Counting subgraphs in locally dense graphs

Domagoj Bradač[*]   Benny Sudakov[*]   Yuval Wigderson[†]

Abstract

A graph G is said to be p-locally dense if every induced subgraph of G with linearly many vertices has edge density at least p. A famous conjecture of Kohayakawa, Nagle, Rödl, and Schacht predicts that locally dense graphs have, asymptotically, at least as many copies of any fixed graph H as are found in a random graph of edge density p.

In this paper, we prove several results around the KNRS conjecture. First, we prove that certain natural gluing operations on H preserve this property, thus proving the conjecture for many graphs H for which it was previously unknown. Secondly, we study a stability version of this conjecture, and prove that for many graphs H, approximate equality is attained in the KNRS conjecture if and only if the host graph G is quasirandom. Finally, we introduce a weakening of the KNRS conjecture, which requires the host graph to be nearly degree-regular, and prove this conjecture for a larger family of graphs. Our techniques reveal a surprising connection between these questions, semidefinite optimization, and the study of copositive matrices.

1   Introduction

1.1   Background

Many of the most basic questions in extremal graph theory ask to understand, for a fixed graph H, how many copies of H can appear in a large graph G with certain constraints. For example, Mantel's theorem states that if $H = K_3$, then G contains *zero* copies of H only if G has at most $\lfloor v(G)^2/4 \rfloor$ edges.

A convenient way of capturing the asymptotic nature of such problems is via the *homomorphism density* $t(H, G)$, which is defined as the probability that a random function $f : V(H) \to V(G)$ is a graph homomorphism, that is, that f maps every edge of H to an edge of G. As $v(G) \to \infty$, asymptotically almost all homomorphisms are injective, so knowing $t(H, G)$ is essentially the same as knowing how many copies of H there are in G. In this language, Mantel's theorem can be stated as $t(K_3, G) = 0$ only if $t(K_2, G) \leq \frac{1}{2}$. More generally, the Erdős–Stone–Simonovits theorem [11, 13] states that $t(H, G) = 0$ only if $t(K_2, G) \leq 1 - \frac{1}{\chi(H)-1} + o(1)$, where the $o(1)$ term tends to 0 as $v(G) \to \infty$ (for fixed H).

Even more generally, one could ask, for a fixed value of $p = t(K_2, G)$, what is the asymptotic minimum value of $t(H, G)$? In general, the answer to this question is extremely complicated [20, 21, 24]. However, for certain bipartite graphs H, this problem has a very clean answer. For example, two applications of the Cauchy–Schwarz inequality imply that $t(C_4, G) \geq p^4$ for any graph G satisfying $t(K_2, G) = p$. Moreover, an elementary computation shows that this bound is asymptotically tight if G is a random graph of edge density p. A famous conjecture of Sidorenko [30] asserts that something similar should happen for *all* bipartite H, namely that

$$t(H, G) \geq p^{e(H)} \tag{1}$$

for all bipartite H, all $p \in [0,1]$, and all G satisfying $t(K_2, G) = p$. Sidorenko's conjecture has been verified for many natural classes of bipartite graphs (see e.g. [5, 7, 8, 15, 16]), but remains wide open in general. Note that again, Sidorenko's conjecture is best possible if true, since a random graph G of edge density p satisfies $t(H, G) = p^{e(H)} + o(1)$ a.a.s.

Moreover, there is a sort of converse to the last statement. Following Chung–Graham–Wilson [4], let us say that a graph G is (p, δ)-*quasirandom* if every set $S \subseteq V(G)$ contains $\frac{p}{2}|S|^2 \pm \delta v(G)^2$ edges. As proved by Thomason [35] and Chung–Graham–Wilson [4], this condition is equivalent (up to a polynomial change in δ) to many other natural notions of "random-like" behavior in graphs. In particular, one famous result of Chung–Graham–Wilson is that if $t(K_2, G) = p$ and $t(C_4, G) \leq p^4 + o(1)$, then G is $(p, o(1))$-quasirandom. The *forcing conjecture* (see [32]) states that, again, such behavior should be more general: for any bipartite H that is not a forest, we have that $t(H, G) = p^{e(H)} + o(1)$ only if G is $(p, o(1))$-quasirandom. It is known (see e.g. [5]) that the forcing conjecture implies Sidorenko's conjecture.

Additionally, the assumption that H is bipartite is crucial for a statement like Sidorenko's conjecture to hold, since $t(H, K_{n,n}) = 0$ whenever H is non-bipartite, despite the fact that $t(K_2, K_{n,n}) = \frac{1}{2}$. Nonetheless, it is tempting to wonder if, by imposing some extra condition on G, one can obtain a sensible statement, along the lines of (1), which holds for all H. Such a statement was conjectured by Kohayakawa, Nagle, Rödl, and Schacht [17]. To state it we first need the following definition.

Definition 1.1. A graph G is called (p, δ)-*locally dense* if, for every $S \subseteq V(G)$ with $|S| \geq \delta|V(G)|$, we have $e(S) \geq p\frac{|S|^2}{2}$.

One can think of this condition as a "one-sided" version of quasirandomness—it does not assert that all large vertex sets have edge density roughly p (as in quasirandomness), but only that all such sets have edge density *at least* p. Simple examples, such as the disjoint union of two cliques, show that this is a strictly weaker condition than quasirandomness. Locally dense graphs and hypergraphs have been extensively studied and used over the years, see e.g. [3, 12, 14, 17, 18, 23, 25, 26, 27, 28, 29]. With this definition, we can state the conjecture of Kohayakawa–Nagle–Rödl–Schacht [17].

Conjecture 1.2 (Kohayakawa–Nagle–Rödl–Schacht [17])**.** *For every graph H and all $p, \varepsilon \in (0, 1)$, there exists some $\delta > 0$ such that the following holds. If G is a (p, δ)-locally dense graph, then $t(H, G) \geq p^{e(H)} - \varepsilon$.*

Informally[1], this conjecture states that if G is a $(p, o(1))$-locally dense graph, then $t(H, G) \geq p^{e(H)} - o(1)$ for any fixed H. For brevity, we will say that a graph H is *KNRS* if Conjecture 1.2 holds for H. Despite a great deal of interest, only a relatively restricted class of graphs are known to be KNRS, which we now briefly summarize.

First, if H is a bipartite graph satisfying Sidorenko's conjecture, then H is KNRS. Indeed, Sidorenko's conjecture is equivalent to Conjecture 1.2 in case $\varepsilon = 0$ and $\delta = 1$. As observed by Kohayakawa, Nagle, Rödl, and Schacht [17], a simple inductive argument shows that all cliques (and, more generally, all complete multipartite graphs) are KNRS. They also noted that the main result from [6] implies that all line graphs of hypercubes are KNRS. Next, answering a question raised by Kohayakawa–Nagle–Rödl–Schacht, Reiher [23] proved that all odd cycles are KNRS. Finally, Lee [18] showed that all unicyclic graphs and all graphs obtained from a cycle by adding a chord are KNRS. Additionally, he showed that graphs obtained by certain gluing operations are KNRS, and in particular settled Conjecture 1.2 for all graphs on at most 5 vertices.

---

[1]Shortly, we will turn to the language of graphons, where these informal asymptotic notions can be made rigorous.

1.2 Our results

In this paper, we prove a number of results around Conjecture 1.2. Our first main result shows that certain "highly symmetric" graphs, which are obtained from smaller KNRS graphs by a certain gluing operation, are themselves KNRS.

Definition 1.3. Let H_1, H_2 be two graphs, let I be an independent set in H_1 and let a be a vertex in $V(H_1) \setminus I$. We define the graph $H_1 \ltimes_I^a H_2$ as follows. Let $A_1, \ldots, A_{v(H_2)}$ be copies of H_1 with the set I identified and otherwise disjoint. Furthermore, place a copy of H_2 on the set of $v(H_2)$ vertices corresponding to a in each of $A_1, \ldots, A_{v(H_2)}$.
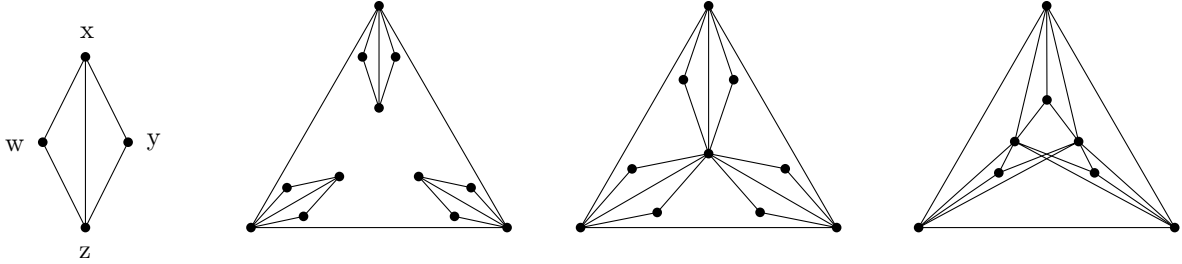

Figure 1: Let D denote the first graph, which is a labelled diamond. The remaining graphs are, from left to right: $D \ltimes_\varnothing^x K_3, D \ltimes_{\{z\}}^x K_3$ and $D \ltimes_{\{y,w\}}^x K_3$.

Theorem 1.4. *Suppose that H_1, H_2 are KNRS, I is an independent set in H_1 and $a \in V(H_1) \setminus I$. Then $H_1 \ltimes_I^a H_2$ is KNRS.*

This theorem already captures many of the known partial results about Conjecture 1.2. For example, if H_1 is the star with t leaves[2], I is the set of its leaves and a is the center, then $H_1 \ltimes_I^a H_2$ is obtained from H_2 by adding t vertices connected to all vertices in H_2; thus repeated applications of Theorem 1.4 imply that all complete multipartite graphs are KNRS. Similarly, applying it with $H_1 = P_\ell$ and $H_2 = K_2$, where a and b are the two endpoints of the path[3] and $I = \{b\}$, we obtain that the odd cycle $C_{2\ell+1}$ is KNRS, recovering the result of Reiher [23]. By adding an apex vertex (namely applying the theorem with $H_1 = K_2$), we find that all wheels are KNRS. Theorem 1.4 also implies that more complex graphs, such as the ones depicted in Figure 1, are KNRS.

The techniques used to prove Theorem 1.4 are based on those developed by Reiher [23] to prove that odd cycles are KNRS. However, we recast the entire problem, including a key lemma of Reiher's, in the more analytic setting of graphons, which we believe makes the proofs simpler and more transparent. Moreover, the true power of this perspective is even more pronounced in our subsequent results, which demonstrate surprising connections between Conjecture 1.2 and seemingly unrelated topics such as semidefinite optimization. To state our next main result, we make the following definition.

Definition 1.5. A graph H is called *KNRS-forcing* if, for every $p, \alpha \in (0, 1)$, there exist $\delta, \varepsilon > 0$ such that the following holds. For every (p, δ)-locally dense graph G, it holds that $t(H, G) > p^{e(H)} + \varepsilon$, unless G is (p, α)-quasirandom.

---

[2]An application of Jensen's inequality shows that all stars are Sidorenko, hence KNRS.

[3]Blakley and Roy [2] proved (in a different language) that all paths satisfy Sidorenko's conjecture, and hence are KNRS.

Thus, a graph is KNRS-forcing if it is KNRS, and, moreover, the only locally dense graphs which asymptotically minimize $t(H, G)$ are themselves quasirandom. Just as Sidorenko's conjecture is strengthened by the forcing conjecture, we pose the following strengthening of Conjecture 1.2.

Conjecture 1.6. *If H is not a forest, then H is KNRS-forcing.*

While we are not able to resolve Conjecture 1.6 in general, we are able to prove it for certain natural graph families.

