View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Warwick Research Archives Portal Repository

A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/113462

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications

Theory of Combinatorial Limits and Extremal Combinatorics

by

Taísa Lopes Martins

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Mathematics Institute

September 2018

THE UNIVERSITY OF WARWICK

To my mom,
the bravest and fiercest angel.

Abstract

In the past years, techniques from different areas of mathematics have been successfully applied in extremal combinatorics problems. Examples include applications of number theory, geometry and group theory in Ramsey theory and analytical methods to different problems in extremal combinatorics.

By providing an analytic point of view of many discrete problems, the theory of combinatorial limits led to substantial results in many areas of mathematics and computer science, in particular in extremal combinatorics.

In this thesis, we explore the connection between combinatorial limits and extremal combinatorics. In particular, we prove that extremal graph theory problems may have unique optimal solutions with arbitrarily complex structure, study a property closely related to Sidorenko's conjecture, one of the most important open problems in extremal combinatorics, and prove a 30-year old conjecture of Győri and Tuza regarding decomposing the edges of a graph into triangles and edges.

Acknowledgments

First, I would like to thank God for always being present in my life and walking with me through this journey which is called a PhD. I'm thankful for all the experiences and opportunities. Second, I would like to thank my family for their infinite support, concern and love.

I'm grateful for all the love and support that I have received from everyone throughout these years. All of it was key to keep me motivated and make maths and life fun. In particular, thank you to all my house mates for giving me a place to call home outside of Brazil. I'm very thankful to all of you for being super nice and giving me such a happy environment.

I wouldn't have started my PhD without the help of Diana Sasaki, Hugo Nobrega, Luís Felipe Ignácio, Paloma Lima, Bruno Buss and Rafael Lopes (Bola). Thank you very much for all your support with my PhD application. Thank you for reading/revising my proposal and the emails to my supervisor, helping me with english, giving me emotional support and/or encouragement.

I also would like to thank the Science without Borders program, a brazilian government program, for funding me throughout my whole PhD and the European Research Council for funding my trips to summer schools, to attend conferences, and to visit different researchers to work on joint projects.

I'm very grateful to my officemates, especially Ronja Kühne, Massimo Iberti, Yanitsa Pehova and Jakub Sliačan for always making the office a warm place, sharing their lives, frustrations, happiness, food (particularly coffee, chocolate and biscuits) and knowledge. Thank you for all the math and non-math conversations.

To all my co-authors, I'm thankful for all the discussions and collaboration.

I have learned something new from each one of you.

From my research group, I would like to thank Jacob Cooper, Jordan Venters, Jakub Sosnovec, Ping Hu, Maryam Sharifzadeh, Katherine Staden and Hong Liu. In particular, thank you Jordan for always helping me when needed, all the pool matches and math discussions. Jake, Katherine, and Maryam; thank you for being yourselves. I will cherish you three forever!

My two favorite Brazilians in Warwick, Rafael Barbosa and Victor Quintela. Thank you Rafael for being my best friend in UK and Victor for all the times we cooked together, ordered pizza and went for a beer.

I'm very grateful to Dan Král', my supervisor. Thank you for giving me so many opportunities in my PhD, sharing your knowledge, all the collaboration and all the chocolate in your office. I would also like to thank Oleg Pikurko and John Talbot for being my examiners and the Warwick staff for all their work, especially Rimi and Carole.

Finally, I would like to thank four very special people: Bruno Buss, Paloma Lima, Denise Requiel and Marcia Martins. Bruno, thank you very much for your friendship. You always believe in me, support me and know how to help me whenever I feel lost. Paloma, one of my best friends, thank you for always knowing how to make me feel better, all our trips, our 3-hour skype sessions and all your support. Denise Requiel, thank you for all your messages, love and skype sessions; they made all the difference. Finally, thank you mom for the unconditional support, your faith in me and for moving mountains to help me. I will miss you forever.

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

Contents

Chapter 1

Introduction

A typical problem in extremal combinatorics asks to maximize or minimize the size of a set given certain constraints. Two of the most classical results in the area are Mantel's theorem (1907) and Turán's theorem (1941), a generalization of the first. Turán's theorem states that any K_k-free n-vertex graph has at most $\left(1-\frac{1}{k-1}\right)\binom{n}{2}$ edges. Examples of extremal questions are 'Given a set, what is the maximum size of a subset without some particular substructure?','What is the largest number of edge-disjoint copies of a graph H in a n-vertex graph with m edges?' and 'What is the minimum density of a graph H in a graph with edge density p?'.

Two well known problems of the last type are the minimal density of triangles in graphs and Sidorenko's Conjecture. It is easy to see that there are graphs with edge density $p \leq 0.5$ with no triangles just by taking a balanced bipartite graph with edge density $2p$ between its parts. For $p = 1$, we have a complete graph and we have that the density of triangles is 1. The minimal density of triangles problem was open for around forty years and several researchers, among them Bollobás [9], Fisher [30], Goodman [40], Lovász and Simonovits [66, 67], made improvements towards the answer but only recently it was solved by Razborov [78] using the flag algebra method. More details on the flag algebra method and an application of it are giving in Chapter 4.

Likewise, Sidorenko's Conjecture is one of the most famous problems in extremal combinatorics. The conjecture asserts that the density of every bipartite graph is minimized by a quasirandom graph with the same edge density. In Chapter 3, we give an introduction to the topic and study a stronger version of the conjecture.

Going back to extremal combinatorics, the asymptotic version of extremal problems are also interesting on their own. Quite often they are the first step towards understanding the specific problem but they can also show the general behavior by not taking into account lower order terms.

One of the motivations behind the development of the theory of combinatorial limits was to create a theory that deals with asymptotic behavior of combinatorial structures and potentially, using tools from different branches of mathematics, be able to say something/solve extremal problems. The theory made new connections between analysis, combinatorics, computer science, ergodic theory, group theory and probability theory.

Graph limits can be mostly divided into two regimes: the sparse one and the dense one. The best understood branch is limits of dense graphs which is the one that we will be concern in this thesis. We say that a sequence of graphs converge if all its subgraph densities converge. The limit object corresponding to a convergent sequence of dense graphs is called a graphon which is defined as a symmetric measurable function from the unit square to the unit interval. For a comprehensive introduction to the theory of graph limits, we refer the reader to the monograph of Lovász [65].

When the limits of convergent sequences of dense graphs are uniquely determined by finitely many density constraints, we call them finitely forcible graph limits. Such objects are closely related to problems in extremal combinatorics and they correspond to unique extremal configurations of problems from extremal graph theory. Indeed, extremal graph theory questions can be cast as optimization problems over the graph limit space with optimal solutions being the extremal points. We delve further into this theory in Chapter 2.

Closely related to graph limits is the flag algebra method, mentioned earlier. The method developed by Razborov provides a uniform framework for standard counting techniques used in extremal combinatorics and its application resulted in substantial progress on many long standing open problems in the area, e.g. [23, 29, 35, 38, 39, 48, 51, 60, 62, 64]. We give a brief introduction to the flag algebra method in Chapter 4, where we also show an application of the method in extremal graph theory.

Next, we give a short description of the main topics covered in the thesis. Further details are given in the introduction of its respective chapter.

Universality of finitely forcible graphons. We devise a unified framework to construct finitely forcible graphons with complex properties, such as non-compactness or large regularity partitions by showing that every graphon is a subgraphon of some finitely forcible graphon. The paper is available on arXiv and it was accepted for publication on Advances in Mathematics [21].

Triangle and edge decomposition. We prove a conjecture of Györi and Tuza which states that the edges of every n-vertex graph G can be decomposed into edges and triangles graphs $C_1, \dots, C_k$ such that $|C_1| + \dots + |C_k| \le (1/2 + o(1))n^2$. The paper is available on arxiv and it was accepted for publication at Combinatorics, Probability and Computing [57].

Step Sidorenko property of graphs. We study the step Sidorenko property of a graph H. We show that many bipartite graphs fail to have the step Sidorenko property and use our results to show the existence of a bipartite edge-transitive graph that is not weakly norming; this answers a question of Hatami [Israel J. Math. 175 (2010), 125–150]. The paper is available on arxiv and it was accepted for publication on the Journal of Combinatorial Theory, Series A [61].

This thesis is based on joint work with my coauthors: J. W. Cooper, D. Kráľ, B. Lidický, P. P. Pach, Y. Pehova and M. Wrochna.

1.1 Preliminaries

In this section, we introduce part of the notation used throughout this thesis. Notation used only in a particular chapter can be found in the corresponding chapter.

The set of integers from 1 to k will be denoted by $[k]$, the set of all positive integers by $\mathbb{N}$ and the set of all non-negative integers by $\mathbb{N}_0$. All measures considered in this paper are the Borel measures on $\mathbb{R}^d$, $d \in \mathbb{N}$. If a set $X \subseteq \mathbb{R}^d$ is measurable, then $|X|$ denotes its measure, and if X and Y are two measurable sets, then we write $X \sqsubseteq Y$ if $|X \setminus Y| = 0$.

In general, we follow standard graph theory notation. All graphs considered are simple and without loops. We sometimes consider graphs with vertices and edges assigned non-negative weights; when this is the case, we refer to such a graph as a *weighted graph*. The order of a graph G, i.e., its number of vertices, will be denoted by $|G|$ and the size of a graph G, i.e., its number of edges, by $\|G\|$.

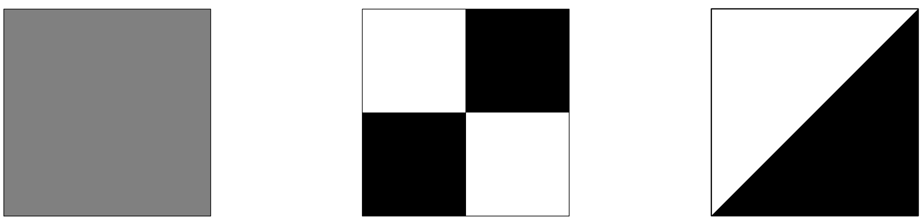
1.1.1 Graph limits

The theory of graph limits offers analytic tools to study large graphs. We present here only those notions that we need further, and refer the reader to the monograph of Lovász [65] on the subject for a comprehensive introduction to the theory. Graph limits also generated new tools and perspectives on many problems in mathematics and computer science. For example, the flag algebra method of Razborov [77], which bears close connections to convergent sequences of dense graphs, catalyzed progress on many important problems in extremal combinatorics, e.g. [1–3, 41, 49, 58, 60, 75–79]. In relation to computer science, the theory of graph limits shed new light on property and parameter testing algorithms [70].

Given two graphs H and G, the *density* of H in G is the probability that a uniformly chosen $|H|$-tuple of vertices of G induces a subgraph isomorphic to H; the density of H in G is denoted by $d(H, G)$. We adopt the convention that if $|H| > |G|$, then $d(H, G) = 0$.

A sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is *convergent* if the sequence $d(H, G_n)$ converges for every graph H. We will require that the orders of graphs in a convergent sequence tend to infinity. A convergent sequence of graphs can be associated with an analytic limit object, which is called a graphon. A *graphon* is a symmetric measurable function W from the unit square $[0, 1]^2$ to the unit interval $[0, 1]$, where *symmetric* refers to the property that $W(x, y) = W(y, x)$ for all $x, y \in [0, 1]$. One can think (although very imprecisely) of a graphon as a continuous version of the adjacency matrix of a graph and view the values of $W(x, y)$ as the density between different parts of a large graph represented by W.

To aid the transparency of our ideas, we often include a visual representation of graphons that we consider: the domain of a graphon W is represented as a unit square $[0, 1]^2$ with the origin $(0, 0)$ in the top left corner, and the values of W are represented by an appropriate shade of gray (ranging from white to black), with 0 represented by white and 1 by black. As an example, the following graphons are depicted in Figure 1.1: W_1 which value is $1/2$ almost everywhere; W_2 with value 1 for almost every pair (x, y) where $x \leq 0.5$, $y > 0.5$ and $x > 0.5$, $y \leq 0.5$, and 0, otherwise; and W_3 with 1 for almost every (x, y) where $1 - x \leq y$ and 0 otherwise.



(a) W_1 (b) W_2 (c) W_3

Figure 1.1: Examples of graphons.

Given a graphon W, a *W-random graph* of order n is a graph obtained from

W by sampling n vertices $v_1, v_2, \ldots, v_n \in [0,1]$ independently and uniformly at random and joining vertices v_i and v_j by an edge with probability $W(v_i, v_j)$ for all $i, j \in [n]$. The density of a graph H in a graphon W, denoted by $d(H, W)$, is the probability that a W-random graph of order $|H|$ is isomorphic to H. Note that the expected density of H in a W-random graph of order $n \geq |H|$ is equal to $d(H, W)$. We say that a convergent sequence $(G_n)_{n\in\mathbb{N}}$ converges to a graphon W if

$$\lim_{n\to\infty} d(H, G_n) = d(H, W)$$

for every graph H. It is not hard to show that if W is a graphon, then the sequence of W-random graphs with increasing orders is convergent with probability one and the graphon W is its limit.

Two graphons W_1 and W_2 are *weakly isomorphic* if $d(H, W_1) = d(H, W_2)$ for every graph H. Borgs, Chayes and Lovász [10] have shown that two graphons W_1 and W_2 are weakly isomorphic if and only if there exist measure preserving maps $\varphi_1, \varphi_2 : [0,1] \to [0,1]$ such that $W_1(\varphi_1(x), \varphi_1(y)) = W_2(\varphi_2(x), \varphi_2(y))$ for almost every $(x, y) \in [0,1]^2$. Graphons that can be uniquely determined up to a weak isomorphism by fixing the densities of a finite set of graphs are called *finitely forcible graphons* and are the central object of Chapter 2. Observe that a graphon W is finitely forcible if and only if there exist graphs $H_1, \ldots, H_k$ such that if a graphon W' satisfies $d(H_i, W') = d(H_i, W)$ for $i \in [k]$, then $d(H, W') = d(H, W)$ for every graph H.

A less obvious characterization of finitely forcible graphons that follows from flag algebra is the following.

Proposition 1. *A graphon W is finitely forcible if and only if there exist graphs $H_1, \ldots, H_k$ and reals $\alpha_1, \ldots, \alpha_k$ such that*

$$\sum_{i=1}^{k} \alpha_i d(H_i, W) \leq \sum_{i=1}^{k} \alpha_i d(H_i, W')$$

for every graphon W' and the equality holds only if W and W' are weakly isomorphic.

Chapter 2

Finitely forcible graph limits are universal

Central to dense graph convergence is the analytic representation of the limit of a convergent sequence of dense graphs, known as a *graphon* [11–13,69]; In this chapter, we are interested in graphons that are uniquely determined (up to isomorphism) by finitely many graph densities, which are called *finitely forcible* graphons. As mentioned in Chapter 1, such graphons are related to various problems from extremal graph theory and from graph theory in general. For example, for every finitely forcible graphon W, there exists a linear combination of graph densities such that the graphon W is its unique minimizer. Another example is the characterization of quasirandom graphs in terms of graph densities by Thomason [86, 87] which is essentially equivalent to stating that the constant graphon is finitely forcible by densities of 4-vertex graphs; also see [15, 80] for further results on quasirandom graphs. Lovász and Sós [68] generalized this characterization by showing that every step graphon, a multipartite graphon with quasirandom edge densities between its parts, is finitely forcible. Other examples of finitely forcible graphons are given in [71].

Early examples of finitely forcible graphons indicated that all finitely forcible graphons might possess a simple structure, as formalized by Lovász and Szegedy, who conjectured the following [71, Conjectures 9 and 10].

Conjecture 1. *The space of typical vertices of every finitely forcible graphon is compact.*

Conjecture 2. *The space of typical vertices of every finitely forcible graphon has finite dimension.*

Both conjectures were disproved through counterexample constructions [36, 37]. A stronger counterexample to Conjecture 2 was found in [20]: Conjecture 2 would imply that the number of parts of a weak ε-regular partition of a finitely forcible graphon is bounded by a polynomial of ε^{-1} but the construction given in [20] almost matches the best possible exponential lower bound from [16].

The purpose of this chapter is to show that finitely forcible graphons can have arbitrarily complex structure. Our main result reads as follows.

Theorem 1. *For every graphon W_F, there exists a finitely forcible graphon W_0 such that W_F is a subgraphon of W_0, and the subgraphon is formed by a $1/14$ fraction of the vertices of W_0.*

Theorem 1 contrasts with [71, Theorem 7.12], which states that the set of finitely forcible graphons is meager in the space of all graphons. In addition, since every finitely forcible graphon is the unique minimizer of some linear combination of densities of subgraphs (see Proposition 1), Theorem 1 also shows that optimal solutions of problems seeking to minimize a linear combination of densities of subgraphs, which are among the simplest stated problems in extremal graph theory, may have unique optimal solutions with highly complex structure.

Theorem 1 also immediately implies that both conjectures presented above are false since we can embed graphons not having the desired properties in a finitely forcible graphon. By considering a graphon containing appropriately scaled copies of graphons corresponding to the lower bound construction of Conlon and Fox from [16], which were described in [20], we also obtain the following.

Corollary 1. *For every non-decreasing function $f : \mathbb{R} \to \mathbb{R}$ tending to infinity, there exist a finitely forcible graphon W and positive reals ε_i tending to 0 such that every weak ε_i-regular partition of W has at least $2^{\Omega\left(\frac{\varepsilon_i^{-2}}{f(\varepsilon_i^{-1})}\right)}$ parts.*

Since every fixed graphon has weak ε-regular partitions with $2^{o(\varepsilon^{-2})}$ parts, Corollary 1 gives the best possible dependance on ε^{-1}.

The proof of Theorem 1 builds on the methods introduced in [37], which were further developed and formalized in [36]. In particular, the proof uses the technique of decorated constraints, which we present in Subsection 2.1.1. The main idea of the proof is the following. The graphon W_F is determined up to a set of measure zero by its density in squares with coordinates being the inverse powers of two. The countable sequence of such densities can be encoded into a single real number between 0 and 1, which will be embedded as the density of a suitable part of the graphon W_0. We then set up the structure of W_0 in a way that this encoding restricts the densities inside another part of W_0 rendering W_F unique up to a set of measure zero. While this approach seems uncomplicated upon first glance, the proof hides a variety of additional ideas and technical details. The reward is a result enabling the embedding of any graphon in a finitely forcible graphon with no additional effort.

2.1 Preliminaries

In this section we introduce additional notation used throughout the chapter. We start by presenting graphon analogues of several graph theoretic notions. The *degree* of a vertex $x \in [0,1]$ is defined as

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

Note that the degree is well-defined for almost all vertices of W and if x is chosen to be a vertex of an n-vertex W-random graph, then its expected degree is $(n-1)\cdot \deg_W(x)$. When it is clear from the context which graphon we are referring to, we will omit the subscript, i.e., we just write $\deg(x)$ instead of $\deg_W(x)$. We define the *neighborhood* $N_W(x)$ of a vertex $x \in [0,1]$ in a graphon W as the set of vertices $y \in [0,1]$ such that $W(x,y) > 0$. In our considerations, we will often analyze a restriction of a graphon to the substructure induced by a pair of measurable subsets A and B of $[0,1]$. If W is a graphon and A is a non-null measurable subset of $[0,1]$, then the *relative degree* of a vertex $x \in [0,1]$ with respect to A is

$$\deg_W^A(x) = \frac{\int_A W(x,y)\ \mathrm{d}y}{|A|},$$

i.e., the measure of the neighbors of x in A normalized by the measure of A. Similarly, $N_W^A(x) = N_W(x) \cap A$ is the *relative neighborhood* of x with respect to A. Note that $\deg_W^A(x)\cdot|A| \le |N_W^A(x)|$ and the inequality can be strict. Again, we drop the subscripts when W is clear from the context.

2.1.1 Finite forcibility and decorated constraints

Decorated constraints have been introduced and developed in [36, 37] as a method of showing finite forcibility of graphons. This method uses the language of the flag algebra method of Razborov, which, as we have mentioned earlier, has had many substantial applications in extremal combinatorics. We now present the notion of decorated constraints, partially following the lines of [36] in our exposition.

A *density expression* is iteratively defined as follows: a real number or a graph are density expressions, and if D_1 and D_2 are density expressions, then so are $D_1 + D_2$ and $D_1 \cdot D_2$. The value of a density expression with respect to a graphon W is the value obtained by substituting for each graph its density in the graphon W. A *constraint* is an equality between two density expressions. A graphon W *satisfies* a constraint if the density expressions on the two sides of the constraints have the same value. If $\mathcal{C}$ is a finite set of constraints such that there exists a unique (up to weak isomorphism) graphon W that satisfies all constraints in $\mathcal{C}$, then the graphon W is finitely forcible [37]; in particular, W can be forced by specifying the densities of graphs appearing in the constraints in $\mathcal{C}$.

A graphon W is said to be *partitioned* if there exist $k \in \mathbb{N}$, positive reals $a_1, \ldots, a_k$ with $a_1 + \cdots + a_k = 1$, and distinct reals $d_1, \ldots, d_k \in [0,1]$, such that the set of vertices in W with degree d_i has measure a_i. The set of all vertices with degree d_i will be referred to as a *part*; the *size* of a part is its measure and its *degree* is the common degree of its vertices. The following lemma was proved in [36, 37].

Lemma 1. *Let $a_1, \ldots, a_k$ be positive real numbers summing to one and let $d_1, \ldots, d_k \in [0,1]$ be distinct reals. There exists a finite set of constraints $\mathcal{C}$ such that any graphon satisfying all constraints in $\mathcal{C}$ is a partitioned graphon with parts of sizes $a_1, \ldots, a_k$ and degrees $d_1, \ldots, d_k$, and every partitioned graphon with parts of sizes $a_1, \ldots, a_k$ and degrees $d_1, \ldots, d_k$ satisfies all constraints in $\mathcal{C}$.*

We next introduce a formally stronger version of constraints, called decorated constraints. Fix $a_1, \ldots, a_k$ and $d_1, \ldots, d_k$ as in Lemma 1. A *decorated* graph is a graph G with $m \le |G|$ distinguished vertices labeled from 1 to m, which are called *roots*, and with each vertex assigned one of the k parts, which is referred to as the decoration of a vertex. Note that the number m can be zero in the definition of a decorated graph, i.e., a decorated graph can have no roots. Two decorated graphs are *compatible* if the subgraphs induced by their roots are isomorphic through an isomorphism preserving the labels (the order of the roots) and the decorations (the assignment of parts). A *decorated constraint* is an equality between two density expressions that contain decorated graphs instead of ordinary graphs and all the decorated graphs appearing in the constraint are compatible.

Consider a partitioned graphon W with parts of sizes $a_1, \ldots, a_k$ and degrees $d_1, \ldots, d_k$, and a decorated constraint C. Let H_0 be the (decorated) graph induced by the roots of the decorated graphs in the constraint; let $v_1, \ldots, v_m$ be the roots of H_0. We say that the graphon W *satisfies* the constraint C if the following holds for almost every m-tuple $x_1, \ldots, x_m \in [0,1]$ such that x_i belongs to the part that v_i is decorated with, $W(x_i, x_j) > 0$ for every edge $v_i v_j$ and $W(x_i, x_j) < 1$ for every non-edge v_iv_j: if each decorated graph H in C is replaced with the probability that a W-random graph is the graph H conditioned on the event that the roots are chosen as the vertices $x_1, \dots, x_m$ and they induce the graph H_0, and that each non-root vertex is randomly chosen from the part of W that is decorated with, then the left and right hand sides of the constraint C have the same value.

We now give an example of evaluating a decorated constraint. Consider a partitioned graphon W, which is depicted in Figure 2.1, with parts A and B each of size 1/2; the graphon W is equal to 1/2 on A^2, to 1/3 on $A \times B$, and to 1 on B^2. Let H be the decorated graph with two adjacent roots both decorated with A and two adjacent non-root vertices v_1 and v_2 both decorated with B such that v_1 is adjacent to only one of the roots and v_2 is adjacent to both roots; the decorated graph H is also depicted in Figure 2.1. If H appears in a decorated constraint, then its value is independent of the choice of the roots in the part A and is always equal to 2/81, which is the probability as defined in the previous paragraph.

Figure 2.1: An example of evaluating a decorated constraint. The root vertices are depicted by squares and the non-root vertices by circles. The graphon is equal to 1/2 on A^2, to 1/3 on $A \times B$, and to 1 on B^2.

Note that the condition on the m-tuple $x_1, \dots, x_m$ is equivalent to that there is a positive probability that a W-random graph with the vertices $x_1, \dots, x_m$ is H_0. Also note that, unlike in the definition of the density of a graph in a graphon, we do not allow permuting any vertices. For example, if W is the graphon (with a single part) that is equal to $p \in [0, 1]$ almost everywhere, then the cherry $K_{1,2}$ with each vertex decorated with the single part of W would be replaced in a decorated constraint with $p^2(1-p)$.

The next lemma, proven in [37], asserts that every decorated constraint is equivalent to a non-decorated constraint.

Lemma 2. *Let $k \in \mathbb{N}$, let $a_1, \dots, a_k$ be positive real numbers summing to one, and let $d_1, \dots, d_k$ be distinct reals between zero and one. For every decorated constraint C, there exists a constraint C' such that any partitioned graphon W with parts of sizes $a_1, \dots, a_k$ and degrees $d_1, \dots, d_k$ satisfies C if and only if it satisfies C'.*

In particular, if a graphon W is a unique partitioned graphon up to weak isomorphism that satisfies a finite collection of decorated constraints, then it is a unique graphon satisfying a finite collection of ordinary constraints by Lemmas 1 and 2, and hence W is finitely forcible.

We will visualize decorated constraints using the convention from [20], which we now describe and have already used in Figure 2.1. The root vertices of decorated graphs in a decorated constraint will be depicted by squares and the non-root vertices by circles; each vertex will be labeled with its decoration, i.e., the part that it should be contained in. The roots will be in all the decorated graphs in the constraint in the same mutual position, so it is easy to see the correspondence of the roots of different decorated graphs in the same constraint. A solid line between two vertices represents an edge, and a dashed line represents a non-edge. The absence of a line between two root vertices indicates that the decorated constraint should hold for both the root graph containing this edge and not containing it. Finally, the absence of a line between a non-root vertex and another vertex represents the sum of decorated graphs with this edge present and without this edge. If there are k such lines absent, the figure represents the sum of 2^k possible decorated graphs with these edges present or absent.

We finish this subsection with two auxiliary lemmas. The first is a lemma stated in [20], which essentially states that if a graphon W_0 is finitely forcible in its own right, then it may be forced on a part of a partitioned graphon without altering the structure of the rest of the graphon.

Lemma 3. *Let $k \in \mathbb{N}$, $m \in [k]$, let $a_1, \dots, a_k$ be positive real numbers summing to one, and let $d_1, \dots, d_k$ be distinct reals between zero and one. If W_0 is a finitely forcible graphon, then there exists a finite set $\mathcal{C}$ of decorated constraints such that any partitioned graphon W with parts of sizes $a_1, \dots, a_k$ and degrees $d_1, \dots, d_k$ satisfies C if and only if there exist measure preserving maps $\varphi_0 : [0,1] \to [0,1]$ and $\varphi_m : [0, a_m] \to A_m$ such that $W(\varphi_m(xa_m), \varphi_m(ya_m)) = W_0(\varphi_0(x), \varphi_0(y))$ for almost every $(x, y) \in [0,1]^2$, where A_m is the m-th part of W.*

Note that Lemma 2 implies that the set $\mathcal{C}$ of decorated constraints from Lemma 3 can be turned into a set of ordinary (i.e., non-decorated) constraints.

The second lemma is implicit in [71, proof of Lemma 3.3]; its special case has been stated explicitly in, e.g., [20, Lemma 8].

Lemma 4. *Let $X, Z \subseteq \mathbb{R}$ be two measurable non-null sets, and let $F : X \times Z \to [0,1]$*

be a measurable function. If there exists $C \in \mathbb{R}$ *such that*

$$\int_Z F(x,z)F(x',z)\ dz = C$$

for almost every $(x, x') \in X^2$*, then*

$$\int_Z F(x,z)^2\ dz = C$$

for almost every $x \in X$*.*

2.1.2 Regularity partitions and step functions

A *step function* $W : [0,1]^2 \to [-1,1]$ is a measurable function such that there exists a partition of $[0,1]$ into measurable non-null sets $U_1, \ldots, U_k$ that W is constant on $U_i \times U_j$ for every $i, j \in [k]$. A non-negative symmetric step function is a *step graphon*. If W is a step function (in particular, W can be a step graphon) and A and B two measurable subsets of $[0,1]$, then the *density* $d_W(A,B)$ *between* A *and* B is defined to be

$$d_W(A,B) = \int\limits_{A\times B} W(x,y)\,\mathrm{d}x\,\mathrm{d}y\ .$$

We will omit W in the subscript if W is clear from the context. Note that it always holds that $|d(A,B)| \le |A| \cdot |B|$. A step function W' *refines* a step function W with parts $U_1, \ldots, U_k$, if each part of W' is a subset of one of the parts of W and the density $d_W(U_i, U_j)$ between U_i and U_j is equal to the weighted average of the densities between the pairs of those parts of W' that are subsets of U_i and U_j, respectively.

