Graph limits and their applications in extremal combinatorics*

Daniel Král'†

Abstract

Combinatorial limits provide an analytic way to represent large discrete objects, and are closely related to the flag algebra method, which led to solving several long-standing open problems in extremal combinatorics. These lecture notes are primarily focused on limits of dense graphs, which form the best understood case in the theory of combinatorial limits, and the applications of the flag algebra method in extremal graph theory and to study of structural properties of graph limits.

1 Introduction

The theory of combinatorial limits has opened new exciting links between analysis, combinatorics, computer science, ergodic theory, group theory and probability theory. The techniques have been developed to some extent independently for dense discrete objects and sparse discrete objects. By a dense discrete object, we mean an object such that a random subobject of a constant size carries some non-trivial structure. For example in the case of graphs, a random subgraph of a constant size has some edges with positive probability. We will be concerned with the convergence and limit representations of graphs and dense graphs in particular. However, many of the results presented further can be translated to other discrete objects, e.g., permutations [29,30,35] or partial orders [28,31]. We also refer the reader to a recent monograph by Lovász [36], where the theory of graph limits is treated in a more detailed and thorough way.

The theory of limits of dense graphs evolved in a series of papers by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi [11–13, 39, 40]. The closely related flag algebra method, which was introduced by Razborov [47], resulted in

*These lecture notes have been prepared for the São Paulo School of Advanced Science on Algorithms, Combinatorics and Optimization held in July 2016.

†Mathematics Institute, DIMAP and Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK. E-mail: d.kral@warwick.ac.uk.

substantial progress on many long standing open problems in extremal combinatorics, e.g. [2–5, 24–26, 33, 34, 45–49]. We present basic results from this theory in Section 2. Limits of dense graphs also led to new views on existing concepts in mathematics and computer science, e.g., on graph quasirandomness as we illustrate in Section 3. We present the flag algebra method in Section 4, where we also give some simple applications of the method in extremal graph theory. We also use the method to prove several results on so-called finitely forcible graph limits in Section 5.

The theory of limits of sparse graphs, such as graphs of bounded degree, is less developed. Several notions of convergence of such graphs were proposed and in general, the sparse graph convergence is considered to be significantly less understood than the convergence of dense graphs. Still, the area of sparse limits offers one of the most fundamental open problems on graph limits: the conjecture of Aldous and Lyons [1]. This conjecture gives a necessary and sufficient condition on a local neighborhood distribution to correspond to a sequence of graphs, and is essentially equivalent to Gromov's question whether all countable discrete groups are sofic. We will cover basic notions on sparse graph convergence in Section 6.

2 Dense graph convergence

In this section, we present basic results related to the convergence of dense graphs. We start by emphasizing the distinction between the concept of *convergence* and the concept of *limit representation*. The former refers to the property that elements forming a sequence of discrete objects are similar to each other. It is natural to consider whether a convergent sequence of discrete objects can be equipped with a limit representation, which captures essential properties of the elements in the sequence, i.e., the properties related to the convergence of the sequence. In the case of dense graph convergence, such a limit object is called graphon. However, it is possible to talk about convergence of discrete objects without having any particular limit representation in mind.

We now define the notion of convergence of dense graphs. If G and H are two graphs, the *density* of H in G, denoted by $d(H, G)$, is the probability that a randomly chosen subset of $|H|$ vertices of G induces a subgraph isomorphic to H, where $|H|$ is the order of H, i.e., its number of vertices. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *convergent* if the sequence of densities $d(H, G_n)$ converges for every graph H. In what follows, we will only consider convergent sequences $(G_n)_{n\in\mathbb{N}}$ of graphs such that the number of vertices of G_n tends to infinity.

Simple examples of convergent graph sequences include the sequence of complete graphs K_n, the sequence of complete bipartite graphs $K_{n,n}$ with parts of equal size, the sequence of complete bipartite graphs $K_{\lfloor \alpha n\rfloor,n}$ with part sizes converging to a particular ratio $\alpha \in (0, 1)$. A less trivial example of a convergent sequence of graphs is the sequence of Erdős-Rényi random graphs $G_{n,p}$. Recall that the *Erdős-Rényi random graph* $G(n,p)$, $n \in \mathbb{N}$ and $p \in [0,1]$, is the graph with n vertices such that any two of its vertices are joined by an edge with probability p independently of all the other pairs of vertices. The convergence of this sequence of graphs can be shown using Azuma-Hoeffding inequality and Borel-Cantelli lemma; we leave the details as an exercise.

Exercise 1. *Let $p \in [0,1]$ and let G_n be the Erdős-Rényi random graph $G(n,p)$. Show that the sequence $(G_n)_{n\in\mathbb{N}}$ is convergent with probability one.*

Another example of a convergent sequence of graphs is any sequence of sparse graphs: if $(G_n)_{n\in\mathbb{N}}$ is a sequence of graphs such that the number of edges of G_n is $o(|G_n|^2)$, then the density $d(H,G_n)$ of any non-edgeless graph H converges to zero and the density $d(H,G_n)$ of any edgeless graph H converges to one. Hence, the sequence $(G_n)_{n\in\mathbb{N}}$ is convergent. This is the reason why the notions presented in this section are of interest in relation to sequences of dense graphs, i.e., graphs with the number of edges quadratic in the number of their vertices. We discuss notions of convergent more appropriate in the sparse setting in Section 6.