Theorem 1.7. *All cliques, cycles, and wheels are KNRS-forcing. Also, if H_1 is KNRS-forcing, then so is $H_1 \ltimes_I^a H_2$ for all independent sets $I \subseteq V(H_1)$, all $a \in V(H_1) \setminus I$, and all KNRS graphs H_2.*

In fact, we are able to prove that $H_1 \ltimes_I^a H_2$ is KNRS-forcing even in certain instances when H_1 is not; we defer the precise statement to Theorem 3.2. We remark that our proof that odd cycles are KNRS-forcing is closely related to the fact that the Schatten norms are strictly convex on the positive-semidefinite cone. In particular, the KNRS-forcing condition—which roughly says that $t(H, \cdot)$ has a unique global minimum—can be deduced in this instance from the fact that a strictly convex function has a unique global minimum. We discuss this in greater detail later, but we believe that such connections are an important feature of our techniques.

Finally, we also pose a weakening of Conjecture 1.2 which we can, perhaps surprisingly, prove for a larger class of graphs. Let us say that an n-vertex graph G is (p, δ)-*nearly regular* if all but at most δn vertices of G have degree in the interval $[(p - \delta)n, (p + \delta)n]$.

Conjecture 1.8. *For every graph H and all $p, \varepsilon \in (0, 1)$, there exists some $\delta > 0$ such that the following holds. If G is a (p, δ)-locally dense (p, δ)-nearly regular graph, then $t(H, G) \geq p^{e(H)} - \varepsilon$.*

We say H is *regular-KNRS* if Conjecture 1.8 holds for H. In other words, to check that H is regular-KNRS, it suffices to prove that $t(H, G) \geq p^{e(H)} - o(1)$ over all nearly regular locally dense graphs G. It is known that in the setting of Sidorenko's conjecture, adding such a restriction is immaterial: H is Sidorenko if and only if $t(H, G) \geq p^{e(H)}$ for all (p, δ)-nearly regular graphs G with $t(K_2, G) = p$. A simple proof of a slightly weaker statement can be found in [16, Lemma 3.3], and stronger statements are proved in [10, 34]. Because of this, it is natural to expect that H is KNRS if and only if it is regular-KNRS, but we have been unable to prove this. However, since being regular-KNRS is a weaker condition, we were able to verify it for many more graphs.

Theorem 1.9. *The following classes of graphs are regular-KNRS.*

- *The ℓ-subdivision of a regular-KNRS graph, for any $\ell \geq 1$. In particular, all balanced subdivisions of cliques are regular-KNRS.*

- *Any graph obtained from a regular-KNRS graph by gluing forests to its vertices.*

- *Any generalized theta graph, consisting of an arbitrary number internally vertex-disjoint paths, of arbitrary lengths, between two fixed vertices.*

- *The graph $H = \langle\!\!\!\!\square\!\!\!\!\rangle$, obtained by gluing a triangle to two opposite edges of C_4.*

In particular, this result implies that all unicyclic graphs and all graphs obtained from a cycle by adding a chord are regular-KNRS, so the examples above include the graphs which Lee [18] proved are KNRS, as well as generalizations of them; however, in contrast to Lee, we are only able to prove the weaker condition of regular-KNRS. Additionally, H_0 was raised by Lee

[18] as the smallest example of a graph which is not known to be KNRS; again our techniques are not able to prove this, but are able to show the weaker condition of being regular-KNRS.

In the analytic perspective we take, the regular-KNRS problem turns out to be substantially more well-behaved. At a high level, the reason is that locally dense graphs correspond to *copositive matrices*, a class of matrices which is very poorly behaved. However, if one imposes the almost regularity condition, one instead obtains a positive semidefinite matrix, and thus we are able to use the rich theory of PSD matrices to prove results such as Theorem 1.9. In particular, certain operations which preserve the PSD cone, such as tensor product, Hadamard product, and matrix powers, correspond to useful operations on graphs. Because of this, we believe that Conjecture 1.8 is a natural setting in which to work and to obtain partial results, with the hopes of ultimately using similar techniques to resolve the full Conjecture 1.2.

The rest of this paper is organized as follows. In Section 2, we collect a number of general lemmas that we will use in our proofs, mostly related to the translation of these problems to the language of graphons. Proofs of many of these lemmas, as well as a more in-depth discussion of graph limits, are deferred to Appendix A. The proofs of Theorems 1.4 and 1.7 are given in Section 3. We prove results about regular-KNRS graphs, including Theorem 1.9, in Section 4. Finally, in Section 5, we collect a few examples which rule out natural avenues towards proving Conjecture 1.2.

2 Preliminaries

In this section, we collect a few preliminary definitions and results that we will need in the proofs of our main theorems. We begin by introducing a weaker forcing notion, which will be useful in certain of our proofs, and which is interesting in its own right.

Definition 2.1. A graph H is called *density-forcing* if, for every $p, \alpha \in (0, 1)$, there exist $\delta, \varepsilon > 0$ such that the following holds. If G is a (p, δ)-locally dense graph with at least $(p + \alpha)v(G)^2/2$ edges, then $t(H, G) > p^{e(H)} + \varepsilon$.

It is not hard to show that if H is density-forcing, it is also KNRS. Thus, loosely speaking, a graph H is density-forcing if it is KNRS and the asymptotic minimizers all have edge density p; this is weaker than being KNRS-forcing, which says that the asymptotic minimizers are p-quasirandom.

2.1 Graphons

For the rest of the paper, we will work in the setting of graphons, rather than of graphs, as this makes a number of analytic arguments substantially simpler. In this section we simply recall the basic definitions, and state Lemma 2.6, which restates all of the questions we are studying in the language of graphons. All of the results we need about graphons are essentially standard, but we provide detailed proofs in Appendix A. For a more thorough introduction to the theory of graph limits, see [19] or [36, Chapter 4].

Definition 2.2. Let (Ω, Σ, μ) be an atomless standard probability space. A *kernel* on Ω is a bounded symmetric measurable function $W : \Omega \times \Omega \to \mathbb{R}$. A kernel $W : \Omega \times \Omega \to [0, 1]$ is called a *graphon*.

As is standard in functional analysis, we will view two kernels that agree almost everywhere as the same object.

Usually when working with graphons (or kernels), one just sets Ω to be the interval $[0,1]$ endowed with Lebesgue measure, which is without loss of generality[4] as every atomless standard probability space is isomorphic to $[0,1]$. However, in some places in this paper, it will be convenient to work with more general probability spaces. When there is no confusion, we will use $|U|$ to denote the measure of a measurable subset $U \subseteq \Omega$, and write $\int_\Omega f \, \mathrm{d}x$ rather than $\int_\Omega f \, \mathrm{d}\mu(x)$ for the integral of a function $f : \Omega \to \mathbb{R}$ with respect to μ.

We now define some quantities associated to graphons. The first is a graphon analogue of the degree sequence of a graph.

Definition 2.3. If W is a graphon, its *degree function* is the function $d_W : \Omega \to [0,1]$ defined by

$$d_W(x) := \int_\Omega W(x,y) \, \mathrm{d}y.$$

W is called *p-regular* if $d_W = p$ a.e.

The second is the notion of homomorphism density, which extends naturally from graphs to kernels.

Definition 2.4. Given a graph H, its homomorphism density in a kernel W is given by

$$t_\mu(H, W) := \int_{\Omega^{V(H)}} \prod_{uv \in E(H)} W(x_u, x_v) \cdot \prod_{v \in V(H)} \mathrm{d}\mu(x_v). \tag{2}$$

When the measure μ is clear from the context, we will often omit the subscript. Additionally, for distinct $v_1, \ldots, v_k \in V(H)$ and any $x_1, \ldots, x_k \in \Omega$, we write

$$t(H, W \mid x_{v_1} = x_1, \ldots, x_{v_k} = x_k) := \int_{\Omega^{V(H) \setminus \{v_1, \ldots, v_k\}}} \prod_{uv \in E(H)} W(x_u, x_v) \cdot \prod_{v \in V(H) \setminus \{v_1, \ldots, v_k\}} \mathrm{d}x_v,$$

where on the right-hand side we plug in x_i for every instance of x_{v_i}.

To every graph G, one can associate a graphon W_G (see Definition A.1) such that $t(H, W_G) = t(H, G)$ for any graph H, hence this really is a generalization of the homomorphism density of graphs (see e.g. [36, Section 4.3] for more details).

Similarly, the definition of a locally dense graph translates naturally to that of a locally dense graphon. Because graphons already capture the asymptotic nature of the problem, the parameter δ from Definition 1.1 does not appear in the following definition.

Definition 2.5. A graphon W on (Ω, Σ, μ) is called *p-locally dense* if, for every $U \subseteq \Omega$ we have that

$$\iint_{U \times U} W(x, y) \, \mathrm{d}x \, \mathrm{d}y \geq p|U|^2.$$

Locally dense graphons precisely capture the asymptotic theory of locally dense graphs (see Lemma A.5 for a precise statement). In particular, the following result shows that all of the properties we are studying are equivalently captured by the behavior of homomorphism counts in locally dense graphons.

Lemma 2.6. *Let H be a graph.*

---

[4]In fact, a graphon on an arbitrary probability space can be converted into one on $[0,1]$ with essentially no loss of generality; see [19, Chapter 13].

(a) H is KNRS if and only if $t(H, W) \geq p^{e(H)}$ for every p and every p-locally dense graphon W.

(b) H is KNRS-forcing if and only if $t(H, W) > p^{e(H)}$ for every p and every p-locally dense graphon W, unless $W = p$ a.e.

(c) H is density-forcing if and only if $t(H, W) > p^{e(H)}$ for every p and every p-locally dense graphon W with $\iint_{\Omega \times \Omega} W > p$.

2.2 An infinitary version of Reiher's lemma

Reiher [23] proved a very important lemma, which has been used in many of the results on locally dense graphs; we will also use it multiple times throughout this paper. Informally, it states that a locally dense graph also satisfies a "fractional" version of the locally dense property in Definition 1.1. Formally, it states the following.

Lemma 2.7. *Let G be a (p, δ)-locally dense graph with n vertices. Let $f : V(G) \to [0, 1]$ be a function satisfying $\sum_{v \in V(G)} f(v) \geq \delta n$. Then*

$$\sum_{xy \in E(G)} f(x) f(y) \geq \frac{p}{2} \left(\sum_{x \in V(G)} f(x) \right)^2 - n.$$

Note that if we let f be the indicator function of a subset $S \subseteq V(G)$, then we precisely recover the condition in Definition 1.1, apart from the lower-order error term. Naturally, there is an infinitary analogue of Lemma 2.7. We use the standard notation $\|f\|_1$ to denote the L^1 norm of $f : \Omega \to \mathbb{R}$, that is, $\|f\|_1 := \int_\Omega |f(x)| \, \mathrm{d}x$.

Lemma 2.8. *Let W be a graphon on Ω. W is p-locally dense if and only if for every bounded measurable function $f : \Omega \to [0, \infty)$, we have*

$$\iint_{\Omega \times \Omega} f(x) W(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y \geq p \|f\|_1^2. \tag{3}$$

An immediate but important corollary of Lemma 2.8 is that the property of being locally dense is invariant under changing the probability measure on Ω. To state this result properly, we need to introduce a bit of extra notation. As we will now be dealing with different measures on Ω, all of our integrals will be written with respect to a given measure, e.g. we will write $\iint_{\Omega \times \Omega} W(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$. Similarly, all L^p norms will be computed with respect to a given measure, and we denote, for example, $\|f\|_{1,\mu}$ for $\int_\Omega |f(x)| \, \mathrm{d}\mu(x)$. Let us also say that a graphon W is p-*locally dense with respect to* μ if Definition 2.5 holds for the measure μ, that is, if $\iint_{U \times U} W(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) \geq p \mu(U)^2$ for all measurable sets $U \subseteq \Omega$.