We next recall the notion of the cut norm. If W is a step function, then the *cut norm* of W, denoted by $||W||_\square$, is

$$\sup_{A,B\subseteq[0,1]} \left| \int_{A\times B} W(x,y)\,\mathrm{d}x\,\mathrm{d}y \right|\ ,$$

where the supremum is taken over all measurable subsets A and B of $[0,1]$. The supremum in the definition is always attained and the cut norm induces the same topology on the space of step functions as the L_1-norm; this can be verified following the lines of the analogous arguments for graphons in [65, Chapter 8]. It can be shown that if H is a k-vertex graph and W and W' are two graphons, then

$$\left|d(H,W) - d(H,W')\right| \le \binom{k}{2} ||W - W'||_{\square} \, .$$

We will say that two graphons W and W' are *ε-close* if $||W - W'||_{\square} \le \varepsilon$.

A partition of $[0,1]$ into measurable non-null sets $U_1, \ldots, U_k$ is said to be *ε-regular* if

$$\left| d(A,B) - \sum_{i,j \in [k]} \frac{d(U_i,U_j)}{|U_i||U_j|} |U_i \cap A||U_j \cap B| \right| \le \varepsilon$$

for every two measurable subsets A and B of $[0,1]$. In other words, the step graphon W' with parts $U_1, \ldots, U_k$ that is equal to $\frac{d(U_i,U_j)}{|U_i||U_j|}$ on $U_i \times U_j$ is ε-close to W in the cut norm metric. In particular, the step graphon W' determines the densities of k-vertex graphs in W up to an additive error of $\binom{k}{2}\varepsilon$.

The Weak Regularity Lemma of Frieze and Kannan [34] extends to graphons as follows (see [65, Section 9.2] for further details): for every $\varepsilon > 0$, there exists $K \le 2^{O\left(\varepsilon^{-2}\right)}$, which depends on ε only, such that every graphon has an ε-regular partition with at most K parts. This dependence of K on ε is best possible up to a constant factor in the exponent [16]. We will need a slightly stronger version of this statement, which we formulate as a proposition; its proof is an easy modification of a proof of the standard version of the statement, e.g., the one presented in [65, Section 9.2].

Proposition 2. *For every $\varepsilon > 0$ and $k \in \mathbb{N}$, there exists $K \in \mathbb{N}$ such that for every graphon W and every partition $U_1, \ldots, U_k$ of $[0,1]$ into disjoint measurable non-null sets, there exist an ε-regular partition $U'_1, \ldots, U'_{K'}$ of $[0,1]$ with $K' \le K$ such that every part U'_i, $i \in [K']$, is a subset of one of the parts $U_1, \ldots, U_k$.*

For a step function W, we define $d(\Gamma_4, W)$ to be the following integral:

$$d(\Gamma_4, W) = \int_{[0,1]^4} W(x,y)W(x',y)W(x,y')W(x',y') \, \mathrm{d}x \, \mathrm{d}x' \, \mathrm{d}y \, \mathrm{d}y' \, .$$

Note that if W is a graphon, then

$$d(\Gamma_4, W) = \frac{1}{3} d(C_4, W) + \frac{1}{3} d(K_4^-, W) + d(K_4, W) \, ,$$

where K_4^- is the graph obtained from K_4 by removing one of its edges. In particular, $d(\Gamma_4, W)$ can be understood as the density of non-induced C_4 in the graphon W, since it is equal to the expected density of non-induced copies of C_4 in a W-random graph. If W is a step function, then $d(\Gamma_4, W) \le 4||W||_\square$. However, the converse also holds: $d(\Gamma_4, W) \ge ||W||_\square^4$; we refer e.g. to [65, Section 8.2], where a proof for symmetric step functions W is given and this proof readily extends to the general case. Lemma 7, which we present further, aims at a generalization of this statement to step graphons. Before we can state this lemma, we need to prove two auxiliary lemmas, which we state for matrices rather than step functions for simplicity.

Lemma 5. *Let M be a $K \times K$ real matrix and let $i, j \in [K]$. Define N to be the following $K \times K$ matrix:*

$$N_{x,y} = \begin{cases} \frac{M_{i,y}+M_{j,y}}{2} & \text{if } x = i \text{ or } x = j, \text{ and} \\ M_{x,y} & \text{otherwise.} \end{cases}$$

It holds that $Tr\, MM^TMM^T \ge Tr\, NN^TNN^T$.

Proof. Set $M(x,y)$, $x, y \in [K]$, to be the following quantity:

$$M(x,y) = \sum_{z=1}^{K} M_{x,z}M_{y,z} \ ,$$

and define $N(x,y)$, $x, y \in [K]$, in the analogous way. Observe that

$$\text{Tr}\ MM^TMM^T - \text{Tr}\ NN^TNN^T = \sum_{x,y=1}^{K} M(x,y)^2 - N(x,y)^2 \ .$$

We now analyze the difference on the right hand side of the equality by grouping the terms on the right hand side into disjoint sets such that the sum of the terms in each set is non-negative.

The terms with $x, y \in [K] \setminus \{i, j\}$ form singleton sets; note that $M(x, y) = N(x, y)$ for each such term. Fix $x \in [K] \setminus \{i, j\}$ and consider the two terms corresponding to $y = i$ and $y = j$. It follows that

$$M(x,i)^2 + M(x,j)^2 - N(x,i)^2 - N(x,j)^2 =$$

$$M(x,i)^2 + M(x,j)^2 - 2\left(\sum_{z=1}^{K} M_{xz}\frac{M_{i,z}+M_{j,z}}{2}\right)^2 =$$

$$M(x,i)^2 + M(x,j)^2 - \frac{1}{2}\left(M(x,i) + M(x,j)\right)^2 =$$

$$\frac{1}{2}M(x,i)^2 + \frac{1}{2}M(x,j)^2 - M(x,i)M(x,j) = \frac{1}{2}\left(M(x,i) - M(x,j)\right)^2 \ .$$

Hence, the sum of any pair of such terms is non-negative. The analysis of the terms with $y \in [K] \setminus \{i, j\}$ and $x = i$ or $x = j$ is symmetric.

The remaining four terms that have not been analyzed are the terms corresponding to the following pairs (x, y): (i, i), (i, j), (j, i) and (j, j). In this case, we obtain the following:

$$M(i,i)^2 + 2M(i,j)^2 + M(j,j)^2 - N(i,i)^2 - 2N(i,j)^2 - N(j,j)^2 =$$

$$M(i,i)^2 + 2M(i,j)^2 + M(j,j)^2 - 4\left(\frac{M(i,i) + 2M(i,j) + M(j,j)}{4}\right)^2 =$$

$$\frac{1}{4}\left(M(i,i) - M(j,j)\right)^2 + \frac{1}{2}\left(M(i,i) - M(i,j)\right)^2 + \frac{1}{2}\left(M(j,j) - M(i,j)\right)^2 .$$

Hence, the sum of these four terms is also non-negative, and the lemma follows. □

The next lemma follows by repeatedly applying Lemma 5 to pairs of rows of the matrix M with indices from the same set A_i and to pairs of rows of the matrix M^T with indices from the same set B_i, and considering the limit matrix N.

Lemma 6. *Let M be a $K \times K$ real matrix. Further, let $X_1, \ldots, X_k$ be a partition of $[K]$ into k disjoint sets and let $Y_1, \ldots, Y_\ell$ be a partition of $[K]$ into ℓ disjoint sets. Define the $K \times K$ matrix N as follows. If $x \in X_i$, $y \in Y_j$, then*

$$N_{x,y} = \frac{1}{|X_i| \cdot |Y_j|} \sum_{x' \in X_i, y' \in Y_j} M_{x',y'} .$$

It holds that $\mathrm{Tr}\, MM^TMM^T = \mathrm{Tr}\, M^TMM^TM \geq \mathrm{Tr}\, NN^TNN^T = \mathrm{Tr}\, N^TNN^TN$.

The following auxiliary lemma can be viewed as an extension of [65, Lemma 8.12], which states that $d(\Gamma_4, W) \geq ||W||_\square^4$ for every graphon W, from the zero graphon to general step graphons (consider the statement for W_0 being the zero graphon). We remark that we have not tried to obtain the best possible dependence on the parameter ε in the statement of the lemma. The lemma also holds in a more general setting, where the parts of graphons are not required to be of the same size.

Lemma 7. *Let W_0 be a step graphon with all parts of the same size, and W a step graphon refining W_0 such that all parts of W have the same size. If $||W - W_0||_\square \geq \varepsilon$, then $d(\Gamma_4, W) \geq d(\Gamma_4, W_0) + \varepsilon^4/8$.*

Proof. Since $||W - W_0||_\square \geq \varepsilon$, there exist two measurable subsets A and B of $[0, 1]$ such that

$$\left| \int_{A \times B} W(x,y) - W_0(x,y) \,\mathrm{d}x\, \mathrm{d}y \right| \geq \varepsilon . \tag{2.1}$$

Let U be one of the parts of the graphon W. Depending whether $\int_{U\times B} W-W_0\,\mathrm{d}x\,\mathrm{d}y$ is positive or negative, replacing A with either $A\cup U$ or $A\setminus U$ does not decrease the integral in (2.1). Hence, we can assume that each part of W is either a subset of A or is disjoint from A, and the same holds with respect to B (but different parts U of W may be contained in A and B).

Let k be the number of parts of W_0 and K the number of parts W. Further, let M be the $K\times K$ matrix such that the entry $M_{i,j}$, $i,j\in K$, is the density of W between its i-th and the j-th parts, and let P be the $K\times K$ matrix such that $P_{i,j}$, $i,j\in K$, is the density of W_0 between the i-th and the j-th parts of W. Let U_i, $i\in[k]$, be the subset of $[K]$ containing the indices of the parts of W contained in the i-th part of W_0. Observe that both matrices M and P are symmetric and the matrix P is constant on each submatrix indexed by pairs from $U_i\times U_j$ for some $i,j\in[k]$. Since $d(\Gamma_4,W)=\mathrm{Tr}\ M^4$ and $d(\Gamma_4,W_0)=\mathrm{Tr}\ P^4$, our goal is to show that $\mathrm{Tr}\ M^4-\mathrm{Tr}\ P^4\geq\varepsilon^4/8$. Finally, let A' be the indices of parts of W contained in A, and let B' be the indices of parts of W contained in B. Observe that (2.1) yields that the sum of the entries of the matrix $M-P$ with the indices in $A'\times B'$ is either at least ε or at most $-\varepsilon$.

Let N be the matrix from the statement of Lemma 6 for the matrix M, $X_i=\{i\}$, $i\in[K]$, and $Y_j=U_j$, $j\in[k]$. Let ε_1 be the sum of the entries of the matrix $M-N$ with the indices in $A'\times B'$, and let ε_2 be the sum of the entries of the matrix $N-P$ with the indices in $A'\times B'$. Note that $|\varepsilon_1+\varepsilon_2|\geq\varepsilon$, which implies that $|\varepsilon_1|+|\varepsilon_2|$ is at least ε. By Lemma 6, it holds that $\mathrm{Tr}\ M^4-\mathrm{Tr}\ NN^TNN^T\geq 0$. Since P^T can be obtained from the matrix N^T by applying Lemma 6 with $X_i=\{i\}$, $i\in[K]$, and $Y_j=U_j$, $j\in[k]$, it follows that $\mathrm{Tr}\ N^TNN^TN-\mathrm{Tr}\ P^4=\mathrm{Tr}\ NN^TNN^T-\mathrm{Tr}\ P^4\geq 0$.

We now show that $\mathrm{Tr}\ M^4-\mathrm{Tr}\ NN^TNN^T\geq\varepsilon_1^4$. Let $Q=M-N$. We now want to analyze the entries of the matrix $(N+\alpha Q)(N+\alpha Q)^T$ for $\alpha\in[0,1]$. Fix $x,y\in[K]$ and observe that the entry in the x-th row and the y-th column of the matrix $(N+\alpha Q)(N+\alpha Q)^T$ is equal to

$$\sum_{j=1}^{k}\sum_{z\in U_j}(N+\alpha Q)_{x,z}(N+\alpha Q)_{y,z}\ .$$

The definition of the matrix N implies that

$$\sum_{z\in U_j}Q_{x,z}=\sum_{z\in U_j}Q_{y,z}=0$$

for every $j\in[k]$. It also holds that $N_{x,z}=N_{x,z'}$ and $N_{y,z}=N_{y,z'}$ for any z

and z' from the same set U_j, $j \in [k]$, which implies that the entry of the matrix $(N + \alpha Q)(N + \alpha Q)^T$ in the x-th row and the y-th column is

$$\sum_{z=1}^{K} N_{x,z}N_{y,z} + \alpha^2 Q_{x,z}Q_{y,z} \; .$$

Hence, we conclude that $(N + \alpha Q)(N + \alpha Q)^T = NN^T + \alpha^2 QQ^T$. It follows that

$$\begin{aligned} &\mathrm{Tr}\ (N + \alpha Q)(N + \alpha Q)^T(N + \alpha Q)(N + \alpha Q)^T = \\ &\mathrm{Tr}\ NN^TNN^T + 2\alpha^2 \mathrm{Tr}\ NN^TQQ^T + \alpha^4 \mathrm{Tr}\ QQ^TQQ^T \; . \end{aligned} \tag{2.2}$$

By Lemma 6 applied with $M = N + \alpha Q$ and the same sets X_i and Y_j as earlier,

$$\mathrm{Tr}\ (N + \alpha Q)(N + \alpha Q)^T(N + \alpha Q)(N + \alpha Q)^T - \mathrm{Tr}\ NN^TNN^T \geq 0$$

for every $\alpha \geq 0$, which implies that Tr $NN^TQQ^T \geq 0$. In particular, we obtain from (2.2) for $\alpha = 1$ that

$$\begin{aligned} &\mathrm{Tr}\ M^4 - \mathrm{Tr}\ NN^TNN^T = \\ &\mathrm{Tr}\ (N + \alpha Q)(N + \alpha Q)^T(N + \alpha Q)(N + \alpha Q)^T - \mathrm{Tr}\ NN^TNN^T \geq \\ &\mathrm{Tr}\ QQ^TQQ^T \; . \end{aligned} \tag{2.3}$$

Since the cut-norm of the step graphon corresponding to Q is at least ε_1, it follows that Tr $QQ^TQQ^T \geq \varepsilon_1^4$.

Applying the symmetric argument to the matrices P^T and $N^T = N$, we obtain that

$$\mathrm{Tr}\ NN^TNN^T - \mathrm{Tr}\ P^4 \geq \mathrm{Tr}\ (N - P)(N - P)^T(N - P)(N - P)^T \geq \varepsilon_2^4 \; . \tag{2.4}$$

Since Tr M^4 − Tr $NN^TNN^T \geq 0$ and Tr N^TNN^TN − Tr $P^4 \geq 0$, we obtain from (2.3) and (2.4) using $|\varepsilon_1| + |\varepsilon_2| \geq \varepsilon$ that Tr M^4 − Tr $P^4 \geq \varepsilon_1^4 + \varepsilon_2^4 \geq \varepsilon^4/8$, as desired. □

2.2 General setting of the proof of Theorem 1

In this section, we provide a general overview of the structure of the graphon W_0 from Theorem 1 and the proof of Theorem 1. The visualization of the structure of the graphon W_0 can be found in Figure 2.2. The proof of Theorem 1 is spread through Sections 2.2–2.5, with this section containing its initial steps.

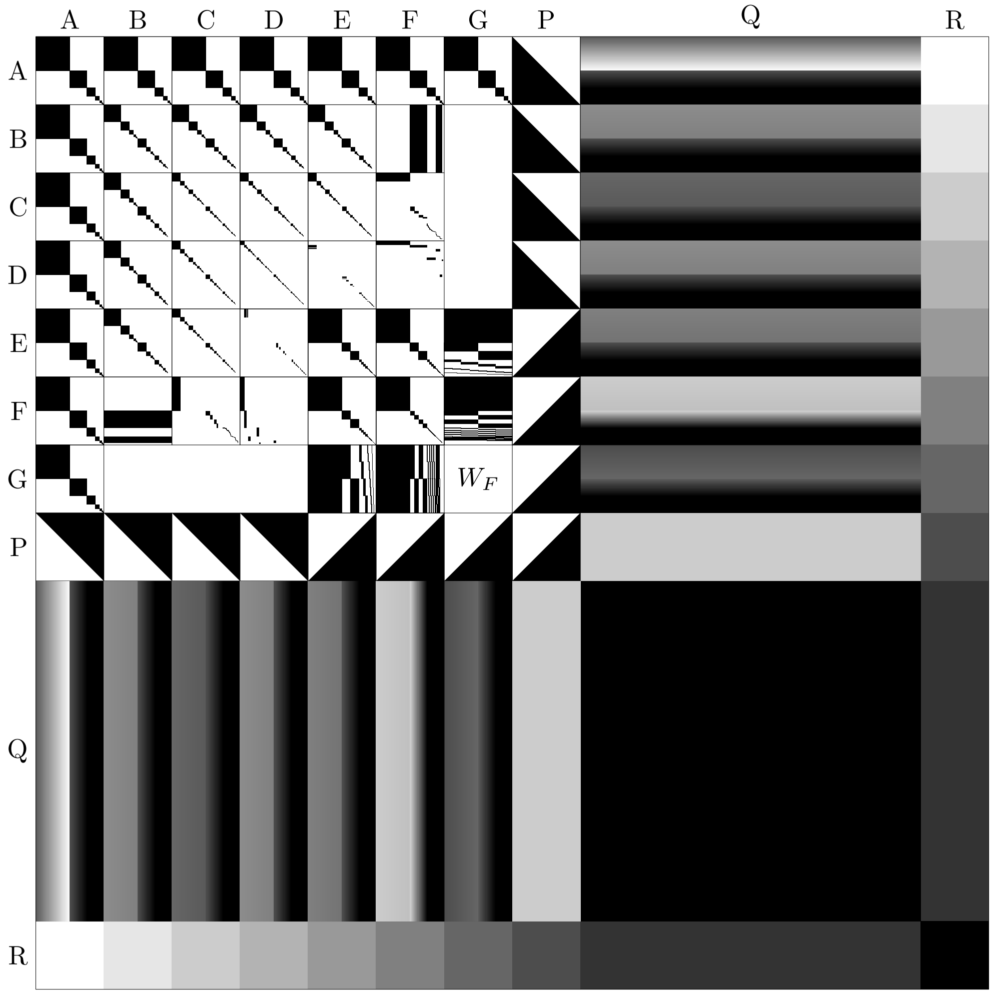


Figure 2.2: The sketch of the graphon W_0 from Theorem 1.

Fix a graphon W_F. The graphon W_0 is a partitioned graphon with 10 parts denoted by capital letters from A to R. Each part except for Q has size $1/14$, and the size of Q is $5/14$. If $X, Y \in \{A, \dots, G, P, Q, R\}$ are two parts, the restriction of the graphon W_0 to $X \times Y$ will be referred to as the *tile* $X \times Y$. The graphon W_F will be contained in the tile $G \times G$ of the graphon W_0. The degrees of the parts (i.e., the degrees of the vertices forming the parts) are given in Table 2.1; the degree of Q is at least $5/14 + 8/252$, i.e., larger than the degree of any other part, and will be fixed later in the proof.

Part	A	B	C	D	E	F	G	P	Q	R
Degree	$\frac{90}{252}$	$\frac{91}{252}$	$\frac{92}{252}$	$\frac{93}{252}$	$\frac{94}{252}$	$\frac{95}{252}$	$\frac{96}{252}$	$\frac{97}{252}$	$> \frac{98}{252}$	$\frac{77}{252}$

Table 2.1: The degrees of the vertices in the parts of the graphon W_0 from the proof of Theorem 1.

Rather than giving a complex definition of the graphon W_0 at once, we decided to present the particular details of the structure of W_0 together with the decorated constraints fixing the structure of W_0 in Sections 2.2–2.5. Table 2.2 gives references to subsections where the individual tiles of the graphon W_0 are considered and the corresponding decorated constraints are given.

We now start the proof of the finite forcibility of the graphon W_0. Let W be a graphon that satisfies the constraints from Lemma 1 with respect to the sizes and degrees of the parts of W_0 and that satisfies all the decorated constraints given in Sections 2.2–2.5. It will be obvious that the graphon W_0 also satisfies these constraints. So, if we show that W is weakly isomorphic to W_0, then we will have established that W_0 is finitely forcible. We will achieve this goal by constructing a measure preserving map $g : [0, 1] \to [0, 1]$ such that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in [0, 1]^2$.

Let $A, \dots, G, P, Q, R$ be the parts of the graphon W. To make a clear distinction between the parts of W and W_0, we will use $A_0, \dots, G_0, P_0, Q_0, R_0 \subseteq [0, 1]$ to denote the subintervals forming the parts of W_0. The Monotone Reordering Theorem [65, Proposition A.19] implies that, for every $X \in \{A, \dots, G, P, Q, R\}$, there exist a measure preserving map $\varphi_X : X \to [0, |X|)$ and a non-decreasing function $\tilde{f}_X : [0, |X|) \to \mathbb{R}$ such that

$$\tilde{f}_X(\varphi_X(x)) = \deg_W^P(x) = \frac{1}{|P|} \int_P W(x, y) \, \mathrm{d}y$$

for almost every $x \in X$. The function g maps the vertex $x \in X$, $X \in \{A, \dots, G, P, Q, R\}$,

	A	B	C	D	E	F	G	P	Q	R
A	2.2.2	2.2.2	2.2.2	2.2.2	2.2.2	2.2.2	2.2.2	2.2.1	2.5.1	2.5.2
B	2.2.2	2.2.3	2.2.3	2.2.3	2.2.3	2.3.1	2.5	2.2.1	2.5.1	2.5.2
C	2.2.3	2.2.3	2.2.3	2.2.3	2.2.3	2.4.3	2.5	2.2.1	2.5.1	2.5.2
D	2.2.2	2.2.3	2.2.3	2.2.3	2.3.3	2.3.2	2.5	2.2.1	2.5.1	2.5.2
E	2.2.2	2.2.3	2.2.3	2.3.3	2.4.1	2.4.1	2.4.2	2.2.1	2.5.1	2.5.2
F	2.2.2	2.3.1	2.4.3	2.3.2	2.4.1	2.4.1	2.4.2	2.2.1	2.5.1	2.5.2
G	2.2.2	2.5	2.5	2.5	2.4.2	2.4.2	2.4.4	2.2.1	2.5.1	2.5.2
P	2.2.1	2.2.1	2.2.1	2.2.1	2.2.1	2.2.1	2.2.1	2.2.1	2.5.1	2.5.2
Q	2.5.1	2.5.1	2.5.1	2.5.1	2.5.1	2.5.1	2.5.1	2.5.1	2.5	2.5.2
R	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2	2.5.2

Table 2.2: Subsections where the structure of the tiles are presented and the related decorated constraints then given.

of W to the vertex $\eta_X(\varphi_X(x)/|X|)$ where η_X is the bijective linear map from $[0,1)$ to the part X_0 of the graphon W_0 of the form $\eta_X(x) = |X_0|\cdot x + c_X$ for some $c_X \in [0,1]$ (we intentionally define η_X in this way, instead of defining η_X as a linear measure preserving map from $[0,|X_0|)$ to X_0, since this definition simplifies our exposition later). In addition, we define a function $f_X : X \to [0,1]$ as $f_X(x) = \tilde{f}_X(\varphi_X(x))$ for every $x \in X$. Clearly, g is a measure preserving map from $[0,1]$ to $[0,1]$; hence, we "only" need to show that $W(x,y) = W_0(g(x),g(y))$ for almost every $(x,y) \in [0,1]^2$.

2.2.1 Coordinate system

In this subsection, we analyze the tile $P\times P$ and the tiles $P\times X$ (and the symmetric tiles $X\times P$) where $X \in \{A,\dots,G\}$. The *half-graphon* $W_\triangle$ is the graphon such that $W_\triangle(x,y)$ is equal to 1 if $x+y \geq 1$ and equal to 0 otherwise; the half-graphon is finitely forcible as shown by Diaconis and Janson [74] and Lovász and Szegedy [71]. Consider the decorated constraints from Lemma 3 forcing the tile $P\times P$ to be weakly isomorphic to the half-graphon. This implies that $\tilde{f}_P(x) = \varphi_P(x)/|P|$ for every $x \in [0,|P|)$, where φ_P and $\tilde{f}_P$ are the functions from the Monotone Reordering Theorem used to define the function g. Lemma 3 and the finite forcibility of the half-graphon yield that $W(x,y) = W_0(g(x),g(y))$ for almost every $(x,y) \in P^2$.

Next consider the decorated constraints depicted in Figure 2.3 and fix $X \in \{A,\dots,G\}$. The first constraint in Figure 2.3 implies that $W(x,y) \in \{0,1\}$ for almost every $(x,y) \in P\times X$ and that $N^X_W(x) \sqsubseteq N^X_W(x')$ or $N^X_W(x') \sqsubseteq N^X_W(x)$ for almost every pair $(x,x') \in P^2$. It follows that there exists a function $h_X : P \to [0,1]$ such that it holds for almost every $(x,y) \in P\times X$ that $W(x,y) = 1$ if and

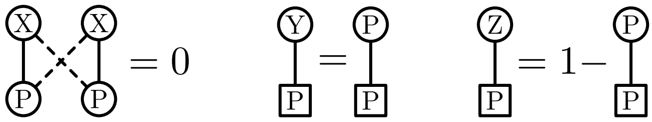


Figure 2.3: Decorated constraints forcing the tiles $X \times P$ where $X \in \{A, \dots, G\}$, $Y \in \{E, F, G\}$ and $Z \in \{A, B, C, D\}$.

only if $\varphi_X(y)/|X| \geq 1 - h_X(x)$. If $X \in \{E, F, G\}$, then the second constraint in Figure 2.3 implies that $\deg^X(x) = |N^X(x)| = \deg^P(x)$ for almost every $x \in P$, i.e., $h_X(x) = f_P(x)$. Since it holds that $W(x, y) = 1$ if and only if $\varphi_X(y)/|X| \geq 1 - h_X(x)$ for almost every $(x, y) \in P \times X$, we obtain that $\tilde{f}_X(y) = \varphi_X(y)/|X|$ for $y \in [0, |X|)$, $W(x, y) = 1$ for almost every $(x, y) \in P \times X$ with $f_P(x) + f_X(y) \geq 1$ and $W(x, y) = 0$ for almost every $(x, y) \in P \times X$ with $f_P(x) + f_X(y) < 1$. It follows that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in P \times X$, where $X \in \{E, F, G\}$. The analogous argument using the third constraint in Figure 2.3 implies that $\deg^X(x) = |N^X(x)| = |X| - \deg^P(x)$ for almost every $x \in P$, which yields that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in P \times X$, where $X \in \{A, B, C, D\}$.

We conclude this subsection by observing that $\deg^P_W(x) = f_X(x)$ for almost every $x \in X$, where $X \in \{A, \dots, G\} \cup \{P\}$. In particular, we may interpret the relative degree of a vertex with respect to P as its coordinate. Also observe that $N^P_W(x) \sqsubseteq N^P_W(x')$ for almost every pair $(x, x') \in X \times X$ such that $f_X(x) \leq f_X(x')$.

2.2.2 Checker tiles

We now consider the tiles $A \times X$ where $X \in \{A, \dots G\}$. The argument follows the lines of the analogous argument presented in [20, 36, 37], however, we include the details for completeness. The *checker graphon* W_C is obtained as follows: let $I_k = [1 - 2^{-k}, 1 - 2^{-(k+1)})$ for $k \in \mathbb{N}_0$ and set $W_C(x, y)$ equal to 1 if $(x, y) \in \bigcup_{k=0}^{\infty} I_k^2$, i.e., both x and y belong to the same I_k, and equal to 0 otherwise. The checker graphon W_C is depicted in Figure 2.4. We remark that we present an iterated version of this definition in Subsection 2.2.3. We set $W_0(\eta_A(x), \eta_X(y)) = W_C(x, y)$ for $x, y \in [0, 1)^2$ where $X \in \{A, \dots G\}$.

Consider the decorated constraints in Figure 2.5, which we claim to force the structure of the tile $A \times A$. The first constraint in Figure 2.5 implies that there exists a collection $\mathcal{J}_A$ of disjoint measurable non-null subsets of A such that the following holds for almost every $(x, y) \in A \times A$: $W(x, y) = 1$ if and only if x and y belong to the same set $J \in \mathcal{J}_A$, and $W(x, y) = 0$ otherwise.

The second constraint in Figure 2.5 implies that almost every triple $(x, x', x'') \in$

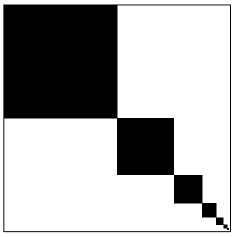


Figure 2.4: The checker graphon W_C.

