$$\sum_{i=1}^{k'-1} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i[h]) \geq \frac{(1-\delta_{k'-1})^{e(H)+mn+\ell}}{(k'-1)^{e(H)+mn+\ell-1}}.$$

It follows using (76) and $\|h\|_1 \geq \alpha_{k'}$ that

$$\sum_{i=1}^{k'-1} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i) \geq \|h\|_1^{v(H)+m+n+\ell-1}\frac{(1-\delta_{k'-1})^{e(H)+mn+\ell}}{(k'-1)^{e(H)+mn+\ell-1}} \geq \left(\frac{1}{k'}\right)^{e(H)+mn+\ell-1},$$

which implies that (77) holds. In the rest of the proof, we assume that $\alpha_{\delta_{k'-1}/2}(W_i) < \alpha_{k'}$ for every $i \in [k']$.

Next suppose that there exists $i \in [k']$ such that the density p_i of W_i is less than $1/2k$; by symmetry, we may assume that $i = k'$. It follows that the density of one of the graphons $W_1, \ldots, W_{k'-1}$, say the graphon W_1, is at least

$$\frac{1 - 1/2k - \delta_{k'}}{k'-1} \geq \frac{1 - 1/2k' - 1/4k'}{k'-1} = \frac{4k'-3}{4(k'-1)k'} \geq \frac{1+1/4k}{k'}.$$

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

We obtain using Corollary 4.8 (if $\|W_1 - p_1\|_\square \leq \varepsilon_0$) or Theorem 5.2 (if $\|W_1 - p_1\|_\square > \varepsilon_0$) that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_1) \geq \left(\frac{1+1/4k}{k'}\right)^{e(H)+mn+\ell} \geq k\left(\frac{1}{k'}\right)^{e(H)+mn+\ell} \geq k'\left(\frac{1}{k'}\right)^{e(H)+mn+\ell},$$

which implies that (77) (the middle inequality above follows from (75)).

It remains to consider the case that $p_i \geq 1/2k$ and $\alpha_{\delta_{k'-1}/2}(W_i) \leq \alpha_{k'}$ for every $i \in [k']$. In this case, we apply to each W_i, $i \in [k']$, either Corollary 4.8 (if $\|W_i - p_i\|_\square \leq \varepsilon_0$) or Theorem 5.2 (if $\|W_i - p_i\|_\square > \varepsilon_0$) to get that

$$t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i) \geq p_i^{e(H)+mn+\ell}.$$

Hence, we conclude that

$$\sum_{i=1}^{k'} t(H^\bullet \oplus K^\bullet_{m|\ell,n}, W_i) \geq \sum_{i=1}^{k'} p_i^{e(H)+mn+\ell} \geq k'\left(\frac{p_1 + \cdots + p_{k'}}{k'}\right)^{e(H)+mn+\ell},$$

where the last inequality follows from Jensen's inequality. It follows that (77) holds, which completes the proof of the claim. □

Since the statement of Claim 6.3 for $k' = k$ is actually stronger than the statement of the theorem, the proof of theorem is now completed.

7. Conclusion

We finish with briefly presenting another connection of our result to Sidorenko's conjecture. A simple example showing that a non-bipartite graph H does not have the Sidorenko property is given when the host graph G is bipartite; in such case $t(H, G) = 0$. Kohayakawa *et al.* [KNRS10] conjectured that having a large spare part in the host graph G is the only obstacle for $t(H, G) \geq t(K_2, G)^{e(H)}$ to hold. Recall that a graph is (ρ, d)-*dense* if any subset of at least $\rho \cdot v(G)$ vertices of G induces a subgraph with density at least d; we refer to § 2.1 for the relation to the notions studied in this paper. The conjecture of Kohayakawa *et al.* [KNRS10] states that any graph has the Sidorenko property with respect to $(o(1), d)$-dense host graphs.

CONJECTURE 7.1 (Kohayakawa *et al.* [KNRS10]). Let H be a graph, and let η and p be positive reals. There exists a positive real ρ such that $t(H, G) \geq (1 - \eta)p^{e(H)}$ for every sufficiently large (ρ, p)-dense graph G.