Given a weight function $w : \Omega \to [0, \infty)$ with $\int_\Omega w(x) \, \mathrm{d}\mu(x) = 1$, we can obtain a new probability measure ν_w on Ω by integrating w, namely the measure of a set U is defined as $\nu_w(U) := \int_U w(x) \, \mathrm{d}\mu(x)$. Equivalently, ν_w can be defined by saying that for any bounded function $f : \Omega \to \mathbb{R}$, we have $\int_\Omega f(x) \, \mathrm{d}\nu_w(x) = \int_\Omega f(x) w(x) \, \mathrm{d}\mu(x)$. With this setup, we can now state the result that changing the measure preserves local density.

Lemma 2.9. *Let (Ω, Σ, μ) be an atomless standard probability space, and let W be a graphon on Ω. Let $w : \Omega \to [0, \infty)$ be a bounded weight function with $\int_\Omega w(x) \, \mathrm{d}\mu(x) = 1$, and let ν_w be the associated probability measure on Ω. If W is p-locally dense with respect to μ, then it is p-locally dense with respect to ν_w.*

In our proofs, the following consequence of Lemma 2.9 will be particularly useful. Both Lemmas 2.9 and 2.10 are proved in Appendix A.

Lemma 2.10. *Let* (Ω, Σ, μ) *be an atomless standard probability space, and let* W *be a graphon on* Ω. *Let* $w : \Omega \to [0, \infty)$ *be a bounded weight function. If* H *is a KNRS graph, then*

$$\int_{\Omega^{V(H)}} \prod_{v \in V(H)} w(x_v) \prod_{uv \in E(H)} W(x_u, x_v) \prod_{v \in V(H)} \mathrm{d}\mu(x_v) \geq \|w\|_1^{v(H)} p^{e(H)}.$$

2.3   Kernels as linear and bilinear operators

A kernel can be thought of as an infinite-dimensional generalization of a matrix, and just as matrices act as linear and bilinear operators, so do kernels. We collect here the important properties and definitions that we will need about such actions; a detailed introduction can be found in [19, Section 7.5], and the functional analysis background can be found in, for example, [33, Chapter 4] or [22, Chapter VI].

Given a kernel W on Ω and a parameter $1 \leq p \leq \infty$, we can define a linear operator $T_W : L^p(\Omega) \to L^p(\Omega)$ by

$$(T_W f)(x) := \int_\Omega W(x, y) f(y) \, \mathrm{d}y.$$

The fact that W is bounded and that Ω is a probability space implies that $T_W f \in L^p$ whenever $f \in L^p$. A function $f \in L^2$ is called an *eigenfunction* of T_W with *eigenvalue* λ if $T_W f = \lambda f$ a.e. It follows from standard Hilbert–Schmidt theory that T_W is a self-adjoint compact operator on L^2, hence the spectral theorem implies that there exist real numbers $\lambda_1, \lambda_2, \ldots$ and functions $f_1, f_2, \cdots \in L^2$ such that each f_i is an eigenfunction of T_W with eigenvalue λ_i, and such that the set $\{f_i\}$ forms an orthonormal basis of L^2. In particular, this implies that there is a spectral decomposition of W as

$$W(x, y) \sim \sum_{i=1}^{\infty} \lambda_i f_i(x) f_i(y).$$

Here, the symbol $\sim$ denotes that the series on the right-hand side may not converge for a.e. x, y, but that the right-hand side converges to $W(x, y)$ in L^2. We note for future reference that, although a priori the eigenfunctions f_i are only in L^2, one can show that eigenfunctions associated to a non-zero eigenvalue are bounded; see [19, Proposition 7.17] for a proof. The following result, an analogue of one of the basic results in spectral graph theory, shows that the eigenvalues of a graphon can be used to compute the homomorphism count of any cycle.

Theorem 2.11 ([19, (7.22)]). *Let* W *be a kernel and let* $\lambda_1, \lambda_2, \ldots$, *be its sequence of eigenvalues with* $|\lambda_1| \geq |\lambda_2| \geq \ldots$. *Then, for any* $k \geq 2$,

$$t(C_k, W) = \sum_{i=1}^{\infty} \lambda_i^k.$$

Recall that a matrix $M \in \mathbb{R}^{n \times n}$ is called *copositive* if $v^T M v \geq 0$ for all vectors $v \in \mathbb{R}^n$ all of whose coordinates are non-negative. Extending the definition of copositivity from matrices, let us say that a kernel W is *copositive* if, for every bounded $f : \Omega \to [0, \infty)$, we have

$$\iint_{\Omega \times \Omega} f(x) W(x, y) f(y) \geq 0.$$

More restrictively, we say that W is *positive semidefinite* if the same condition holds for every bounded $f : \Omega \to \mathbb{R}$, rather than only for bounded non-negative functions.

Recall that L^2 is endowed with a natural inner product, given by $\langle f, g \rangle = \int_\Omega f(x)g(x)\,\mathrm{d}x$. With this notation, we can more concisely say that W is copositive if $\langle f, T_W f \rangle \geq 0$ for every bounded non-negative f, and that it is positive semidefinite if $\langle f, T_W f \rangle \geq 0$ for every bounded real-valued f. We will need the following two basic lemmas; the first is a generalization of a well-known fact about matrices, while the second connects locally dense graphons to copositive kernels. Proofs of both are given in Appendix A.

Lemma 2.12. *A kernel W is positive semidefinite if and only if all eigenvalues of T_W are non-negative.*

Lemma 2.13. *A graphon W is p-locally dense if and only if the kernel $W - p$ is copositive.*

In general, being positive semidefinite is a much stronger condition than being copositive, and the space of positive semidefinite kernels is much more well-behaved than the space of copositive kernels. Nonetheless, there is a simple condition which can be used to show that a copositive kernel is positive semidefinite.

Lemma 2.14. *Let W be a kernel. Suppose that there exists a bounded function $g : \Omega \to \mathbb{R}$ with $\inf_\Omega g > 0$ and $T_W g = 0$ a.e. If W is copositive, then it is positive semidefinite.*

Proof. Let $f : \Omega \to \mathbb{R}$ be a bounded function. Since $\inf g > 0$, for sufficiently large $C \in \mathbb{R}$, we have $f + Cg \geq 0$. Moreover, $f + Cg$ is bounded, so the copositivity of W implies that

$$0 \leq \langle f + Cg, T_W(f + Cg) \rangle = \langle f, T_W f \rangle + 2C \langle f, T_W g \rangle + C^2 \langle T_W g, T_W g \rangle = \langle f, T_W f \rangle.$$

As f was arbitrary, W is positive semidefinite. $\square$

We will mostly use this lemma in the form of the following corollary.

Corollary 2.15. *Let W be a p-locally dense, p-regular graphon. Then $W - p$ is positive semidefinite.*

Proof. If we let g denote the constant 1 function, then the p-regularity of W is equivalent to saying that $T_W g = pg$. Hence $T_{W-p} g = 0$. As $W - p$ is copositive by Lemma 2.13, Lemma 2.14 implies that $W - p$ is positive semidefinite. $\square$

Corollary 2.15 is, essentially, the reason why we are able to prove substantially more about regular-KNRS graphs than KNRS graphs. Indeed, since the space of PSD kernels is well-behaved, many arguments work for proving that a given graph is regular-KNRS but fail for proving that it is KNRS.

Given two kernels W_1, W_2, their *operator product* $W_1 \circ W_2$ is defined by

$$(W_1 \circ W_2)(x, y) := \int_\Omega W_1(x, z) W_2(z, y)\,\mathrm{d}z.$$

Note that $W_1 \circ W_2$ is another kernel, as the boundedness of W_1, W_2 implies that $W_1 \circ W_2$ is bounded as well. This definition extends the standard definition of matrix multiplication, and is chosen so that $T_{W_1 \circ W_2} = T_{W_1} \circ T_{W_2}$, where $\circ$ on the right-hand side denotes function composition. We denote the *operator power*, defined as the k-fold operator product of W with itself, by $W^{\circ k}$. The main result we will need about operator products is that W and $W^{\circ k}$ have the same eigenfunctions, and that the eigenvalues of $W^{\circ k}$ are given by $\{\lambda_i^k\}$, where $\{\lambda_i\}$ are the eigenvalues of W.

Another important operation is the *tensor product*. Given two kernels W_1, W_2 on Ω, their tensor product $W_1 \otimes W_2$ is a kernel on $\Omega \times \Omega$ defined by

$$(W_1 \otimes W_2)((x_1, x_2), (y_1, y_2)) := W_1(x_1, y_1) W_2(x_2, y_2).$$

It is not hard to show, extending the well-known property of matrix tensor products, that the eigenvalues of $W_1 \otimes W_2$ are given by $\{\lambda_i \mu_j\}_{i,j=1}^{\infty}$, where $\{\lambda_i\}, \{\mu_j\}$ are the eigenvalues of W_1, W_2, respectively. The most useful property of the tensor product is that graph homomorphism densities are multiplicative over it, that is, for every graph H and every pair of kernels W_1, W_2, we have

$$t(H, W_1 \otimes W_2) = t(H, W_1) t(H, W_2). \tag{4}$$

We denote the *tensor power*, defined as the k-fold tensor product of W with itself, by $W^{\otimes k}$.

Lemma 2.16. *If W is a positive semidefinite kernel, then $W^{\circ k}, W^{\otimes k}$ are positive semidefinite for all $k \geq 1$.*

Proof. This follows immediately from Lemma 2.12 and the characterization of the eigenvalues of $W^{\circ k}$ and $W^{\otimes k}$. $\qquad\square$

In addition to matrix powers and tensor products, another matrix operation that preserves positive-semidefiniteness is the Hadamard product. In the world of kernels, this corresponds to the pointwise product, namely given kernels W_1, W_2, we define the Hadamard product $W_1 \odot W_2$ by $(W_1 \odot W_2)(x, y) := W_1(x, y) W_2(x, y)$. Since W_1, W_2 are measurable and bounded, so is $W_1 \odot W_2$, hence it is also a kernel. As with matrices, the Hadamard product preserves the positive semidefinite cone.

Lemma 2.17. *If W_1, W_2 are positive semidefinite kernels, then so is $W_1 \odot W_2$.*

The analogous statement for matrices follows immediately from the fact that the tensor product preserves positive-semidefiniteness, since the Hadamard product of two matrices is a principal submatrix of their tensor product. It does not seem that one can make such an argument for kernels, so our proof of Lemma 2.17 is rather long and is given in Appendix A.

3 Gluing operations

In this section, we prove that various graphs are KNRS and KNRS-forcing. We begin with a strengthening of Reiher's theorem [23], proving that all cycles are KNRS-forcing.

Theorem 3.1. *For any $k \geq 3$, the k-cycle is KNRS-forcing.*

Proof. If k is even, the result was proved in the seminal paper of Chung, Graham and Wilson [4]. Hence, we will assume that k is odd and write $k = 2\ell + 1$. By Lemma 2.6(b), it is enough to prove that for any p-locally dense graphon W, it holds that $t(H, W) = p^{e(H)}$ if and only if $W = p$ a.e. So let W be a p-locally dense graphon with $t(H, W) = p^{e(H)}$.

First we show that W is p-regular. Indeed, if not, then by a result of Sidorenko [30], we have $\|W^{\circ \ell}\|_1 = t(P_\ell, W) > p^\ell$, where P_ℓ is the path with ℓ edges. In other words, if W is not p-regular, we have strictly more ℓ-paths than expected. Note that we may count k-cycles by counting pairs of ℓ-paths starting from the same vertex and whose endpoints are joined by an edge. Formally,

$$
\begin{aligned}
t(C_k, W) &= \int_\Omega \left(\int_{\Omega \times \Omega} W^{\circ \ell}(x, y) W^{\circ \ell}(x, z) W(y, z) \, \mathrm{d}y \, \mathrm{d}z \right) \mathrm{d}x \\
&\geq \int_\Omega (d_{W^{\circ \ell}}(x))^2 p \, \mathrm{d}x \geq \|W^{\circ \ell}\|_1^2 p \geq p^{2\ell + 1} = p^k,
\end{aligned}
$$

where in the first inequality we used Lemma 2.8 with the weight function $w(y) = W^{\circ \ell}(x, y)$, in the second inequality we used Cauchy-Schwarz and in the third inequality we used that the ℓ-path is Sidorenko.