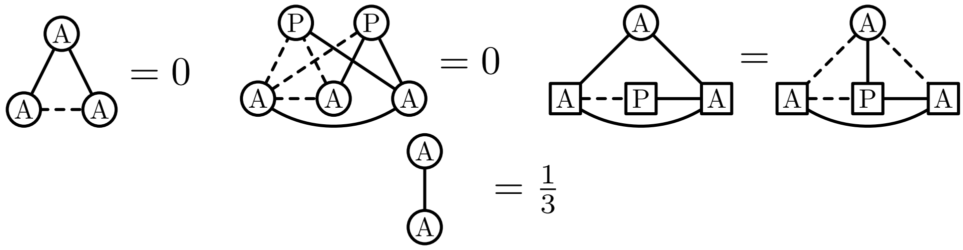


Figure 2.5: The decorated constraints forcing the structure of the tile $A \times A$.

A^3 satisfies that if x and x'' belong to the same set $J \in \mathcal{J}_A$ and $f_A(x) < f_A(x') < f_A(x'')$, then x' also belongs to the set J (since x and x' cannot be non-adjacent). This implies that for every $J \in \mathcal{J}_A$, there exists an open interval J' such that J and $f_A^{-1}(J')$ differ on a null set. Let $\mathcal{J}'_A$ be the collection of these open intervals for different sets $J \in \mathcal{J}_A$; since f_A is a measure preserving map and the sets in $\mathcal{J}_A$ are disjoint, the intervals in $\mathcal{J}'_A$ must be disjoint.

The third constraint in Figure 2.5 implies that almost every pair $(x, x') \in A^2$ satisfies that if x and x' belong to the same set $J \in \mathcal{J}_A$ and $f_A(x) < f_A(x')$, then $|N_W^A(x) \cap N_W^A(x')| = |J|$ is the measure of the set Y of the points $y \in A$ such that $y \notin J$ and $f_A(y) > f_P(x'')$ for almost every $x'' \in P$ with $f_A(x) < f_P(x'') < f_A(x')$. Observe that if J is fixed and $J = f_A^{-1}(J')$ for $J' \in \mathcal{J}'_A$, then the set Y differs from $f_A^{-1}([\sup J', 1))$ on a null set. It follows that the measure $|J| = |J'|$ is equal to $1 - \sup J'$. Hence, each interval in $\mathcal{J}'_A$ is of the form $(1-2\gamma, 1-\gamma)$ for some $\gamma \in (0, 1/2]$; let Γ be the set of all the values of γ for that there is a corresponding interval in $\mathcal{J}'_A$. Note that if $\gamma \in \Gamma$, then $\Gamma \cap (\gamma/2, \gamma) = \emptyset$, which implies in particular that the set Γ is countable. Let γ_k be the k-th largest value in the set Γ and in case that Γ is finite, set $\gamma_k = 0$ for $k > |\Gamma|$. It follows that

$$\frac{1}{|A|^2} \int\limits_{A \times A} W(x, y) \, \mathrm{d}x \, \mathrm{d}y = \sum_{J' \in \mathcal{J}'_A} \left(\sup J' - \inf J'\right)^2 = \sum_{k \in \mathbb{N}} \gamma_k^2 \, .$$

The last constraint in Figure 2.5 implies that the integral on the left hand side of the above equality is equal to $1/3$, which is possible only if $\gamma_k = 2^{-k}$ for every $k \in \mathbb{N}$. It follows that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in A^2$.

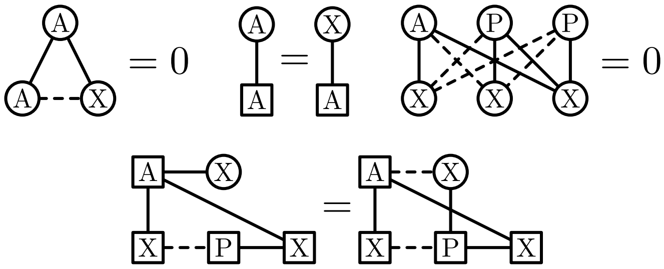


Figure 2.6: The decorated constraints forcing the structure of the tiles $A \times X$ where $X \in \{B, \ldots, G\}$.

We now consider the decorated constraints from Figure 2.6. Fix $X \in \{B, \ldots, G\}$. The first constraint in Figure 2.6 implies that for every $J \in \mathcal{J}_A$, there exists a measurable set $Z(J) \subseteq X$ such that the following holds for almost every pair $(x, y) \in A \times X$: $W(x, y) = 1$ if $x \in J$ and $y \in Z(J)$, and $W(x, y) = 0$ otherwise. Note that the sets $Z(J)$ need not be disjoint. The second constraint in Figure 2.6 yields that $\deg_W^A(x) = \deg_W^X(x)$ for almost every $x \in A$, which implies that the sets J and $Z(J)$ have the same measure. The third constraint implies that the following holds for almost every triple $(y, y', y'') \in X^3$: if $f_P(y) < f_P(y') < f_P(y'')$, $y \in Z(J)$ and $y'' \in Z(J)$, then $y' \in Z(J)$. Consequently, for every $Z(J)$, there exists an open interval $Z'(J)$ such that $Z(J)$ differs from the set $g_X^{-1}(Z'(J))$ on a null set. Finally, the last constraint in Figure 2.6 yields that the following holds for almost every $x \in J$: the measure of $N_W^X(x) = Z(J)$, which is $|Z(J)| = |Z'(J)|$, is equal to the measure of the set containing all $y \notin Z(J)$ with $f_X(y) \geq \sup Z'(J)$. It follows that the interval $Z'(J)$ is equal to $(1 - 2\gamma, \gamma)$ for some $\gamma \in (0, |X|/2]$. Since the measures of J and $Z'(J)$ are the same, it must hold that $Z'(J) = J'$ where $J' \in \mathcal{J}'_A$ is the interval corresponding to J. It follows that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in A \times X$.

2.2.3 Iterated checker tiles

The checker graphon W_C represents a large graph formed by disjoint complete graphs on the $1/2, 1/4, 1/8, \ldots$ fractions of its vertices. We now present a family of iterated checker graphons. Informally speaking, we start with the checker graphon W_C and at each iteration, we paste a scaled copy of W_C on each clique of the current graphon. The formal definition is as follows. Fix $k \in \mathbb{N}_0$. If $k = 0$,

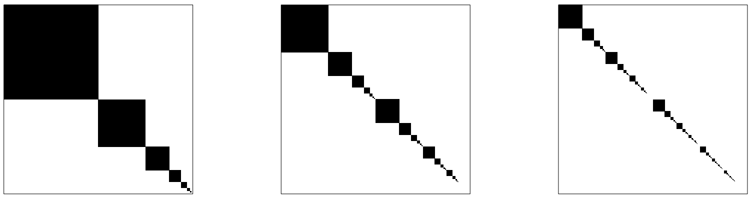


Figure 2.7: The iterated checker graphons W_C^0, W_C^1 and W_C^2.

define I_{j_0}, $j_0 \in \mathbb{N}_0$, to be the interval

$$I_{j_0} = \left[1 - 2^{-j_0}, 1 - 2^{-j_0-1}\right).$$

If $k > 0$, we define $I_{j_0,\dots,j_k}$ for $(j_0, \dots, j_k) \in \mathbb{N}_0^k$ as

$$I_{j_0,\dots,j_k} = \left[\sup I_{j_0,\dots,j_{k-1}} - 2^{-j_k}|I_{j_0,\dots,j_{k-1}}|, \sup I_{j_0,\dots,j_{k-1}} - 2^{-j_k-1}|I_{j_0,\dots,j_{k-1}}|\right).$$

The *k-iterated checker graphon* W_C^k is then defined as follows: $W_C^k(x, y)$ is equal to 1 if there exists a $(k+1)$-tuple $(j_0, \dots, j_k) \in \mathbb{N}_0^k$ such that both x and y belong to the interval $I_{j_0,\dots,j_k}$, and it is equal to 0 otherwise. The iterated checker graphons W_C^0, W_C^1 and W_C^2 are depicted in Figure 2.7. Note that $W_C^0 = W_C$ and the definition of I_{j_0} coincides with that given in Subsection 2.2.2. We will also refer to an interval $I_{j_0,\dots,j_k}$ as to a *k-iterated binary interval*.

For $X \in \{B, C\}$ and $Y \in \{X, \dots, E\}$, we set

$$W_0(\eta_X(x), \eta_Y(y)) = \begin{cases} W_C^1(x, y) & \text{if } X = B, \text{ and} \\ W_C^2(x, y) & \text{if } X = C \end{cases}$$

for all $x, y \in [0, 1)^2$. We also set the tile $D \times D$ to be such that

$$W_0(\eta_D(x), \eta_D(y)) = W_C^3(x, y)$$

for all $x, y \in [0, 1)^2$. This also defines the values of W_0 in the symmetric tiles, i.e., the values for the tile $X \times Y$ determine the values for the tile $Y \times X$.

Consider the decorated constraints depicted in Figures 2.8 and 2.9. We first analyze the structure of the tile $B \times B$, then all the tiles $B \times Y$, $Y \in \{B, \dots, E\}$, then the tile $C \times C$, then all the tiles $C \times Y$, $Y \in \{C, \dots, E\}$, before finishing with the tile $D \times D$. Fix (X, Y) to be one of the pairs $(A, B), (B, C)$ or (C, D). We assume that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in X \times X$ and almost every $(x, y) \in X \times Y$, and our goal is to show that $W(x, y) = W_0(g(x), g(y))$ for

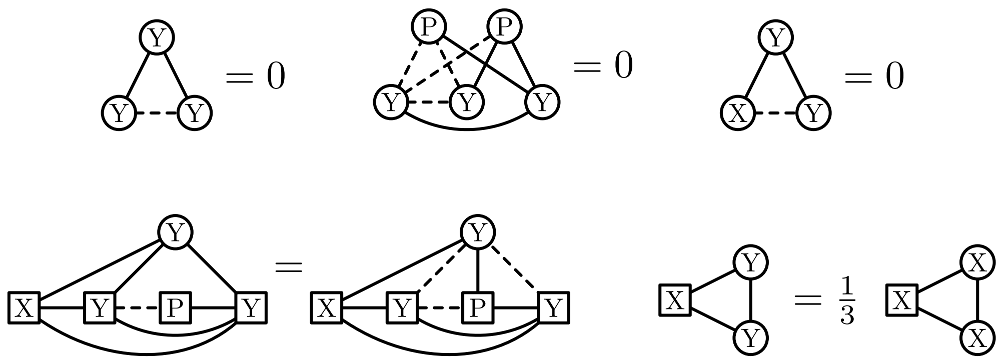


Figure 2.8: The decorated constraints forcing the structure of the tiles B^2, C^2, and D^2, where $(X, Y) \in \{(A, B), (B, C), (C, D)\}$.

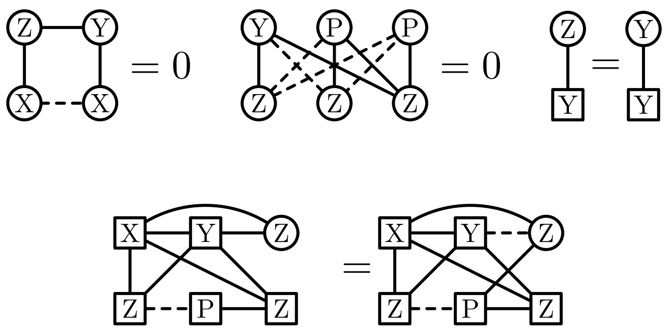


Figure 2.9: The decorated constraints forcing the structure of the iterated checker graphons on the non-diagonal tiles, where $(X, Y) \in \{(A, B), (B, C)\}$ and $Z \in \{C, D, E, F\}$ if $X = A$ and $Z \in \{D, E, F\}$ if $X = B$.

almost every $(x,y) \in Y \times Y$.

The first two constraints on the first line in Figure 2.8 imply that there exists a collection $\mathcal{J}'_Y$ of disjoint open intervals such that the following holds for almost every $(x,y) \in Y^2$: $W(x,y)$ is equal to 1 if and only if $f_Y(x)$ and $f_Y(y)$ belong to the same interval $J' \in \mathcal{J}'_Y$, and it is equal to 0 otherwise. The third constraint on the first line in Figure 2.8 yields that each interval in $\mathcal{J}'_Y$ is a subinterval of an interval in $\mathcal{J}'_X$.

The first constraint on the second line in Figure 2.8 yields that the following holds for almost every triple $(x,y,y') \in X \times Y \times Y$ such that $f_Y(y)$ and $f_Y(y')$ are from the same interval $J'_Y \in \mathcal{J}'_Y$ and $f_X(x)$ is from the interval $J'_X \in \mathcal{J}'_X$ that is a superinterval of J'_Y: the measure of J'_Y (which is equal to the left hand side of the equality) is the same as the measure of the set of all y'' such that $f_Y(y'') \in J'_X$ and $f_Y(y'') > \sup J'_Y$ (which is equal to the right hand side). It follows that

$$J'_Y = (\sup J'_X - 2\gamma, \sup J'_X - \gamma)$$

for some $\gamma \in (0, |J'_X|/2]$. The very last constraint in Figure 2.8 yields for every $J'_X \in \mathcal{J}'_X$ that

$$\sum_{J'_Y \in \mathcal{J}'_Y, J'_Y \subseteq J'_X} |J'_Y|^2 = \frac{1}{3}|J'_X|^2.$$

However, this is only possible if the set $\mathcal{J}'_Y$ contains all intervals of the form $(\sup J'_X - 2\gamma, \sup J'_X - \gamma)$ for every $J'_X \in \mathcal{J}'_X$ and every $\gamma = |J'_X| \cdot 2^{-i}$, $i \in \mathbb{N}$. It follows that $W(x,y) = W_0(g(x), g(y))$ for almost every $(x,y) \in Y \times Y$.

We continue to fix a pair $(X,Y) \in \{(A,B),(B,C)\}$, but in addition we now fix $Z \in \{Y,\ldots,E\} \setminus \{Y\}$ where $Y \in \{B,C\}$. Our next goal is to show that $W(y,z) = W_0(g(y), g(z))$ for almost every $(y,z) \in Y \times Z$, which is achieved using the decorated constrains given in Figure 2.9. The first constraint in Figure 2.9 implies that it holds for almost every $y \in Y$ that $f_Z(N^Z_W(y)) \sqsubseteq J'_X$ where J'_X is the unique interval of $\mathcal{J}'_X$ containing $f_Y(y)$. The second constraint in Figure 2.9 yields that for almost every $y \in Y$, there exists an interval J_y such that $N^Z_W(y)$ and $f_Z^{-1}(J_y)$ differ on a null set, $W(y,z) = 1$ for almost every $z \in f_Z^{-1}(J_y)$, and $W(y,z) = 0$ for almost every $z \in Z \setminus f_Z^{-1}(J_y)$. The third constraint yields that $\deg^Y_W(y) = \deg^Z_W(y)$ for almost every $y \in Y$, i.e., the measure of J_y is the same as the measure of the interval in $\mathcal{J}_Y$ containing $f_Y(y)$.

Finally, the last constraint in Figure 2.9 implies that almost every quadruple $x \in X$, $y \in Y$, $z, z' \in Z$ such that $f_Z(z) < f_Z(z')$, $f_Z(z)$ and $f_Z(z')$ belong to the interval J_y, which is a subinterval of $J'_X \in \mathcal{J}'_X$ with $f_X(x) \in J'_X$, satisfies that the measure of $N_W^Z(y)$ (note that $N_W^Z(y)$ is a subset of $f_Z^{-1}(J'_X)$) and the measure of all $z' \in f_Z^{-1}(J'_X) \setminus N_W^Z(y)$ with $f_Z(z') > \sup J_y$ are equal. In particular, the interval J_y is of the form $(\sup J'_X - 2\gamma, \sup J'_X - \gamma)$ for almost every $y \in Y$, where J'_X is the unique interval of $\mathcal{J}'_X$ containing $f_Y(y)$. Hence, the interval J_y is equal to the interval in $\mathcal{J}'_Y$ containing $f_Y(y)$ for almost every $y \in Y$. It follows that $W(y,z) = W_0(g(y), g(z))$ for almost every $(y,z) \in Y \times Z$.

2.3 Encoding the target graphon

In this section, we describe how the densities in dyadic squares of the graphon W_F are wired in a single binary sequence, which will be encoded in the tile $B \times F$. To achieve this, we need to fix a mapping φ from $\mathbb{N}_0^4$ to $\mathbb{N}_0$. Let us define this mapping as follows. The 4-tuples (a,b,c,d) with the same sum $s = a+b+c+d$ of their entries are injectively mapped to the numbers between $\binom{s+3}{4}$ and $\binom{s+4}{4} - 1$ in the lexicographic order. For example, $\varphi(0,0,0,1) = 1$ and $\varphi(0,1,0,0) = 3$.

2.3.1 Encoding dyadic square densities

The tile $B \times F$ encodes the edge densities on all dyadic squares of W_F. Let $I^d(s)$ be the interval $\left[\frac{s}{2^d}, \frac{s+1}{2^d}\right)$, and define for $d, s, t \in \mathbb{N}_0$ the value $\delta(d,s,t)$ as

$$\delta(d,s,t) = 2^{2d} \cdot \int_{I^d(s) \times I^d(t)} W_F(x,y) \,\mathrm{d}x \,\mathrm{d}y$$

if $0 \le s, t \le 2^d - 1$, and $\delta(d,s,t) = 0$, otherwise. If W_F is the one graphon, i.e., W_F is equal to 1 almost everywhere, we fix $r = 1$. Otherwise, we fix $r \in [0,1)$ to be the unique real satisfying that

$$\delta(d,s,t) = \sum_{p=0}^{\infty} 2^{-p} r_{\varphi(d,s,t,p)+1} \text{ , and} \tag{2.5}$$

that for all $d, s, t \in \mathbb{N}_0$, the value of $r_{\varphi(d,s,t,p)+1}$ is equal to zero for infinitely many $p \in \mathbb{N}_0$, where r_k is the k-th bit in the standard binary representation of r (with the first bit following immediately the decimal point). The standard binary representation is the unique representation with infinitely many digits equal to zero If W_F is the one graphon, we set $r_k = 1$ for every $k \in \mathbb{N}$. Observe that r is not a multiple of an inverse power of two unless W_F is the zero graphon or the one graphon ($r \in \{0,1\}$ in these two cases).

We now define $W_0(\eta_B(x), \eta_F(y)) = r_{k+1}$ for $x \in [0,1]$ and $y \in I_k$, $k \in \mathbb{N}_0$,

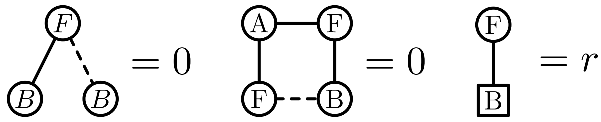


Figure 2.10: The decorated constraints forcing the structure of the tile $B \times F$.

and force the corresponding structure of the tile $B \times F$. Consider the decorated constraints depicted in Figure 2.10. The first constraint implies that $\deg_W^B(x) \in \{0,1\}$ for almost every $x \in F$. In particular, W is $\{0,1\}$-valued almost everywhere on $B \times F$. The second constraint implies that for every $k \in \mathbb{N}_0$ and for almost every $x, x' \in f_F^{-1}(I_k)$, $\deg_W^B(x) = \deg_W^B(x')$. Let r'_k be the common degree $\deg_W^B(x)$ of the vertices $x \in f_F^{-1}(I_{k-1})$, $k \in \mathbb{N}$. The last constraint in the figure implies that

$$\sum_{k\in\mathbb{N}} 2^{-k} r_k = \sum_{k\in\mathbb{N}} 2^{-k} r'_k \ .$$

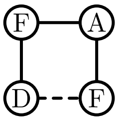
Since r is not a non-zero multiple of an inverse power of two unless $r \in \{0,1\}$, it follows that $r_k = r'_k$ for all $k \in \mathbb{N}$. If $r \in \{0,1\}$, it follows that $r_k = r'_k = r$ trivially. We conclude that $W(x,y) = W_0(g(x), g(y))$ for almost every $(x,y) \in B \times F$.

2.3.2 Matching tile

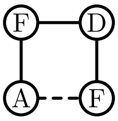
In this subsection, we introduce and analyze the tile $D \times F$. This tile is supposed to link the 4-fold indexing to linear indexing. Formally, we define $W_0(\eta_D(x), \eta_F(y))$ to be equal to 1 if $x \in I_{a,b,c,d}$ and $y \in I_{\varphi(a,b,c,d)}$ for some $(a,b,c,d) \in \mathbb{N}_0^4$ and to be equal to 0, otherwise.

Consider the decorated constraints in Figure 2.11. The first constraint implies that W is $\{0,1\}$-valued almost everywhere in $D \times F$ and that for almost every $x \in D$, it holds that $N_W^F(x) = \cup_{k \in K_x} f_F^{-1}(I_k)$ up to a null set for some $K_x \subseteq \mathbb{N}_0$. The second constraint implies that for almost every vertex of D, the set K_x has cardinality 0 or 1. The third constraint yields that for every $(a,b,c,d) \in \mathbb{N}_0^4$, the set K_x is the same for almost all $x \in D$ with $f_D(x) \in I_{a,b,c,d}$. Finally, the last constraint in the first line implies that the sets K_x and K_y are disjoint for almost all $x, y \in D$ with $f_D(x)$ and $f_D(y)$ from different 3-iterated binary intervals.

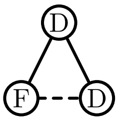
Let $\tau(a,b,c,d)$ be the common degree $\deg_W^F(x)$ of vertices $x \in f_D^{-1}(I_{a,b,c,d})$. If K_x is empty for almost all $x \in f_D^{-1}(I_{a,b,c,d})$, then $\tau(a,b,c,d) = 0$; otherwise, $\tau(a,b,c,d)$ is 2^{-k-1}, where k is the unique integer contained in K_x for almost all $x \in f_D^{-1}(I_{a,b,c,d})$. Note that the non-zero values of $\tau(a,b,c,d)$ are distinct for distinct



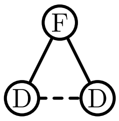
$= 0$



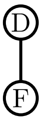
$= 0$



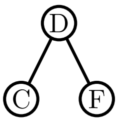
$= 0$



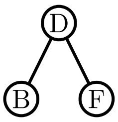
$= 0$



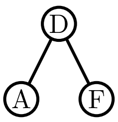
$$= \sum_{a,b,c,d=0}^{\infty} 2^{-a-b-c-d-4} \cdot 2^{-\varphi(a,b,c,d)-1}$$



$$= \sum_{a,b,c,d=0}^{\infty} 2^{-a-b-c-d-4} \cdot 2^{-a-b-c-3} \cdot 2^{-\varphi(a,b,c,d)-1}$$



$$= \sum_{a,b,c,d=0}^{\infty} 2^{-a-b-c-d-4} \cdot 2^{-a-b-2} \cdot 2^{-\varphi(a,b,c,d)-1}$$



$$= \sum_{a,b,c,d=0}^{\infty} 2^{-a-b-c-d-4} \cdot 2^{-a-1} \cdot 2^{-\varphi(a,b,c,d)-1}$$

Figure 2.11: The decorated constraints forcing the structure of the tile $D \times F$.

$(a, b, c, d) \in \mathbb{N}_0^4$. Observe that the edge density in the tile $D \times F$ is the following:

$$\int_{D\times F} W(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{(a,b,c,d)\in\mathbb{N}_0^4} |I_{a,b,c,d}|\tau(a,b,c,d) = \sum_{s\in\mathbb{N}_0} 2^{-(s+4)} \sum_{\substack{(a,b,c,d)\in\mathbb{N}_0^4\\ a+b+c+d=s}} \tau(a,b,c,d).$$

The constraint in the second line in Figure 2.11 yields the following:

$$\sum_{s\in\mathbb{N}_0} 2^{-(s+4)} \sum_{\substack{(a,b,c,d)\in\mathbb{N}_0^4\\ a+b+c+d=s}} \tau(a,b,c,d) = \sum_{s\in\mathbb{N}_0} 2^{-(s+4)} \sum_{\substack{(a,b,c,d)\in\mathbb{N}_0^4\\ a+b+c+d=s}} 2^{-\varphi(a,b,c,d)-1}\ .$$

Since the non-zero values of $\tau(a, b, c, d)$ are mutually distinct, this equality can hold only if

$$\{\tau(a,b,c,d) \text{ s.t. } a+b+c+d=s\} = \{2^{-\varphi(a,b,c,d)-1} \text{ s.t. } a+b+c+d=s\}$$

for every $s \in \mathbb{N}_0$.

The constraint in the third line in Figure 2.11 implies that

$$\sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-b-c-3}\cdot\tau(a,b,c,d) = \sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-b-c-3}\cdot 2^{-\varphi(a,b,c,d)-1}\ .$$

Since it holds for every $s \in \mathbb{N}_0$ that

$$\{\tau(a,b,c,d) \text{ s.t. } a+b+c+d=s\} = \{2^{-\varphi(a,b,c,d)-1} \text{ s.t. } a+b+c+d=s\}\ ,$$

we get that the following holds for all $d \in \mathbb{N}_0$ and $s \in \mathbb{N}_0$:

$$\{\tau(a,b,c,d) \text{ s.t. } a+b+c=s\} = \{2^{-\varphi(a,b,c,d)-1} \text{ s.t. } a+b+c=s\}\ .$$

Similarly, the constraint in the fourth line implies that

$$\sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-b-2}\cdot\tau(a,b,c,d) = \sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-b-2}\cdot 2^{-\varphi(a,b,c,d)-1}\ ,$$

which yields that it holds for all $c, d, s \in \mathbb{N}_0$ that

$$\{\tau(a,b,c,d) \text{ s.t. } a+b=s\} = \{2^{-\varphi(a,b,c,d)-1} \text{ s.t. } a+b=s\}\ .$$

Finally, the constraint in the fifth line implies that

$$\sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-1}\cdot\tau(a,b,c,d) = \sum_{(a,b,c,d)\in\mathbb{N}_0^4} 2^{-a-b-c-d-4}\cdot 2^{-a-1}\cdot 2^{-\varphi(a,b,c,d)-1} ,$$

which implies that $\tau(a,b,c,d) = 2^{-\varphi(a,b,c,d)-1}$ for all $a,b,c,d \in \mathbb{N}_0$. It follows that $W(x,y) = W_0(g(x), g(y))$ for almost every $(x,y) \in D \times F$.

2.3.3 Collating dyadic square densities

The tile $D\times E$ is designed to group the values of $\delta(d,s,t)$. We set $W_0(\eta_D(x), \eta_E(y)) = r_{\varphi(d,s,t,p)+1}$ for all $x \in I_{d,s,t,p}$, $y \in I_{d,s,t}$ and $(d,s,t,p) \in \mathbb{N}_0^4$, and we set $W_0(\eta_D(x), \eta_E(y))$ to be zero elsewhere. An example of a tile with this structure is depicted in Figure 2.12. Note that the density of the square $\eta_D(I_{d,s,t}) \times \eta_E(I_{d,s,t})$ is equal to $\delta(d,s,t)$.

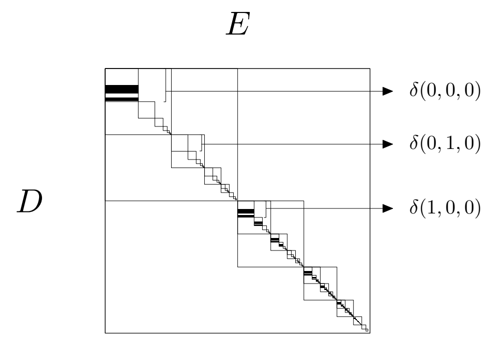


Figure 2.12: An example of the $D \times E$ tile.

Consider the decorated constraints depicted in the Figure 2.13. The first constraint implies that $W(x,y) = 0$ for almost every (x,y) such that $x \in f_D^{-1}(I_{d,s,t})$, $y \in f_E^{-1}(I_{d',s',t'})$ and $(d,s,t) \neq (d',s',t')$. The second constraint yields that for almost every $x \in D$ such that $x \in f_D^{-1}(I_{d,s,t})$, $\deg_W^E(x)$ is either 0 or $2^{-d-s-t-3}$. In particular, $W(x,y) \in \{0,1\}$ for almost every $(x,y) \in D \times E$.

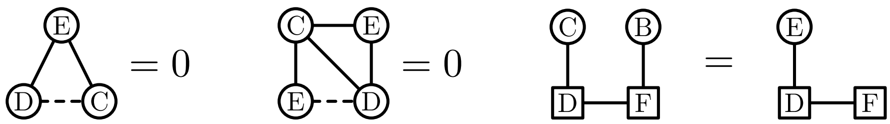


Figure 2.13: The decorated constraints forcing the structure of the tile $D \times E$.

We now analyze the last decorated constraint depicted in the Figure 2.13. This constraint implies that the following holds for almost every choice of a D-root

x and an F-root y such that $f_D(x) \in I_{d,s,t,p}$ and $f_F(y) \in I_{\varphi(d,s,t,p)}$:

$$2^{-d-s-t-3} \cdot r_{\varphi(d,s,t,p)+1} = \deg_W^E(x) \ .$$

It follows that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in D \times E$.

2.4 Forcing the target graphon

In this section, we force the densities in each dyadic square of the tile $G \times G$ to be as in the graphon W_F and we argue that the graphon inside the tile is the graphon W_F. To achieve this, we first need to set up some auxiliary structures.