A convergent sequence of graphs can be represented by an analytic object called a graphon. A *graphon* is a symmetric measurable function $W : [0,1]^2 \to [0,1]$, where symmetric stands for the property that $W(x,y) = W(y,x)$ for all $x,y \in [0,1]$. One can think of a graphon as a continuous analogue of the adjacency matrix of a graph; this analogy provides a good first intuition when working with graphons, however, the matter is more complex than it may seem at the first sight. This analogy leads to the following definition. If W is a graphon, then a *W-random graph of order n* is the random graph obtained by sampling n points $x_1,\dots,x_n$ independently and uniformly in the unit interval $[0,1]$ and joining the i-th vertex and the j-th vertex of the graph by an edge with probability $W(x_i,x_j)$. Note that if W is the graphon equal to $p \in [0,1]$ for all $x,y \in [0,1]$, then the W-random graph of order n is the Erdős-Rényi random graph $G(n,p)$.

We define the *density* of a graph H in a graphon W to be the probability that the W-random graph of order $|H|$ is isomorphic to H; this probability is denoted by $d(H,W)$. The following holds:

$$d(H,W) = \frac{|H|!}{|\mathrm{Aut}(H)|} \int\limits_{[0,1]^{|H|}} \prod_{v_iv_j \in E(H)} W(x_i,x_j) \prod_{v_iv_j \notin E(H)} (1 - W(x_i,x_j))\,\mathrm{d}x_1 \cdots x_{|H|}$$

where $V(H) = \{v_1,\dots,v_{|H|}\}$ and $\mathrm{Aut}(H)$ is the automorphism group of H. We say that a graphon W is the *limit* of a convergent sequence $(G_n)_{n\in\mathbb{N}}$ of graphs if

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

for every graph H.

Graphons are usually depicted in the unit square with values being different shades of gray with white representing zero and black representing one; the origin

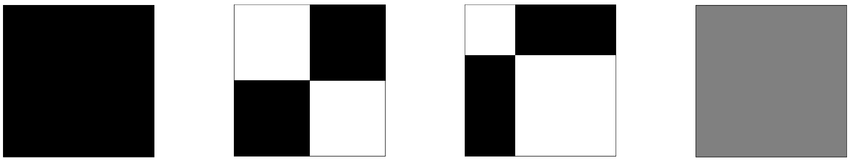

Figure 1: Graphons representing the sequences $(K_n)_{n\in\mathbb{N}}$, $(K_{n,n})_{n\in\mathbb{N}}$, $(K_{n,2n})_{n\in\mathbb{N}}$ and $(G(n,1/2))_{n\in\mathbb{N}}$.

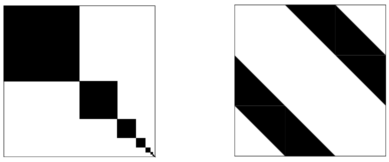

Figure 2: Graphons from Exercise 3.

of the coordinate system is usually in the top left corner to follow the analogy with adjacency matrices. In Figure 1, it is possible to find graphons representing some of the convergent sequence of graphs that we have presented earlier.

Exercise 2. *Show that the graphons depicted in Figure 1 are limits of the sequences $(K_n)_{n\in\mathbb{N}}$, $(K_{n,n})_{n\in\mathbb{N}}$, $(K_{n,2n})_{n\in\mathbb{N}}$ and $(G(n,1/2))_{n\in\mathbb{N}}$.*

Exercise 3. *Find convergent sequences of graphs such that the graphons depicted in Figure 2 are their limits.*

It is natural to ask whether every convergent sequence of graphs has a limit, whether this limit is unique (if it exists), and whether every graphon is a limit of a convergent sequence of graphs. We start with the last of these questions, which is the simplest to answer.

Theorem 1. *Let W be a graphon and let G_n be a W-random graph of order n, $n\in\mathbb{N}$. The sequence $(G_n)_{n\in\mathbb{N}}$ is convergent and the graphon W is its limit with probability one.*

Proof. Fix a graph H and an integer n such that $n \geq |H|$. The probability that a particular $|H|$-tuple of vertices of G_n induces a copy of H is $d(H,W)$. The linearity of expectation implies that the expected number of copies of H in G_n is equal to $d(H,W)\binom{n}{|H|}$. We next estimate the probability of a large deviation from this expected value. Let X_i, $i = 0,\dots,n$, be the random variable equal to the expected number of copies of H after the first i choices of the vertices of G_n are made in the interval $[0,1]$ and the edges between the first i vertices are fixed when constructing a W-random graph. Observe that X_n is just the number of copies of H in G_n and X_0 is equal to $d(H,W)\binom{n}{|H|}$.