By assumption $t(C_k, W) = p^k$ so each of the above inequalities must be tight. In particular, $t(P_\ell, W) = p^\ell$, which by a result of Sidorenko [30], implies that W is p-regular.

Let us write $W = p + M$. Note that $\langle M, 1 \rangle = 0$ so 1 is an eigenfunction of W with eigenvalue p. Let $1 = f_1, f_2, f_3, \ldots$, be eigenfunctions of W with corresponding eigenvalues $p = \lambda_1, \lambda_2, \lambda_3, \ldots$ such that $\{f_i\}_{i \in \mathbb{N}}$ forms an orthornomal basis. Note that for $i \geq 2$,

$$\langle W, f_i \rangle = \langle p, f_i \rangle + \langle M, f_i \rangle = \langle M, f_i \rangle,$$

since f_i is orthogonal to the all one function. Hence, f_i is also an eigenfunction of M with the same eigenvalue. By Corollary 2.15, M is positive semidefinite, which implies $\lambda_i \geq 0$, for all $i \geq 1$. Finally, by Theorem 2.11, we have

$$t(C_k, W) = \sum_{i=1}^\infty \lambda_i^k = p^k + \sum_{i=2}^\infty \lambda_i^k \geq p^k.$$

Equality implies that $\lambda_i = 0$ for all $i \geq 2$, and thus $W = p$ a.e. as claimed. $\qquad \square$

We remark that there is an alternative way of viewing the proof of Theorem 3.1. For a kernel W and an integer $k \geq 3$, let $f_k(W) = \sum_{i=1}^\infty \lambda_i^k$, where $\{\lambda_i\}_{i=1}^\infty$ are the eigenvalues of W. From Theorem 2.11, we know that $f_k(W) = t(C_k, W)$. The function f_k, when k is odd, is not a convex function on the space of all kernels, but it is strictly convex when restricted to the cone of PSD kernels; this is closely related to the statement that the Schatten p-norm, for $p > 1$, is strictly convex. A basic fact of convex optimization is that strictly convex functions have a unique local minimum (which is also the global minimum). This implies that $t(C_k, \cdot)$ has a unique local and global minimum in the cone of PSD kernels, hence that a stability result holds for the count of C_k. As it is not hard to identify this minimum, this stability result in turn implies that C_k is KNRS-forcing.

While this connection may appear somewhat superficial—and we found it easiest to prove Theorem 3.1 without speaking about strictly convex functions at all—we are hopeful that further such connections will be uncovered. In particular, it seems possible that techniques from convex and semidefinite optimization could be used to prove results around Conjecture 1.2.

We now recall the key definition of our gluing operation (see Figure 1 for an illustration).

Definition 1.3. Let H_1, H_2 be two graphs, let I be an independent set in H_1 and let a be a vertex in $V(H_1) \setminus I$. We define the graph $H_1 \ltimes_I^a H_2$ as follows. Let $A_1, \ldots, A_{v(H_2)}$ be copies of H_1 with the set I identified and otherwise disjoint. Furthermore, place a copy of H_2 on the set of $v(H_2)$ vertices corresponding to a in each of $A_1, \ldots, A_{v(H_2)}$.

Theorem 3.2. *Suppose that H_1, H_2 are KNRS, I is and independent set in H_1 and $a \in V(H_1) \setminus I$. Then $H_1 \ltimes_I^a H_2$ is KNRS. Moreover, if H_1 is KNRS-forcing or H_2 is density-forcing and $H_1 \ltimes_I^a K_2$ is KNRS-forcing, then $H_1 \ltimes_I^a H_2$ is KNRS-forcing.*

Proof. We first prove that $H_1 \ltimes_I^a H_2$ is KNRS and the forcing statement will follow by examining the equality case. Note that $e(H_2 \ltimes_I^a H_2) = v(H_2)e(H_1) + e(H_2)$. Let W be a p-locally dense graphon. Let $k = |I|$ and let us label the vertices of I by $1, \ldots, k$. For $z_1, \ldots, z_k, y \in \Omega$, let $w_{z_1, \ldots, z_k}(y) = t(H_1, W \mid x_1 = z_1, \ldots, x_k = z_k, x_a = y)$, that is, $w_{x_1, \ldots, x_k}(y)$ counts the number of copies of H_1 with vertex i embedded into x_i for $i \in [k]$ and vertex a embedded into y. Observe

that

$$t(H_1 \ltimes_I^a H_2, W \mid x_1 = z_1, \ldots, x_k = z_k)$$

$$= \int_{\Omega^{V(H_2)}} \prod_{v \in V(H_2)} w_{z_1,\ldots,z_k}(x_v) \prod_{uv \in E(H_2)} W(x_u, x_v) \prod_{v \in V(H_2)} \mathrm{d}x_v$$

$$\geq \|w_{z_1,\ldots,z_k}\|_1^{v(H_2)} p^{e(H_2)},$$

where in the inequality we used Lemma 2.10 and the assumption that H_2 is KNRS. Since H_1 is KNRS,

$$t(H_1, W) = \int_{\Omega^k} \|w_{z_1,\ldots,z_k}\|_1 \prod_{i=1}^{k} \mathrm{d}z_i \geq p^{e(H_1)}.$$

Integrating over the choice of $z_1, \ldots, z_k$, applying Lemma 2.10 and Jensen's inequality for the function $x^{v(H_2)}$ on $[0, \infty)$, we have

$$t(H_1 \ltimes_I^a H_2, W) = \int_{\Omega^k} t(H_1 \ltimes_I^a H_2, W \mid x_1 = z_1, \ldots, x_k = z_k) \prod_{i=1}^{k} \mathrm{d}z_i$$

$$\geq \int_{\Omega^k} \|w_{z_1,\ldots,z_k}\|_1^{v(H_2)} p^{e(H_2)} \prod_{i=1}^{k} \mathrm{d}z_i \geq t(H_1, W)^{v(H_2)} p^{e(H_2)} \geq p^{v(H_2)e(H_1)+e(H_2)},$$

showing that $H_1 \ltimes_I^a H_2$ is KNRS.

Finally, to show the forcing statements, suppose that $t(H_1 \ltimes_I^a H_2, W) = p^{v(H_2)e(H_1)+e(H_2)}$. Then, in the above chain of inequalities, equality holds throughout. In particular, $t(H_1, W) = p^{e(H_1)}$. Since H_1 is KNRS-forcing, we obtain that $W = p$ a.e. Hence $H_1 \ltimes_I^a H_2$ is also KNRS-forcing.

Finally, let us show that if H_2 is density-forcing and $H_1 \ltimes_I^a K_2$ is KNRS-forcing, then $H_1 \ltimes_I^a H_2$ is KNRS-forcing. In this case, we have

$$t(H_1 \ltimes_I^a H_2, W \mid x_1 = z_1, \ldots, x_k = z_k) = \|w_{z_1,\ldots,z_k}\|_1^{v(H_2)} p^{e(H_2)},$$
$$\text{for almost all } z_1, \ldots, z_k \text{ such that } \|w_{z_1,\ldots,z_k}\|_1 > 0. \tag{5}$$

Using (5) and the equality case of Jensen's inequality, it follows that

$$\|w_{z_1,\ldots,z_k}\|_1 = p^{e(H_1)}, \text{ for almost all } z_1, \ldots, z_k. \tag{6}$$

Consider $z_1, \ldots, z_k$ for which $\|w_{z_1,\ldots,z_k}\|_1 = p^{e(H_1)}$ and $t(H_1 \ltimes_I^a H_2, W \mid x_1 = z_1, \ldots, x_k = z_k) = p^{v(H_2)e(H_1)+e(H_2)}$. Let $\nu_{z_1,\ldots,z_k}(y) := w_{z_1,\ldots,z_k}(y)/\|w_{z_1,\ldots,z_k}\|_1$ and note that $\nu_{z_1,\ldots,z_k}(y)$ is a probability measure on W. Observe that

$$t(H_1 \ltimes_I^a H_2, W, \mid x_1 = z_1, \ldots, x_k = z_k)$$

$$= \int_{\Omega^{V(H_2)}} \prod_{v \in V(H_2)} w_{z_1,\ldots,z_k}(x_v) \prod_{uv \in E(H_2)} W(x_u, x_v) \prod_{v \in v(H_2)} \mathrm{d}\mu(x_v)$$

$$= \|w_{z_1,\ldots,z_k}\|_1^{v(H_2)} \cdot t_{\nu_{z_1,\ldots,z_k}}(H_2, W),$$

implying that $t_{\nu_{z_1,\ldots,z_k}}(H_2, W) = p^{e(H_2)}$. By Lemma 2.9, the graphon W is p-locally dense with respect to the measure $\nu_{z_1,\ldots,z_k}$. Since H_2 is density-forcing, we have

$$p = \int_{\Omega \times \Omega} W(y, y') \, d\nu_{z_1, \ldots, z_k}(y) \, d\nu_{z_1, \ldots, z_k}(y')$$

$$= \frac{1}{\|w_{z_1, \ldots, z_k}\|_1^2} \int_{\Omega \times \Omega} w_{z_1, \ldots, z_k}(y) w_{z_1, \ldots, z_k}(y') W(y, y') \, d\mu(y) \, d\mu(y')$$

$$= \frac{1}{\|w_{z_1, \ldots, z_k}\|_1^2} t(H_1 \ltimes_I^a K_2, W \mid x_1 = z_1, \ldots, x_k = z_k)$$

$$= p^{-2e(H_1)} t(H_1 \ltimes_I^a K_2, W \mid x_1 = z_1, \ldots, x_k = z_k).$$

In other words, for almost all $z_1, \ldots, z_k$, we have $t(H_1 \ltimes_I^a K_2, W \mid x_1 = z_1, \ldots, x_k = z_k) = p^{2e(H_1)+1}$, so

$$t(H_1 \ltimes_I^a K_2, W) = \int_{\Omega^k} t(H_1 \ltimes_I^a K_2, W \mid x_1 = z_1, \ldots, x_k = z_k) \prod_{i=1}^{k} dz_i = p^{2e(H_1)+1} = p^{e(H_1 \ltimes_I^a K_2)}.$$

Because $H_1 \ltimes_I^a K_2$ is KNRS-forcing, we obtain that $W = p$ a.e., so $H_1 \ltimes_I^a H_2$ is KNRS-forcing as well. $\qquad\square$

As a special case of Theorem 3.2, if H_1 is KNRS-forcing, then $H_1 \ltimes_I^a K_2$ is KNRS-forcing as well. The converse, however, need not hold, e.g. K_2 is clearly not KNRS-forcing, but denoting by x, y the endpoints of K_2, the graph $K_2 \ltimes_{\{x\}}^y K_2 \cong K_3$ is KNRS-forcing.

Corollary 3.3. *For any $k \geq 3$, the k-clique and the k-wheel are KNRS and KNRS-forcing.*

Proof. The proof of the first statement is by induction on k. By Theorem 3.1, K_3 is KNRS and KNRS-forcing and assume, by induction, that K_{k-1} is KNRS and KNRS-forcing. Let x, y be the two endpoints of K_2. We have that $K_k \cong K_2 \ltimes_{\{x\}}^y K_{k-1}$ for all k, showing by the previous theorem that K_k is KNRS. Moreover, as $K_2 \ltimes_{\{x\}}^y K_2 \cong K_3$ is KNRS-forcing and K_{k-1} is KNRS-forcing by induction, we have that K_k is KNRS-forcing.