2.4.1 Dyadic square indices

We start with the tiles $E \times E$, $E \times F$ and $F \times F$, which represent splitting the 0-iterated binary interval I_k into 2^k and 2^{2k} equal length parts. Formally, $W_0(\eta_E(x), \eta_E(y))$ is equal to 1 for $x, y \in [0, 1)$ if x and y belong to the same 0-iterated binary interval I_k and

$$\left\lfloor \frac{x - \min I_k}{|I_k|} \cdot 2^k \right\rfloor = \left\lfloor \frac{y - \min I_k}{|I_k|} \cdot 2^k \right\rfloor \ ,$$

and it is equal to 0 otherwise. Similarly, $W_0(\eta_F(x), \eta_F(y))$ is equal to 1 for $x, y \in [0, 1)$ if x and y belong to the same 0-iterated binary interval I_k and

$$\left\lfloor \frac{x - \min I_k}{|I_k|} \cdot 2^{2k} \right\rfloor = \left\lfloor \frac{y - \min I_k}{|I_k|} \cdot 2^{2k} \right\rfloor \ ,$$

and it is equal to 0 otherwise. An illustration can be found in Figure 2.14. Finally, we set $W_0(\eta_E(x), \eta_F(y)) = W_0(\eta_E(x), \eta_E(y))$ for all $x, y \in [0, 1)$.

Figure 2.14: Representation of the tiles $E \times E$ and $F \times F$.

Fix $X \in \{E, F\}$ and consider the decorated constraints given in Figure 2.15. The three constraints on the first line in Figure 2.15 imply that $W(x, y) \in \{0, 1\}$ for almost every pair $(x, y) \in X \times X$ and that there exists a collection of disjoint open intervals $\mathcal{J}_X$, which are subintervals of 0-iterated binary intervals I_k, such that

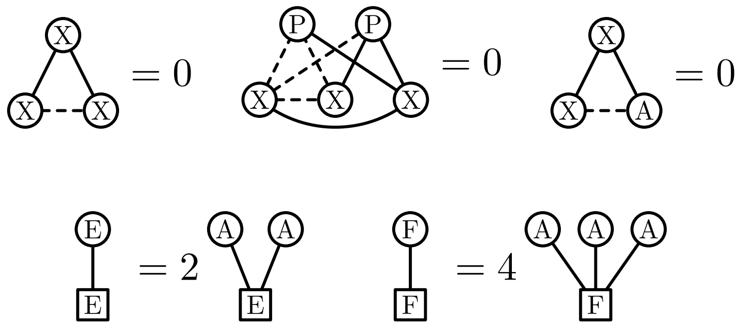


Figure 2.15: The decorated constraints forcing the tiles $E \times E$ and $F \times F$, where $X \in \{E, F\}$.

$W(x, y) = 1$ if and only if $f_X(x)$ and $f_X(y)$ belong to the same interval $J \in \mathcal{J}_X$ (except for a subset of $X \times X$ of measure zero).

If $X = E$, then the first constraint on the second line in Figure 2.15 implies that $\deg_W^E(x) = 2^{-2k-1}$ for almost every $x \in f_E^{-1}(J_k)$, i.e., if $J \in \mathcal{J}_E$ and $J \subseteq I_k$, then $|J| = 2^{-k}|I_k|$. Hence, the set $\mathcal{J}_E$ is formed precisely by the intervals

$$\left(\min I_k + \frac{\ell - 1}{2^k}|I_k|, \min I_k + \frac{\ell}{2^k}|I_k|\right)$$

for $k \in \mathbb{N}_0$ and $\ell \in [2^k]$. Hence, $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in E \times E$. The analogous argument using the last constraint on the second line in Figure 2.15 gives that $\mathcal{J}_F$ is formed precisely by the intervals

$$\left(\min I_k + \frac{\ell - 1}{2^{2k}}|I_k|, \min I_k + \frac{\ell}{2^{2k}}|I_k|\right)$$

for $k \in \mathbb{N}_0$ and $\ell \in [2^{2k}]$, which leads to the conclusion that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in F \times F$.

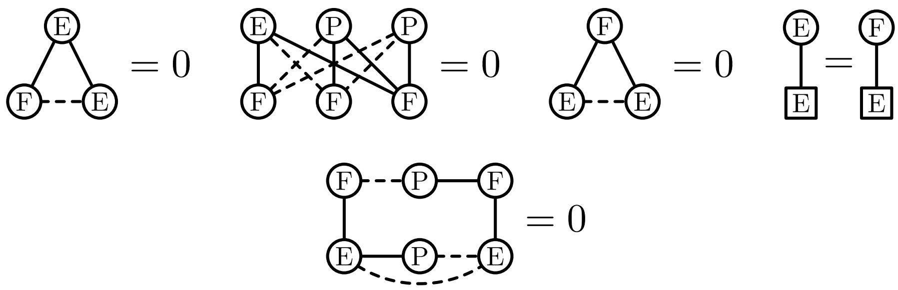


Figure 2.16: The decorated constraints forcing the structure of the tile $E \times F$.

It remains to analyze the tile $E \times F$. Consider the decorated constraints given in Figure 2.16. The first two constraints in Figure 2.16 imply that for every $J \in \mathcal{J}_E$, there exists an open interval $K(J)$ such that the following holds for almost every $(x, y) \in E \times F$: $W(x, y) = 1$ if $f_E(x) \in J$ and $f_F(y) \in K(J)$ for some $J \in \mathcal{J}_E$, and $W(x, y) = 0$ otherwise. The third constraint implies that the intervals $K(J)$ and $K(J')$ are disjoint for $J \neq J'$, and the fourth constraint yields that the measure of $K(J)$ is equal to $|J|$. Finally, the last constraint implies that if an interval $J_1 \in \mathcal{J}_E$ precedes an interval $J_2 \in \mathcal{J}_E$, then $K(J_1)$ precedes the interval $K(J_2)$. We conclude that $K(J) = J$ for every $J \in \mathcal{J}_E$. Consequently, $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in E \times F$.

2.4.2 Referencing dyadic squares

We now describe the tiles $E \times G$ and $F \times G$, which allow referencing particular dyadic squares by the intervals from $\mathcal{J}_E$ and $\mathcal{J}_F$. Formally, $W_0(\eta_E(x), \eta_G(y)) = 1$ for $x, y \in [0, 1)$ if and only if $x \in I_k$ and

$$\left\lfloor \frac{x - \min I_k}{|I_k|} \cdot 2^k \right\rfloor = \left\lfloor y \cdot 2^k \right\rfloor ,$$

and it is equal to 0 otherwise.

In an intuitive level, the above formula express the following function: for every $k \in \mathbb{N}_0$, split I_k on 2^k ordered subintervals of same length (coordinate x) and do the same with the unit interval (coordinate y). For each fixed value of k, the function has value 1, on the product of two intervals with the same position in the ordering given by k, and 0, otherwise.

Similarly, $W_0(\eta_F(x), \eta_G(y)) = 1$ for $x, y \in [0, 1)$ if and only if $x \in I_k$ and

$$\left\lfloor \frac{x - \min I_k}{|I_k|} \cdot 2^{2k} \right\rfloor \equiv \left\lfloor y \cdot 2^k \right\rfloor \pmod{2^k} ,$$

and it is equal to 0 otherwise.

Informally, the tile $F \times G$, is constructed in the following way: for each $k \in \mathbb{N}_0$, copy the function of $E \times G$ in $I_k \times [0, 1]$, scale it by 2^{-k} and make 2^k disjoint copies on $I_k \times [0, 1]$. The tiles are depicted in Figure 2.17.

Fix $X \in \{E, F\}$ and set $Y = A$ if $X = E$, and $Y = E$ if $X = F$. Consider the decorated constraints given in Figure 2.18. The first two constraints in Figure 2.18 imply that for every $J \in \mathcal{J}_X$, there exists an open interval $K^X(J)$ such that the following holds for almost every $(x, y) \in X \times G$: $W(x, y) = 1$ if $f_E(x) \in J$ and $f_F(y) \in K^X(J)$ for some $J \in \mathcal{J}_X$, and $W(x, y) = 0$ otherwise.

Figure 2.17: The tiles $E \times G$ and $F \times G$.

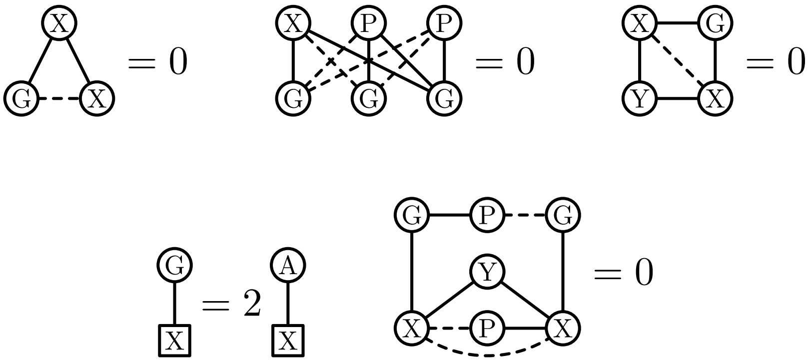


Figure 2.18: The decorated constraints forcing the structure of the tiles $E \times G$ and $F \times G$, where $(X, Y) \in \{(E, A), (F, E)\}$.

If $X = E$, the third constraint on the first line in Figure 2.18 implies that if $J, J' \in \mathcal{J}_E$ and J and J' are subintervals of the same 0-iterated binary interval, then $K^E(J)$ and $K^E(J')$ are disjoint; the second constraint on the second line implies that if J precedes J' inside the same 0-iterated binary interval, then $K^E(J)$ precedes $K^E(J')$. Likewise, if $X = F$, the third constraint on the first line gives that if $J, J' \in \mathcal{J}_F$ and J and J' are subintervals of the same interval contained in $\mathcal{J}_E$, then $K^F(J)$ and $K^F(J')$ are disjoint, and the second constraint on the second line gives that if J precedes J' inside the same interval of $\mathcal{J}_E$, then $K^F(J)$ precedes $K^F(J')$.

Finally, the first constraint on the second line implies that $\deg_W^G(x) = 2\deg_W^A(x)$ for almost every $x \in X$. This implies that if J is a subinterval of a 0-iterated binary interval J_k, then $|K^X(J)| = 2^{-k}$. We conclude that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in X \times G$.

2.4.3 Indexing dyadic squares

We now describe the tile $C \times F$, which allows referencing particular dyadic squares by 2-iterated binary intervals; the tile is depicted in Figure 2.19. Formally, $W_0(\eta_C(x), \eta_F(y)) =$

1 for $x, y \in [0,1)$ if and only if $x \in I_{d,s,t}$, $y \in I_d$, $s < 2^d$, $t < 2^d$, and

$$\left\lfloor \frac{y - \min I_d}{|I_d|} \cdot 2^{2d} \right\rfloor = 2^d \cdot s + t,$$

and it is equal to 0 otherwise. Informally, for each $d \in \mathbb{N}_0$, we pair, in an ordered fashion, each interval $I_{d,s,t}$, where $s < 2^d$, $t < 2^d$, with the ordered intervals produced from splitting I_d in 2^{2d} equal size subintervals. The function has value 1 in each of these pairs, and 0 otherwise.

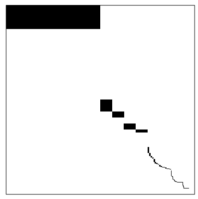


Figure 2.19: The tile $C \times F$.

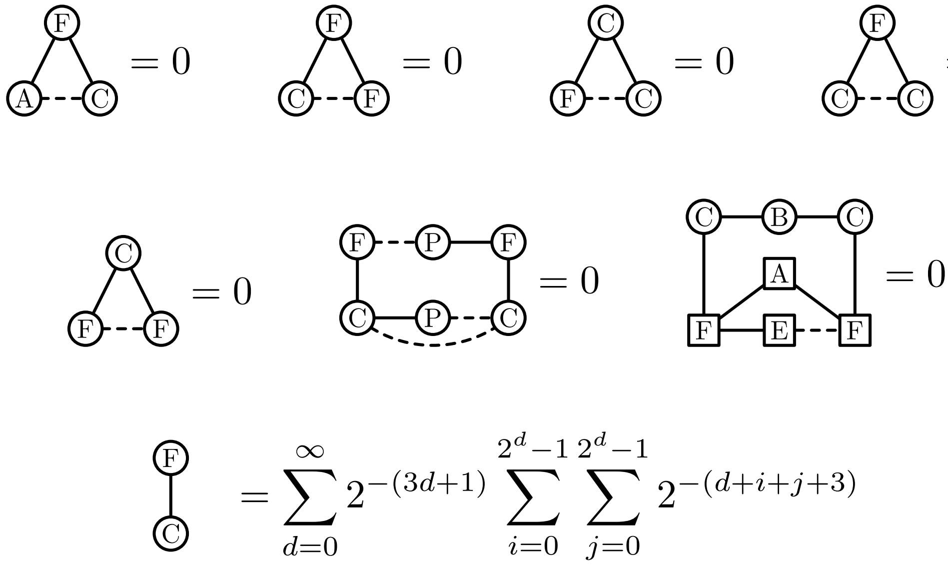


Figure 2.20: The decorated constraints forcing the structure of the tile $C \times F$.

Consider the constrains given in Figure 2.20. The four constraints in the first line in Figure 2.20 imply the following: there exists a function $h : \mathbb{N}_0^3 \to \mathbb{N}_0 \cup \{\infty\}$ such that $h(d,s,t) \in \{0,\dots,2^{2d}-1\} \cup \{\infty\}$ and the following holds for almost every

$(x, y) \in C \times F$: $W(x, y) = 1$ if and only if $f_C(x) \in I_{d,s,t}$ and

$$\left\lfloor \frac{f_F(y) - \min I_d}{|I_d|} \cdot 2^{2d} \right\rfloor = h(d, s, t) ,$$

$W(x, y) = 0$ otherwise. In particular, if $h(d, s, t) = \infty$, then $W(x, y) = 0$ for almost every $x \in f_C^{-1}(I_{d,s,t})$ and $y \in F$.

The first constraint in the second line in Figure 2.20 implies that if $(d, s, t) \neq (d, s', t')$, then $h(d, s, t) \neq h(d, s', t')$ unless $h(d, s, t) = h(d, s', t') = \infty$. The second constraint in the second line then implies that if $h(d, s, t)$ and $h(d, s', t')$ are both different from ∞ and $h(d, s, t) < h(d, s', t')$, then either $s = s'$ and $t < t'$ or $s < s'$. Finally, the last constraint in the second line yields that if $h(d, s, t)$ and $h(d, s', t')$ are both different from ∞ and $s \neq s'$, then $\lfloor h(d, s, t)/2^d \rfloor \neq \lfloor h(d, s', t')/2^d \rfloor$. Consequently, for any d, there are at most 2^d values of s such that $h(d, s, t) \neq \infty$ for some $t \in \mathbb{N}_0$, and for any d and s, there are at most 2^d values of t such that $h(d, s, t) \neq \infty$.

The density of the tile $C \times F$ is equal to the following:

$$\sum_{d=0}^{\infty} 2^{-(3d+1)} \sum_{\substack{s,t \in \mathbb{N}_0 \\ h(d,s,t) \neq \infty}} 2^{-d-s-t-3} \tag{2.6}$$

Since for any d, there are at most 2^d values of s such that $h(d, s, t) \neq \infty$ for some $t \in \mathbb{N}_0$, and for any d and s, there are at most 2^d values of t such that $h(d, s, t) \neq \infty$, the inner sum in (2.6) is at most

$$\sum_{s,t=0}^{2^d-1} 2^{-d-s-t-3} .$$

The constraint on the third line in Figure 2.20 now yields that $h(d, s, t) \neq \infty$ if and only if $s < 2^d$ and $t < 2^d$. Since it holds that if $h(d, s, t) \neq \infty$, $h(d, s', t') \neq \infty$ and $h(d, s, t) < h(d, s', t')$, then either $s = s'$ and $t < t'$ or $s < s'$, it follows that $h(d, s, t) = 2^d \cdot s + t$ for all d, $s < 2^d$ and $t < 2^d$. It follows that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in C \times F$.

2.4.4 Forcing densities

We now focus on the tile $G \times G$, which contains the graphon W_F itself; we define the value $W_0(\eta_G(x), \eta_G(y))$ to be equal to $W_F(x, y)$ for every $(x, y) \in [0, 1)^2$.

Consider the first decorated constraint given in Figure 2.21. Almost every choice of the roots of the constraint satisfies the following: if $x \in F$ is the F-root,

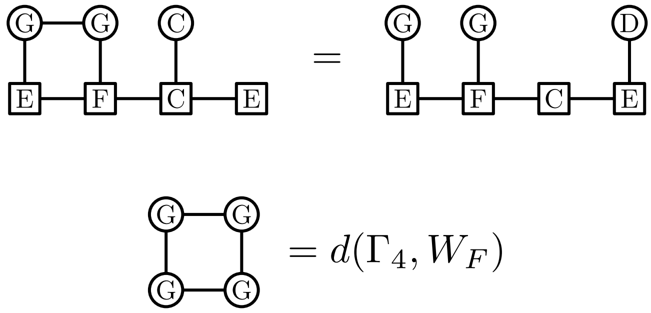


Figure 2.21: The decorated constraint forcing the $G \times G$ tile.

$f_F(x) \in I_d$, $d \in \mathbb{N}_0$, and

$$\left\lfloor \frac{f_F(x) - \min I_d}{|I_d|} \cdot 2^{2d} \right\rfloor = s \times 2^d + t \ ,$$

where $s, t \in \{0, \ldots, 2^d - 1\}$, then the left E-root $y \in E$ satisfies that $f_E(y) \in I_d$ and

$$\left\lfloor \frac{f_E(y) - \min I_d}{|I_d|} \cdot 2^d \right\rfloor = s \ .$$

Moreover, the C-root $y' \in C$ and the right E-root $y'' \in E$ satisfy that $f_C(y') \in I_{d,s,t}$, and $f_E(y'') \in I_{d,s,t}$. The left hand side of the density constraint is then equal to

$$2^{-d-s-t-3} \int_{f_G^{-1}(I^d(s)) \times f_G^{-1}(I^d(t))} W(x, y) \, \mathrm{d}x \, \mathrm{d}y \ ,$$

and the right hand side of the density constraint is equal to

$$2^{-2d} \cdot \deg_W^D(y'') = 2^{-2d} \cdot 2^{-d-s-t-3} \cdot \delta(d, s, t) \ .$$

It follows that

$$2^{2d} \int_{f_G^{-1}(I^d(s)) \times f_G^{-1}(I^d(t))} W(x, y) \, \mathrm{d}x \, \mathrm{d}y = \delta(d, s, t) \ . \tag{2.7}$$

Fix a measurable bijection $\psi : [0, 1] \to G$ such that $|\psi^{-1}(X)| = |X|/|G|$ for every measurable $X \subseteq G$, and define a graphon W_G as $W_G(x, y) = W(\psi(x), \psi(y))$ and a graphon W'_F as $W'_F(x, y) = W_F(f_G(\psi(x)), f_G(\psi(y)))$. Observe that $W'_F(x, y) = W_0(g(\psi(x)), g(\psi(y)))$ for almost every $(x, y) \in G \times G$. Note that the second constraint in Figure 2.21 implies that $d(\Gamma_4, W_G) = d(\Gamma_4, W_F)$, which is equal to $d(\Gamma_4, W'_F)$. We now show that W_G and W'_F are equal almost everywhere.

Suppose that $||W_G - W'_F||_\square = \varepsilon > 0$; note that $\varepsilon \leq 1$. For $d \in \mathbb{N}_0$, define a graphon W^d to be a step graphon with parts $\psi^{-1}\left(f_G^{-1}\left(I^d(k)\right)\right)$, $k = 0, \ldots, 2^d - 1$, such that

$$W^d(x,y) = \delta(d,s,t) \text{ for } x \in \psi^{-1}\left(f_G^{-1}\left(I^d(s)\right)\right) \text{ and } y \in \psi^{-1}\left(f_G^{-1}\left(I^d(t)\right)\right).$$

The sequence $(W^d)_{d\in\mathbb{N}_0}$ forms a martingale on $[0,1]^2$, and Doob's Martingale Convergence Theorem implies that W^d converges uniformly to W'_F. Hence, there exists $d \in \mathbb{N}_0$ such that $||W'_F - W^d||_\square \leq \varepsilon^4/1800$. Apply Proposition 2 to the graphon W_G and the partition $\psi^{-1}\left(f_G^{-1}\left(I^d(k)\right)\right)$, $k \in \{0, \ldots, 2^d - 1\}$, to obtain a step graphon W'_G that refines W^d and is $\varepsilon^4/1800$-close to W_G. Consequently, we get $||W'_G - W^d||_\square \geq \varepsilon - \varepsilon^4/1800 \geq \varepsilon/2$, which implies that

$$d(\Gamma_4, W'_G) - d(\Gamma_4, W^d) \geq \varepsilon^4/128 \tag{2.8}$$

by Lemma 7. On the other hand, the choice of W'_G and W^d implies that

$$\left|d(\Gamma_4, W_G) - d(\Gamma_4, W'_G)\right| \leq \varepsilon^4/300 \text{ and } \left|d(\Gamma_4, W'_F) - d(\Gamma_4, W^d)\right| \leq \varepsilon^4/300\,. \tag{2.9}$$

The inequalities (2.8) and (2.9) now yield that $d(\Gamma_4, W'_F) > d(\Gamma_4, W_G)$. However, this is impossible since $d(\Gamma_4, W'_F) = d(\Gamma_4, W_G)$. Hence, the graphons W_G and W'_F are equal almost everywhere, which implies that $W(x,y)$ and $W_0(g(x), g(y))$ are equal for almost every $(x,y) \in G \times G$.

2.5 Cleaning up

We now finish the description and the argument that the graphon W_0 is finitely forcible. Let us start with the remaining tiles between the parts $A, \ldots, G$. Fix (X,Y) to be one of the pairs (B,G), (C,G), or (D,G) and define $W_0(\eta_X(x), \eta_Y(y)) = 0$ for all $(x,y) \in [0,1)^2$. Clearly, the first decorated constraint in Figure 2.22 forces W to be equal to zero for almost every $(x,y) \in X \times Y$. Hence, we can conclude that $W(x,y) = W_0(g(x), g(y))$ for almost every pair $(x,y) \in (A \cup \cdots \cup G \cup P)^2$.

Similarly, we define $W_0(\eta_Q(x), \eta_Q(y)) = 1$ for all $(x,y) \in [0,1)^2$; this is easy to force by the second constraint in Figure 2.22. Hence, $W(x,y) = W_0(g(x), g(y))$ for almost every pair $(x,y) \in Q \times Q$.

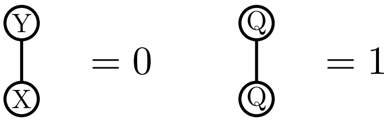


Figure 2.22: The decorated constraint forcing the tiles $X \times Y$, where (X, Y) is one of the pairs (B, G), (C, G) and (D, G), and the decorated constraint forcing the tile $Q \times Q$.

2.5.1 Degree balancing

We use the tiles $Q \times X$, where $X \in \{A, \ldots, G\} \cup \{P\}$, to guarantee that

$$\deg_{W_0}^{A_0 \cup \cdots \cup G_0 \cup P_0 \cup Q_0}(x) = \frac{5}{13}$$

for every vertex $x \in A_0 \cup \cdots \cup G_0 \cup P_0$. It may seem counterintuitive to force the degrees of the vertices in the parts $A, \ldots, G, P$ to be equal; however, it is simpler to begin by enforcing the parts to be degree-regular (with the same degree) and then enforce the different degrees of the parts.

First, note that $\deg_{W_0}^{A_0 \cup \cdots \cup G_0 \cup P_0}(x) \leq \frac{5}{8}$ for every $x \in A_0 \cup \cdots \cup G_0 \cup P_0$. Let $\xi(x) = 1 - \frac{8}{5} \cdot \deg_{W_0}^{A_0 \cup \cdots \cup G_0 \cup P_0}(x)$ for every such x; note that $\xi(x) \in [0, 1]$. We define $W_0(x, y) = \xi(x)$ for every $x \in A_0 \cup \cdots \cup G_0 \cup P_0$ and $y \in Q_0$. Further, we define $W_0(x, y) = 1$ for all $(x, y) \in Q_0^2$.

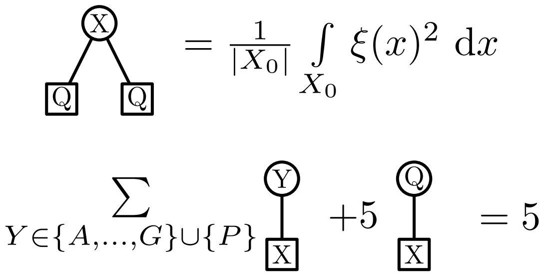


Figure 2.23: The decorated constraints forcing the tiles $Q \times X$ where $X \in \{A, \ldots, G\} \cup \{P\}$.

Fix $X \in \{A, \ldots, G\} \cup \{P\}$ and consider the decorated constraints given in Figure 2.23. The first constraint implies that almost every z and z' from Q satisfy that

$$\int_X W(z, x) W(z', x) \, \mathrm{d}x = \int_{X_0} \xi(x)^2 \, \mathrm{d}x \, .$$

Lemma 4 implies that almost every z from Q satisfies that

$$\int_X W(z,x)^2 \, \mathrm{d}x = \int_{X_0} \xi(x)^2 \, \mathrm{d}x \, .$$

In particular, when z is fixed and $W(z,x)$ is viewed as a function of x, the L_2-norm of the function $W(z,x)$ for almost every $z \in Q$ is the same, and the inner product of the functions $W(z,x)$ and $W(z',x)$ for almost every pair $z, z' \in Q$ is also the same and equal to the L_2-norm. Hence, the Cauchy-Schwarz Inequality yields that there exists a function $h : X \to [0,1]$ such that $W(z,x) = h(x)$ for almost every $x \in X$ and almost every $z \in Q$. It follows that $W(x,z) = h(x)$ for almost every pair $(x,z) \in X \times Q$.

Since $\deg_W^{A\cup\dots\cup G\cup P}(x) = \deg_{W_0}^{A_0\cup\dots\cup G_0\cup P_0}(g(x))$ for almost every $x \in X$, the second constraint in Figure 2.23 implies that $h(x) = \xi(g(x))$ for almost every $x \in X$. It follows that $W(x,y) = W_0(g(x),g(y))$ for almost every pair $(x,y) \in X \times Q$. We now conclude that $W(x,y) = W_0(g(x),g(y))$ for almost every pair $(x,y) \in (A \cup \dots \cup G \cup P \cup Q)^2$.

2.5.2 Degree distinguishing

It remains to define and analyze the tiles $X \times R$, $X \in \{A, \dots, G, P, Q, R\}$. Fix (X,k) to be one of the pairs $(A,0), \dots, (G,6), (P,7), (Q,8)(R,9)$. We define $W_0(x,y) = k/18$ for all $x \in X_0$ and $y \in R_0$. It is easy to check that each vertex of X_0 has the same degree in W_0, and this degree is equal to the one given in Table 2.1.

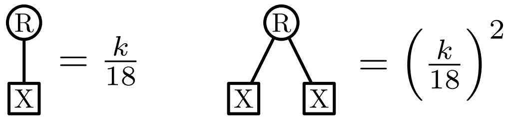


Figure 2.24: The decorated constraints used to force the structure of the tiles $X \times R$ where $(X,k) \in \{(A,0), \dots, (G,6), (P,7), (Q,8), (R,9)\}$.

Consider the two constraints given in Figure 2.24. The first constraint implies that it holds for almost every $x \in X$ that

$$\frac{1}{|R|} \int_R W(x,y) \, \mathrm{d}y = \frac{k}{18} \, ,$$

and the second constraint in Figure 2.24 implies that it holds for almost every pair

$(x, x') \in X^2$ that

$$\frac{1}{|R|} \int_R W(x, y) W(x', y) \, \mathrm{d}y = \left(\frac{k}{18} \right)^2 .$$

We conclude using Lemma 4 that it holds that

$$\frac{1}{|R|} \int_R W(x, y)^2 \, \mathrm{d}y = \left(\frac{k}{18} \right)^2$$

for almost every $x \in X$. The Cauchy-Schwarz Inequality now yields that $W(x, y) = k/18$ for almost every pair $(x, y) \in X \times R$. We can now conclude that $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in X \times R$.

We have shown that if a graphon W satisfies the presented decorated constraints, then $W(x, y) = W_0(g(x), g(y))$ for almost every $(x, y) \in [0, 1]^2$. Since all the presented decorated constraints are satisfied by W_0 and they can be turned into ordinary constraints by Lemma 2, the proof of Theorem 1 is now finished.

2.6 Further remarks

The only constraints used to force the structure of the graphon W_0 that depend on the graphon W_F are the last constraint in Figure 2.10, the last constraint in Figure 2.21 and the first constraint in Figure 2.23. In each of the three constraints, the structure of the graphon W_F influences the numerical value of the right side of the constraint only. Hence, Theorem 1 holds in the following stronger form.