Since the random variables $X_0, \ldots, X_n$ form a martingale, we can apply Azuma-Hoeffding inequality (Theorem 17) with $c_i \le n^{|H|-1}$ and get that

$$\mathbb{P}\left(|X_n - X_0| \ge t\right) \le 2e^{\frac{-t^2}{2n^{2|H|-1}}}$$

for every $t \in \mathbb{R}$. Substituting $t = \varepsilon n^{|H|}$, we get that

$$\mathbb{P}\left(|X_n - X_0| \ge \varepsilon n^{|H|}\right) \le 2e^{-\varepsilon^2 n/2} ,$$

which yields that

$$\mathbb{P}\left(|d(H, G_n) - d(H, W)| \ge |H|!2^{|H|}\varepsilon\right) \le 2e^{-\varepsilon^2 n/2}$$

if $n \ge 2|H|$. Borel-Cantelli lemma implies that the sequence $(d(H, G_n))_{n\in\mathbb{N}}$ is convergent with probability one and its limit is $d(H, W)$. In particular, the sequence $(G_n)_{n\in\mathbb{N}}$ is convergent and the graphon W is its limit with probability one. □

To address the existence of a graphon associated with a convergent sequence of graphs, we will need to recall the notion of graph regularity. We will use the notion of regularity due to Frieze and Kannan [19]; this notion is weaker than the notion of Szemerédi regularity, which is more well-known. Let us define the notion. If G is a graph and S and T two subsets of its vertices, $e(S, T)$ is the number of pairs of vertices $s \in S$ and $t \in T$ joined by an edge and $d(S, T)$ is the corresponding density, i.e., $d(S, T) = \frac{e(S,T)}{|S|\,|T|}$. A partition $V_1, \ldots, V_k$ of a vertex set of a graph G is an *equipartition* if $|\,|V_i| - |V_j|\,| \le 1$ for every $i, j \in [k]$. A partition $V_1, \ldots, V_k$ of the vertex set of a graph G is weakly ε-regular if it is an equipartition and it holds that

$$\left| e(S, T) - \sum_{i,j=1}^{k} d(V_i, V_j)\ |S \cap V_i|\ |T \cap V_j| \right| \le \varepsilon |G|^2$$

for any two subsets S and T of the vertex set of G. Frieze and Kannan [19] proved the following:

Theorem 2. *For every $\varepsilon \in (0, 1)$, there exists $K = 2^{O(\varepsilon^{-2})}$ such that every graph G has a weakly ε-regular partition with at most K parts.*

We will need a strengthening of Theorem 2. A partition $V'_1, \ldots, V'_{k'}$ of a vertex set of a graph G is a *refinement* of a partition $V_1, \ldots, V_k$ if for every $j \in [k']$, there exists $i \in [k]$ such that $V'_j \subseteq V_i$. The following strengthening of Theorem 2 holds.

Theorem 3. *For every $\varepsilon \in (0, 1)$, there exists $K = 2^{O(\varepsilon^{-2})}$ such that every equipartition of the vertex set of a graph G into k parts can be refined to a weakly ε-regular partition with at most $K \cdot k$ parts.*

The following theorem relates weak regular partitions to subgraph densities.

Theorem 4. *For every graph H and every $\delta \in (0,1)$, there exists $\varepsilon \in (0,1)$ such that if G is a graph with at least ε^{-1} vertices and $V_1, \dots, V_k$ is a weakly ε-regular partition of its vertex set, then*

$$\left| d(H,G) - \frac{|H|!}{|\mathrm{Aut}(H)|k^{|H|}} \sum_{i_1,\dots,i_{|H|}=1}^{k} \prod_{v_j v_{j'} \in E(H)} d(V_{i_j}, V_{i_{j'}}) \prod_{v_j v_{j'} \notin E(H)} (1 - d(V_{i_j}, V_{i_{j'}})) \right| \le \delta$$

where $V(H) = \{v_1, \dots, v_{|H|}\}$.

We are now ready to prove that every convergent sequence of graphs has a limit.

Theorem 5 (Lovász and Szegedy [39])**.** *Let $(G_n)_{n\in\mathbb{N}}$ be a convergent sequence of graphs. There exists a graphon W that is a limit of the sequence $(G_n)_{n\in\mathbb{N}}$.*