Similarly, the wheel W_k can be constructed as $K_2 \ltimes_{\{x\}}^y C_k$. Since C_k is KNRS and KNRS-forcing by Theorem 3.1, and since $K_2 \ltimes_{\{x\}}^y K_2 \cong K_3$ is KNRS-forcing, we conclude that W_k is KNRS and KNRS-forcing. $\qquad\square$

4   On the regular-KNRS conjecture

In this section we prove Conjecture 1.8 for a family of graphs which includes many graphs for which Conjecture 1.2 is not known. Note that if W is a p-regular p-locally dense graphon, then by Corollary 2.15, $W - p$ is a 0-regular, positive semidefinite kernel. For any graph H, by expanding (2), we obtain

$$t(H, W) = t(H, p + (W - p)) = \sum_{H' \subseteq H} p^{e(H) - e(H')} t(H', W - p), \tag{7}$$

where H' runs over all spanning subgraphs of H. This motivates the following definition.

Definition 4.1. We say that a graph H is *PSD-nonnegative* if for any positive semidefinite 0-regular kernel W, it holds that $t(H, W) \geq 0$. We say that H is *hereditarily PSD-nonnegative* if every spanning subgraph of H is PSD-nonnegative.

We remark that a stronger notion has been studied in the literature. Namely, a graph H is said to be *positive* if $t(H, W) \geq 0$ for all kernels W. Clearly, every positive graph is in particular PSD-nonnegative. The positive graph conjecture of Antolín Camarena–Csóka–Hubai–Lippner–Lovász [1] predicts a precise characterization of positive graphs (see [9, 31] for recent progress on this conjecture). However, for our purposes we need the notion of hereditary PSD-nonnegativity, which does not seem to have a direct connection to the theory of positive graphs.

Equation (7) immediately implies the following.

Proposition 4.2. *Every hereditarily PSD-nonnegative graph satisfies Conjecture 1.8.*

Proof. Suppose that H is a hereditarily PSD-nonnegative graph. Let W be a p-regular p-locally dense graphon and let $W_0 := W - p$. Note that by Corollary 2.15, W_0 is 0-regular and positive semidefinite. By (7), we have

$$t(H, W) = \sum_{H' \subseteq H} p^{e(H) - e(H')} t(H', W_0) = p^{e(H)} + \sum_{H' \subseteq H, E(H') \neq \varnothing} p^{e(H) - e(H')} t(H', W_0) \geq p^{e(H)},$$

where in the inequality we used the assumption that all spanning subgraphs of H' are PSD-nonnegative. $\qquad\square$

Lemma 4.3. *If W is a positive semidefinite kernel and W' is an arbitrary kernel on the same probability space (Ω, Σ, μ), then the kernel $W' \circ W \circ W'$ is positive semidefinite.*

Proof. Let $f \colon \Omega \to [0, \infty)$ be an arbitrary bounded measurable function. Setting $\Phi(x') = \int_\Omega f(x) W(x, x') \, dx$, we have

$$\int_{\Omega \times \Omega} f(x)(W' \circ W \circ W')(x, y) f(y) \, dx \, dy = \int_{\Omega^4} f(x) W'(x, x') W(x', y') W'(y', y) f(y) \, dx \, dx' \, dy' \, dy$$

$$= \int_{\Omega \times \Omega} \Phi(x') W(x', y') \Phi(y') \, dx' \, dy' \geq 0,$$

where in the last inequality we used that W is positive semidefinite. As f was arbitrary, the statement follows. $\qquad\square$

Given a kernel W and a graph H with two distinct vertices labelled a and b, we define a new kernel $t_{a,b}(H, W)$ by setting

$$t_{a,b}(H, W)(x, y) = t(H, W \mid x_a = x, x_b = y).$$

For $s_1, \ldots, s_t$ with $s_1 \geq 0$ and $s_2, \ldots, s_t \geq 1$, let $\Theta(s_1, \ldots, s_t)$ be the graph consisting of two distinct vertices a and b and t internally vertex disjoint paths between a and b, with the ith path having s_i internal vertices, where if $s_1 = 0$, there is an edge between a and b. We call $\Theta(s_1, \ldots, s_t)$ a *generalized Θ-graph*.

Lemma 4.4. *Let $s_1 \geq 0, s_2 \ldots, s_t \geq 1$, be arbitrary, denote $\Theta = \Theta(s_1, \ldots, s_t)$ and let a, b be the two distinguished vertices of Θ. Then, for any positive semidefinite kernel W, the kernel $t_{a,b}(\Theta, W)$ is positive semidefinite.*

Proof. Indeed, observe that

$$t_{a,b}(\Theta, W)(x, y) = \prod_{i=1}^{t} W^{\circ s_i + 1}(x, y),$$

14

so we can write

$$t_{a,b}(\Theta, W) = \overset{t}{\underset{i=1}{\odot}} W^{\circ s_i + 1}.$$

By Lemma 2.16, $W^{\circ s_i + 1}$ is positive semidefinite for all $i \in [t]$ and by Lemma 2.17, the Hadamard product of two positive semidefinite kernels is positive semidefinite, implying that $t_{a,b}(\Theta, W)$ is positive semidefinite. $\qquad\square$

Lemma 4.5. *If H has a vertex of degree one, then H is PSD-nonnegative.*

Proof. In fact we prove that for any 0-regular kernel W, it holds that $t(H, W) = 0$. Let v be a vertex of degree one in H, let u be its unique neighbour and let $H' = H \setminus \{v\}$. Then for a 0-regular kernel W, we have

$$t(H, W) = \int_{\Omega} t(H', W \mid x_u = x) \int_{\Omega} W(x, y) \, \mathrm{d}y \, \mathrm{d}x = 0,$$

since $\int_{\Omega} W(x, y) \, \mathrm{d}y = 0$ for almost all x. $\qquad\square$

Theorem 4.6. *Any generalized Θ-graph is hereditarily PSD-nonnegative and, in particular, it satisfies Conjecture 1.8.*

Proof. Let H' be a subgraph of a generalized Θ-graph. We need to show that H' is PSD-nonnegative. Indeed, note that every spanning subgraph of a generalized Θ-graph is itself a generalized Θ-graph (potentially plus isolated vertices) or has a vertex of degree one. In the latter case, we are done by Lemma 4.5, so assume that H' is a generalized Θ-graph plus isolated vertices. Homomorphism densities are multiplicative over disjoint unions, so we may remove the isolated vertices without changing $t(H', W)$. By Lemma 4.4, $t_{a,b}(H', W)$ is positive semidefinite, so in particular,

$$t(H', W) = \int_{\Omega \times \Omega} t_{a,b}(H', W)(x, y) \, \mathrm{d}x \, \mathrm{d}y \geq 0,$$

as needed. $\qquad\square$

Theorem 4.7. *Let H be the graph with 6 vertices and 8 edges, where $V(H) = \mathbb{Z}_6$ and $E(H) = \{\{i, i+1\} : 1 \leq i \leq 6\} \cup \{\{1, 5\}\{2, 4\}\}$. H is hereditarily PSD-nonngetive and thus it satisfies Conjecture 1.8.*

Proof. See Figure 2 for an illustration of H and its labelling. We need to show that every spanning subgraph of H is PSD-nonnegative. Let H' be a spanning subgraph of H. It is easy to verify that H' satisfies at least one of the following cases which we handle in order.

- H' has a vertex of degree one. Then H' is PSD-nonnegative by Lemma 4.5.

- H' is a generalized Θ-graph. Then H' is PSD-nonnegative by Theorem 4.6.

- H' is a disjoint union of two triangles. For any kernel W, $t(H', W) = t(K_3, W)^2 \geq 0$.

- H' is two vertex disjoint triangles connected by a single edge. Let x, y denote the two vertices of degree 3 in H'. Let $f(x) = t(K_3, W \mid x_1 = x)$, i.e. $f(x)$ counts triangles containing x. Then $t(H', W) = \iint f(x) W(x, y) f(y) \geq 0$, since W is positive semidefinite.

- $H' \cong H$. Denote $F = H[\{1, 2, 3, 4, 5\}] \setminus \{\{1, 5\}\}$. Let $W_3 = t_{a,b}(K_3, W)$ where a, b are two distinct vertices of a triangle. Note that a triangle is a generalized Θ-graph so by Lemma 4.4, W_3 is a positive semidefinite kernel. Additionally, observe that $t_{1,5}(F, W) = W \odot W_3 \odot W$, so $t_{1,5}(F, W)$ is positive semidefinite by Lemma 4.3. Finally, note that

15

$t_{1,5}(H', W) = W \odot W^{\circ 2} \odot t_{1,5}(F, W)$ which is positive semidefinite by Lemma 2.16 and Lemma 2.17. In particular this implies that $t(H', W) = \int_{\Omega \times \Omega} t_{1,5}(H', W)(x, y) \, dx \, dy \geq 0$, as required. $\qquad \square$
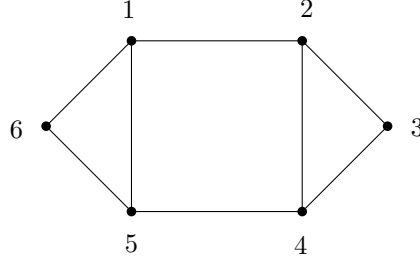

Figure 2: The graph H from Theorem 4.7.

The above proof strategy can be used to prove Conjecture 1.8 for a slightly larger family of graphs, we have not included the most general statement for the sake of clarity of presentation.

Recall that the ℓ-subdivision of a graph H is obtained by replacing each edge by a path of length $\ell + 1$.

Proposition 4.8. *If a graph H is regular-KNRS then for every $\ell \geq 1$, the ℓ-subdivision of H is regular-KNRS.*

Proof. Let H be regular-KNRS and let W be a p-regular p-locally dense graphon. Let $H^{\circ \ell}$ denote the ℓ-subdivision of H. Crucially, note that $t(H^{\circ \ell}, W) = t(H, W^{\circ \ell+1})$. By Corollary 2.15, $W_0 := W - p$ is a 0-regular positive semidefinite kernel. Observe that since W_0 is 0-regular, $p \circ W_0 = W_0 \circ p = 0$ a.e. Hence,

$$W^{\circ \ell+1} = (p + W_0)^{\circ \ell+1} = p^{\circ \ell+1} + W_0^{\circ \ell+1} = p^{\ell+1} + W_0^{\circ \ell+1}.$$

Since W_0 is positive semidefinite, by Lemma 2.16, $W_0^{\circ \ell+1}$ is positive semidefinite and in particular, it is copositive. Furthermore, $W_0^{\circ \ell+1}$ is 0-regular. Indeed, for any $x \in \Omega$, we have

$$\int_\Omega W_0^{\circ \ell+1}(x, y) \, dy = \int_{\Omega^\ell} W_0(x, z_1) \prod_{i=1}^{\ell-1} W_0(z_i, z_{i+1}) \int_\Omega W_0(z_\ell, y) \prod_{i=1}^{\ell} dz_i \, dy.$$

Since W_0 is 0-regular, for almost every choice of z_ℓ, we have $\int_\Omega W_0(z_\ell, y) \, dy = 0$ and so $\int_\Omega W_0^{\circ \ell+1}(x, y) \, dy = 0$ for all x, implying that $W_0^{\circ \ell+1}$ is 0-regular. We conclude that $W^{\circ \ell+1}$ is a $p^{\ell+1}$-locally dense, $p^{\ell+1}$-regular graphon. By the assumption, it follows that

$$t(H^{\circ \ell}, W) = t(H, W^{\circ \ell+1}) \geq (p^{\ell+1})^{e(H)} = p^{e(H^{\circ \ell})},$$

which completes the proof. $\qquad \square$

We remark that in the setting of Sidorenko's conjecture, Conlon, Kim, Lee, and Lee [7] proved a similar result, namely that the 1-subdivision of any KNRS graph is Sidorenko.

Finally, we show that in the setting of Conjecture 1.8, we can use the so-called tensor power trick. We start with a lemma.