Theorem 2. *There exist graphs $H_1, \dots, H_m$ with the following property. For every graphon W_F, there exist a graphon W_0 and reals $\delta_1, \dots, \delta_m \in [0, 1]$ such that W_F is a subgraphon of W_0 that is formed by a $1/14$ fraction of the vertices of W_0 and the graphon W_0 is the only graphon W, up to a weak isomorphism, such that $d(H_i, W) = \delta_i$ for all $i \in [m]$.*

In view of Theorem 2, one can wonder how much the fraction $1/14$ could be improved. Using similar techniques to the ones presented in this chapter, Kráľ, Lovász Jr., Noel and Sosnovec, announced the following strengthening of Theorem 2:

Theorem 3. *[59] For every $\varepsilon > 0$, there exist graphs $H_1, \dots, H_{m_\varepsilon}$ with the following property. For every graphon W_F, there exist a graphon W_0 and reals $\delta_1, \dots, \delta_{m_\varepsilon} \in [0, 1]$ such that W_F is a subgraphon of W_0 that is formed by a $1 - \varepsilon$ fraction of the vertices of W_0 and the graphon W_0 is the only graphon W, up to a weak isomorphism, such that $d(H_i, W) = \delta_i$ for all $i \in [m_\varepsilon]$.*

Notice that unlike Theorem 2, the set of graphs $H_1, \ldots, H_{m_\varepsilon}$ in Theorem 3 depends on $\varepsilon > 0$ and this is necessary as shown in [59].

The construction presented in the proof of Theorem 1 can be viewed as a map from the space of all graphons to the space of finitely forcible graphons. The particular map implied by the proof of Theorem 1 is not continuous with respect to the cut norm topology (and we have not attempted to achieve this property). However, the following weaker statement can be derived from the proof of Theorem 1 since the L_1-distance between the functions defining the graphons W_0 and W_0' is at most ε for a suitable value of k.

Proposition 3. *For every $\varepsilon > 0$, there exists $k \in \mathbb{N}$ such that the following holds. If W_F and W_F' are two graphons such that the densities of their dyadic squares of sizes 2^{-k} agree up to the first k bits after the decimal point in the standard binary representation, then the cut distance between the finitely forcible graphons W_0 and W_0' containing W_F and W_F', respectively, that are constructed in the proof of Theorem 1, is at most ε.*

Using Theorem 2, Proposition 3 and several key new ideas, Grzesik, Král and Lovász Jr. constructed a counterexample [42]to a conjecture of Lovász on the finite forcibility of optimum solutions in extremal graph theory. The conjecture states that every finite feasible set of subgraph density constraints can be extended further by a finite set of density constraints so that the resulting set is satisfied by an asymptotically unique graph. It was considered to be the second most important problem on graph limits and the most important problem on dense graph limits.

Chapter 3

The step Sidorenko property and non-norming edge-transitive graphs

Sidorenko's Conjecture is one of the most important open problems in extremal graph theory. A graph H has the *Sidorenko property* if a quasirandom graph minimizes the density of H among all graphs with the same edge density. The beautiful conjecture of Erdős and Simonovits [82] and of Sidorenko [81] asserts that every bipartite graph has the Sidorenko property (it is easy to see that non-bipartite graphs fail to have the property). In this chapter, we consider a more general property, the step Sidorenko property, and explore the link between this property and weakly norming graphs to show the existence of a bipartite edge-transitive graph that is not weakly norming. This answers a question of Hatami [47] whether such graphs exist.

Sidorenko's Conjecture has been a subject of a great amount of interest inside extremal combinatorics. Sidorenko [81] confirmed the conjecture for trees, cycles and bipartite graphs with one of the sides having at most three vertices; the case of paths is equivalent to the Blakley-Roy inequality for matrices, which was proven in [8]. Additional graphs, such as cubes and bipartite graphs with a vertex complete to the other part, were added to the list of graphs with the Sidorenko property by Conlon, Fox and Sudakov [17], by Hatami [47], and by Szegedy [85]. Recursively described classes of bipartite graphs that have the Sidorenko property were obtained by Conlon, Kim, Lee and Lee [18], by Kim, Lee and Lee [56], by Li and Szegedy [63] and by Szegedy [84]. In particular, Szegedy [84] has described a class of graphs called thick graphs, which are amenable to showing the Sidorenko property using the entropy method argument that he developed. Sidorenko's Conjecture is known to hold in the local sense [65, Proposition 16.27], i.e., a small modification of a quasirandom graph preserving its edge density does not decrease the number of copies of any bipartite graph. A stronger statement of this type, which comes with uniform quantitative bounds, has recently been proven by Fox and Wei [31].

Sidorenko's Conjecture is also related to other well-studied problems in graph theory. We would like to particularly mention the connection to quasirandom graphs. We say that a graph H is *forcing* if all minimizers of the density of H among graphs with the same edge density are quasirandom graphs. Note that if H is forcing, then H has the Sidorenko property. The classical result of Thomason [86], also see [15], says that the cycle of length four is forcing. This result was generalized by Chung, Graham and Wilson [15], who showed that every complete bipartite graph $K_{2,n}$ is forcing, and by Skokan and Thoma [83], who showed that all complete bipartite graphs are forcing. A characterization of forcing graphs was stated as a question by Skokan and Thoma [83] and conjectured by Conlon, Fox and Sudakov [17]: a graph H is forcing if and only if H is bipartite and contains a cycle.

Another graph theoretic notion related to Sidorenko's Conjecture is that of common graphs. A graph H is *common* if a quasirandom graph minimizes the sum of densities of H and the complement of H. An old theorem of Goodman [40] says that the complete graph K_3 is common. The conjecture of Erdős that the complete graph K_4 is also common was disproved by an ingenious construction of Thomason [88]; counterexamples with a simpler structure were found by Franek and Rödl in [33]. Jagger, Šťovíček and Thomason [52] showed that no graph containing K_4 is common. On the other hand, it is known that the graph obtained from K_4 by removing an edge [32] is common and so is the wheel W_5 [48]. The classification of common graphs remains a wildly open problem.

Our results are motivated by the relation of Sidorenko's Conjecture to *weakly norming* graphs, which are of substantial interest in the theory of graph limits. Due to its technical nature, we defer the definition to Section 3.1. Intuitively, these are graphs H such that the density of H in other graphs defines a norm on the space of graphons (graph limits). Chapter 14.1 in Lovász' book [65] gives an introduction to this notion. Every weakly norming graph has the Sidorenko property [47]. However, every weakly norming graph also has a stronger property [65, Proposition 14.13], which we call the *step Sidorenko property.* Informally speaking, a graph H has the step Sidorenko property if a multipartite quasirandom graph minimizes the density among all multipartite graphs with the same density inside and between its parts; we give a formal definition in Section 3.1. It is not hard to find a graph that has the Sidorenko property but not the step Sidorenko property; the cycle of length four with an added pendant edge is an example (see Section 3.1).

In this chapter, we present techniques for showing that a bipartite graph fails to have the step Sidorenko property. Our techniques allow us to show that graphs as simple and symmetric as toroidal grids, i.e., Cartesian products of any number of cycles, do not have the step Sidorenko property. The only exceptions are hypercubes (and single cycles of even length), which were shown to be weakly norming by Hatami [47] (see also [65, Proposition 14.2] for a concise presentation). The fact that most of the toroidal grids are not weakly norming is surprising when contrasted with the result of Conlon and Lee [19] that the incidence graphs of regular polytopes are weakly norming. Since toroidal grids are edge-transitive, this answers in the negative a question of Hatami [47] whether all edge-transitive bipartite graphs are weakly norming.

3.1 Preliminaries

In this section, we introduce notation related to Cartesian product, graph homomorphisms and present notions from the theory of graph limits that we need for our exposition. We also formally define the Sidorenko property, the step Sidorenko property and weakly norming graphs.

The *Cartesian product* of graphs $G_1, \dots, G_k$, denoted $G_1 \square \cdots \square G_k$, is the graph with vertex set equal to the Cartesian product of the vertex sets of $G_1, \dots, G_k$, where two vertices $(u_1, \dots, u_k)$ and $(v_1, \dots, v_k)$ are adjacent if there exists $1 \le i_0 \le k$ such that $u_{i_0} v_{i_0}$ is an edge of G_{i_0} and $u_i = v_i$ for all $i \ne i_0$.

3.1.1 Graph homomorphisms

A *homomorphism* from a graph H to a graph G is a mapping f from $V(H)$ to $V(G)$ such that if vv' is an edge of H, then $f(v)f(v')$ is an edge of G. If f is a homomorphism from H to G, $|f^{-1}(X)|$ for $X \subseteq V(G)$ denotes the number of vertices of H mapped to a vertex in X and $|f^{-1}(X)|$ for $X \subseteq E(G)$ denotes the number of edges mapped to an edge in X; for simplicity, we write $|f^{-1}(x)|$ instead of $|f^{-1}(\{x\})|$.

We will need to consider homomorphisms extending a partial mapping between vertices of H and G and we now introduce notation that will be helpful in this setting. If H is a graph with k distinguished vertices $v_1, \dots, v_k$, then we write $H(v_1, \dots, v_k)$. If $H(v_1, \dots, v_k)$ and $G(v'_1, \dots, v'_k)$ are two graphs with k distinguished vertices, then a homomorphism from $H(v_1, \dots, v_k)$ to $G(v'_1, \dots, v'_k)$ is a homomorphism from H to G that maps v_i to v'_i for $i = 1, \ldots, k$.

We will also consider homomorphisms to graphs with vertex and edge weights. As given earlier, a *weighted graph* is a graph G where each vertex and each edge of G is assigned a non-negative weight; the mapping w from $V(G) \cup E(G)$ assigning the weights will be referred to as a *weight function* of G. The *weight of a homomorphism* f from H to a weighted graph G, denoted $w(f)$, is defined as

$$\prod_{v \in V(H)} w(f(v)) \prod_{vv' \in E(H)} w(f(v)f(v')) = \prod_{v \in V(G)} w(v)^{|f^{-1}(v)|} \prod_{e \in E(G)} w(e)^{|f^{-1}(e)|} .$$

We will often speak about the sum of the weights of homomorphisms from a graph $H(v_1, \ldots, v_k)$ to a weighted graph $G(v'_1, \ldots, v'_k)$; this sum will be denoted by $\hom(H(v_1, \ldots, v_k), G(v'_1, \ldots, v'_k))$ and we will understand it to be zero if no such homomorphism exists.

We also use the just introduced notation for graphs with distinguished vertices when talking about blow-ups of graphs. A *k-blow-up* of a graph $H(v)$ is the graph obtained from H by replacing the vertex v with k new vertices, which we refer to as *clones* of v. The vertices different from v preserve their adjacencies, the clones of v form an independent set and each of them is adjacent precisely to the neighbors of v. Observe that if H is a weighted graph, then if the edges of the k-blow-up of $H(v)$ have the same weight as in H, the vertices of the k-blow-up except for the clones have the same weights as in H and each clone has weight equal to $1/k$ of the weight of v, then the sum of the weights of homomorphisms from G to H and the sum of the weights of homomorphisms from G to the k-blow-up are the same for every graph G.

3.1.2 Graph limits

Let $t(H, G)$ be the normalized number of homomorphisms from a graph H to a graph G, i.e., $t(H, G) = \hom(H, G)/|V(G)|^{|V(H)|}$ where G in $\hom(H, G)$ is understood to have all the vertex and edge weights equal to one.

Recall from the introduction that $d(H, G)$ is the probability that $|H|$ uniformly and independently randomly chosen vertices of G induce a subgraph isomorphic to H. Observe that $t(H, G)$ can be written in terms of $d(H, G)$ and vice-versa. Therefore, an equivalent way to define a convergent sequence is to say that a sequence $(G_n)_{n \in \mathbb{N}}$ of graphs is *convergent* if the sequence $t(H, G_n)$ converges for every graph H.

Throughout this chapter, we will refer to $t(H, G)$ as the *density* of H in G rather than the homomorphism density. All the statements could be cast in terms of the induced density but for simplicity, we use the homomorphism density in this scenario.

As we have seen a convergent sequence of graphs can be represented by a graphon and one can think (although very imprecisely) of a graphon as a continuous version of the adjacency matrix of a graph. Led by this intuition, we can define the (homomorphism) density of a graph H in a graphon W as

$$t(H, W) = \int_{[0,1]^{V(H)}} \prod_{vv' \in E(H)} W(x_v, x_{v'}) \, \mathrm{d}x^{V(H)} .$$

Note that the definition of $t(H, W)$ does not require W to be non-negative and we can define $t(H, f)$ in the same way for any bounded measurable function $f : [0,1]^2 \to \mathbb{R}$.

The density $t(K_2, W)$ of K_2 is equal to the L_1-norm of a graphon W as a function from $[0,1]^2$. This leads to the question which graphs H can be used to define a norm on the space of measurable functions on $[0,1]^2$ or, more restrictively, on the space of graphons. That is, we say that a graph H is *weakly norming* if the function $\|W\|_H = |t(H, W)|^{1/\|H\|}$ is a norm on the space of graphons, i.e., $\|W\|_H = 0$ if and only if W is equal to zero almost-everywhere and the triangle inequality $\|W_1 + W_2\|_H \leq \|W_1\|_H + \|W_2\|_H$ holds for any two graphons W_1 and W_2. Observe that H is weakly norming if and only if $\| \, |f| \, \|_H$ is a norm on the set of all bounded symmetric functions f from $[0,1]^2$ to $\mathbb{R}$ (if we required that $\|f\|_H$, without the absolute value, is a norm on all such functions, we would get the slightly stronger notion of *norming graphs*).

It is not hard to show that every weakly norming graph must be bipartite. Hatami [47] showed stronger statements as corollaries of his characterization of weakly norming graphs as those satisfying a certain Hölder type inequality. First, every weakly norming graph H must be biregular, i.e., all vertices in the same part of its bipartition have the same degree. Second, every subgraph H' of a connected weakly norming graph H must satisfy that

$$\frac{\|H'\|}{|H'| - 1} \leq \frac{\|H\|}{|H| - 1} .$$

Known examples of weakly norming graphs include complete bipartite graphs (in particular, stars), even cycles and hypercubes.

Every weighted graph G with a weight function w that assigns edges weights between 0 and 1 can be associated with a graphon W_G as follows. Each vertex v of G is associated with a measurable set J_v with measure $w(v)/w(V(G))$ in such a way that the sets J_v, $v \in V(G)$, form a partition of the interval $[0,1]$; $w(V(G))$

denotes the sum of the weights of the vertices of G. For $x \in J_v$ and $y \in J_{v'}$, we set $W(x,y) = w(vv')$ if $vv' \in E(G)$ and $W(x,y) = 0$ otherwise (in particular, we set $W(x,y) = 0$ if $v = v'$). It is not hard to observe that $\hom(H,G)$ is equal to $t(H,W_G) \cdot w(V(G))^{|H|}$; in particular, if the sum of the weights of vertices of G is one, then $\hom(H,G) = t(H,W_G)$. This correspondence will allow us to study weakly norming graphs in terms of weighted homomorphisms.

3.1.3 Step Sidorenko property

We now use the language of graph limits to describe the Sidorenko property and to formally define the step Sidorenko property. A graph H has the *Sidorenko property* if

$$t(K_2,W)^{\|H\|} \leq t(H,W) \tag{3.1}$$

for every graphon W. The letft hand side can also be written as $t(H,U_p)$, where $U_p \equiv p$ is the constant graphon with the same edge density $p = t(K_2,W)$ as W. A graph H is *forcing* if it has the Sidorenko property and (3.1) holds with equality only if W is equal to some $p \in [0,1]$ almost everywhere. As we have presented earlier, Sidorenko's Conjecture asserts that every bipartite graph has the Sidorenko property and the Forcing Conjecture asserts that every bipartite graph with a cycle is forcing.

Let $\mathcal{P}$ be a partition of the interval $[0,1]$ into finitely many non-null measurable sets. We now define the *stepping operator*. If W is a graphon, then the graphon $W^{\mathcal{P}}$ is defined for $(x,y) \in [0,1]^2$ as the 'step-wise average':

$$W^{\mathcal{P}}(x,y) = \frac{1}{|J||J'|} \int_{J \times J'} W(s,t) \,\mathrm{d}s\,\mathrm{d}t$$

where J and J' are the unique parts from $\mathcal{P}$ such that $x \in J$ and $y \in J'$, and $|X|$ denotes the measure of a measurable subset $X \subseteq [0,1]$. Note that the graphon $W^{\mathcal{P}}$ is constant on $J \times J'$ for any $J, J' \in \mathcal{P}$, i.e., the graphon $W^{\mathcal{P}}$ is a step graphon.

Let $\mathcal{P}_0$ be the partition with a single part being the interval $[0,1]$ itself. A graph H has the Sidorenko property if and only if $t(H,W^{\mathcal{P}_0}) \leq t(H,W)$ for every graphon W. This motivates the following definition. A graph H has the *step Sidorenko property* if and only if

$$t(H,W^{\mathcal{P}}) \leq t(H,W)$$

for every graphon W and every partition $\mathcal{P}$ of $[0,1]$ into finitely many non-null measurable sets. Since all weakly norming graphs [65, Proposition 14.13] have the step

Sidorenko property, it follows that complete bipartite graphs, even cycles, hypercubes and more generally reflection graphs defined by Conlon and Lee [19] all have the step Sidorenko property. In fact, all connected graphs with the step Sidorenko property are weakly norming [26].

The definition of the step Sidorenko property yields that every graph that has the step Sidorenko property also has the Sidorenko property. However, the converse is not true in general as we now demonstrate in our following example. Let C_4^+ be the 5-vertex graph obtained from a cycle of length four by adding a single vertex adjacent to one of the vertices of the cycle. The graph C_4^+ has the Sidorenko property because, e.g., it is a bipartite graph with a vertex complete to the other part [17]. On the other hand, C_4^+ does not have the step Sidorenko property. Consider the partition $\mathcal{P} = \{[0, \frac{2}{5}), [\frac{2}{5}, 1]\}$ and the graphon W that is defined as follows (the symmetric cases of (x, y) are omitted).

$$W(x,y) = \begin{cases} 0.9 & \text{if } (x,y) \in [0, \frac{1}{5}) \times [0, \frac{1}{5}), \\ 0.85 & \text{if } (x,y) \in [0, \frac{1}{5}) \times [\frac{1}{5}, \frac{2}{5}), \\ 0.2 & \text{if } (x,y) \in [0, \frac{1}{5}) \times [\frac{2}{5}, 1], \\ 1 & \text{if } (x,y) \in [\frac{1}{5}, \frac{2}{5}) \times [\frac{1}{5}, \frac{2}{5}), \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

A straightforward computation yields that

$$\begin{aligned} t(C_4^+, W) &\simeq 0.007453 \text{ and} \\ t(C_4^+, W^{\mathcal{P}}) &\simeq 0.007508 > t(C_4^+, W) . \end{aligned}$$

Hence, the graph C_4^+ does not have the step Sidorenko property.

3.2 Grids

In this section, we demonstrate our techniques from Section 3.3 in a less general setting. We believe that this makes our presentation more accessible.

Intuitively, we consider a graph G with distinguished vertices u_0, u_1, u_2 such that u_0u_1 and u_0u_2 are edges. The idea is to blow-up u_0 into two copies and slightly perturb weights only on edges corresponding to u_0u_1 and u_0u_2, increasing weights of edges for one copy and decreasing it for the other proportionally to a parameter α, resulting in a weighted graph G_α. A partition $\mathcal{P}$ on the corresponding graphon W_α is then defined so that the stepping operator only averages out this perturbation, returning to the original graph: $W_\alpha^{\mathcal{P}} = W_G$. The difference in homomorphism densities $t(H, W_\alpha^{\mathcal{P}}) - t(H, W_\alpha)$ is then analyzed in the limit of small perturbations α: first order changes (those linear in α) cancel out; second order changes result in a condition that can be expressed fairly concisely as positive semidefiniteness of a matrix whose entries count certain constrained homomorphisms.

The more powerful setting in Section 3.3 uses essentially the same idea, only blowing up more vertices, resulting in a larger matrix and allowing to further constraint the homomorphisms we have to count. We turn to choosing the starting weighted graph G and interpreting these counts in later corollaries.

Theorem 4. *Let H be a graph and let G be a weighted graph with three distinguished vertices u_0, u_1 and u_2 such that u_0u_1 and u_0u_2 are edges. For $i, j \in \{1, 2\}$, let M_{ij} be the sum of the weights of homomorphisms from $H(v_0, v_1, v_2)$ to $G(u_0, u_i, u_j)$ summed over all choices of vertices v_0, v_1 and v_2 in H such that v_0v_1 and v_0v_2 are edges, i.e.,*

$$M_{ij} = \sum_{v_0v_1, v_0v_2 \in E(H)} \hom(H(v_0, v_1, v_2), G(u_0, u_i, u_j)) \ .$$

If the (2×2)-matrix M is not positive semidefinite, i.e., $M_{11}M_{22} < {M_{12}}^2$, then H does not have the step Sidorenko property.

Proof. Let w be the weight function of G. We assume that the sum of the weights of vertices of G is one (if needed, we multiply the weights of all vertices by the same constant). Consider the step graphon W_G associated with the weighted graph G. Let J_u be the measurable set corresponding to a vertex u of G and set $\mathcal{P} = \{J_u, u \in V(G)\}$.

Suppose that the matrix M associated with G is not positive semidefinite and fix a vector $a = (a_1, a_2)^T$ such that $a^T M a < 0$. We next define a weighted graph G_α with a parameter $\alpha \geq 0$ as follows. The graph G_α is a 2-blow-up of $G(u_0)$; let u_0^+ and u_0^- be the clones of u_0. Each of the clones u_0^+ and u_0^- has weight $w(u_0)/2$. The weight of the edge $u_0^+u_i$ is $w(u_0u_i)(1 + \alpha a_i)$ and the weight of the edge $u_0^-u_i$ is $w(u_0u_i)(1 - \alpha a_i)$, $i = 1, 2$. The remaining vertices and edges have weights equal to their counterparts in G. Let W_α be the step graphon associated with the weighted graph G_α such that the set corresponding to a vertex $u \neq u_0$ is J_u and the sets corresponding to the vertices u_0^+ and u_0^- are subsets of J_{u_0}. Observe that $W_G = W_\alpha$ for $\alpha = 0$ and that $W_G = W_\alpha^{\mathcal{P}}$ for any α.

Our aim is to show that $t(H, W_\alpha) < t(H, W_G)$ for some $\alpha \in (0, 1)$. To do so, we analyze the density $t(H, W_\alpha)$ as a function of α. Note that $t(H, W_\alpha)$ is actually a polynomial in α. We next wish to determine the coefficients c_1 and c_2 such that

$$t(H,W_\alpha) = t(H,W_G) + c_1\alpha + c_2\alpha^2 + O(\alpha^3) \ . \tag{3.2}$$

The coefficient c_1 can be determined as follows:

$$c_1 = \sum_{v_0v_1\in E(H)} a_1 \hom(H(v_0,v_1),G_0(u_0^+,u_1)) - a_1 \hom(H(v_0,v_1),G_0(u_0^-,u_1)) + a_2 \hom(H(v_0,v_1),G_0(u_0^+,u_2)) - a_2 \hom(H(v_0,v_1),G_0(u_0^-,u_2)) \ .$$

Since $\hom(H(v_0,v_1),G_0(u_0^+,u_i)) = \hom(H(v_0,v_1),G_0(u_0^-,u_i))$ for all edges $v_0v_1 \in E(G)$ and all $i \in \{1,2\}$, we conclude that $c_1 = 0$.

We next analyze the coefficient c_2. In this case, we need to count homomorphisms mapping two edges, say v_0v_1 and $v_0'v_1'$, of H to edges $u_0^+u_i$ and to $u_0^-u_i$ of G_0, $i=1,2$. If $v_0 \neq v_0'$, then the contributions of the homomorphisms mapping the edge v_0v_1 to $u_0^+u_i$ and $u_0^-u_i$ have opposite signs and cancel out. Hence, we obtain the following formula for c_2:

$$c_2 = \sum_{v_0v_1,v_0v_2\in E(H)} \sum_{i,j=1}^{2} a_ia_j \quad \big(\hom(H(v_0,v_1,v_2),G_0(u_0^+,u_i,u_j)) + \hom(H(v_0,v_1,v_2),G_0(u_0^-,u_i,u_j))\big) \ .$$

The definition of the matrix M now yields that

$$c_2 = \sum_{i,j=1}^{2} a_ia_j \cdot M_{ij} = a^T M a < 0 \ .$$

Since $c_1 = 0$ and $c_2 < 0$, we conclude using $W_G = W_\alpha^{\mathcal{P}}$ and (3.2) that $t(H,W_\alpha) < t(H,W_G)$ for small enough $\alpha > 0$. It follows that the graph H does not have the step Sidorenko property. □

The setting of Theorem 4 is sufficient to prove that the only two-dimensional toroidal grid that is weakly norming is $C_4\square C_4$ (note that the toroidal grids $C_\ell\square C_\ell$ with ℓ odd are not Sidorenko, and hence also not weakly norming, because they are not bipartite).

The idea is to consider homomorphisms from the grid to itself: the identity homomorphisms then contributes to the off-diagonal entry of the matrix from Theorem 4. Homomorphisms corresponding to diagonal entries have to "fold" two edges onto one. We show this forces some vertices to be mapped closer together or at certain distinct positions; changing weights in the target grid accordingly allows us

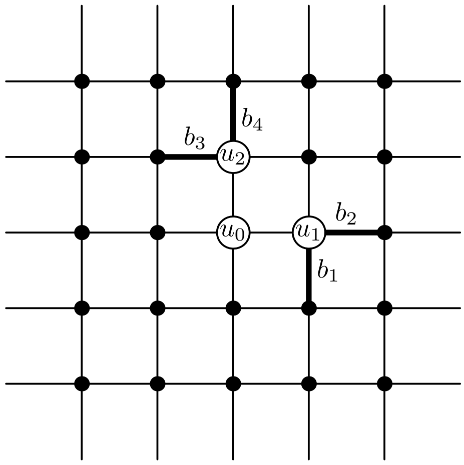


Figure 3.1: Notation used in the proof of Corollary 2. The edges b_1, b_2, b_3 and b_4 are drawn bold.

to make the contribution of such homomorphisms smaller, making the matrix not positive semidefinite.

Corollary 2. *Let $\ell \geq 6$ be an even integer. The Cartesian product $C_\ell \square C_\ell$ does not have the step Sidorenko property.*

Proof. Fix $\ell \geq 6$ and let G and H be both equal to the graph $C_\ell \square C_\ell$; we denote the vertices of G and H by (i, j), $0 \leq i, j \leq \ell - 1$, in such a way that two vertices are adjacent if they agree in one of the coordinates and differ by one in the other (all computations with the entries are computed modulo ℓ throughout the proof). Let u_0 be the vertex $(0, 0)$, u_1 the vertex $(1, 0)$ and u_2 the vertex $(0, 1)$. Further, let b_1 be the edge $(1, 0)(1, -1)$, b_2 the edge $(1, 0)(2, 0)$, b_3 the edge $(0, 1)(-1, 1)$ and b_4 the edge $(0, 1)(0, 2)$ (see Figure 3.1).

We next define the weights of the vertices and the edges of G; to do so, we use a parameter $\gamma \in \mathbb{N}$, which will be fixed later. The weight $w(v)$ of a vertex v is $\gamma^{\mathrm{dist}(u_0,v)}$ for $v \neq u_0, u_1, u_2$, $w(u_0) = \gamma^{-3}$ and $w(u_i) = \gamma^{\mathrm{dist}(u_0,u_i)-3} = \gamma^{-2}$, $i = 1, 2$. The weights of all edges of G are equal to one except for the edges b_1, b_2, b_3 and b_4 that have weight $\gamma^{-1/4}$.

We wish to apply Theorem 4 with the graphs H and G, and the distinguished vertices u_0, u_1 and u_2. Instead of verifying that the matrix M from the statement of Theorem 4 is not positive semidefinite, we consider the matrix M such that

$$M_{ij} = \sum_{v_1, v_2 \in N_H(u_0)} \hom(H(u_0, v_1, v_2), G(u_0, u_i, u_j)) \ .$$

Since H is vertex-transitive, the considered matrix M is positive semidefinite if and only if the matrix from the statement of Theorem 4 is. Observe that $M_{1,1} = M_{2,2}$ and $M_{1,2} = M_{2,1}$.

Consider a homomorphism f from $H(u_0, v_1, v_2)$ to $G(u_0, u_i, u_j)$ for some $i, j \in \{1, 2\}$. Observe that the weight of the homomorphism f is equal to

$$\gamma^{\sum_{v \in V(H)} \mathrm{dist}(u_0, f(v)) - 3\left|f^{-1}(\{u_0,u_1,u_2\})\right| - \frac{1}{4}\left|f^{-1}(\{b_1,b_2,b_3,b_4\})\right|} .$$

Note that if f is the identity, then the weight of f is equal to γ^W where

$$W = \sum_{v \in V(H)} \mathrm{dist}(u_0, v) - 10 .$$

Since the identity is a homomorphism from $H(u_0, u_i, u_j)$ to $G(u_0, u_i, u_j)$ for $i \neq j$, it follows that the entries $M_{1,2}$ and $M_{2,1}$ are of order $\Omega(\gamma^W)$, as functions of γ. We next show that both $M_{1,1}$ and $M_{2,2}$ are of order $o(\gamma^W)$. Since $M_{1,1} = M_{2,2}$, it is enough to argue that $M_{1,1} = o(\gamma^W)$ once we can make γ sufficiently large.