Sidorenko's conjecture would imply that Conjecture 7.1 holds for all bipartite graphs. Among non-bipartite graphs, Conjecture 7.1 is known to hold only for some specific families [KNRS10, Rei14, Lee21, BSW24], e.g. complete multipartite graphs, unicyclic graphs, cycles with a single chord, graphs obtained by gluing complete multipartite graphs in a tree-like way, and graphs obtained by gluing along independent sets in highly symmetric ways. Theorems 4.1 and 5.2 imply that for every graph G with girth at least 50, there exists a connected rooted graph $H^\bullet$ containing G as an induced subgraph and n_0 such that the graph $H^\bullet \oplus K^\bullet_{m|\ell,n}$ satisfies Conjecture 7.1 for all even integers $m, n, \ell \geq n_0$ such that m is divisible by 5 and $n + e(H) \leq \ell \leq mn/4$.

We finish with an open problem motivated by the techniques used in the proof Theorem 4.1, where it was essential that the graph $H^\bullet \oplus K^\bullet_{m|\ell,n}$ has a long path formed by vertices of degree two. In particular, the techniques used in this paper would not establish the existence of a 3-connected high-chromatic common graph. So, it is natural to ask the following.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Problem 7.2. Is it true that for every $\ell \geq 2$ and every $k \geq 3$, there exists an ℓ-chromatic k-connected common graph?

Ko and Lee [KL23] answered the above question in the affirmative by combining our construction and the book product of graphs. They also posed the following more general problem.

Problem 7.3 (Ko and Lee [KL23, Question 3.1]). Is it true that for every $\ell \geq 2$, every $k \geq 3$ and every $g \geq 4$, there exists an ℓ-chromatic k-connected common graph with girth at least g?

Appendix A

We provide a self-contained proof of Lemma 5.1 for completeness.

Proof of Lemma 5.1. Fix $\delta > 0$. We show by induction on r that there exist positive reals ω_r and α_r such that $t(K_r, W) \geq \omega_r$ or $\alpha_\delta(W) \geq \alpha_r$. The statement of the lemma will follow by setting $\omega_0 = \omega_{v(H)}$ and $\alpha_0 = \alpha_{v(H)}$ (note that $t(K_r, W) \leq t(H, W)$ for every r-vertex graph H).

The base of the induction is the case $r = 1$, which holds with $\omega_1 = 1$ and any $\alpha_1 \in (0, 1]$. We next present the induction step when $r \geq 2$. Assuming the existence of ω_{r-1} and α_{r-1}, we show that it is possible to set $\omega_r = (1 - \delta)\delta^{r-1}\omega_{r-1}$ and $\alpha_r = \delta\alpha_{r-1}$.

Consider a graphon W and let A be the set of $x \in [0, 1]$ such that $\deg_W(x) \leq \delta^2$. If $\mu(A) \geq \delta$, then the density of the graphon $W[A]$ is at most $\mu(A)\delta^2/\mu(A)^2 \leq \delta$, which yields that $\alpha_\delta(W) \geq \delta \geq \alpha_r$. Hence, we can assume that $\mu(A) \leq \delta$ in what follows.

Let $f_x : [0, 1] \to [0, 1]$ for $x \in [0, 1]$ be the function defined as $f_x(y) = W(x, y)$. Observe that

$$t(K_r, W) = \int_{[0,1]} \left(\int_{[0,1]} f_x(y)\, dy \right)^{r-1} t(K_{r-1}, W[f_x])\, dx.$$

It follows that if $t(K_{r-1}, W[f_x]) \geq \omega_{r-1}$ for every $x \in [0, 1] \setminus A$, then $t(K_r, W) \geq (1 - \mu(A))\delta^{r-1}\omega_{r-1} \geq \omega_r$. Otherwise, there exists $x \in [0, 1] \setminus A$ such that $\alpha_\delta(W[f_x]) \geq \alpha_{r-1}$, which means that there exist $g : [0, 1] \to [0, 1]$ such that

$$\frac{\int_{[0,1]} f_x(y)g(y)\, dy}{\int_{[0,1]} f_x(y)\, dy} \geq \alpha_{r-1} \quad \text{and} \quad \frac{\int_{[0,1]^2} f_x(y)f_x(z)g(y)g(z)W(y, y)\, dy\, dz}{\left(\int_{[0,1]} f_x(y)g(y)\, dy \right)^2} \leq \delta.$$

It follows that the function $h : [0, 1] \to [0, 1]$ defined as $h(y) = f_x(y)g(y)$ satisfies that $\|h\|_1 \geq \alpha_{r-1}\delta = \alpha_r$ and the density of $W[h]$ is at most δ. Hence, $\alpha_\delta(W) \geq \alpha_r$ as required. This completes the proof of the induction step and so the proof of the lemma. □

Acknowledgements

The authors would like to thank Jon Noel and Sergey Norin for insightful comments on possible ways of constructing common and more generally k-common graphs and many detailed discussions concerning common graphs and graph limits. The authors are also indebted to the anonymous referee for many comments that have helped to improve the presentation of the results and proofs in the paper and led to making the presented proofs more accessible.