Lemma 4.9. *Let W be a p-regular, p-locally dense graphon. Then $W \otimes W$ is a p^2-regular, p^2-locally dense graphon.*

Proof. Let $W_0 := W - p$. We can write $W \otimes W$ as follows:

$$W \otimes W = (W_0 + p) \otimes (W_0 + p) = W_0 \otimes W_0 + W_0 \otimes p + p \otimes W_0 + p^2,$$

where p^2 is the all-p^2 kernel on $\Omega \times \Omega$. By Corollary 2.15, W_0 is 0-regular and positive semidefinite. Clearly, the all-p graphon is positive semidefinite. Recall that the tensor product of two positive semidefinite kernels is itself positive semidefinite and that the sum of two positive semidefinite kernels is positive semidefinite. Hence, $W \otimes W - p^2 = W_0 \otimes W_0 + W_0 \otimes p + p \otimes W_0$ is positive semidefinite and, in particular, it is copositive. By Lemma 2.13, $W \otimes W$ is p^2-locally dense. It remains to verify that $W \otimes W$ is p^2-regular. Indeed, for almost all $(x,y) \in \Omega \times \Omega$, we have

$$\int_{\Omega \times \Omega} (W \otimes W)((x,y),(x',y')) \, \mathrm{d}x' \, \mathrm{d}y' = \int_{\Omega \times \Omega} W(x,x')W(y,y') \, \mathrm{d}x' \, \mathrm{d}y' = p^2. \qquad \square$$

Proposition 4.10. *Let H be a graph and suppose there is an absolute constant $c > 0$ such that $t(H, W) \geq cp^{e(H)}$ for every p-locally dense p-regular graphon W. Then the same holds with $c = 1$, that is, H is regular-KNRS.*

Proof. Let H, c be as in the statement and let W be an arbitrary p-locally dense p-regular graphon. Inductively applying Lemma 4.9, we obtain that for any $k \geq 0$, the graphon $W^{\otimes 2^k}$ is p^{2^k}-locally dense and p^{2^k}-regular. By (4) and the assumption, we have

$$t(H, W) = \left(t(H, W^{\otimes 2^k}) \right)^{1/2^k} \geq \left(cp^{2^k \cdot e(H)} \right)^{1/2^k} = c^{1/2^k} p^{e(H)}.$$

Taking $k \to \infty$, we obtain $t(H, W) \geq p^{e(H)}$, as claimed. $\qquad \square$

5 Negative results

In this section we collect counterexamples to several natural avenues of proving Conjecture 1.2.

Proposition 5.1. *For all $p \in (0, 1/3)$, there is a p-locally dense graphon W satisfying the followng:*

1. *$t(K_2, W) = \frac{9p}{8}$.*

2. *For any graphon W' which is not 0 a.e., $W - W'$ is not p-locally dense.*

3. *$W \otimes W$ is not p^2-locally dense.*

4. *For any $\ell \geq 5$, $W^{\circ \ell}$ is not p^ℓ-locally dense.*

Let us briefly comment on the statement of Proposition 5.1. The first two points imply the existence of a $(p, o(1))$-locally dense n-vertex graph with at least $(\frac{9p}{8} - o(1))\binom{n}{2}$ edges such that if we remove any set of $\Omega(n^2)$ edges, the graph is no longer $(p, o(1))$-locally dense. In particular, this presents a difficulty in establishing that a KNRS graph H is density forcing. Indeed, suppose such an example did not exist, so for any $(p, o(1))$ locally-dense graph G with $(p + \varepsilon)\binom{n}{2}$ edges, there is a set F of $\Omega(n^2)$ edges such that $G \setminus F$ is $(p, o(1))$-locally dense. It is easy to show that there are at least $\Omega(n^{v(H)})$ copies of H using edges of F, and thus we would have $t(H, G) \geq \Omega(n^{v(H)}) + t(H, G \setminus F) \geq (1 + \delta)n^{v(H)}p^{e(H)}$ for some $\delta > 0$, which would imply that H is density forcing.

The third point gives a counterexample to the direct way of trying to apply the tensor power trick to the setting of Conjecture 1.2. We remark that in the setting of Sidorenko's conjecture,

the tensor power trick allows one to reduce to the case when G is nearly regular; thus the third point also shows the difficulty of proving that Conjecture 1.8 implies Conjecture 1.2. Finally, the fourth point illustrates why we cannot prove an analogue of Proposition 4.8 without the assumption of p-regularity.

Proof of Proposition 5.1. Let W be the graphon arising from the matrix $A = (a_{ij})_{i,j}$ defined as

$$A = \begin{pmatrix} 3p & 0 & 0 & 0 \\ 0 & p & 2p & 2p \\ 0 & 2p & p & 2p \\ 0 & 2p & 2p & p \end{pmatrix}$$

In other words, let $\Omega = [0,1)$ endowed with the Lebesgue measure and for $i \in [4]$, denote $I_i = [(i-1)/4, i/4)$. Then $W(x,y)$ is given by a_{ij} where $x \in I_i, y \in I_j$. It is simple to verify that W is p-locally dense and that $t(K_2, W) = \frac{9p}{8}$.

To see Property 2, note that for $2 \le j \le 4$, the set $[0,1] \setminus I_j$ has density exactly p in W, and every pair $(x,y) \in [0,1]^2$ is contained in one of these sets.

To prove Property 3, consider the set $(I_1 \times I_2) \cup (I_2 \times I_1) \cup (I_2 \times I_2) \cup (I_2 \times I_3) \subseteq [0,1]^2$. This set has density $\frac{3}{4}p^2$ in $W \otimes W$.

Finally, for Property 4, note that for $x \in I_1$, we have $d_{W^{\circ \ell}}(x) = (3p/4)^\ell$. Indeed, $d_{W^{\circ \ell}}(x)$ counts the number of walks from x and if $x \in I_1$, any such walk stays in I_1. Hence,

$$\int_{I_1 \times I_1} W^{\circ \ell}(x,y)\,\mathrm{d}x\,\mathrm{d}y \le \int_{I_1} d_{W^{\circ \ell}}(x)\,\mathrm{d}x = \frac{3^\ell}{4^{\ell+1}}p^\ell.$$

For $\ell \ge 5$, we have $\frac{3^\ell}{4^{\ell+1}}p^\ell < \frac{1}{16}p^\ell = |I_1|^2 p^\ell$, so for $\ell \ge 5$, $W^{\circ \ell}$ is not p^ℓ-locally dense. $\qquad \square$

Finally, it might be tempting to conjecture that every graph is PSD-nonnegative, which would provide a direction for proving Conjecture 1.8. This however is not the case.

Proposition 5.2. *There exists a graph H and a positive-semidefinite 0-regular kernel W such that $t(H,W) < 0$.*

Proof. Let H be the 6-vertex graph in Figure 3 and let W be the kernel defined by the matrix

$$\begin{pmatrix} 18 & -12 & 12 & -12 & -6 \\ -12 & 35 & 28 & -19 & -32 \\ 12 & 28 & 56 & -44 & -52 \\ -12 & -19 & -44 & 35 & 40 \\ -6 & -32 & -52 & 40 & 50 \end{pmatrix}$$

as in the proof of Proposition 5.1. This example was found by computer search. It is straightforward to check that W is a positive semidefinite 0-regular kernel and that $t(H,W) < 0$. $\qquad \square$
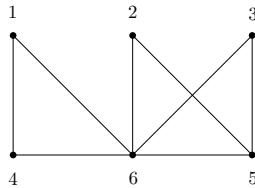

Figure 3: The graph H in Proposition 5.2 which is not PSD-nonnegative.

References

[1] O. Antolín Camarena, E. Csóka, T. Hubai, G. Lippner, and L. Lovász, Positive graphs, *European J. Combin.* **52** (2016), 290–301.

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

[3] M. Bucić, J. W. Cooper, D. Kráľ, S. Mohr, and D. Munhá Correia, Uniform Turán density of cycles, *Trans. Amer. Math. Soc.* **376** (2023), 4765–4809.

[4] F. R. K. Chung, R. L. Graham, and R. M. Wilson, Quasi-random graphs, *Combinatorica* **9** (1989), 345–362.

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

[6] D. Conlon, H. Hàn, Y. Person, and M. Schacht, Weak quasi-randomness for uniform hypergraphs, *Random Structures Algorithms* **40** (2012), 1–38.

[7] 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.

[8] D. Conlon and J. Lee, Sidorenko's conjecture for blow-ups, *Discrete Anal.* (2021), Paper No. 2, 13.

[9] D. Conlon, J. Lee, and L. Versteegen, Around the positive graph conjecture, 2024. Preprint available at arXiv:2404.17467.

[10] L. N. Coregliano and A. A. Razborov, Biregularity in Sidorenko's conjecture, 2021. Preprint available at arXiv:2108.06599.

[11] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.

[12] P. Erdős and V. T. Sós, On Ramsey-Turán type theorems for hypergraphs, *Combinatorica* **2** (1982), 289–295.

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

[14] R. Glebov, D. Kráľ, and J. Volec, A problem of Erdős and Sós on 3-graphs, *Israel J. Math.* **211** (2016), 349–366.

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

[16] J. H. Kim, C. Lee, and J. Lee, Two approaches to Sidorenko's conjecture, *Trans. Amer. Math. Soc.* **368** (2016), 5057–5074.

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

[18] J. Lee, On some graph densities in locally dense graphs, *Random Structures Algorithms* **58** (2021), 322–344.

[19] L. Lovász, *Large networks and graph limits*, *American Mathematical Society Colloquium Publications*, vol. 60, American Mathematical Society, Providence, RI, 2012.

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

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

[22] M. Reed and B. Simon, *Methods of modern mathematical physics. I: Functional analysis*, second ed., Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1980.

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

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

[25] C. Reiher, Extremal problems in uniformly dense hypergraphs, *European J. Combin.* **88** (2020), 103117, 22.

[26] C. Reiher, V. Rödl, and M. Schacht, On a generalisation of Mantel's theorem to uniformly dense hypergraphs, *Int. Math. Res. Not. IMRN* (2018), 4899–4941.

[27] C. Reiher, V. Rödl, and M. Schacht, On a Turán problem in weakly quasirandom 3-uniform hypergraphs, *J. Eur. Math. Soc. (JEMS)* **20** (2018), 1139–1159.

[28] V. Rödl and A. Ruciński, Threshold functions for Ramsey properties, *J. Amer. Math. Soc.* **8** (1995), 917–942.

[29] V. Rödl, A. Ruciński, and M. Schacht, Ramsey properties of random graphs and Folkman numbers, *Discuss. Math. Graph Theory* **37** (2017), 755–776.

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

[31] A. Sidorenko, On positive hypergraphs, *European J. Combin.* **106** (2022), Paper No. 103574, 5.

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

[33] E. M. Stein and R. Shakarchi, *Real analysis: Measure theory, integration, and Hilbert spaces*, *Princeton Lectures in Analysis*, vol. 3, Princeton University Press, Princeton, NJ, 2005.

[34] B. Szegedy, Sparse graph limits, entropy maximization and transitive graphs, 2015. Preprint available at arXiv:1504.00858.

[35] A. Thomason, Pseudorandom graphs, in *Random graphs '85 (Poznań, 1985)*, *North-Holland Math. Stud.*, vol. 144, North-Holland, Amsterdam, 1987, 307–331.

[36] Y. Zhao, *Graph theory and additive combinatorics—exploring structure and randomness*, Cambridge University Press, Cambridge, 2023.

A More on graphons

A.1 Graph limits

As usual, given a function $f : \Omega \to \mathbb{R}$ and a real number $p \geq 1$, we define the L^p norm of f as

$$\|f\|_p := \left(\int_\Omega |f(x)|^p \, \mathrm{d}x \right)^{1/p},$$

assuming this integral is finite. We also define $\|f\|_\infty$ to be the essential supremum of $|f|$. These definitions naturally extend to kernels, viewed as functions on the probability space $\Omega \times \Omega$; concretely, we define

$$\|W\|_p := \left(\iint_{\Omega \times \Omega} |W(x,y)|^p \, \mathrm{d}x \, \mathrm{d}y \right)^{1/p},$$

and similarly $\|W\|_\infty$ is the essential supremum of $|W|$. Note that since kernels are defined to be bounded, $\|W\|_p$ is finite for all $1 \leq p \leq \infty$.