We show that every homomorphism f from $H(u_0, v_1, v_2)$ to $G(u_0, u_1, u_1)$ has weight at most $\gamma^{W-\frac{1}{2}}$; this will imply that $M_{1,1} = o(\gamma^W)$. Fix a homomorphism f from $H(u_0, v_1, v_2)$ to $G(u_0, u_1, u_1)$ with weight at least γ^W. By symmetry, we may assume that $v_1 = (1, 0)$ and $v_2 \in \{(-1, 0), (0, 1)\}$. Note that $\left|f^{-1}(\{u_0, u_1, u_2\})\right| \geq 3$. Since f is a homomorphism, any shortest path from u_0 to v is mapped by f to a walk of at most length $\mathrm{dist}(u_0, v)$ from $f(u_0) = u_0$ to $f(v)$, it follows that $\mathrm{dist}(u_0, f(v)) \leq \mathrm{dist}(u_0, v)$ for every vertex v. Also observe that the parities of $\mathrm{dist}(u_0, f(v))$ and $\mathrm{dist}(u_0, v)$ are the same since the graph $G = H$ is bipartite. Since the weight of f is at least γ^W, the following holds: $\left|f^{-1}(\{u_0, u_1, u_2\})\right| = 3$, $\mathrm{dist}(u_0, f(v)) = \mathrm{dist}(u_0, v)$ for every vertex v of H and $\left|f^{-1}(\{b_1, b_2, b_3, b_4\})\right| \leq 4$. Since $\left|f^{-1}(\{u_0, u_1, u_2\})\right| = 3$, no vertex other than u_0, v_1 and v_2 is mapped by f to any of u_0, u_1 and u_2; in particular, no vertex is mapped to u_2.

To finish the proof, we distinguish two cases based on whether $v_2 = (-1, 0)$ or $v_2 = (0, 1)$. We start with analyzing the case $v_2 = (-1, 0)$. Let $i \in \{1, 2\}$ and let v be a neighbor of v_i different from $(0, 0)$ and $v_i + v_i$. If $f(v) = (1, 1)$ or $f(v) = (2, 0)$, then the common neighbor of $(0, 0)$ and v different from v_i must be mapped to u_1 or u_2, which is impossible. Hence, $f(v) = (1, -1)$. Since the choice of i and v was arbitrary, it follows that all the four edges $(1, 0)(1, 1)$, $(1, 0)(1, -1)$, $(-1, 0)(-1, 1)$ and $(-1, 0)(-1, -1)$ are mapped to the edge b_1; in particular, no other edge is mapped to b_1 or b_2. This implies that the vertex $(2, 0)$ is mapped by f to $(1, 1)$. It follows that the vertex $(2, 1)$, which is a common neighbor of $(1, 1)$ and

$(2,0)$, must be mapped to the unique common neighbor $u_1 = (1,0)$ of the vertices $f((1,1)) = (1,-1)$ and $f((2,0)) = (1,1)$, which is impossible. This finishes the analysis of the case $v_2 = (-1,0)$.

It remains to analyze the case that $v_2 = (0,1)$. If the vertex $(1,-1)$ was mapped to $(2,0)$ or $(1,1)$, then the vertex $(0,-1)$, which is a common neighbor of $(1,-1)$ and $(0,0)$, would have to be mapped to $(1,0)$ or $(0,1)$, which is impossible. Hence, the vertex $(1,-1)$ is mapped by f to itself and the vertex $(0,-1)$ is also mapped to itself. Since swapping coordinates is a symmetry mapping v_1 and v_2 between each other, a symmetric argument yields that the vertex $(-1,0)$ is mapped to $(0,-1)$.

Next, if the vertex $(2,0)$ was mapped to the vertex $(1,1)$, then the vertex $(2,-1)$, which is a common neighbor of $(2,0)$ and $(1,-1)$, would have to be mapped to $(1,0)$, which is impossible. It follows that the vertex $(2,0)$ must be mapped to $(2,0)$ or $(1,-1)$. We conclude that the edge b_1 is mapped to itself and the edge b_2 to either b_1 or b_2. A symmetric argument yields that the edge b_3 is mapped to b_1 and the edge b_4 to b_1 or b_2. In particular, no other edges of G are mapped to any of the edges b_1, b_2, b_3 and b_4. This implies that the vertex $(1,1)$ is mapped by f to itself. Consequently, the vertex $(2,0)$ is also mapped to itself (otherwise, the vertex $(2,1)$ would have to be mapped to $(1,0)$).

We now prove the following statement for $r = 1, \ldots, \ell/2 - 1$ by induction on r: all the vertices $(r,1)$, $(r,-1)$ and $(r+1,0)$ are mapped by f to themselves. We have already established this statement for $r = 1$, so it remains to present the induction step. Fix $r = 2, \ldots, \ell/2 - 1$ and assume that all the vertices $(r-1,1)$, $(r-1,-1)$ and $(r,0)$ are mapped to themselves. The vertex $(r,1)$, which is a common neighbor of $(r-1,1)$ and $(r,0)$, must be mapped to a common neighbor of $(r-1,1)$ and $(r,0)$ at the distance $r+1$ from $(0,0)$. However, the only such vertex is $(r,1)$. A symmetric argument yields that the vertex $(r,-1)$ is mapped to itself. Since the vertex $(r+1,0)$ must be mapped to a neighbor of $(r,0)$ at distance $r+1$ from $(0,0)$, it can only be mapped to one of the vertices $(r,1)$, $(r+1,0)$ and $(r,-1)$. By symmetry, it is enough to exclude that it is mapped to $(r,1)$. If this was the case, then the vertex $(r+1,-1)$, which is a common neighbor of $(r,-1)$ and $(r+1,0)$, must be mapped to $(r,0)$, which is impossible. Hence, the vertex $(r+1,0)$ is mapped to itself, concluding the proof of the statement.

We have just shown that the vertex $(\ell/2,0) = (-\ell/2,0)$ is mapped to itself; earlier, we have shown that the vertex $(-1,0)$ is mapped to $(0,-1)$. However, the path $(-1,0)(-2,0)\cdots(-\ell/2,0)$ must be mapped by f to a walk with at most $\ell/2$ vertices but there is no such walk between the vertices $(0,-1)$ and $(-\ell/2,0)$. Hence, there is no homomorphism from $H(u_0, v_1, v_2)$ to $G(u_0, u_1, u_1)$ with weight at least γ^W. $\square$

3.3 General Condition

We now present our general technique for establishing that certain graphs do not have the step Sidorenko property. One difference is that instead of considering only two neighbors of a distinguished vertex u_0, we can choose any number of neighbors $u_1, \ldots, u_k$, giving a larger matrix. More importantly, we are able to restrict counted homomorphism to only those that map the neighborhood of each u_i bijectively (to the neighborhood of the image of u_i, or a chosen subset of it).

The proof extends the arguments presented in the proof of Theorem 4. The main new idea is that by blowing up u_i, and appropriately choosing weights on copies of the edges to its neighbors, we can obtain an expression that is counting homomorphisms to the original graph, but with a weight that is an arbitrary function of how many neighbors of u_i map to each neighbor of the image of u_i. We choose this function to enforce that exactly one neighbor of u_i (or exactly zero) must map to each neighbor of its image.

Theorem 5. *Let H be a graph and let G be a weighted graph with $k+1$ distinguished vertices $u_0, u_1, \ldots, u_k$ such that $u_0u_1, \ldots, u_0u_k$ are edges and $u_1, \ldots, u_k$ form an independent set. Further, let U_i, $i = 1, \ldots, k$, be a subset of neighbors of u_i containing u_0, and let M be the $(k \times k)$-matrix such that the entry M_{ij} is the sum of the weights of homomorphisms from $H(v_0, v_1, v_2)$ to $G(u_0, u_i, u_j)$, where the sum runs over all choices of vertices v_0, v_1 and v_2 in H, such that the neighbors of v_1 are one-to-one mapped to U_i and the neighbors of v_2 to U_j. If the matrix M is not positive semidefinite, then H does not have the step Sidorenko property.*

Proof. Suppose that the matrix M is not positive semidefinite and fix a vector a such that $a^T M a < 0$. Let w be the weight function of G. As in the proof of Theorem 4, we assume that the sum of the weights of vertices of G is one. Similarly, we assume that the weight of each edge is at most $1/2$ (if needed, we can multiply the weights of all edges by the same constant).

We next define a weighted graph $G_{\varepsilon,\alpha}$, which is parameterized by $\varepsilon > 0$ and $\alpha \in \mathbb{R}$. The structure of the graph is independent of ε and α and is the following. Let n be the number of vertices of H. We consider the 3-blow-up of a vertex u_0 and $\left(n^{|U_i|-1} + 1\right)$-blow-up of a vertex u_i. The three clones of u_0 will be denoted by u_0', u_0^+ and u_0^-; one of the $n^{|U_i|-1} + 1$ clones of u_i will be denoted by u_i' and the remaining ones by $u_{i,j_1,\dots,j_{|U_i|-1}}$ where $1 \le j_1, \dots, j_{|U_i|-1} \le n$. We next remove every edge going from the vertex $u_{i,j_1,\dots,j_{|U_i|-1}}$ to a vertex outside the set U_i that is not u_0^+ or u_0^-, i.e., the vertex $u_{i,j_1,\dots,j_{|U_i|-1}}$ is adjacent to u_0^+, u_0^- and the vertices of $U_i \setminus \{u_0\}$.

The weight of the vertex u_0' is $(1-2\varepsilon)w(u_0)$ and the weight of each of the vertices u_0^+ and u_0^- is $\varepsilon w(u_0)$. The weight of the vertex u_i' is $(1-n^{|U_i|-1}\varepsilon)w(u_i)$ and the weight of each of the vertices $u_{i,j_1,\dots,j_{|U_i|-1}}$ is $\varepsilon w(u_i)$. The remaining vertices of $G_{\varepsilon,\alpha}$ have the same weights as in G.

Before defining the weights of the edges, we define an auxiliary matrix B. The matrix B has n rows and n columns and $B_{ij} = 2^{(i-1)(j-1)}$. Note that B is a Vandermonde matrix. Since the matrix B is invertible, there exists a vector b such that $Bb = (0,1,0,\dots,0)^T$. The weight of the edge between u_0^+ and $u_{i,j_1,\dots,j_{|U_i|-1}}$ is equal to

$$w(u_0u_i)\left(1 + a_i\alpha \prod_{m=1}^{|U_i|-1} b_{j_m}\right),$$

and the weight of the edge between u_0^- and $u_{i,j_1,\dots,j_{|U_i|-1}}$ is equal to

$$w(u_0u_i)\left(1 - a_i\alpha \prod_{m=1}^{|U_i|-1} b_{j_m}\right).$$

The weights of the edges incident with u_0' and the remaining edges incident with u_0^+ and u_0^- are equal to the weights of their counterparts in G. Fix $i \in \{1,\dots,k\}$ and let $z_1,\dots,z_{|U_i|-1}$ be the vertices of U_i different from u_0. The weight of the edge between the vertices $u_{i,j_1,\dots,j_{|U_i|-1}}$ and z_m is equal to $2^{j_m-1}w(u_iz_m)$. The weights of the edges incident with the vertex u_i' are the same as the weights of their counterparts in G. We have just defined the weights of all edges incident with at least one clone. The weights of the remaining edges are the same as in G.

We analyze $t(H, W_{\varepsilon,\alpha})$ as a function of α for $\alpha, \varepsilon \in (0,1)$. In particular, we will show that

$$t(H, G_{\varepsilon,\alpha}) = t(H, G_{\varepsilon,0}) + c_\varepsilon \varepsilon^3 \alpha^2 + O(\varepsilon^4\alpha^2) \tag{3.3}$$

for a coefficient c_ε, which we will estimate. Since the coefficient c_ε depends on ε, it is important to emphasize that the constants hidden in big O notation in (3.3) are independent of ε and α, i.e., the equality (3.3) represents that there exists $K > 0$, which is independent of ε, and a coefficient c_ε for every $\varepsilon \in (0,1)$ such that the value of $t(H, G_{\varepsilon,\alpha})$ differs from $t(H, G_{\varepsilon,0}) + c_\varepsilon\varepsilon^3\alpha^2$ by at most $K\varepsilon^4\alpha^2$ for every $\alpha \in (0,1)$.

We now proceed with analyzing the function $t(H, W_{\varepsilon,\alpha})$. As in the proof of

Theorem 4, we observe that $t(H, W_{\varepsilon,\alpha})$ is a polynomial in α and the linear terms in α cancel out by pairing homomorphisms using u_0^+ and those using u_0^-. Hence, only quadratic and higher order terms remain. To estimate c_ε, we need to consider the terms corresponding to homomorphisms mapping exactly three vertices of H to the vertices of $G_{\varepsilon,\alpha}$ with weight ε and these vertices must induce a 2-edge path with the middle vertex mapped to u_0^+ or to u_0^- (the contribution of other homomorphisms cancels out by pairing those using u_0^+ and those using u_0^-, similarly as in the proof of Theorem 4). We arrive at the following identity.

$$\begin{aligned} c_\varepsilon \varepsilon^3 \quad = \quad & \sum_{v_0v_1, v_0v_2 \in E(H)} \sum_{i,i'=1}^{k} \sum_{j \in [n]^{|U_i|-1}} \sum_{j' \in [n]^{|U_{i'}|-1}} a_i a_{i'} \prod_{m=1}^{|U_i|-1} b_{j_m} \prod_{m=1}^{|U_{i'}|-1} b_{j'_m} \cdot \\ & \cdot \Big(\hom(H(v_0, v_1, v_2), G_{\varepsilon,0}(u_0^+, u_{i,j_1,\dots,j_{|U_i|-1}}, u_{i',j'_1,\dots,j'_{|U_{i'}|-1}})) + \\ & \hom(H(v_0, v_1, v_2), G_{\varepsilon,0}(u_0^-, u_{i,j_1,\dots,j_{|U_i|-1}}, u_{i',j'_1,\dots,j'_{|U_{i'}|-1}})) \Big) \end{aligned}$$

It follows that

$$\begin{aligned} \lim_{\varepsilon \to 0} c_\varepsilon = & \sum_{\substack{v_0v_1 \\ v_0v_2} \in E(H)} \sum_{i,i'=1}^{k} \sum_{h} 2 a_i a_{i'} w(h) \sum_{j \in [n]^{|U_i|-1}} \sum_{j' \in [n]^{|U_{i'}|-1}} \\ & \prod_{m=1}^{|U_i|-1} b_{j_m} 2^{(j_m-1)h(v_1 \hookrightarrow z_m)} \prod_{m=1}^{|U_{i'}|-1} b_{j'_m} 2^{(j'_m-1)h(v_2 \hookrightarrow z'_m)} \end{aligned}$$

where the sum is taken over all homomorphisms h from H to G such that $h(v_0) = u_0$, $h(v_1) = u_i$ and $h(v_2) = u_{i'}$, and $w(h)$ denotes the weight of the homomorphism h, $h(v_1 \hookrightarrow z_m)$ denotes the number of neighbors of v_1 mapped to $z_m \in U_i$ and $h(v_2 \hookrightarrow z'_m)$ denotes the number of neighbors of v_2 mapped to $z'_m \in U_{i'}$. Observe that b was chosen so that the expression

$$\sum_{j_1,\dots,j_{|U_i|-1}=1}^{n} \prod_{m=1}^{|U_i|-1} b_{j_m} 2^{(j_m-1)h(v_1 \hookrightarrow z_m)} = \prod_{m=1}^{|U_i|-1} \sum_{j_m=1}^{n} b_{j_m} 2^{(j_m-1)h(v_1 \hookrightarrow z_m)}$$

is one if $h(v_1 \hookrightarrow z_m) = 1$ and it is zero otherwise. Hence, it follows that

$$\lim_{\varepsilon \to 0} c_\varepsilon = \sum_{v_0v_1, v_0v_2 \in E(H)} \sum_{i,i'=1}^{k} \sum_{h} a_i a_{i'} w(h)$$

where the sum is taken over homomorphisms h from H to G such that $h(v_0) = u_0$, $h(v_1) = u_i$, $h(v_2) = u_{i'}$, all neighbors of v_1 are one-to-one mapped to U_i and all neighbors of v_2 are one-to-one mapped to $U_{i'}$. The definition of the matrix M now implies that

$$\lim_{\varepsilon \to 0} c_\varepsilon = \sum_{i,i'=1}^{k} M_{ii'} a_i a_{i'} = a^T M a < 0 \,. \tag{3.4}$$

The expressions (3.3) and (3.4) imply that there exist $\varepsilon > 0$ and $\alpha > 0$ such that $t(H, G_{\varepsilon,\alpha}) < t(H, G_{\varepsilon,0})$. Fix such ε and α for the rest of the proof.

Consider the graphons W_0 and W_α associated with the weighted graphs $G_{\varepsilon,0}$ and $G_{\varepsilon,\alpha}$, respectively. Let J_u be the measurable set corresponding to the vertex u of $G_{\varepsilon,0}$; we can assume that the measurable set corresponding to the vertex u of $G_{\varepsilon,\alpha}$ is also J_u. Let $\mathcal{P}$ be the partition of $[0, 1]$ formed by $J_{u_0^+} \cup J_{u_0^-}$ and J_u, $u \neq u_0^+, u_0^-$. Observe that $W_0 = W_\alpha^{\mathcal{P}}$. Since $t(H, W_0) = t(H, G_{\varepsilon,0})$ and $t(H, W_\alpha) = t(H, G_{\varepsilon,\alpha})$, we conclude that the graph H does not have the step Sidorenko property. □

Theorem 5 yields immediately the following corollary, which rules out many non-biregular graphs to be weakly norming. Note that the assumptions of the corollary are easy to verify.

Corollary 3. *Let H be a graph and $\mathcal{D}_H$ the set of degrees of its vertices. Further let M be the matrix with rows and columns indexed by the elements of $\mathcal{D}_H$ such that the entry $M_{dd'}$ is equal to the number of 2-edge paths from a vertex of degree d to a vertex of degree d' in H. If the matrix M is not positive semidefinite, then H does not have the step Sidorenko property.*

Proof. We can assume without loss of generality that H is bipartite; if not, H does not even have the Sidorenko property. Let $n = |H|$, let $d_1 < \cdots < d_k$ be the degrees of vertices of H, i.e., $\mathcal{D}_H = \{d_1, \ldots, d_k\}$, and let $D = d_1 + \cdots + d_k$. We next construct a weighted bipartite graph G_ε with weights depending on a parameter $\varepsilon > 0$. One part of G_ε has $k+1$ vertices, which are denoted by $u_1, \ldots, u_{k+1}$, and the other part has $D - k + 1$ vertices. One of the vertices of the second part is denoted by u_0 and the remaining $D - k$ vertices are split into disjoint sets $U_1, \ldots, U_k$ such that $|U_i| = d_i - 1$, $i = 1, \ldots, k$. The vertices u_0 and u_{k+1} have weight one, each of the vertices u_i has weight $\varepsilon^{\frac{1}{|U_i|}}$ and each vertex contained in a set U_i has weight $\varepsilon^{\frac{1}{|U_i|}}/(|U_i| - 1)!$, $i = 1, \ldots, k$. The weights of all edges of G_ε are equal to one.

We will apply Theorem 5 with the weighted graph G_ε, vertices $u_0, \ldots, u_k$ and sets $U_1 \cup \{u_0\}, \ldots, U_k \cup \{u_0\}$. Let M_ε be the matrix from the statement of Theorem 5 for the graph G_ε. Fix $i, j \in \{1, \ldots, k\}$ and a 2-edge path $v_1 v_0 v_2$ such that the degree of v_1 is d_i and the degree of v_2 is d_j. Let h be a mapping such that $h(v_0) = u_0$, $h(v_1) = u_i$ and $h(v_2) = u_j$. The mapping h can be extended to $(|U_i| - 1)!(|U_j| - 1)!$ homomorphisms from H to G such that

- the neighbors of v_1 are one-to-one mapped to $U_i \cup \{u_0\}$,
- the neighbors of v_2 are one-to-one mapped to $U_j \cup \{u_0\}$, and
- all other vertices of H are mapped to u_0 or to u_{k+1}.

Each such homomorphism has weight $\frac{\varepsilon^2}{(|U_i|-1)!(|U_j|-1)!}$, i.e., their total weight is ε^2. Any other extensions of h to a homomorphism from H to G such that the neighbors of v_1 are one-to-one mapped to $U_i \cup \{u_0\}$ and the neighbors of v_2 to $U_j \cup \{u_0\}$ has weight at most ε^{2+1/d_k}. We conclude that the entry of the matrix M_ε in the i-th row and the j-th column is equal to $M_{ij}\varepsilon^2 + O(\varepsilon^{2+1/d_k})$. It follows that there exists $\varepsilon > 0$ such that the matrix M_ε is not positive semidefinite. Theorem 5 now yields that H does not have the step Sidorenko property. □

The weights of vertices and edges of the graph G in Theorem 5 can be set to lower the weight of specific homomorphisms, as we did in Corollary 2. We first formalize the ideas used there, so that we can focus on just the existence of very restricted homomorphisms, without counting or weights.

Lemma 8. *Let H be a vertex-transitive graph. Let u_0, u_1 and u_2 be (distinct) distinguished vertices in H such that u_0u_1 and u_0u_2 are edges. Suppose that for each distinct neighbors v_1 and v_2 of u_0, there is no homomorphism f from $H(u_0, v_1, v_2)$ to $H(u_0, u_1, u_1)$ that simultaneously satisfies the following:*

- *neighbors of v_i are one-to-one mapped to neighbors of u_1 for $i = 1, 2$,*
- *distances from u_0 are preserved, i.e., $\mathrm{dist}(v, u_0) = \mathrm{dist}(f(v), u_0)$ for each $v \in V(H)$, and*
- *no vertex other than u_0, v_1 and v_2 is mapped to any of u_0, u_1 and u_2.*

Then H does not have the step Sidorenko property.

Proof. We start with constructing a weighted graph G_γ where the weights depend on a parameter $\gamma \in \mathbb{N}$. The graph G_γ is obtained from H by setting $w(v) := \gamma^{\mathrm{dist}(u_0,v)-1}$ for $v \in \{u_0, u_1, u_2\}$ and $w(v) := \gamma^{\mathrm{dist}(u_0,v)}$ for each vertex $v \neq u_0, u_1, u_2$. The weights of all edges of G_γ are one. We apply Theorem 5 to H and G_γ with the distinguished vertices u_0, u_1 and u_2. Since H is vertex-transitive, we will analyze the matrix M such that M_{ij} is the sum of weights of homomorphisms from $H(u_0, v_1, v_2)$ to

$G_\gamma(u_0, u_i, u_j)$ such that the neighbors of v_1 and v_2 are mapped one-to-one to the neighbors of u_i and u_j, respectively, where the sum runs over all choices of v_1 and v_2 in H. Note that the matrix from the statement Theorem 5 is the considered matrix M with each entry multiplied by $|G|$, in particular, it is enough to show that the considered matrix M is not positive semidefinite for some γ.

Let $W := \sum_{v\in V(H)} \mathrm{dist}(v, u_0) - 3$. We show that $M_{1,1} = o(\gamma^W)$, $M_{1,2} = M_{2,1} = \Omega(\gamma^W)$ and $M_{2,2} = O(\gamma^W)$ (as functions of the parameter γ). Hence, if γ is large enough, the matrix M is not positive semidefinite and H does not have the step Sidorenko property by Theorem 5.

By the definition, the entry $M_{1,2}$ contains a summand corresponding to the identity homomorphism from $H(u_0, v_1, v_2)$ to $G_\gamma(u_0, u_1, u_2)$; the weight of this summand is exactly γ^W. It follows $M_{1,2} = M_{2,1} = \Omega(\gamma^W)$.

Consider a homomorphism f contributing to the sum defining the entry $M_{i,i}$ for $i \in \{1, 2\}$. Observe that f satisfies $\left|f^{-1}(\{u_0, u_1, u_2\})\right| \geq 3$ (at least the three vertices u_0, v_1 and v_2 are mapped to u_0 and u_i) and $\mathrm{dist}(u_0, f(v)) \leq \mathrm{dist}(u_0, v)$ for every vertex v (a shortest walk from u_0 to v is mapped by f to a walk of at most the same length from u_0 to $f(v)$). Hence, it holds that $w(f(v)) \leq w(v)$ for every vertex v, and the equality holds for all vertices v if and only if $\mathrm{dist}(u_0, f(v)) = \mathrm{dist}(u_0, v)$ for every vertex v of H and $\left|f^{-1}(\{u_0, u_1, u_2\})\right| = 3$. In particular, the equality does not hold for any homomorphism f contributing to the sum defining the entry $M_{1,1}$. It follows that each summand in the sum defining the entry $M_{1,1}$ is of order $O(\gamma^{W-1})$ and each summand in the sum defining the entry $M_{2,2}$ is of order $O(\gamma^W)$. Since the number of the summands is independent of γ, we conclude that $M_{1,1} = o(\gamma^W)$ and $M_{2,2} = O(\gamma^W)$. □

We conclude by using Lemma 8 to show that all multidimensional grids other than hypercubes are not weakly norming.

Corollary 4. *Let $k \geq 2$. The Cartesian product $C_{\ell_1} \square \cdots \square C_{\ell_k}$ has the step Sidorenko property if and only if the length of each cycle in the product is four, i.e., $\ell_1 = \cdots = \ell_k = 4$.*

Proof. Let $H = C_{\ell_1} \square \cdots \square C_{\ell_k}$. By symmetry, we can assume that ℓ_1 is the largest and ℓ_2 is the smallest among $\ell_1, \ldots, \ell_k$. If $\ell_1 = \cdots = \ell_k = 4$, the graph H is isomorphic to the $2k$-dimensional hypercube graph, which is weakly norming, see [47] and [65, Proposition 14.2]; this implies implies that H has the step Sidorenko property [65, Proposition 14.13]. If ℓ_i is odd for some i, then the graph H is not bipartite, which implies that it fails to even have the Sidorenko property. Hence, we can assume that all ℓ_i are even and $\ell_1 > 4$.

We will view the vertices of H as the elements of $\mathbb{Z}_{\ell_1} \times \cdots \times \mathbb{Z}_{\ell_k}$ and perform all computations involving the i-th coordinate modulo ℓ_i. Let e_i be the i-th unit vector. Note that two vertices of H are adjacent if their difference is equal to e_i or $-e_i$ for some $i = 1, \ldots, k$. Also observe that if v is a vertex of H and $\ell_i > 4$, then v is the only common neighbor of $v + e_i$ and $v - e_i$.

We apply Lemma 8 with $u_0 = (0, \ldots, 0)$ and $u_i = e_i$ for $i = 1, 2$. Suppose that for some distinct vertices v_1 and v_2, there is a homomorphism f from $H(u_0, v_1, v_2)$ to $H(u_0, e_1, e_1)$ contradicting the assumption of Lemma 8, i.e.,

1. the neighbors of v_i are one-to-one mapped to neighbors of e_1, for $i = 1, 2$,
2. $\mathrm{dist}(u_0, v) = \mathrm{dist}(u_0, f(v))$ for each $v \in V(H)$, and
3. no vertex other than u_0, v_1 and v_2 is mapped to any of the vertices u_0, e_1 and e_2.

We will show that the existence of such a homomorphism f leads to a contradiction. By symmetry, we can assume that $v_1 = e_{i_1}$ for some i_1 and either $v_2 = -e_{i_1}$ or $v_2 = e_{i_2}$ for some $i_2 \neq i_1$.

Note that the neighbors of v_1 are one-to-one mapped to the neighbors of e_1, and let i' be such that $f(e_{i_1} + e_{i'}) = e_1 + e_1$. If $i' \neq i_1$, both common neighbors of u_0 and $e_{i_1} + e_{i'}$, which are e_{i_1} and $e_{i'}$, must be mapped to the unique common neighbor of u_0 and $e_1 + e_1$, which is the vertex e_1 (note that $\ell_1 > 4$). However, this would contradict 3. Hence, $i' = i_1$, i.e., $f(v_1 + v_1) = f(e_{i_1} + e_{i_1}) = e_1 + e_1$. It follows that there exists a bijection π between $\{\pm e_{i'} \mid i' \neq i_1\}$ and $\{\pm e_{j'} \mid j' \neq 1\}$ such that $f(e_{i_1} + e) = e_1 + \pi(e)$ for $e \in \{\pm e_{i'} \mid i' \neq i_1\}$. Observe that a symmetric argument to the one that we have just presented yields that $f(v_2 + v_2) = e_1 + e_1$.