Conflicts of interest

None.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Financial support

DK has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no 648509). This publication reflects only its authors' view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains. JV is supported by grant 23-06815M of the Grant Agency of the Czech Republic. The first and second authors were also supported by the MUNI Award in Science and Humanities (MUNI/I/1677/2018) of the Grant Agency of Masaryk University. FW is supported by NSF grants DMS-2404167 and DMS-2401414.

Journal information

Compositio Mathematica is owned by the Foundation Compositio Mathematica and published by the London Mathematical Society in partnership with Cambridge University Press. All surplus income from the publication of *Compositio Mathematica* is returned to mathematics and higher education through the charitable activities of the Foundation, the London Mathematical Society and Cambridge University Press.

References

AS16 N. Alon and J. H. Spencer, *The probabilistic method*, Wiley Series in Discrete Mathematics and Optimization, fourth edition (John Wiley & Sons, Hoboken, NJ, 2016).

BR65 G. R. Blakley and P. Roy, *A Hölder type inequality for symmetric matrices with nonnegative entries*, Proc. Amer. Math. Soc. **16** (1965), 1244–1245.

BSW24 D. Bradač, B. Sudakov and Y. Wigderson, *Counting subgraphs in locally dense graphs*, Preprint (2024), arXiv:2406.12418.

BR80 S. A. Burr and V. Rosta, *On the Ramsey multiplicities of graphs—problems and recent results*, J. Graph Theory **4** (1980), 347–361.

CGMS23 M. Campos, S. Griffiths, R. Morris and J. Sahasrabudhe, *An exponential improvement for diagonal ramsey*, Preprint (2023), arXiv:2303.09521.

Cla92 L. Clark, *The minimum number of subgraphs in a graph and its complement*, J. Graph Theory **16** (1992), 451–458; MR 1185009.

Con09 D. Conlon, *A new upper bound for diagonal Ramsey numbers*, Ann. of Math. (2) **170** (2009), 941–960.

CFS10 D. Conlon, J. Fox and B. Sudakov, *An approximate version of Sidorenko's conjecture*, Geom. Funct. Anal. **20** (2010), 1354–1366.

CFS15 D. Conlon, J. Fox and B. Sudakov, Recent developments in graph Ramsey theory, in *Surveys in combinatorics 2015*, London Mathematical Society Lecture Note Series, vol. 424 (Cambridge University Press, Cambridge, 2015), 49–118; MR 3497267.

CKLL18 D. Conlon, J. H. Kim, C. Lee and J. Lee, *Some advances on Sidorenko's conjecture*, J. Lond. Math. Soc. (2) **98** (2018), 593–608.

CL17 D. Conlon and J. Lee, *Finite reflection groups and graph norms*, Adv. Math. **315** (2017), 130–165.

CL21 D. Conlon and J. Lee, *Sidorenko's conjecture for blow-ups*, Discrete Anal. (2021), 2.

CKM18 J. W. Cooper, D. Kráľ and T. L. Martins, *Finitely forcible graph limits are universal*, Adv. Math. **340** (2018), 819–854.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

CHL23 E. Csóka, T. Hubai and L. Lovász, *Locally common graphs*, J. Graph Theory **102** (2023), 472–483.

Erd59 P. Erdős, *Graph theory and probability*, Canad. J. Math. **11** (1959), 34–38; MR 102081.

Erd62 P. Erdős, *On the number of complete subgraphs contained in certain graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl. **7** (1962), 459–464.

ES83 P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192; MR 726456.

Fox08 J. Fox, *There exist graphs with super-exponential Ramsey multiplicity constant*, J. Graph Theory **57** (2008), 89–98.

FW17 J. Fox and F. Wei, *On the local approach to Sidorenko's conjecture*, in *The European conference on combinatorics, graph theory and applications (EUROCOMB'17)*, Electronic Notes in Discrete Mathematics, vol. 61, eds M. Drmota, M. Kang, C. Krattenthaler and J. Nešetřil (2017), 459–465.

Goo59 A. W. Goodman, *On sets of acquaintances and strangers at any party*, Amer. Math. Monthly **66** (1959), 778–783.