More important in the study of graphons is the *cut norm*, which is defined for an arbitrary kernel as

$$\|W\|_\square := \sup_{S,T \subseteq \Omega} \left| \iint_{S \times T} W(x,y) \, \mathrm{d}x \, \mathrm{d}y \right|,$$

where the supremum runs over all pairs of measurable subsets $S, T \subseteq \Omega$. It is well-known [19, Lemma 8.10] that we can equivalently define the cut norm as

$$\|W\|_\square = \sup_{f,g:\Omega \to [0,1]} \left| \iint_{\Omega \times \Omega} f(x) W(x,y) g(y) \, \mathrm{d}x \, \mathrm{d}y \right|, \tag{8}$$

where the supremum runs over all pairs of measurable functions $f, g : \Omega \to [0,1]$.

The cut norm, like any norm, defines a natural metric on the space of kernels, but for the study of graphons it is more useful to define an "unlabeled" version of this metric. If $\varphi : \Omega \to \Omega$ is an invertible measure-preserving map, then we denote by W^φ the kernel given by $W^\varphi(x,y) := W(\varphi(x), \varphi(y))$. The *cut distance* between two kernels W_1, W_2 is then defined as

$$\delta_\square(W_1, W_2) := \inf_\varphi \|W_1 - W_2^\varphi\|_\square,$$

where the infimum runs over all measure-preserving invertible maps $\varphi : \Omega \to \Omega$.

As it turns out, every graph can naturally be viewed as a graphon, as we now define.

Definition A.1. Let G be a graph on vertex set $[n]$, and let Ω be an atomless standard probability space. Arbitrarily partition Ω into measurable sets $I_1, \ldots, I_n$, with $|I_i| = 1/n$ for all i (this is possible as Ω is atomless). For $x \in \Omega$, define $\iota(x)$ to be the unique index i such that $x \in I_i$. We then define a graphon $W_G : \Omega \times \Omega \to [0,1]$ by

$$W_G(x,y) := \begin{cases} 1 & \text{if } (\iota(x), \iota(y)) \in E(G), \\ 0 & \text{otherwise.} \end{cases}$$

Note that this definition depends both on the identification between $V(G)$ and $[n]$ and on the choice of partition of Ω, but these differences turn out to be immaterial, and we denote by W_G any graphon that arises in this way.

We now turn to the notion of graph limits.

Definition A.2. Let $G_1, G_2, \ldots$ be a sequence of graphs. We say that this sequence *converges* to a graphon W if $\delta_\square(W_{G_n}, W) \to 0$ as $n \to \infty$, and we write $G_n \to W$.

The most important basic result about graphons is that they are the limit objects for sequences of graphs.

Theorem A.3 (see [19, Theorem 11.21, Lemma 10.18, Proposition 11.32, and Corollary 11.34]). *If W is a graphon, then there exist graphs $G_1, G_2, \ldots$ such that $G_n \to W$.*

Conversely, if $G_1, G_2, \ldots$ is an arbitrary sequence of graphs with $v(G_n) \to \infty$, then there is a subsequence $G_{n_1}, G_{n_2}, \ldots$ converging to some graphon W.

A crucial property about homomorphism densities is that they characterize convergence of graph sequences.

Theorem A.4 (see e.g. [36, Theorem 4.3.7]). *A sequence of graphs $G_1, G_2, \ldots$ converges to a graphon W if and only if $t(H, G_n) \to t(H, W)$ for every graph H.*

A.2 Locally dense graphons

In this section, we prove various basic results about locally dense graphons, and ultimately prove Lemma 2.6.

Note that, given a (p, δ)-locally dense graph G, its associated graphon W_G is not p-locally dense for any $p > 0$ as, for example, the sets corresponding to a single vertex have measure $1/v(G)$ but the density of W_G inside them is zero. Hence, to make a meaningful connection, we must look at the limit of a sequence of locally dense graphs.

Lemma A.5. *Let W be a graphon on Ω. W is p-locally dense if and only if there exists a sequence of graphs $G_1, G_2, \ldots$ with $G_n \to W$ such that G_n is (p_n, δ_n)-locally dense, where $p_n \to p$ and $\delta_n \to 0$ as $n \to \infty$.*

Proof. Suppose first that W is p-locally dense. By Theorem A.3, there exists a sequence of graphs (G_n) with $G_n \to W$. So it suffices to prove that G_n is (p_n, δ_n)-locally dense, where $p_n = p + o(1)$ and $\delta_n = o(1)$, and all $o(1)$ terms tend to 0 as $n \to \infty$.

Recall that by the definition of convergence, we have that $\delta_\square(W_{G_n}, W) \to 0$. From the definitions of $\delta_\square$ and W_G, we see that we may select a partition of Ω into sets $I_1, \ldots, I_{v(G_n)}$, each of measure $1/v(G_n)$, so that the resulting graphon W_{G_n} satisfies $\|W_{G_n} - W\|_\square \to 0$. Let $\delta_n = \|W_{G_n} - W\|_\square^{1/3}$. We claim that G_n is $(p - \delta_n, \delta_n)$-locally dense, which suffices since $\delta_n \to 0$ as $n \to \infty$. Suppose for contradiction that G_n is not $(p - \delta_n, \delta_n)$-locally dense; this means that there exists $S \subseteq V(G_n)$ with $|S| \geq \delta_n v(G_n)$, such that $e_{G_n}(S) < (p - \delta_n)|S|^2/2$.

Let $U = \bigcup_{i \in S} I_i$. Then from the definition of W_{G_n}, we see that

$$\iint_{U \times U} W_{G_n}(x, y)\, \mathrm{d}x\, \mathrm{d}y = \frac{2e_{G_n}(S)}{v(G_n)^2} < (p - \delta_n)\frac{|S|^2}{v(G_n)^2} = (p - \delta_n)|U|^2.$$

On the other hand, as W is p-locally dense, we have that $\iint_{U \times U} W(x, y)\, \mathrm{d}x\, \mathrm{d}y \geq p|U|^2$. Combining these inequalities with the definition of the cut norm, we find that

$$\|W_{G_n} - W\|_\square \geq \iint_{U \times U} (W - W_{G_n})(x, y)\, \mathrm{d}x\, \mathrm{d}y > \delta_n|U|^2 \geq \delta_n^3,$$

which is a contradiction since our choice of δ_n implies that $\delta_n^3 = \|W_{G_n} - W\|_\square$.

It remains to prove the converse, so suppose that G_n is a sequence of (p_n, δ_n)-locally dense graphs, such that $p_n \to p, \delta_n \to 0$, and $G_n \to W$ as $n \to \infty$. As above, we may pick graphons

W_{G_n} such that $\|W_{G_n} - W\|_\square \to 0$ as $n \to \infty$. We may assume that $p > 0$, otherwise the statement is trivial as any graphon is 0-locally dense. Note that since $\delta_n \to 0$, we have that $v(G_n) \to \infty$. Suppose for contradiction that W is not p-locally dense, and let $U \subseteq \Omega$ be a measurable set such that $\iint_{U \times U} W < p|U|^2$. Let $\delta = |U|$, and let $\varepsilon > 0$ such that $\iint_{U \times U} W \le (p - \varepsilon)|U|^2$. We now define functions $f_n : V(G_n) \to [0, 1]$ as follows. Recall that W_{G_n} is obtained by partitioning Ω into sets $I_1, \ldots, I_{v(G_n)}$, each of measure $1/v(G_n)$. For $1 \le i \le v(G_n)$, we define $f_n(i) := |U \cap I_i|/|I_i|$ to be the fraction of I_i contained in U. Note that this definition immediately implies that

$$\frac{1}{v(G_n)^2} \sum_{(i,j) \in E(G_n)} f_n(i) f_n(j) = 2 \iint_{U \times U} W_{G_n}(x, y) \, \mathrm{d}x \, \mathrm{d}y. \tag{9}$$

Since $\delta_n \to 0$, we may pick N_1 such that $\delta_n \le \delta$ for all $n \ge N_1$. If $\delta_n \le \delta = |U|$, then $\sum_{v \in V(G_n)} f_n(v) = \delta v(G_n) \ge \delta_n v(G_n)$. Hence, for $n \ge N_1$, Lemma 2.7 and (9) imply that

$$\iint_{U \times U} W_{G_n}(x, y) \, \mathrm{d}x \, \mathrm{d}y \ge p_n |U|^2 - \frac{1}{v(G_n)}.$$

However, as $p_n \to p$ and $v(G_n) \to \infty$, we may pick N_2 such that for all $n \ge N_2$, we have $p_n |U|^2 - \frac{1}{v(G_n)} \ge (p - \frac{\varepsilon}{2})|U|^2$. Thus, for all $n \ge \max\{N_1, N_2\}$, we find that

$$\|W_{G_n} - W\|_\square \ge \iint_{U \times U} (W_{G_n} - W)(x, y) \, \mathrm{d}x \, \mathrm{d}y \ge \left(p - \frac{\varepsilon}{2} - (p - \varepsilon)\right)|U|^2 = \frac{\varepsilon}{2}|U|^2.$$

However, the left-hand side tends to 0 as $n \to \infty$, a contradiction. $\qquad\square$

Given Lemma A.5, it is not hard to prove Lemma 2.6.

Proof of Lemma 2.6. All three parts are proved in essentially the same way, by combining Theorems A.3 and A.4 and Lemma A.5. As such, we omit some details that are repeated.

(a) Suppose first that H is KNRS. If W is a p-locally dense graphon, then by Lemma A.5 there exists a sequence of graphs $G_n \to W$ which are $(p + o(1), o(1))$-locally dense. But then the fact that H is KNRS implies that $t(H, G_n) \ge p^{e(H)} - o(1)$. Since $t(H, W) = \lim_n t(H, G_n)$ by Theorem A.4, we find that $t(H, W) \ge p^{e(H)}$. For the converse, suppose that H is not KNRS. Then the definition of KNRS fails for some fixed $p, \varepsilon \in (0, 1)$. Thus, we may pick a sequence G_n of $(p, o(1))$-locally dense graphs with $t(H, G_n) \le p^{e(H)} - \varepsilon$. By Theorem A.3, there is a convergent subsequence, which necessarily converges to a p-locally dense graphon W by Lemma A.5. Hence $t(H, W) \le \limsup_n t(H, G_n) \le p^{e(H)} - \varepsilon$, a contradiction.

(b) This is proved in exactly the same way, with the added ingredient that a graphon W is equal to p a.e. if and only if it is the limit of a sequence of $(p + o(1), o(1))$-quasirandom graphs [19, Example 11.37].

(c) This is proved in exactly the same way, except now using the fact that if $G_n \to W$, then $t(K_2, G_n) \to t(K_2, W)$ and thus the edge density of W equals the asymptotic edge density of G_n. $\qquad\square$

A.3   Reiher's lemma and new measures

We begin by proving Lemma 2.8.

Proof of Lemma 2.8. The "if" direction is immediate, for if we set f to be the indicator function of a set U, the condition (3) is precisely the statement that $\iint_{U \times U} W \geq p|U|^2$. So it remains only to prove the "only if" direction.

Let W be a p-locally dense graphon, and fix a bounded measurable function $f : \Omega \to [0, \infty)$. Note that (3) is invariant under rescaling of f, so we may assume that f has codomain $[0, 1]$. There is nothing to prove if $f = 0$ a.e., so we may assume that $\int_\Omega f(x) \, \mathrm{d}x = \delta > 0$. By Lemma A.5, we may pick a sequence of graphs $G_1, G_2, \ldots$ such that G_n is (p_n, δ_n)-locally dense, where $p_n \to p$ and $\delta_n \to 0$, such that $\|W_{G_n} - W\|_\square \to 0$ as $n \to \infty$.