To exclude the case that $v_2 = -e_{i_1}$, let $e = \pi^{-1}(e_2)$, i.e., $f(e_{i_1} + e) = e_1 + e_2$. Note that $e \neq \pm e_{i_1}$. It follows that the vertex e, which is a common neighbor of u_0 and $e_{i_1} + e$, must be mapped to a common neighbor of u_0 and $e_1 + e_2$, i.e., either to e_1 or to e_2. The first case would contradict 3, hence e is mapped to e_2, meaning $v_2 = e$. We conclude that $v_2 = e_{i_2}$ for some $i_2 \neq i_1$ and that $f(e_{i_1} + e_{i_2}) = e_1 + e_2$.

Suppose that $\ell_2 = 4$ and recall that $f(v_2 + v_2) = e_1 + e_1$. If additionally $\ell_{i_2} = 4$, then $-e_{i_2}$, which is a common neighbor of u_0 and $e_{i_2} + e_{i_2}$, must be mapped to the unique common neighbor of u_0 and $e_1 + e_1$, i.e., to the vertex e_1; this is impossible by 3. Hence, $\ell_{i_2} \neq 4$.

Let us call two vertices v and v' *close* if they have at least two common neighbors. Observe that two close distinct neighbors v and v' of e_{i_1} must be mapped to close neighbors of e_1; otherwise, all common neighbors of v and v' would be mapped to e_{i_1}, contradicting 3. Since the neighborhood of e_{i_1} is one-to-one mapped to the neighborhood of e_1 and the number of pairs of close neighbors of e_{i_1} is the same as the number of pairs of close neighbors of e_1, it follows that pairs of close neighbors of e_{i_1} are one-to-one mapped to pairs of close neighbors of e_1 and pairs of non-close neighbors of e_{i_1} are one-to-one mapped to pairs of non-close neighbors of e_1. Since $\ell_{i_2} \neq 4$, the neighbors $e_{i_1} + e_{i_2}$ and $e_{i_1} - e_{i_2}$ of e_{i_1} are not close. On the other hand, since $\ell_2 = 4$, the vertex $f(e_{i_1} + e_{i_2}) = e_1 + e_2$ has a common neighbor other than e_1 with each neighbor of e_1. In particular, $f(e_{i_1} + e_{i_2})$ and $f(e_{i_1} - e_{i_2})$ are close, which is impossible. We conclude that $\ell_2 \neq 4$. Since ℓ_2 is the smallest among $\ell_1, \ldots, \ell_k$, it follows that each ℓ_i is at least six.

As the final step of the proof of the corollary, we prove the following statement for $r = 1, \ldots, \ell_{i_1}/2$ by induction on r:

$$\begin{aligned} f((r-1)e_{i_1}) &= (r-1)e_1, \quad f(re_{i_1}) = re_1, \text{ and} \\ f(re_{i_1} + e) &= re_1 + \pi(e) \text{ for } e \in \{\pm e_{i'} \mid i' \neq i_1\}. \end{aligned} \tag{3.5}$$

The case $r = 1$ follows from the definition of i_1 and π. We assume that the above statement holds for r and prove it for $r+1 \leq \ell_{i_1}/2$. We first show that $f((r+1)e_{i_1}) = (r+1)e_1$. Note that $f(re_{i_1} + e_{i_1})$ cannot be $re_1 - e_1$ by 2. If $f(re_{i_1} + e_{i_1})$ is $re_1 + e_j$ for some $j \neq 1$, then the common neighbor $re_{i_1} + e_{i_1} + \pi^{-1}(-e_j)$ of $re_{i_1} + e_{i_1}$ and $re_{i_1} + \pi^{-1}(-e_j)$ must be mapped to the unique common neighbor of $re_1 + e_j$ and $re_1 - e_j$, which is re_1, contradicting 2. An analogous argument excludes that $f(re_{i_1} + e_{i_1})$ is $re_1 - e_j$ for some $j \neq 1$. Since the vertex $f((r+1)e_{i_1})$ must be a neighbor of $f(re_{i_1}) = re_1$, it follows that $f((r+1)e_{i_1}) = (r+1)e_1$.

We next analyze $f((r+1)e_{i_1} + e)$ for $e \neq \pm e_{i_1}$. Since the vertex $(r+1)e_{i_1} + e = re_{i_1} + e_{i_1} + e$ is a common neighbor of $re_{i_1} + e_{i_1}$ and $re_{i_1} + e$, it must be mapped to a common neighbor of $re_1 + e_1$ and $re_1 + \pi(e)$, i.e., to re_1 or $re_1 + e_1 + \pi(e)$. Since the former is excluded by 2, it follows that $f((r+1)e_{i_1} + e) = (r+1)e_1 + \pi(e)$. This concludes the proof of (3.5).

The statement (3.5) implies that $f(\ell_{i_1}/2 \cdot e_{i_1}) = \ell_{i_1}/2 \cdot e_1$, in particular $\ell_{i_1} \geq \ell_1$ by 2. Since the path $u_0, -e_{i_1}, -2e_{i_1}, \ldots, -\ell_{i_1}/2 \cdot e_{i_1}$ must be mapped to a path from u_0 to $f(-\ell_{i_1}/2 \cdot e_{i_1}) = f(\ell_{i_1}/2 \cdot e_{i_1}) = \ell_{i_1}/2 \cdot e_1$ and the vertices of the path must be mapped to vertices at distances $0, 1, \ldots, \ell_{i_1}/2$ from u_0 by 2, the path can be mapped only to the path $u_0, e_1, 2e_1, \ldots, \ell_{i_1}/2 \cdot e_1$ or, if $\ell_1 = \ell_{i_1}$, to the path $u_0, -e_1, -2e_1, \ldots, -\ell_{i_1}/2 \cdot e_1$ The former case is impossible since $-e_{i_1}$ cannot be mapped to e_1 by 3. It follows that $\ell_1 = \ell_{i_1}$ and $f(-e_{i_1}) = -e_1$. Hence, the vertex $e_{i_2} - e_{i_1} \neq u_0$, which is a common neighbor of e_{i_2} and $-e_{i_1}$, must be mapped to the unique common neighbor of $f(e_{i_2}) = e_1$ and $f(-e_{i_1}) = -e_1$, which is u_0. However, this contradicts 3. We conclude there is no homomorphism f satisfying 1–3. Lemma 8 now implies that H does not have the step Sidorenko property. $\square$

3.4 Further remarks

Corollary 2 and Corollary 4 give an infinite class of edge-transitive graphs that are not weakly norming, which answers in the negative a question of Hatami [47]. Conlon and Lee [19, Conjecture 6.3] present a large class of weakly norming graphs, which they call reflection graphs, and conjecture that a bipartite graph is weakly norming if and only if it is edge-transitive under a subgroup of its automorphism group (generated by so called 'cut involutions'). In particular, this would imply that all weakly norming graphs are edge-transitive.

Finally, it is natural to wonder about the Forcing Conjecture in the setting of the step Sidorenko property. Let us say that a graph H has the *step forcing property* if and only if

$$t(H, W^{\mathcal{P}}) \leq t(H, W)$$

for every graphon W and every partition $\mathcal{P}$ of $[0,1]$ into finitely many non-null measurable sets and the equality holds if and only if $W^{\mathcal{P}}$ and W are equal almost everywhere. All even cycles have the step forcing property. Graphs with the step forcing property are related to the proof of the existence of graphons via weak* limits given by Doležal and Hladký [25]; in particular, if H has the step forcing property, minimizing the entropy of W in the arguments given in [25] can be replaced by maximizing $t(H, W)$.

Chapter 4

Decomposing graphs into edges and triangles

Results on the existence of edge-disjoint copies of specific subgraphs in graphs are a classical theme in extremal graph theory. Motivated by the following result of Erdős, Goodman and Pósa [27], we study the problem of covering edges of a given graph by edge-disjoint complete graphs.

Theorem 6 (Erdős, Goodman and Pósa [27]). *The edges of every n-vertex graph can be decomposed into at most $\lfloor n^2/4 \rfloor$ complete graphs.*

In fact, they proved the following stronger statement.

Theorem 7 (Erdős, Goodman and Pósa [27]). *The edges of every n-vertex graph can be decomposed into at most $\lfloor n^2/4 \rfloor$ copies of K_2 and K_3.*

The bounds given in Theorems 6 and 7 are best possible as witnessed by complete bipartite graphs with parts of equal sizes.

Theorem 6 actually holds in a stronger form that we now present. Chung [14], Győri and Kostochka [45], and Kahn [53], independently, proved a conjecture of Katona and Tarján asserting that the edges of every n-vertex graph can be covered with complete graphs $C_1, \ldots, C_\ell$ such that the sum of their orders is at most $n^2/2$. In fact, the first two proofs yield a stronger statement, which implies Theorem 6 and which we next state as a separate theorem. To state the theorem, we define $\pi_k(G)$ for a graph G to be the minimum integer m such that the edges of G can be decomposed into complete graphs $C_1, \ldots, C_\ell$ of order at most k with $|C_1| + \cdots + |C_\ell| = m$, and we let $\pi(G) = \min_{k \in \mathbb{N}} \pi_k(G)$.

Theorem 8 (Chung [14]; Győri and Kostochka [45]). *Every n-vertex graph G satisfies $\pi(G) \leq n^2/2$.*

Observe that Theorem 8 indeed implies the existence of a decomposition into at most $\lfloor n^2/4 \rfloor$ complete graphs. McGuinnes [72, 73] extended these results by showing that decompositions from Theorems 6 and 8 can be constructed in the greedy way, which confirmed a conjecture of Winkler of this being the case in the setting of Theorem 6.

In view of Theorem 7, it is natural to ask whether Theorem 8 holds under the additional assumption that all complete graphs in the decomposition are copies of K_2 and K_3, i.e., whether $\pi_3(G) \leq n^2/2$. Győri and Tuza [46] provided a partial answer by proving that $\pi_3(G) \leq 9n^2/16$ and conjectured the following.

Conjecture 3 (Győri and Tuza [89, Problem 40])**.** *Every n-vertex graph G satisfies $\pi_3(G) \leq (1/2 + o(1))n^2$.*

We prove this conjecture. Our result also solves [89, Problem 41], which we state as Corollary 5. We remark that we stated the conjecture in the version given by Győri in several of his talks and by Tuza in [89, Problem 40]; the paper [46] contains a version with a different lower order term.

We would also like to mention a closely related variant of the problem suggested by Erdős, where the cliques in the decomposition have weights one less than their orders. Formally, define $\pi^-(G)$ for a graph to be the minimum m such that the edges of a graph G can be decomposed into complete graphs $C_1, \ldots, C_\ell$ with $(|C_1| - 1) + \cdots + (|C_\ell| - 1) = m$. Erdős asked, see [89, Problem 43], whether $\pi^-(G) \leq n^2/4$ for every n-vertex graph G. This problem remains open and was proved for K_4-free graphs only recently by Győri and Keszegh [43, 44]. Namely, they proved that every K_4-free graph with n vertices and $\lfloor n^2/4 \rfloor + k$ edges contains k edge-disjoint triangles.

4.1 Preliminaries

We follow the terminology presented in Chapter 1. We review here some less standard notation necessary for this chapter and briefly introduce the flag algebra method.

In our arguments, we often consider fractional decompositions. A fractional k-decomposition of a graph G is an assignment of non-negative real weights to complete subgraphs of order at most k such that the sum of the weights of the complete subgraphs containing any edge e is equal to one. The weight of a fractional k-decomposition is the sum of the weights of the complete subgraphs multiplied by their orders, and the minimum weight of a fractional k-decomposition of a graph G is denoted by $\pi_{k,f}(G)$. Observe that $\pi_{k,f}(G) \leq \pi_k(G)$ for every graph G.

4.1.1 Flag algebra method

The flag algebra method introduced by Razborov [77] has changed the landscape of extremal combinatorics. It has been applied to many long-standing open problems, e.g. [2–7, 22, 24, 28]. The method is designed to analyze asymptotic behavior of substructure densities and we now briefly describe it.

We start by introducing some necessary notation. The family of all finite graphs is denoted by $\mathcal{F}$ and the family of graphs with ℓ vertices by $\mathcal{F}_\ell$. If F and G are two graphs, then $p(F,G)$ is the probability that $|F|$ distinct vertices chosen uniformly at random among the vertices of G induce a graph isomorphic to F; if $|F| > |G|$, we set $p(F,G) = 0$. A type is a graph with its vertices labeled with $1,\ldots,|\sigma|$ and a σ-flag is a graph with $|\sigma|$ vertices labeled by $1,\ldots,|\sigma|$ such that the labeled vertices induce a copy of σ preserving the vertex labels. In the analogy with the notation for ordinary graphs, the set of all σ-flags is denoted by $\mathcal{F}^\sigma$ and the set of all σ-flags with exactly ℓ vertices by $\mathcal{F}^\sigma_\ell$.

We next extend the definition of $p(F,G)$ to σ-flags and generalize it to pairs of graphs. If F and G are two σ-flags, then $p(F,G)$ is the probability that $|F|-|\sigma|$ distinct vertices chosen uniformly at random among the unlabeled vertices of G induce a copy of the σ-flag F; if $|F| > |G|$, we again set $p(F,G) = 0$. Let F and F' be two σ-flags and G a σ-flag with at least $|F|+|F'|-|\sigma|$ vertices. The quantity $p(F,F';G)$ is the probability that two disjoint $|F|-|\sigma|$ and $|F'|-|\sigma|$ subsets of unlabeled vertices of G induce together with the labeled vertices of G the σ-flags F and F', respectively. It holds [77, Lemma 2.3] that

$$p(F,F';G) = p(F,G)\cdot p(F',G) + o(1) \tag{4.1}$$

where $o(1)$ tends to zero with $|G|$ tending to infinity.

Let $\vec{F} = [F_1,\ldots,F_t]$ be a vector of σ-flags, i.e., $F_i \in \mathcal{F}^\sigma$. If M is a $t\times t$ positive semidefinite matrix, it follows from (4.1), see [77], that

$$0 \le \sum_{i,j=1}^{t} M_{ij}p(F_i,G)p(F_j,G) = \sum_{i,j=1}^{t} M_{ij}p(F_i,F_j;G) + o(1). \tag{4.2}$$

The inequality (4.2) is usually applied to a large graph G with a randomly chosen labeled vertices in a way that we now describe. Fix σ-flags F and F' and a graph G. We now define a random variable $p(F,F';G^\sigma)$ as follows: label $|\sigma|$ vertices of G with $1,\ldots,|\sigma|$ and if the resulting graph G' is a σ-flag, then $p(F',F';G^\sigma) = p(F,F';G')$; if G' is not a σ-flag, then $p(F_i,F_j;G^\sigma) = 0$. The expected value of $p(F,F';G^\sigma)$

can be expressed as a linear combination of densities of $(|F|+|F'|-|\sigma|)$-vertex subgraphs of G [77], i.e., there exist coefficients α_H, $H \in \mathcal{F}_{|F|+|F'|-|\sigma|}$, such that

$$\mathbb{E}\, p(F, F'; G^\sigma) = \sum_{H \in \mathcal{F}_{|F|+|F'|-|\sigma|}} \alpha_H \cdot p(H, G) \tag{4.3}$$

for every graph G. It can be shown that $\alpha_H = \mathbb{E}\, p(F, F'; H^\sigma)$.

Let $\vec{F} = [F_1, \ldots, F_t]$ be a vector of ℓ-vertex σ-flags and let M be a $t \times t$ positive semidefinite matrix. The equality (4.3) yields that there exist coefficients α_H such that

$$\mathbb{E} \sum_{i,j=1}^{t} M_{ij} p(F_i, F_j; G^\sigma) = \sum_{H \in \mathcal{F}_{2\ell-|\sigma|}} \alpha_H \cdot p(H, G) \tag{4.4}$$

for every graph G, which combines with (4.2) to

$$0 \leq \sum_{H \in \mathcal{F}_{2\ell-|\sigma|}} \alpha_H \cdot p(H, G) + o(1) \tag{4.5}$$

for every graph G, where

$$\alpha_H = \sum_{i,j=1}^{t} M_{ij} \cdot \mathbb{E}\, p(F_i, F_j; H^\sigma)$$

In particular, the coefficients α_H depend only on the choice of $\vec{F}$ and M.

4.2 Main result

We start with proving the following lemma using the flag algebra method.

Lemma 9. *Let G be a weighted graph with all edges of weight one. It holds that*

$$\mathbb{E}_W \pi_{3,f}(G[W]) \leq 21 + o(1)$$

where W is a uniformly chosen random subset of seven vertices of G.

Proof. We use the flag algebra method to find coefficients c_U, $U \in \mathcal{F}_7$, such that

$$0 \leq \sum_{U \in \mathcal{F}_7} c_U \cdot p(U, G) + o(1) \tag{4.6}$$

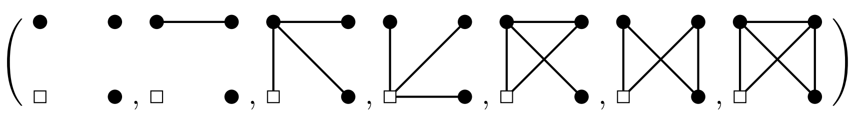
and

$$\pi_{3,f}(U) + c_U \leq 21 \tag{4.7}$$

for every $U \in \mathcal{F}_7$. The statement of the lemma would then follow from (4.6) and (4.7) using $\sum_{U \in \mathcal{F}_7} p(U,G) = 1$ as we next show.

$$\begin{aligned}
\mathbb{E}_W \pi_{3,f}(G[W]) &= \sum_{U \in \mathcal{F}_7} \pi_{3,f}(U) \cdot p(U,G) \\
&\leq \sum_{U \in \mathcal{F}_7} (\pi_{3,f}(U) + c_U) \cdot p(U,G) + o(1) \\
&\leq \sum_{U \in \mathcal{F}_7} 21 \cdot p(U,G) + o(1) = 21 + o(1).
\end{aligned}$$

We now focus on finding the coefficients c_U, $U \in \mathcal{F}_7$, satisfying (4.6) and (4.7). Let σ_1 be a flag consisting of a single vertex labeled with 1 and consider the following vector $\vec{F} = (F_1, \ldots, F_7)$ of σ_1-flags from $\mathcal{F}_4^{\sigma_1}$ (the single labeled vertex is depicted by a white square and the remaining vertices by black circles).

$\vec{F} =$ 

Let M be the following 7×7-matrix.

$$M = \frac{1}{12 \cdot 10^9} \begin{pmatrix}
1800000000 & 2444365956 & 640188285 & -1524146769 & 1386815580 & -732139362 & -129387078 \\
2444365956 & 4759879134 & 1177441152 & -1783771230 & 2546923788 & -1397639394 & -143552208 \\
640188285 & 1177441152 & 484273772 & -317303211 & 1038156300 & -591902130 & -6783162 \\
-1524146769 & -1783771230 & -317303211 & 1558870290 & -651906630 & 305728704 & 154602378 \\
1386815580 & 2546923788 & 1038156300 & -651906630 & 2285399634 & -1283125950 & -10755036 \\
-732139362 & -1397639394 & -591902130 & 305728704 & -1283125950 & 734039016 & -1621938 \\
-129387078 & -143552208 & -6783162 & 154602378 & -10755036 & -1621938 & 23860164
\end{pmatrix}.$$

The matrix M is a positive semidefinite matrix with rank six; the eigenvector corresponding to the zero eigenvalue is $(1, 0, 3, 1, 0, 3, 0)$. Let

$$c_U = \sum_{i,j=1}^{7} M_{ij} \mathbb{E}\, p(F_i, F_j; U^{\sigma_1}) \ .$$

The inequality (4.5) implies that

$$0 \leq \sum_{U \in \mathcal{F}_7} c_U \cdot p(U,G) + o(1),$$

which establishes (4.6). The inequality (4.7) is verified with computer assistance by evaluating the coefficient c_U and the quantity $\pi_{3,f}(U)$ for each $U \in \mathcal{F}_7$. Since

$|\mathcal{F}_7| = 1044$, we do not list c_U and $\pi_{3,f}(U)$ here. The computer programs that we used and their outputs have been made available on arXiv as ancillary files and are also available at `http://orion.math.iastate.edu/lidicky/pub/tile23`. □

The following lemma can be derived from the result of Haxell and Rödl [50] on fractional triangle decompositions or from a more general result of Yuster [93].

Lemma 10. *Let G be a graph with n vertices. It holds that $\pi_3(G) \leq \pi_{3,f}(G)+o(n^2)$.*

We now use Lemmas 9 and 10 to prove our main result.

Theorem 9. *Every n-vertex graph G satisfies $\pi_3(G) \leq (1/2 + o(1))n^2$.*

Proof. Fix an n-vertex graph G. By Lemma 10, it is enough to show that $\pi_{3,f}(G) \leq (1/2 + o(1))n^2$.

Fix an optimal fractional 3-decomposition of $G[W]$ for every 7-vertex subset $W \subseteq V(G)$, and set the weight $w(e)$ of an edge e to the sum of its weights in the optimal fractional 3-decomposition of $G[W]$ with $e \subseteq W$ multiplied by $\binom{n-2}{5}^{-1}$, and the weight $w(t)$ of a triangle t to the sum its weights in the optimal fractional 3-decomposition of $G[W]$ with $t \subseteq W$ also multiplied by $\binom{n-2}{5}^{-1}$. Since each edge e of G is contained in $\binom{n-2}{5}$ subsets W, we have obtained a fractional 3-decomposition of G. The weight of this decomposition is equal to

$$\frac{1}{\binom{n-2}{5}} \sum_{W \in \binom{V(G)}{7}} \pi_{3,f}(G[W]) \leq \frac{\binom{n}{7}}{\binom{n-2}{5}}(21 + o(1)) = n^2/2 + o(n^2) ,$$

where the inequality follows from Lemma 9. We conclude that $\pi_{3,f}(G) \leq n^2/2 + o(n^2)$, which completes the proof. □

The next corollary follows directly from Theorem 9.

Corollary 5. *Every n-vertex graph with $n^2/4 + k$ edges contains $2k/3 - o(n^2)$ edge-disjoint triangles.*

4.3 Alternative proof

In this section we present the original proof of Theorem 9 which combined the flag algebra method and regularity method arguments. In particular, we proved the fractional relaxation of Conjecture 3 in the setting of weighted graphs and with an additional restriction on its support; this statement was then combined with a blow-up lemma for edge-decompositions recently proved by Kim, Kühn, Osthus and

Tyomkyn [55]. It was then brought to our attention that the results from [50, 93] allow obtaining our main result directly from the fractional relaxation, which is the proof that we presented earlier and submitted to the journal. We believe that the argument combining the flag algebra method and the blow-up lemma of Kim et al. [55] can be of independent interest and so we present the original proof of our result and its idea here.

We start by reviewing some non-standard definitions necessary for the proof.

4.3.1 Designs

An (n, q, r, λ)-design is a collection $\mathcal{B}$ of q-element subsets of an n-element set such that every r-element subset is in exactly λ elements of $\mathcal{B}$. When λ is equal to one, the design is called a Steiner system. Designs do not exist for all choices of the parameters n, q, r and λ. In particular, the parameters must satisfy that $\binom{q-i}{r-i}$ divides $\lambda\binom{n-i}{r-i}$ for every $0 \leq i \leq r-1$. It was a long-standing open problem whether these necessary divisibility conditions are also sufficient for the existence of a design when n is large. The case where $r = 2$ was solved by Wilson in a series of papers [90–92] in the 1970's. However, the whole problem was settled only recently in a breakthrough paper by Keevash [54].

4.3.2 Regularity method

In this subsection, we review the basic notions related to the Szemerédi Regularity Lemma and the blow-up lemma for edge-decompositions of Kim, Kühn, Osthus and Tyomkyn [55].

We start with presenting three definitions that we use further in our exposition. Let G be a graph and V and W two disjoint subsets of its vertices. The *density* of the pair (V, W) is equal to

$$d(V, W) := \frac{e(V, W)}{|V||W|},$$

where $e(V, W)$ is the number of edges between V and W.

Let G be a graph, V and W two disjoint subsets of its vertices, and $\varepsilon \in (0, 1)$. We say that the pair (V, W) is *ε-regular* if the following holds for all subsets $V' \subset V$ and $W' \subset W$ with $|V'| \geq \varepsilon|V|$ and $|W'| \geq \varepsilon|W|$:

$$\left|d(V, W) - d(V', W')\right| \leq \varepsilon.$$

Let G be a graph, V and W two disjoint subsets of its vertices, and $\varepsilon \in (0, 1)$.

We say that the pair (V, W) is *ε-super-regular* if

- (V, W) is ε-regular,
- every vertex of V has at least $(d(V, W) - \varepsilon)|W|$ and at most $(d(V, W) + \varepsilon)|W|$ neighbors in W, and
- every vertex of W has at least $(d(V, W) - \varepsilon)|V|$ and at most $(d(V, W) + \varepsilon)|V|$ neighbors in V.

The Szemerédi Regularity Lemma reads as follows.

Lemma 11 (Regularity Lemma)**.** *For every real $\varepsilon > 0$ and integer $k_0 > 0$, there exists an integer K such that the vertices of every graph G with at least k_0 vertices can be partitioned into $k + 1$ subsets $V_0, \ldots, V_k$ where $k_0 \leq k \leq K$ such that*

- $|V_0| \leq \varepsilon|G|$,
- *the sets $V_1, \ldots, V_k$ have the same size, and*
- *all but at most εk^2 pairs (V_i, V_j) are ε-regular.*

Any partition $V_0, \ldots, V_k$ with the three properties given in Lemma 11 is called an ε-regular partition.

Let G be a graph and $V_0, \ldots, V_k$ an ε-regular partition. The *regularity graph* R_G with respect to the partition $V_0, \ldots, V_k$ is the graph with k vertices such that the i-th and the j-th vertex, $1 \leq i, j \leq k$, are adjacent if and only if (V_i, V_j) is an ε-regular pair.

The following result was proven by Kim, Kühn, Osthus and Tyomkyn [55, Theorem 1.3]; we state the result in a version for non-spanning subgraphs, which is equivalent to the original statement.

Theorem 10. *For all $0 < d_0, \alpha_0 \leq 1$ and $\Delta, r \in \mathbb{N}$ there exist $\varepsilon_0 > 0$ and $n_0 \in \mathbb{N}$ such that the following holds for all $n \geq n_0$. Let $H_1, \ldots, H_s$ be r-partite graphs such that each of them has r parts, each of size at most n, and its maximum degree is at most Δ. If G is an r-partite graph with parts of sizes n such that every pair of its parts is ε_0-super-regular with density at least d_0, and $||H_1|| + \cdots + ||H_s|| \leq (1 - \alpha_0)||G||$, then G contains edge-disjoint copies of $H_1, \ldots, H_s$.*

The following proposition is a direct corollary of Theorem 10.

Proposition 4. *For every $\alpha \in (0, 1)$ and every $d \in (0, 1]$, there exists $\varepsilon > 0$ and $N \in \mathbb{N}$ with the following property. If G is a graph and V_1, V_2 and V_3 disjoint*

n-vertex subsets of its vertices, $n \geq N$, such that (V_i, V_j) is an ε-regular pair with density at least d for $1 \leq i < j \leq 3$, then G contains at least $dn^2 - \alpha n^2$ edge-disjoint triangles with one vertex in V_1, one in V_2 and one in V_3.

Proof. Let $\varepsilon = \varepsilon_0/3$ and $N = \lceil n_0/(1-2\varepsilon) \rceil$, where ε_0 and n_0 are the values from Theorem 10 applied with $r = 3$, $\Delta = 2$, $d_0 = d/2$ and $\alpha_0 = \alpha/4$. We can assume that $\varepsilon \leq \frac{\alpha}{8}$, $d - 4\varepsilon \geq d_0$ and $n_0 \geq 4/\alpha$.

For $i = 1, \ldots, 3$, let V_i' be the set of all vertices $v \in V_i$ such that v has at least $(d(V_i, V_j) - \varepsilon)|V_j|$ and at most $(d(V_i, V_j) + \varepsilon)|V_j|$ neighbors in V_j, $j \neq i$. Since all the pairs (V_i, V_j) are ε-regular, it follows that $|V_i'| \geq (1 - 2\varepsilon)|V_i|$. Let V_i'' be any $\lceil (1-2\varepsilon)n \rceil$-element subset of V_i'.

Let G' be the subgraph of G with the vertex set $V_1'' \cup V_2'' \cup V_3''$ and all edges between V_i'' and V_j'' with $i \neq j$. Note that every pair (V_i'', V_j'') is ε_0-super-regular with density at least $d - 4\varepsilon$. Set $H_i = K_3$, where $i = 1, \ldots, s$ and

$$s = \lceil (d - 4\varepsilon - \alpha/2)n^2 \rceil \leq (d - 4\varepsilon - \alpha/2)n^2 + 1 \leq (d - 4\varepsilon - \alpha_0)n^2 \ .$$

Theorem 10 now implies that G' has at least $s \geq (d - 4\varepsilon - \alpha/2)n^2 \geq (d - \alpha)n^2$ edge-disjoint triangles. $\square$

4.3.3 Main result

We start by proving two auxiliary results: the first one, a fractional version of Conjecture 3 with and additional restriction; and the second one, a simple application of the probabilistic method which we include it for completeness.