GRS13 R. L. Graham, B. L. Rothschild and J. H. Spencer, *Ramsey theory*, Wiley Series in Discrete Mathematics and Optimization (John Wiley & Sons, Hoboken, NJ, 2013).

GLLV22 A. Grzesik, J. Lee, B. Lidický, and J. Volec, *On tripartite common graphs*, Combin. Probab. Comput. **1** (2022), 907–923.

HKKV23 R. Hancock, D. Král', M. Krnc and J. Volec, *Towards characterizing locally common graphs*, Random Structures Algorithms **62** (2023), 181–218.

Hat10 H. Hatami, *Graph norms and Sidorenko's conjecture*, Israel J. Math. **175** (2010), 125–150.

HHK$^+$12 H. Hatami, J. Hladký, D. Kráľ, S. Norine and A. Razborov, *Non-three-colourable common graphs exist*, Combin. Probab. Comput. **21** (2012), 734–742.

Hin74 N. Hindman, *Finite sums from sequences within cells of a partition of N*, J. Combin. Theory Ser. A **17** (1974), 1–11; MR 349574.

JST96 C. Jagger, P. Štovíček and A. Thomason, *Multiplicities of subgraphs*, Combinatorica **16** (1996), 123–141.

KL23 S. Ko and J. Lee, *Common graphs with arbitrary connectivity and chromatic number*, J. Combin. Theory Ser. B **162** (2023), 223–230.

KNRS10 Y. Kohayakawa, B. Nagle, V. Rödl and M. Schacht, *Weak hypergraph regularity and linear hypergraphs*, J. Combin. Theory Ser. B **100** (2010), 151–160.

KNN$^+$22 D. Král', J. A. Noel, S. Norin, J. Volec and F. Wei, *Non-bipartite k-common graphs*, Combinatorica **42** (2022), 87–114.

Lee21 J. Lee, *On some graph densities in locally dense graphs*, Random Structures Algorithms **58** (2021), 322–344; MR 4201799.

Lov11 L. Lovász, *Subgraph densities in signed graphons and the local Simonovits-Sidorenko conjecture*, Electron. J. Combin. **18** (2011), 127.

Lov12 L. Lovász, *Large networks and graph limits*, vol. 60 (AMS Colloquium Publications, 2012).

Ram30 F. P. Ramsey, *On a problem of formal logic*, Proc. Lond. Math. Soc. (3) **2** (1930), 264–286.

Raz07 A. A. Razborov, *Flag algebras*, J. Symb. Logic **72** (2007), 1239–1282.

Rei14 C. Reiher, *Counting odd cycles in locally dense graphs*, J. Combin. Theory Ser. B **105** (2014), 1–5.

Sah23 A. Sah, *Diagonal Ramsey via effective quasirandomness*, Duke Math. J. **172** (2023), 545–567.

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press

Sid93 A. Sidorenko, *A correlation inequality for bipartite graphs*, Graphs Combin. **9** (1993), 201–204.

Sid96 A. Sidorenko, *Randomness friendly graphs*, Random Structures Algorithms **8** (1996), 229–241.

Sid86 A. F. Sidorenko, *Extremal problems in graph theory and functional-analytic inequalities*, in *Proc. of the All-Union Seminar on Discrete Mathematics and its Applications* (Moskov. Gos. Univ., Mekh.-Mat. Fak., Moscow, 1986), 99–105 (in Russian); MR 930303.

Sid89 A. F. Sidorenko, *Cycles in graphs and functional inequalities*, Mat. Zametki **46** (1989), 72–79.

Sid91 A. F. Sidorenko, *Inequalities for functionals generated by bipartite graphs*, Diskret. Mat. **3** (1991), 50–65.

Sze14 B. Szegedy, *An information theoretic approach to Sidorenko's conjecture*, Preprint (2014), arXiv:1406.6738.

Sze75 E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, collection of articles in memory of Juriĭ Vladimirovič Linnik.

Tho89 A. Thomason, *A disproof of a conjecture of Erdős in Ramsey theory*, J. Lond. Math. Soc. (2) **39** (1989), 246–255.

Daniel Kráľ dkral@fi.muni.cz
Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic

Jan Volec jan@ucw.cz
Department of Theoretical Computer Science, Faculty of Information Technology, Czech Technical University in Prague, Thákurova 9, 160 00 Prague, Czech Republic

Fan Wei fan.wei@duke.edu
Duke Mathematics, Campus Box 90320, Durham, NC 27708, USA

https://doi.org/10.1112/S0010437X24007681 Published online by Cambridge University Press