Proof. Fix a convergent sequence $(G_n)_{n\in\mathbb{N}}$, and set $\varepsilon_\ell = 2^{-\ell}$ for $\ell \in \mathbb{N}$. For every graph G_n in the sequence, fix a weakly ε_1-regular partition $V_1^{n,1}, \dots, V_{k_{n,1}}^{n,1}$ of its vertex set; such a partition exists by Theorem 2. Suppose that we have already fixed a weakly ε_ℓ-regular partition $V_1^{n,\ell}, \dots, V_{k_{n,\ell}}^{n,\ell}$ of G_n for some $n \in \mathbb{N}$ and $\ell \in \mathbb{N}$. By Theorem 3, there exists a weakly $\varepsilon_{\ell+1}$-regular partition $V_1^{n,\ell+1}, \dots, V_{k_{n,\ell+1}}^{n,\ell+1}$ of G_n that is a refinement of the partition $V_1^{n,\ell}, \dots, V_{k_{n,\ell}}^{n,\ell}$. By reordering the sets in the partition, we can assume that if $V_i^{n,\ell+1} \subseteq V_j^{n,\ell}$, $V_{i'}^{n,\ell+1} \subseteq V_{j'}^{n,\ell}$ and $i < i'$, then it holds that $j \le j'$. We will refer to this property as the *ordering property*. Note that Theorems 2 and 3 yield the existence of a constant K_ℓ, $\ell \in \mathbb{N}$, such that $k_{n,\ell} \le K_\ell$ for every $n \in \mathbb{N}$ and every $\ell \in \mathbb{N}$.

For every $n \in \mathbb{N}$ and $\ell \in \mathbb{N}$, associate the graph G_n with a $k_{n,\ell} \times k_{n,\ell}$-matrix $A^{n,\ell}$ such that the entry $A_{ij}^{n,\ell}$ is equal to $d(V_i^{n,\ell}, V_j^{n,\ell})$. Next choose a subsequence $(G'_n)_{n\in\mathbb{N}}$ of the sequence $(G_n)_{n\in\mathbb{N}}$ such that it holds for every $\ell \in \mathbb{N}$ that

- all but finitely values of $k_{n,\ell}$ are the same, and
- the matrices $A^{n,\ell}$ coordinate-wise converge.

Note that $k_{n,\ell}$ can have only values between 1 and K_ℓ, which implies that it is possible to choose a subsequence satisfying the first of the two properties. For such a subsequence, all but finitely many matrices $A^{n,\ell}$ have the same size and since their coordinates are reals between 0 and 1, it is possible to choose a subsequence of the former subsequence that also satisfies the second property. So, the subsequence $(G'_n)_{n\in\mathbb{N}}$ indeed exists.

Let k_ℓ be the value that appears infinitely often among the values $k_{n,\ell}$ subsequence $(G'_n)_{n\in\mathbb{N}}$. Further, let A^ℓ be the $k_\ell \times k_\ell$-matrix that is the coordinate-wise limit of the matrices $A^{n,\ell}$ for the subsequence $(G'_n)_{n\in\mathbb{N}}$. Theorem 4 implies that the following holds for every graph H:

$$\lim_{n\to\infty} d(H,G'_n) = \lim_{\ell\to\infty} \frac{|H|!}{|\mathrm{Aut}(H)|k_\ell^{|H|}} \sum_{i_1,\dots,i_{|H|}=1}^{k_\ell} \prod_{v_jv_{j'}\in E(H)} A^\ell_{i_j,i_{j'}} \prod_{v_jv_{j'}\notin E(H)} (1-A^\ell_{i_j,i_{j'}})$$

where $V(H) = \{v_1,\dots,v_{|H|}\}$. Since $(G'_n)_{n\in\mathbb{N}}$ is a subsequence of the sequence $(G_n)_{n\in\mathbb{N}}$, it follows that

$$\lim_{n\to\infty} d(H,G_n) = \lim_{\ell\to\infty} \frac{|H|!}{|\mathrm{Aut}(H)|k_\ell^{|H|}} \sum_{i_1,\dots,i_{|H|}=1}^{k_\ell} \prod_{v_jv_{j'}\in E(H)} A^\ell_{i_j,i_{j'}} \prod_{v_jv_{j'}\notin E(H)} (1-A^\ell_{i_j,i_{j'}}) \ . \tag{1}$$

The matrices A^ℓ yield random variables X_ℓ on $[0,1)^2$ defined as follows:

$$X_\ell(x,y) = A^\ell_{\lfloor x\cdot k_\ell\rfloor+1,\lfloor y\cdot k_\ell\rfloor+1} \ .$$

By the ordering property, the random variables X_ℓ, $\ell\in\mathbb{N}$, form a martingale. Hence, Corollary 18 implies that there exists a measurable function from $[0,1]^2$ to $[0,1]$ such that

$$W(x,y) = \lim_{\ell\to\infty} X_\ell(x,y)$$

for almost every $(x,y)\in[0,1)^2$. Observe that the following holds for every m and every $J\subseteq[m]^2$:

$$\int\limits_{[0,1]^m} \prod_{jj'\in J} W(x_j,x_{j'})\,\mathrm{d}x_1\cdots x_m = \lim_{\ell\to\infty} \int\limits_{[0,1)^m} \prod_{jj'\in J} X_\ell(x_j,x_{j'})\,\mathrm{d}x_1\cdots x_m \ . \tag{2}$$