Recall that W_{G_n} is defined by a partition of Ω into sets $I_1, \ldots, I_{v(G_n)}$, each of measure $1/v(G_n)$. We define a function $f_n : V(G_n) \to [0, 1]$ by setting $f_n(i)$ to be the average value of f on the set I_i. Note that

$$\sum_{i \in V(G_n)} f_n(i) = v(G_n) \int_\Omega f(x) \, \mathrm{d}x = \delta v(G_n)$$

and that

$$\sum_{ij \in E(G)} f_n(i) f_n(j) = \frac{1}{2} v(G_n)^2 \iint_{\Omega \times \Omega} f(x) W_{G_n}(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y.$$

Let N_1 be chosen so that $\delta_n \leq \delta$ for all $n \geq N_1$. Applying Lemma 2.7 to G_n for $n \geq N_1$, we find that

$$p_n \|f\|_1^2 = \frac{p_n}{v(G_n)^2} \left(\sum_{i \in V(G_n)} f_n(i) \right)^2 \leq \frac{2}{v(G_n)^2} \sum_{ij \in E(G)} f_n(i) f_n(j) + \frac{2}{v(G_n)}$$

$$= \frac{2}{v(G_n)} + \iint_{\Omega \times \Omega} f(x) W_{G_n}(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y.$$

On the other hand, by (8), we have that

$$\left| \iint_{\Omega \times \Omega} f(x)(W_{G_n} - W)(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y \right| \leq \|W_{G_n} - W\|_\square \overset{n \to \infty}{\longrightarrow} 0.$$

Combining these two estimates, plus the facts that $p_n \to p$ and $v(G_n) \to \infty$ as $n \to \infty$, gives the claimed result. $\square$

We remark that we were unable to find a "direct" proof of Lemma 2.8, that is, a proof that does not use Lemma 2.7. The difficulty is that Reiher's proof of Lemma 2.7 uses the compactness of the unit ball in $\mathbb{R}^n$, whereas the unit ball in $L^2(\Omega)$ is not compact. There are standard ways of overcoming such difficulties (e.g. working in the weak topology, where the unit ball is compact by the Banach–Alaoglu theorem), but we were unable to make such techniques work in this context. It would be interesting to find a direct proof of Lemma 2.8.

As mentioned before, Lemma 2.9 is a direct consequence of Lemma 2.8.

Proof of Lemma 2.9. It suffices to prove that for every bounded $f : \Omega \to [0, \infty)$, we have

$$\iint_{\Omega \times \Omega} f(x) W(x, y) f(y) \, \mathrm{d}\nu_w(x) \, \mathrm{d}\nu_w(y) \geq p \|f\|_{1, \nu_w}^2.$$

By the definition of ν_w, the left-hand side equals

$$\iint_{\Omega \times \Omega} w(x) f(x) W(x, y) w(y) f(y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y),$$

and the right-hand side equals $p \left(\int_\Omega w(x) f(x) \, \mathrm{d}\mu(x) \right)^2$. However, if we define $g = wf$, then g is a bounded function (as both f and w are bounded), so Lemma 2.8 implies that

$$\iint_{\Omega \times \Omega} w(x) f(x) W(x, y) w(y) f(y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y) = \iint_{\Omega \times \Omega} g(x) W(x, y) g(y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$$

$$\geq p \|g\|_{1,\mu}^2 = p \left(\int_\Omega g(x) \, \mathrm{d}\mu(x) \right)^2$$

$$= p \left(\int_\Omega w(x) f(x) \, \mathrm{d}\mu(x) \right)^2.$$

This is precisely the claimed inequality. $\qquad \square$

As a consequence of Lemma 2.9, we can prove Lemma 2.10.

Proof of Lemma 2.10. If $\|w\|_1 = 0$, the inequality holds trivially, so we may assume $\|w\|_1 > 0$. Define $w'(x) = w(x)/\|w\|_1$ and let ν denote the probability measure associated to w'. By Lemma 2.9, W is p-locally dense with respect to ν. Since H is KNRS, we have

$$\int_{\Omega^{V(H)}} \prod_{uv \in E(H)} W(x_u, x_v) \prod_{v \in V(H)} w(x_v) \prod_{v \in V(H)} \mathrm{d}\mu(x_v) =$$

$$= \|w\|_1^{v(H)} \int_{\Omega^{V(H)}} \prod_{uv \in E(H)} W(x_u, x_v) \prod_{v \in V(H)} \mathrm{d}\nu(x_v) \geq \|w\|_1^{v(H)} p^{e(H)}.$$

$\qquad \square$

A.4 Kernels as linear and bilinear operators

In this section we prove the remaining lemmas from Section 2.3. We begin with Lemma 2.12.

Proof of Lemma 2.12. First suppose that $\lambda < 0$ for some eigenvalue λ of T_W, and let f be the corresponding eigenfunction with $\|f\|_2 = 1$. As stated above, f is bounded, and we have that $\langle f, T_W f \rangle = \lambda \|f\|_2^2 < 0$, showing that W is not positive semidefinite.

Conversely, suppose that all eigenvalues of T_W are non-negative, and fix a bounded function $f : \Omega \to \mathbb{R}$; note that this implies $f \in L^2$. Since the eigenfunctions of T_W form an orthonormal basis of L^2, we have that $\langle f, T_W f \rangle = \sum_{i=1}^\infty \lambda_i \langle f, f_i \rangle^2 \geq 0$, hence W is positive semidefinite. $\quad \square$

Next, we prove Lemma 2.13.

Proof of Lemma 2.13. For any bounded non-negative function f, we have that

$$\iint_{\Omega \times \Omega} f(x)(W - p)(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y = \iint_{\Omega \times \Omega} f(x) W(x, y) f(y) - p \|f\|_1^2,$$

since by Fubini's theorem $\iint_{\Omega \times \Omega} f(x) f(y) \, \mathrm{d}x \, \mathrm{d}y = (\int_\Omega f(x) \, \mathrm{d}x)^2$. But by Lemma 2.8, W is p-locally dense if and only if $\iint_{\Omega \times \Omega} f(x) W(x, y) f(y) \, \mathrm{d}x \, \mathrm{d}y \geq p \|f\|_1^2$ for all such functions, which by the above is equivalent to the statement that $W - p$ is copositive. $\qquad \square$

Finally, we prove Lemma 2.17. We remark that this result is fairly easy and well-known in case the kernels W_1, W_2 are assumed to be continuous functions $[0, 1]^2 \to [0, 1]$ (in this case it follows from Mercer's theorem, or from an appropriate infinitary analogue of the argument that the Hadamard product of two matrices is a principal submatrix of their tensor product). However, as we do not assume any continuity, the proof becomes somewhat lengthier.

Proof of Lemma 2.17. We write the spectral decompositions of W_1, W_2 as

$$W_1(x, y) \sim \sum_{i=1}^{\infty} \lambda_i f_i(x) f_i(y), \qquad W_2(x, y) \sim \sum_{j=1}^{\infty} \mu_j g_j(x) g_j(y)$$

where $\sim$ denotes convergence in $L^2(\Omega \times \Omega)$. In these spectral decompositions, we may assume that $\lambda_i, \mu_j \neq 0$ for all i, j (by simply ommitting the terms corresponding to zero eigenvalues), and hence we may assume that f_i, g_j are bounded for all i, j. Note that, by omitting the zero eigenvalues, we may no longer assume that $\{f_i\}$ forms an orthonormal basis of $L^2(\Omega)$ (as the eigenfunctions corresponding to zero eigenvalues are not spanned), but luckily we will not need them to form an orthonormal basis. Note that, by Lemma 2.12, we have that $\lambda_i, \mu_j > 0$ for all i, j.

Define $W = W_1 \odot W_2$. For positive integers K, M, let

$$W_1^K(x, y) := \sum_{i=1}^{K} \lambda_i f_i(x) f_i(y) \qquad \text{and} \qquad W_2^M := \sum_{j=1}^{M} \mu_j g_j(x) g_j(y).$$

Let $W^{KM} := W_1^K \odot W_2^M$, and note that we may write

$$W^{KM}(x, y) = \sum_{i=1}^{K} \sum_{j=1}^{M} \lambda_i \mu_j h_{ij}(x) h_{ij}(y),$$

where $h_{ij}(x) := f_i(x) g_j(x)$ is a bounded measurable function $\Omega \to \mathbb{R}$. In particular, this decomposition shows that W^{KM} is positive-semidefinite for all K, M, since for any bounded $a : \Omega \to \mathbb{R}$, we have

$$\iint_{\Omega \times \Omega} a(x) W^{KM}(x, y) a(y) \, \mathrm{d}x \, \mathrm{d}y = \sum_{i=1}^{K} \sum_{j=1}^{M} \lambda_i \mu_j \iint_{\Omega \times \Omega} a(x) h_{ij}(x) a(y) h_{ij}(y) \, \mathrm{d}x \, \mathrm{d}y$$

$$= \sum_{i=1}^{K} \sum_{j=1}^{M} \lambda_i \mu_j \left(\int_{\Omega} a(x) h_{ij}(x) \, \mathrm{d}x \right)^2 \geq 0,$$

where the final inequality holds since every summand is non-negative. We also note here for future reference that the L^2-limit of positive semidefinite kernels is positive semidefinite. Indeed, suppose that $X_1, X_2, \ldots, X$ are kernels with $\|X - X_k\|_{L^2(\Omega \times \Omega)} \to 0$, and suppose that each X_k is positive semidefinite. It is well-known (e.g. [33, Proposition 5.5(ii)]) that the $L^2 \to L^2$ operator norm of T_X is upper-bounded by $\|X\|_{L^2(\Omega \times \Omega)}$. Applying this to $X - X_k$, as well as Cauchy–Schwarz, implies that for any bounded $a : \Omega \to \mathbb{R}$, we have

$$|\langle a, T_X a \rangle - \langle a, T_{X_k} a \rangle| = |\langle a, T_{X - X_k} a \rangle| \leq \|a\|_2 \|T_{X - X_k} a\|_2 \leq \|a\|_2^2 \|T_{X - X_k}\|_{L^2 \to L^2}$$

$$\leq \|a\|_2^2 \|X - X_k\|_{L^2(\Omega \times \Omega)} \overset{k \to \infty}{\longrightarrow} 0.$$

In other words, $\langle a, T_{X_k} a \rangle \to \langle a, T_X a \rangle$ as $k \to \infty$, implying that $\langle a, T_X a \rangle \geq 0$, and hence that X is positive semidefinite. This shows, as claimed, that the limit of positive semidefinite kernels is positive semidefinite; we will shortly use this observation twice.

For a fixed K, we now claim that $W^{KM} \to W_1^K \odot W_2$ as $M \to \infty$, where the convergence is in $L^2(\Omega \times \Omega)$. Indeed, we have that

$$\|W^{KM} - W_1^K \odot W_2\|_2 = \|W_1^K(W_2^M - W_2)\|_2 \leq \|W_1^K\|_{\infty} \|W_2^M - W_2\|_2.$$

Since W_1^K is a finite sum of bounded functions, it is bounded, so $\|W_1^K\|_\infty$ is some finite number depending only on K. However, as $M \to \infty$, we have that $\|W_2^M - W_2\|_2 \to 0$, by the definition of W_2^M. This shows that, for any fixed K, we have $W^{KM} \to W_1^K \odot W_2$ in L^2. In particular, as we showed that each W^{KM} is positive semidefinite, we conclude that $W_1^K \odot W_2$ is positive semidefinite for all K.

We now argue in almost the same way that $W_1^K \odot W_2 \to W$ in L^2; as above, this implies that W is positive semidefinite, which is what we wanted to prove. So it remains to prove the convergence, which holds since

$$\|W_1^K \odot W_2 - W\|_2 = \|W_2(W_1^K - W_1)\|_2 \le \|W_2\|_\infty \|W_1^K - W_1\|_2,$$

and we know that $\|W_2\|_\infty$ is finite as W_2 is a kernel, and that $W_1^K \to W_1$ in L^2 by the definition of W_1^K. $\qquad\qquad\square$