Theorem 11. *Every n-vertex weighted graph G has a fractional 3-decomposition of weight at most $n^2/2 + o(n^2)$ such that each edge is contained in at most five triangles with positive weight.*

Proof. We can assume $\binom{7}{2}$ divides $\binom{n}{2}$ and 6 divides $n - 1$ (if this were not the case, we would just add at most 42 isolated vertices to G). It follows that there exists $(n, 7, 2, 1)$-design. Let m be the number of edges of G and $d_1, \ldots, d_m$ their weights in the non-decreasing order; set $d_0 = 0$. Let G_i, $1 \leq i \leq m$, be the spanning unweighted subgraph of G formed exactly by the edges of weight at least d_i.

We construct a fractional 3-decomposition of G using the following random procedure. We first choose a $(n, 7, 2, 1)$-design $\mathcal{B}$ uniformly at random among all $(n, 7, 2, 1)$-designs on the vertex set of G; it follows that every 7-vertex subset is included in $\mathcal{B}$ with the same probability, which is equal to $\frac{n(n-1)}{42} \cdot \binom{n}{7}^{-1}$. Note that each pair of vertices of G is included in exactly one set contained in $\mathcal{B}$.

Fix an optimal fractional 3-decomposition of $G_i[B]$ for every subset B in $\mathcal{B}$ and every $i = 1, \ldots, m$. For every edge e of the graph G, we consider the unique subset of B containing both end vertices of e and define $w_i(e)$, $i = 1, \ldots, m$, to be the weight of e in the fractional 3-decomposition of $G_i[B]$ if the weight of e in G is at least d_i and to be zero otherwise. We next define weights $w_i(t)$ for each triangle t of the graph G. If there is a subset B in $\mathcal{B}$ containing all the three end vertices of t and the weights of all three edges of t are at least d_i, $i = 1, \ldots, m$, then $w_i(t)$ is the weight of t in the fractional 3-decomposition of $G_i[B]$. Otherwise, $w_i(t)$ is equal to zero.

We set the weight $w(e)$ of an edge e of G to be

$$w(e) = \sum_{i=1}^{m} (d_i - d_{i-1}) w_i(e)$$

and the weight $w(t)$ of a triangle of G to be

$$w(t) = \sum_{i=1}^{m} (d_i - d_{i-1}) w_i(t) \; .$$

The definition of the graphs G_i yield that w is a fractional 3-decomposition of G. Moreover, if $w(t) > 0$ for a triangle t of G, then all the three vertices of t lie in the common subset B in $\mathcal{B}$. In particular, each edge of G is contained in at most five triangles of positive weight.

We now show that the expected weight of the fractional 3-decomposition w is at most $n^2/2 + o(n^2)$. We use that every 7-vertex subset of vertices is included in $\mathcal{B}$ with the same probability, which implies that

$$\mathbb{E} \sum_{e} 2w(e) + \mathbb{E} \sum_{t} 3w(t) = \sum_{i=1}^{m} (d_i - d_{i-1}) \frac{n(n-1)}{42} \, \mathbb{E}_U \pi_{3,f}(G_i[U]), \qquad (4.8)$$

where U is a uniform random subset of seven vertices of G. We next use Lemma 9 to derive the following from (4.8).

$$\begin{aligned} \mathbb{E} \sum_{e} 2w(e) + \mathbb{E} \sum_{t} 3w(t) &\leq \sum_{i=1}^{m} (d_i - d_{i-1}) \frac{n(n-1)}{42} (21 + o(1)) \\ &= \sum_{i=1}^{m} (d_i - d_{i-1}) \frac{n^2}{2} + o(n^2) \\ &= (d_m - d_0) \left(\frac{n^2}{2} + o(n^2) \right) \leq \frac{n^2}{2} + o(n^2) \; . \end{aligned}$$

Hence, the expected weight of the fractional 3-decomposition w is at most $n^2/2 + o(n^2)$. □

Lemma 12. *For every integer $r \in \mathbb{N}$ and reals $\varepsilon \in (0, 1/4)$ and $\delta \in (0, 1)$, there exists n_0 such that the following holds. For every graph G, every ε-regular pair (V, W) of vertices of G with $|V| = |W| \geq n_0$, and all non-negative reals $d_1, \dots, d_r$ such that $d_1 + \cdots + d_r \leq d(V, W)$, there exists a partition $E_1, \dots, E_r$ of the edges between V and W such that the pair (V, W) when restricted to the edges in E_i, $i = 1, \dots, r$, is an 3ε-regular with density at least $d_i - \delta$.*

We use the Chernoff Bound to prove the lemma, which we now state for reference.

Proposition 5 (Chernoff Bound)**.** *Let X be the sum of n independent random zero-one variables, each being one with probability p. It holds*

$$\mathbb{P}[|X - pn| \geq a] < 2e^{-\frac{a^2}{3pn}}$$

for every real $a \in \mathbb{R}$.

We are now ready to prove Lemma 12.

Proof of Lemma 12. Fix r, ε and δ, and consider a graph G together with an ε-regular pair (V, W) and reals $d_1, \dots, d_r$ as in the statement of the lemma. We can assume without loss of generality that $d_1 + \cdots + d_r = d(V, W)$ and that $\delta \leq \varepsilon$. Also let $n = |V| = |W|$.

We randomly partition the edges between V and W into sets $E_1, \dots, E_r$ in such a way that each edge is included in E_i with probability $p_i = \frac{d_i}{d(V,W)}$ independently of the other edges. The probability that E_i contains fewer than $(d_i - \delta)n^2$ edges or more than $(d_i + \delta)n^2$ edges is at most

$$2e^{-\frac{\delta^2 n^4}{3p_i n^2}} \leq 2e^{-\frac{\delta^2 n^2}{3}} \tag{4.9}$$

by Proposition 5. Next consider subsets $V' \subseteq V$ and $W' \subseteq W$ with $|V'|, |W'| \geq 3\varepsilon n$. The probability that the density of the pair (V', W') restricted to E_i differs from $p_i d(V', W')$ by more than ε is at most

$$2e^{-\frac{\varepsilon^2 |V'|^2 |W'|^2}{3p_i d(V',W')|V'||W'|}} \leq 2e^{-\frac{\varepsilon^2 |V'||W'|}{3}} \leq 2e^{-3\varepsilon^4 n^2} \tag{4.10}$$

by Proposition 5. Since the pair (V, W) is ε-regular, it holds that $|d(V, W) - d(V', W')| \leq \varepsilon$. It follows that the probability that the density of the pair (V', W')

restricted to E_i differs from d_i by more than 2ε is at most $2e^{-3\varepsilon^4 n^2}$. The union bound applied with the estimate (4.10) yields that the probability that there exist such subsets V' and W' for some i is at most

$$r \cdot 2^{2n+1} \cdot e^{-3\varepsilon^3 n^2} \ . \tag{4.11}$$

We now choose n_0 such that each of the estimates (4.9) and (4.11) is at most $1/2r$ for every $n \geq n_0$. Hence, there is a positive probability that every E_i, $i = 1, \ldots, r$, contains between $(d_i - \delta)n^2$ and $(d_i + \delta)n^2$ edges (inclusively), i.e., the density of (V, W) restricted to E_i is between $d_i - \delta$ and $d_i + \delta$, and that all subsets $V' \subseteq V$ and $W' \subseteq W$, $|V'|, |W'| \geq 3\varepsilon n$, satisfy that the density of the pair (V', W') restricted to E_i differs from d_i by at most 2ε. Since such a partition satisfies that the pair (V, W) restricted to E_i is 3ε-regular (we use that $\delta \leq \varepsilon$) for every $i = 1, \ldots, r$, the statement of the lemma follows. □

We are now ready to prove the main result of the paper.

Theorem 12. *Every n-vertex graph G satisfies that $\pi_3(G) \leq (1/2 + o(1))n^2$.*

Proof. We show that for every $\delta > 0$, there exists N such that $\pi_3(G) \leq n^2/2 + \delta n^2$ for every graph G with $n \geq N$ vertices. Fix $\delta > 0$. We can assume without loss of generality that δ^{-1} is an integer.

Let ε_a and N_a be the values of ε and N from Proposition 4 applied for $\alpha = \delta/20$ and $d = a\delta/20$ where $a = 1, \ldots, 20\delta^{-1}$. Next set

$$\varepsilon = \min\{\delta/20, \varepsilon_1/3, \ldots, \varepsilon_{20\delta^{-1}}/3\} \ .$$

Let n_f be such that the $o(n^2)$ term in Theorem 11 is at most $\delta n^2/20$ for all $n \geq n_f$. We apply the Szemerédi Regularity Lemma (Lemma 11) with ε and $k_0 = \max\{20\delta^{-1}, n_f\}$ to get an integer K and Lemma 12 with $r = 6$, ε and $\delta/20$ to get an integer n_0, and set N to be any integer larger than $n_0 K(1-\varepsilon)^{-1}$ and larger than $N_a K(1-\varepsilon)^{-1}$ for $a = 1, \ldots, 20\delta^{-1}$.

Let G be a graph with $n \geq N$ vertices. By the Szemerédi Regularity Lemma, there exists an ε-regular partition $V_0, \ldots, V_k$ of the vertex set of G, where $k_0 \leq k \leq K$. Let R_G be the regularity graph with respect to the partition $V_0, \ldots, V_k$ and let v_i be the vertex of R_G corresponding to the part V_i, $i = 1, \ldots, k$. If (V_i, V_j) is ε-regular, assign the edge joining $v_i v_j$ the weight equal to $d(V_i, V_j)$.

By Theorem 11, the graph R_G has a fractional 3-packing of total weight at most $k^2/2 + \delta k^2/20$ (since $k \geq n_f$). Fix such a fractional 3-packing, let $w(t)$ be the weight of a triangle t of R_G in the packing and $w(e)$ the weight of an edge e.

Consider an edge v_iv_j of R_G. By Theorem 11, there are at most five triangles t containing v_iv_j with $w(t) > 0$. Lemma 12 yields that there exist disjoint subsets E^t_{ij} of the edges between V_i and V_j, where t ranges through the at most five triangles containing v_iv_j with $w(t) > 0$, such that E^t_{ij} contains at least $(w(t) - \delta/20)|V_i||V_j|$ edges and the pair (V_i, V_j) restricted to E^t_{ij} is 3ε-regular. Fix such sets E^t_{ij} for all ε-regular pairs (V_i, V_j).

Let n_V be the number of vertices contained in each of the parts $V_1, \ldots, V_k$; note that $n_V \geq n_0$ by the choice of N. For every triangle $t = v_iv_{i'}v_{i''}$ with $w(t) > 0$, we construct a large family of edge-disjoint triangles with edges from $E^t_{ii'}$, $E^t_{ii''}$ and $E^t_{i'i''}$. Let a be the largest integer such that $w(t) \geq (a+1)\delta/20$. Note that $n_V \geq N_a$ and that each of the sets $E^t_{ii'}$, $E^t_{ii''}$ and $E^t_{i'i''}$ has density at least $a\delta/20$ between the corresponding vertex parts. We apply Proposition 4 for the sets V_i, $V_{i'}$ and $V_{i''}$ with edges from $E^t_{ii'}$, $E^t_{ii''}$ and $E^t_{i'i''}$ and with $\alpha = \delta/20$ and $d = a\delta/20$. This yields a family of at least $dn_V^2 - \alpha n_V^2 \geq (w(t) - \delta/10)n_V^2$ edge-disjoint triangles with edges from $E^t_{ii'}$, $E^t_{ii''}$ and $E^t_{i'i''}$. Consider such a family of at least $(w(t) - \delta/10)n_V^2$ and at most $w(t)n_V^2$ triangles for each triangle t with $w(t) > 0$ and let $\mathcal{T}$ be the union of all such families for t with $w(t) > 0$. Note that the number of triangles contained in $\mathcal{T}$ is at most

$$\sum_t w(t)n_V^2 \leq \frac{n^2}{k^2} \sum_t w(t) \; . \tag{4.12}$$

Since each edge v_iv_j of R_G is contained in at most five triangles with positive weight, we obtain that if (V_i, V_j) is an ε-regular pair, then the triangles contained in $\mathcal{T}$ cover all but at most $(w(v_i, v_j) + \delta/2)n_V^2$ edges between V_i and V_j.

We next estimate the number of edges that are not between (V_i, V_j) forming an ε-regular pair. There are three kinds of such edges: those incident with a vertex from V_0, those with both end vertices inside V_i for some $i = 1, \ldots, k$ and those between parts V_i and V_j, $1 \leq i < j \leq k$, such that (V_i, V_j) is not ε-regular. The number of edges incident with a vertex from V_0 is at most

$$|V_0|n \leq \varepsilon n^2 \leq \delta n^2/20 \; . \tag{4.13}$$

The number of edges with both end vertices inside the same part V_i for some $i = 1, \ldots, k$ is at most

$$k\binom{n_V}{2} \leq \frac{n^2}{2k} \leq \frac{n^2}{2k_0} \leq \delta n^2/40 \; . \tag{4.14}$$

Finally, the number of edges between parts V_i and V_j, $1 \leq i < j \leq k$, such that

(V_i, V_j) is not ε-regular is at most

$$\varepsilon k^2 n_V^2 \le \varepsilon n^2 \le \delta n^2/20 \; . \tag{4.15}$$

Using (4.13), (4.14) and (4.15), we conclude that the number of edges not contained in a triangle in $\mathcal{T}$ is at most

$$\begin{aligned} \frac{5\delta n^2}{40} + \sum_e (w(v_i, v_j) + \delta/2) n_V^2 &\le \frac{\delta n^2}{8} + \frac{\delta n^2}{4} + \frac{n^2}{k^2} \sum_e w(v_i, v_j) \\ &= \frac{3\delta n^2}{8} + \frac{n^2}{k^2} \sum_e w(v_i, v_j) \; . \end{aligned} \tag{4.16}$$

Since the total weight of the fractional 3-packing of R_G is at most $k^2/2 + \delta k^2/20$, we get from (4.12) and (4.16) that the triangles from $\mathcal{T}$ and the edges not covered by $\mathcal{T}$ (viewed as complete graphs of order two) form a 3-packing of G of total weight at most

$$\frac{3\delta n^2}{4} + \frac{n^2}{k^2} \left(\sum_e 2w(v_i, v_j) + \sum_t 3w(t) \right) \le \frac{3\delta n^2}{4} + \frac{n^2}{k^2} \left(\frac{k^2}{2} + \delta \frac{k^2}{20} \right) \le \frac{n^2}{2} + \delta n^2 \; .$$

The proof of the theorem is now finished. □

4.4 Further remarks

We tried to prove Lemma 9 in the non-fractional setting, i.e., to show that $\mathbb{E}_W \pi_3(G[W]) \le 21 + o(1)$. Unfortunately, the computation with 7-vertex flags yields only that $\mathbb{E}_W \pi_3(G[W]) \le 21.588 + o(1)$. We would like to remark that if it were possible to prove Lemma 9 in the non-fractional setting, we would be able to prove Theorem 9 without using additional results as a blackbox: we would consider a random $(n, 7, 2, 1)$-design on the vertex set of an n-vertex graph G as in the alternative proof in Section 4.3 and apply the non-fractional version of Lemma 9 to this design.

Finally, we would also like to mention two open problems related to our main result. Theorem 9 asserts that $\pi_3(G) \le n^2/2 + o(n^2)$ for every n-vertex graph G. However, it could be true (cf. the remark after Problem 41 in [89]) that $\pi_3(G) \le n^2/2 + 2$ for every n-vertex graph G. The second problem that we would like to mention is a possible generalization of Corollary 5, which is stated in [89] as Problem 42. Fix $r \ge 4$. Does every n-vertex graph with $\frac{r-2}{2r-2} n^2 + k$ edges contain $\frac{2}{r} k - o(n^2)$ edge-disjoint complete graphs of order r?

Bibliography

[1] R. Baber: *Turán densities of hypercubes*, preprint arXiv:1201.3587 (2012).

[2] R. Baber and J. Talbot: *Hypergraphs do jump*, Combin. Probab. Comput. **20** (2011), 161–171.

[3] R. Baber and J. Talbot: *A solution to the 2/3 conjecture*, SIAM J. Discrete Math. **28** (2014), 756–766.

[4] J. Balogh, P. Hu, B. Lidický and H. Liu: *Upper bounds on the size of 4- and 6-cycle-free subgraphs of the hypercube*, European J. Combin. **35** (2014), 75–85.

[5] J. Balogh, P. Hu, B. Lidický and F. Pfender: *Maximum density of induced 5-cycle is achieved by an iterated blow-up of 5-cycle*, European J. Combin. **52** (2016), 47–58.

[6] J. Balogh, P. Hu, B. Lidický, F. Pfender, J. Volec and M. Young: *Rainbow triangles in three-colored graphs*, J. Combin. Theory Ser. B **126** (2017), 83–113.

[7] J. Balogh, P. Hu, B. Lidický, O. Pikhurko, B. Udvari and J. Volec: *Minimum number of monotone subsequences of length 4 in permutations*, Combin. Probab. Comput. **24** (2015), 658–679.

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

[9] B. Bollobás: *Relations between sets of complete subgraphs*, Proc. 5th British Comb. Conf. (ed. C.St.J.A. Nash-Williams, J. Sheehan), Utilitas Math. (1976), 79–84.

[10] C. Borgs, J. Chayes and L. Lovász: *Moments of two-variable functions and the uniqueness of graph limits*, Geom. Funct. Anal. **19** (2010), 1597–1619.

[11] C. Borgs, J. Chayes, L. Lovász, V. Sós and K. Vesztergombi: *Graph limits and parameter testing*, Proceedings of the 38rd Annual ACM Symposium on the Theory of Computing (STOC) (2006), 261–270.

[12] C. Borgs, J. Chayes, L. Lovász, V. Sós and K. Vesztergombi: *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.

[13] C. Borgs, J. Chayes, L. Lovász, V. Sós and K. Vesztergombi: *Convergent sequences of dense graphs II. Multiway cuts and statistical physics*, Ann. of Math. **176** (2012), 151–219.

[14] F. R. K. Chung: *On the decomposition of graphs*, SIAM J. Algebraic Discrete Methods **2** (1981), 1–12.

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

[16] D. Conlon and J. Fox: *Bounds for graph regularity and removal lemmas*, Geom. Funct. Anal. **22** (2012), 1191–1256.

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

[18] D. Conlon, J. H. Kim, C. Lee and J. Lee: *Some advances on Sidorenko's conjecture*, to appear in J. Lond. Math. Soc. .

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

[20] J. W. Cooper, T. Kaiser, D. Král' and J. A. Noel: *Weak regularity and finitely forcible graph limits*, to appear in Trans. Amer. Math. Soc. .

[21] J. W. Cooper, D. Král' and T. L. Martins: *Finitely forcible graph limits are universal*, to appear in Adv. Math. .

[22] L. N. Coregliano and A. A. Razborov: *On the density of transitive tournaments*, J. Graph Theory **85** (2017), 12–21.

[23] J. Cummings, D. Král', F. Pfender, K. Sperfeld, A. Treglown and M. Young: *Monochromatic triangles in three-coloured graphs*, J. Combin. Theory Ser. B **103** (2013), 489–503.

[24] S. Das, H. Huang, J. Ma, H. Naves and B. Sudakov: *A problem of Erdős on the minimum number of k-cliques*, J. Combin. Theory Ser. B **103** (2013), 344–373.

[25] M. Doležal and J. Hladký: *Cut-norm and entropy minimization over weak* limits*, preprint arXiv:1701.09160 (2017).

[26] M. Doležal, J. Grebík, J. Hladký, I. Rocha and V. Rozhoň: *Cut distance identifying graphon parameters over weak* limits*, preprint arXiv:1809.03797 (2018).

[27] P. Erdős, A. W. Goodman and L. Pósa: *The representation of a graph by set intersections*, Canad. J. Math. **18** (1966), 106–112.

[28] C. Even-Zohar and N. Linial: *A note on the inducibility of 4-vertex graphs*, Graphs Combin. **31** (2015), 1367–1380.

[29] V. Falgas-Ravry, E. Marchant, O. Pikhurko and E. R. Vaughan: *The codegree threshold for 3-graphs with independent neighborhoods*, SIAM J. Discrete Math. **29** (2015), 1504–1539.

[30] D. Fisher: *Lower bounds on the number of triangles in a graph*, J. Graph Theory **13** (1989), 505–512.

[31] J. Fox and F. Wei: *On the local approach to Sidorenko's conjecture*, in: Electronic Notes in Discrete Mathematics **61** (2017), 459–465 (Proceedings of EuroComb'17).

[32] F. Franek and V. Rödl: *Ramsey problem on multiplicities of complete subgraphs in nearly quasirandom graphs*, Graphs Combin. **8** (1992), 299–308.

[33] F. Franek and V. Rödl: 2*-colorings of complete graphs with a small number of monochromatic K_4 subgraphs*, Discrete Math. **114** (1993), 199–203. Combinatorics and algorithms (Jerusalem, 1988).

[34] A. Frieze and R. Kannan: *Quick approximation to matrices and applications*, Combinatorica **19**, 175–220.

[35] E. Gethner, L. Hogben, B. Lidický, F. Pfender, A. Ruiz and M. Young: *On crossing numbers of complete tripartite and balanced complete multipartite graphs*, J. Graph Theory **84** (2017), 552–565.

[36] R. Glebov, T. Klimošová and D. Král': *Infinite dimensional finitely forcible graphon*, to appear in Proceedings of the LMS .

[37] R. Glebov, D. Král' and J. Volec: *Compactness and finite forcibility of graphons*, preprint arXiv:1309.6695 (2013).

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

[39] X. Goaoc, A. Hubard, R. de Joannis de Verclos, J.-S. Sereni and J. Volec: *Limits of order types*, in: 31st International Symposium on Computational Geometry, *LIPIcs. Leibniz Int. Proc. Inform.*, volume 34, 300–314.

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

[41] A. Grzesik: *On the maximum number of five-cycles in a triangle-free graph*, J. Combin. Theory Ser. B **102** (2012), 1061–1066.

[42] A. Grzesik, D. Král' and L. M. Lovász: *Elusive extremal graphs*, preprint arXiv:1807.01141 (2018).

[43] E. Győri and B. Keszegh: *On the number of edge-disjoint triangles in K_4-free graphs*, preprint arXiv:1506.03306 (2015).

[44] E. Győri and B. Keszegh: *On the number of edge-disjoint triangles in K_4-free graphs*, Electronic Notes in Discrete Mathematics **61** (2017), 557–560.

[45] E. Győri and A. V. Kostochka: *On a problem of G. O. H. Katona and T. Tarján*, Acta Math. Acad. Sci. Hungar. **34** (1979), 321–327 (1980).

[46] E. Győri and Z. Tuza: *Decompositions of graphs into complete subgraphs of given order*, Studia Sci. Math. Hungar. **22** (1987), 315–320.

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

[48] H. Hatami, J. Hladký, D. Král', S. Norine and A. Razborov: *Non-three-colourable common graphs exist*, Combin. Probab. Comput. **21** (2012), 734–742.

[49] H. Hatami, J. Hladký, D. Král', S. Norine and A. Razborov: *On the number of pentagons in triangle-free graphs*, J. Combin. Theory Ser. A **120** (2013), 722–732.

[50] P. E. Haxell and V. Rödl: *Integer and fractional packings in dense graphs*, Combinatorica **21** (2001), 13–38.

[51] J. Hladký, D. Král' and S. Norin: *Counting flags in triangle-free digraphs*, Combinatorica **37** (2017), 49–76.

[52] C. Jagger, P. Šťovíček and A. Thomason: *Multiplicities of subgraphs*, Combinatorica **16** (1996), 123–141.

[53] J. Kahn: *Proof of a conjecture of Katona and Tarján*, Period. Math. Hungar. **12** (1981), 81–82.

[54] P. Keevash: *The existence of designs*, preprint arXiv:1401.3665 (2014).

[55] J. Kim, D. Kühn, D. Osthus and M. Tyomkyn: *A blow-up lemma for approximate decompositions*, preprint arXiv:1604.07282 (2016).

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

[57] D. Král', B. Lidický, T. L. Martins and Y. Pehova: *Decomposing graphs into edges and triangles*, to appear in Combin. Probab. Comput. .

[58] D. Král', C. Liu, J. Sereni, P. Whalen and Z. Yilma: *A new bound for the* 2/3 *conjecture*, Combin. Probab. Comput. **22** (2013), 384–393.

[59] D. Král', L. M. Lovász, J. Noel and J. Sosnovec: *Finitely forcible graphons with an almost arbitrary structure*, in preparation.

[60] D. Král', L. Mach and J.-S. Sereni: *A new lower bound based on Gromov's method of selecting heavily covered points*, Discrete Comput. Geom. **48** (2012), 487–498.

[61] D. Král', T. L. Martins, P. P. Pach and M. Wrochna: *Step Sidorenko property and non-norming edge-transitive graphs*, to appear in J. Combin. Theory Ser. A .

[62] D. Král' and O. Pikhurko: *Quasirandom permutations are characterized by 4-point densities*, Geom. Funct. Anal. **23** (2013), 570–579.

[63] J. L. X. Li and B. Szegedy: *On the logarithimic calculus and Sidorenko's conjecture*, to appear in Combinatorica .

[64] B. Lidický and F. Pfender: *Semidefinite programming and Ramsey numbers*, preprint arXiv:1704.03592 (2017).

[65] L. Lovász: Large Networks and Graph Limits, *Colloquium Publications*, volume 60 (2012).

[66] L. Lovász and M. Simonovits: *On the number of complete subgraphs of a graph*, Combinatorics, Proc. 5th British Comb. Conf. (1976), 439–441.

[67] L. Lovász and M. Simonovits: *On the number of complete subgraphs of a graph II*, Studies in Pure Math., To the memory of P. Turán (ed. P. Erdős) (1983), 459–495.

[68] L. Lovász and V. Sós: *Generalized quasirandom graphs*, J. Combin. Theory Ser. B. **98** (2008), 146–163.

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

[70] L. Lovász and B. Szegedy: *Testing properties of graphs and function*, Israel J. Math. **178** (2010), 113–156.

[71] L. Lovász and B. Szegedy: *Finitely forcible graphons*, J. Combin. Theory Ser. B. **101** (2011), 269–301.

[72] S. McGuinness: *The greedy clique decomposition of a graph*, J. Graph Theory **18** (1994), 427–430.

[73] S. McGuinness: *Greedy maximum-clique decompositions*, Combinatorica **14** (1994), 335–343.

[74] S. H. P. Diaconis and S. Janson: *Threshold graph limits and random threshold graphs*, Internet Math. **5** (2009), 267–318.

[75] O. Pikhurko and A. Razborov: *Asymptotic structure of graphs with the minimum number of triangles*, Combin. Probab. Comput. **26** (2017), 138–160.

[76] O. Pikhurko and E. Vaughan: *Minimum number of k-cliques in graphs with bounded independence number*, Combin. Probab. Comput. **22** (2013), 910–934.

[77] A. Razborov: *Flag algebras*, J. Symbolic Logic **72** (2007), 1239–1282.

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

[79] A. Razborov: *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discrete Math. **24** (2010), 946–963.

[80] V. Rödl: *On universality of graphs with uniformly distributed edges*, Discrete Math. **59** (1986), 125–134.

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

[82] M. Simonovits: *Extremal graph problems, degenerate extremal problems, and supersaturated graphs*, in: Progress in graph theory (Waterloo, Ont., 1982), 419–437.

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

[84] B. Szegedy: *An information theoretic approach to Sidorenko's conjecture*, preprint arXiv:1406.6738 (2014).

[85] B. Szegedy: *On Sidorenko's conjecture for determinants and Gaussian Markov random fields*, preprint arXiv:1701.03632 (2017).

[86] A. Thomason: *Pseudorandom graphs*, in: Random graphs '85 (Poznań, 1985), *North-Holland Math. Stud.*, volume 144, 307–331.

[87] A. Thomason: *Random graphs, strongly regular graphs and pseudorandom graphs*, C. Whitehead (ed.), Surveys in Combinatorics 1987, London Mathematical Society Lecture Note Series **123** (1987), 173–195.

[88] A. Thomason: *A disproof of a conjecture of Erdős in Ramsey theory*, J. London Math. Soc. **39** (1989), 246–255.

[89] Z. Tuza: *Unsolved Combinatorial Problems, Part I*, BRICS Lecture Series LS-01-1 (2001).

[90] R. M. Wilson: *An existence theory for pairwise balanced designs I. Composition theorems and morphisms*, Journal of Combinatorial Theory, Series A **13** (1972), 220 – 245.

[91] R. M. Wilson: *An existence theory for pairwise balanced designs II. The structure of PBD-closed sets and the existence conjectures*, Journal of Combinatorial Theory, Series A **13** (1972), 246 – 273.

[92] R. M. Wilson: *An existence theory for pairwise balanced designs, III: Proof of the existence conjectures*, Journal of Combinatorial Theory, Series A **18** (1975), 71 – 79.

[93] R. Yuster: *Integer and fractional packing of families of graphs*, Random Structures & Algorithms **26** (2005), 110–118.