Since it also holds for every $\ell\in\mathbb{N}$, every $m\in\mathbb{N}$ and every $J\subseteq[m]^2$ that

$$\frac{1}{k_\ell^m} \sum_{i_1,\dots,i_m=1}^{k_\ell} \prod_{jj'\in J} A^\ell_{i_j,i_{j'}} = \int\limits_{[0,1)^m} \prod_{jj'\in J} X_\ell(x_j,x_{j'})\,\mathrm{d}x_1\cdots x_m \ ,$$

it follows that

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

by (1) and (2). □

It remains to consider the uniqueness of a limit of a convergent sequence of graphs. We will say that two graphons W_1 and W_2 are weakly isomorphic if $d(H,W_1) = d(H,W_2)$ for every graph H, i.e., the graphons W_1 and W_2 are limits of the same sequences of graphs. For example, the graphons depicted Figure 3 are weakly isomorphic; they both are a limit of the sequence $(K_{n,n})_{n\in\mathbb{N}}$ of complete

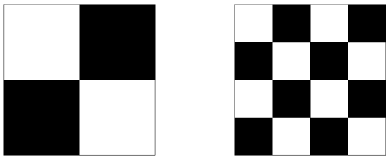

Figure 3: Two weakly isomorphic graphons.

bipartite graphs with parts of equal sizes. In particular, the graphon representing a convergent sequence of graphs is not unique.

The following is a general way of constructing weakly isomorphic graphons. Let φ be a measure preserving map from $[0,1]$ to $[0,1]$. If W is a graphon, let $W^\varphi(x,y) = W(\varphi(x),\varphi(y))$. Observe that the graphons W and W^φ are weakly isomorphic. For example, consider the following measure preserving map:

$$\varphi(x) = \begin{cases} 2x & \text{if } x \geq 1/2, \\ 2x-1 & \text{otherwise.} \end{cases}$$

If W_1 and W_2 are the two graphons depicted in Figure 3, then $W_2 = W_1^\varphi$.

Borgs, Chayes and Lovász [10] have shown that the above way of constructing weakly isomorphic graphons is in a certain sense the only way of obtaining weakly isomorphic graphons.

Theorem 6 (Borgs, Chayes and Lovász [10])**.** *If W_1 and W_2 are weakly isomorphic graphons, then there exist measure preserving maps φ_1 and φ_2 such that the graphons $W_1^{\varphi_1}$ and $W_2^{\varphi_2}$ are equal almost everywhere.*

3 Graph quasirandomness

In this section, we look at graph quasirandomness as studied in [15, 50, 51] to illustrate the concepts introduced in the previous section. Additional notation needs to be introduced. Let $h(H,G)$ denote the density of non-induced copies of H in G, and let $h_p(H)$ denote the expected value of $h(H,G(n,p))$ for $n \geq |H|$. Note that that $h(H,G)$ can be expressed as a linear combination of the values $d(H',G)$, where H' ranges over all supergraphs of H with the same number of vertices. The following is a classical result by Thomason [50].

Theorem 7. *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs. If the limit of $h(K_2,G_n)$ exists and is equal to $h_p(K_2)$ and the limit of $h(C_4,G_n)$ exists and is equal to $h_p(C_4)$ for some $p \in [0,1]$, then the limit of $h(H,G_n)$ exists for every graph H and is equal to $h_p(H)$.*

It is straightforward to cast Theorem 7 in the language of graphons. To do so, one needs to extend the definition of $h(H,G)$ to graphons: $h(H,W)$ is defined to be the expected number of non-induced copies of H in a W-random graph of order $|H|$.

Theorem 8. *Let W be a graphon. If $h(K_2,W)=h_p(K_2)$ and $h(C_4,W)=h_p(C_4)$ for some $p\in[0,1]$, then W is equal to p almost everywhere.*

Proof. Fix a graphon W such that $h(K_2,W)=h_p(K_2)$ and $h(C_4,W)=h_p(C_4)$ for some $p\in[0,1]$. Define $w:[0,1]\to[0,1]$ as follows

$$w(z)=\int\limits_{[0,1]} W(z,x)\mathrm{d}x\ .$$

One can think of the value $w(z)$ as of the degree of $z\in[0,1]$; indeed, if $z\in[0,1]$ is one of the points chosen during a construction of a W-random graph with n vertices, then $w(z)\cdot(n-1)$ is the expected degree of the vertex corresponding to z. Since $h(K_2,W)=h_p(K_2)=p$, we obtain that

$$\int\limits_{[0,1]} w(z)\mathrm{d}z=p\ . \tag{3}$$

Using Cauchy-Schwartz inequality, we obtain from (3) that

$$\int\limits_{[0,1]} w(z)^2\mathrm{d}z\cdot\int\limits_{[0,1]} 1\mathrm{d}z\geq\left(\int\limits_{[0,1]} w(z)\mathrm{d}z\right)^2=p^2\ . \tag{4}$$

We next compute the following integral.

$$\int\limits_{[0,1]^2}\left(\int\limits_{[0,1]} W(x,z)W(y,z)\mathrm{d}z-p^2\right)^2\mathrm{d}xy=$$

$$\int\limits_{[0,1]^4} W(x,z)W(y,z)W(x,z')W(y,z')\mathrm{d}xyzz'-2p^2\int\limits_{[0,1]^3} W(x,z)W(y,z)\mathrm{d}xyz+p^4=$$

$$\int\limits_{[0,1]^4} W(x,z)W(y,z)W(x,z')W(y,z')\mathrm{d}xyzz'-2p^2\int\limits_{[0,1]} w(z)^2\mathrm{d}z+p^4\ .$$

Since the equality $h(C_4,W)=h_p(C_4)$ implies that the first term is equal to p^4 and the whole integral is non-negative, it follows that

$$\int\limits_{[0,1]} w(z)^2\mathrm{d}z\leq p^2\ .$$

Hence, the inequality (4) holds with equality, which is only possible if $w(z)$ is a multiple of the constant one function almost everywhere. We conclude that $w(z) = p$ for almost every $z \in [0,1]$.

In addition, since the integral

$$\int\limits_{[0,1]^2} \left(\int\limits_{[0,1]} W(x,z)W(y,z)\mathrm{d}z - p^2 \right)^2 \mathrm{d}xy$$

is equal to zero, it follows that for almost every pair $x, y \in [0,1]^2$,

$$\int\limits_{[0,1]} W(x,z)W(y,z)\mathrm{d}z = p^2 \ .$$

Lemma 9, which we prove further, implies that

$$\int\limits_{[0,1]} W(x,z)^2\mathrm{d}z = p^2 \tag{5}$$

for almost every $x \in [0,1]$. Since $w(x) = p$ for almost every $x \in [0,1]$, we get that

$$\int\limits_{[0,1]} W(x,z)\mathrm{d}z = p \tag{6}$$

for almost every $x \in [0,1]$. By Cauchy-Schwartz inequality, we have that

$$p^2 = \left(\int\limits_{[0,1]} W(x,z)\mathrm{d}z \right)^2 \leq \left(\int\limits_{[0,1]} W(x,z)^2\mathrm{d}z \right) \left(\int\limits_{[0,1]} 1\mathrm{d}z \right) = p^2$$

for almost every $x \in [0,1]$. Since the equality holds if and only if $W(x,z)$ is a multiple of the constant one function, we conclude that $W(x,z) = p$ for almost every pair $x, z \in [0,1]^2$. $\square$

It remains to prove Lemma 9. The argument is essentially contained in the proof of Lemma 3.3 in [38].

Lemma 9. *If $F : [0,1]^2 \to [0,1]$ satisfies*

$$\int\limits_{[0,1]} F(x,z)F(y,z)\,dz = \xi$$

for almost every pair $x, y \in [0,1]^2$, then it holds that

$$\int\limits_{[0,1]} F(x,z)^2\,dz = \xi$$

for almost every $x \in [0,1]$.

Proof. Define $F_x(z) = F(x,z)$. Let T be the set of functions $f \in L_2[0,1]$ such that the set

$$N_\varepsilon(f) = \left\{ x \in [0,1] \text{ such that } \int\limits_{[0,1]} |F_x(z) - f(z)| \mathrm{d}z \leq \varepsilon \right\}$$

has positive measure for every $\varepsilon > 0$. By the assumption of the lemma, it holds for any two functions $f, g \in T$ that

$$\int\limits_{[0,1]} f(z)g(z)\mathrm{d}z = \xi \; .$$

In particular, it holds that

$$\int\limits_{[0,1]} f(z)^2\mathrm{d}z = \xi$$

for every function $f \in T$.

We finish the proof of the lemma by showing that $F_x \in T$ for almost every $x \in [0,1]$. Fix a countable dense subset S of $L_2[0,1]$ and consider the union U of all the sets $N_\varepsilon(f)$ for $f \in S$ and rational $\varepsilon > 0$ that have zero measure. The set of all $x \in [0,1]$ such that $F_x \in U$ has zero measure (it is a countable union of zero measure sets). On the other hand, if $F_x \notin T$, then there exists $\varepsilon > 0$ such that $N_\varepsilon(F_x)$ has zero measure. Since the set S is dense in $L_2[0,1]$, there exists $f \in S$ and rational $\varepsilon' > 0$ such that $F_x \in N_{\varepsilon'}(f)$ and $N_{\varepsilon'}(f) \subseteq N_\varepsilon(F_x)$. It follows that $F_x \in U$. □

We finish this section with two exercises; hints how to solve the two exercises can be found just after them.

Exercise 4. *Prove that it holds that*

$$\frac{1}{3}d(K_{1,2}, W) + d(K_3, W) \geq d(K_2, W)^2$$

for every graphon W.

Exercise 5. *Prove that $d(\overline{K_3}, W) + d(K_3, W) \geq 1/4$ for every graphon W.*

To solve Exercise 4, observe that

$$\int\limits_{[0,1]} w(z)^2\mathrm{d}z = \frac{1}{3}d(K_{1,2}, W) + d(K_3, W)$$

where $w(z)$ is defined as in the proof of Theorem 8. To solve Exercise 5, expand the integral on the right hand side of the following equality

$$d(\overline{K_3}, W) = \int\limits_{[0,1]^3} (1 - W(x,y))(1 - W(x,z))(1 - W(y,z))\mathrm{d}xyz \ ,$$

express some of the obtained terms using $d(K_2, W)$ and $d(K_3, W)$ and bound the remaining term using arguments analogous to those in the proof of Theorem 8.

4 The flag algebra method

The flag algebra method of Razborov [47] led to progress on many well-known problems in extremal combinatorics. One of the first applications of the method was a solution of a long-standing open problem on the minimum density of triangles in a graph with given edge-density [49]. The method is closely related to the concepts that we have presented in Sections 2 and 3 but it does not build on these concepts directly. In particular, the flag algebra method can be applied to any convergent sequence of dense objects regardless whether there is a suitable analytic representation of the limit. To make our exposition more accessible, we will introduce the method using the language that we have developed in Section 2 and 3.

Let $\mathcal{A}$ be the algebra of formal linear combinations of graphs with real coefficients with the natural operations of addition and multiplication by a scalar. If W is a graphon and $\sum_i \alpha_i H_i \in \mathcal{A}$ is an element of $\mathcal{A}$, we can define a mapping $f_W : \mathcal{A} \to \mathbb{R}$ as

$$f_W\left(\sum_i \alpha_i H_i\right) := \sum_i \alpha_i d(H_i, W) \ .$$

Our aim is to define a multiplication between elements of $\mathcal{A}$ in a way that f_W is an algebra homomorphism of $\mathcal{A}$ preserving addition and multiplication, i.e., we wish that $f_W(a \times a') = f_W(a) \cdot f_W(a')$ for any $a, a' \in \mathcal{A}$. To do so, we define the product of two graphs H_1 and H_2 and extend the definition linearly to the whole $\mathcal{A}$.

We first motivate the definition. Suppose that G is a large graph. The product $d(H_1, G) \cdot d(H_2, G)$ is the probability that $|H_1|$ randomly chosen vertices of G induce H_1 and $|H_2|$ randomly chosen vertices of G induce H_2. If G is large, the two sets of randomly chosen vertices are disjoint with probability close to one. Consequently, we can approximate this probability by taking $|H_1|+|H_2|$ randomly chosen vertices, splitting them into $|H_1|$-tuple and $|H_2|$-tuple and considering the probability that the first tuple induces H_1 and the second tuple induces H_2. This leads us to the following definition of the product of H_1 and H_2:

$$H_1 \times H_2 := \sum_H \frac{|\{(A,B)|V(H) = A \cup B, H[A] \cong H_1 \text{ and } H[B] \cong H_2\}|}{\binom{|H_1|+|H_2|}{|H_1|}} H$$

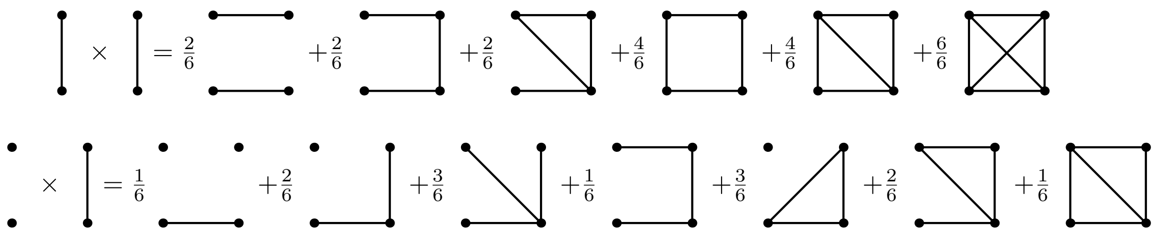

Figure 4: Two examples of multiplication of graphs.

where the sum is taken over all graphs H with $|H| = |H_1| + |H_2|$. The product operation is extended to the whole $\mathcal{A}$ linearly. See Figure 4 for examples. It is routine to verify that $f_W(H_1 \times H_2) = f_W(H_1) \cdot f_W(H_2)$ (the argument is similar to the one presented in the proof of Lemma 10 further).

Exercise 6. *Determine $K_2 \times K_3$ and $\overline{K_2} \times K_3$.*

The algebra homomorphism f_W satisfies another important property, which we state in the next lemma.

Lemma 10. *Let H be a graph with k vertices and W a graphon. For every graph H' with $\ell > k$ vertices, define $\alpha_{H'}$ to be the number of induced copies of H in H'. It holds that*

$$f_W(H) = f_W\left(\sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} H'\right)$$

where the sum runs over all graphs H' with ℓ vertices.

Proof. Let $(G_n)_{n\in\mathbb{N}}$ be a convergent sequence of graphs such that W is its limit. We will show that

$$d(H, G_n) = \sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} d(H', G_n) \tag{7}$$

for every graph G_n with at least ℓ vertices. By the definition of convergence, we get that

$$d(H, W) = \sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} d(H', W)\ ,$$

which implies the statement of the lemma.

To prove (7) consider the following way of choosing a random subset of k vertices of G_n: first choose randomly ℓ vertices of G_n and among those choose randomly k vertices. The subgraph induced by the ℓ randomly chosen vertices is isomorphic to H' with probability $d(H', G_n)$. The probability that the subgraph induced by the final k vertices is isomorphic to H conditioned on the subgraph induced by the ℓ vertices being isomorphic to H' is $\alpha_{H'}/\binom{\ell}{k}$. Hence, the right hand side of (7) is indeed the probability that a random subset of k vertices of G_n induces a subgraph isomorphic to H. $\square$

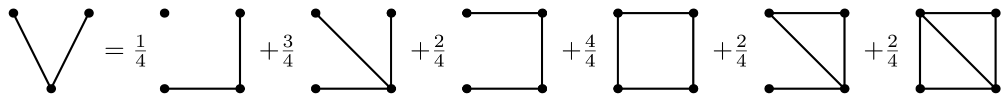

Figure 5: The graph $K_{1,2}$ expressed as a combination of 4-vertex graphs.

Lemma 10 allows us to view the elements H and the sum

$$\sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} H'$$

as the same elements of $\mathcal{A}$ since f_W always has the same value for them. With this view in mind, we can say that the sum

$$\sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} H'$$

expresses H as a linear combination of ℓ-vertex subgraphs. See Figure 5 for an example. It would also be possible to consider the subalgebra $\mathcal{A}'$ generated by expressions of the form

$$H - \sum_{H'} \frac{\alpha_{H'}}{\binom{\ell}{k}} H'$$

for all k-vertex graphs H and all $\ell > k$ and consider the factor algebra $\mathcal{A}/\mathcal{A}'$. Since $\mathcal{A}'$ lies in the kernel of f_W f_W yields a homomorphism from the factor-algebra $\mathcal{A}/\mathcal{A}'$ to $\mathbb{R}$.

Exercise 7. *Express K_3 as a linear combination of 5-vertex graphs.*

We now generalize the definition of the algebra $\mathcal{A}$ to rooted graphs. We first deal with the case of graphs with a single root where it is simpler to explain the main ideas. Let $\mathcal{A}^\bullet$ to be an algebra of formal linear combinations of rooted graphs, i.e., graphs with a single distinguished vertex referred to as the root. The mapping $f_W : \mathcal{A} \to \mathbb{R}$ was defined using subgraph densities in a graphon W, or equivalently subgraph densities in a large graph in a sequence of graphs converging to W. In the case of $\mathcal{A}^\bullet$, we will think of choosing a root vertex randomly and defining an analogous mapping $f_W^\bullet : \mathcal{A}^\bullet \to \mathbb{R}$. However, we will now not deal with a single mapping $f_W^\bullet$ but a probability distribution on such mappings. For $x_0 \in [0,1]$ and a rooted graph H with vertices $v_0, \ldots, v_k$ with v_0 being the root, define

$$f_W^{x_0}(H) = \frac{k!}{|\mathrm{Aut}^\bullet(H)|} \int\limits_{[0,1]^k} \prod_{v_i v_j \in E(H)} W(x_i, x_j) \prod_{v_i v_j \notin E(H)} (1 - W(x_i, x_j)) \,\mathrm{d}x_1 \cdots x_k$$

and extend $f_W^{x_0}$ to the whole $\mathcal{A}^\bullet$ linearly; here, $\mathrm{Aut}^\bullet(H)$ denotes the subgroup of $\mathrm{Aut}(H)$ containing the automorphisms of H fixing the root. The probability distributions on mappings from $\mathcal{A}^\bullet$ to $\mathbb{R}$ is obtained by choosing $x_0 \in [0,1]$ randomly.

To illustrate the just defined notion, let W be the third graphon in Figure 1. Let $e \in \mathcal{A}^\bullet$ be K_2 rooted at one of its vertices and let $c \in \mathcal{A}^\bullet$ be $K_{1,2}$ rooted at the vertex of degree two. Observe that $f_W^\bullet(e) = 1/3$ with probability $2/3$ and $f_W^\bullet(e) = 2/3$ with probability $1/3$. Likewise, $f_W^\bullet(c) = 1/9$ with probability $2/3$ and $f_W^\bullet(c) = 4/9$ with probability $1/3$. However, $f_W^\bullet(e+c) = 4/9$ with probability $2/3$ and $f_W^\bullet(e+c) = 10/9$ with probability $1/3$.

It is possible to define multiplication of rooted graphs such that $f_W^\bullet : \mathcal{A}^\bullet \to \mathbb{R}$ respects both addition and multiplication, and to find an analogue of Lemma 10. However, we do so later in a full generality considering graphs rooted at arbitrary subgraphs. Our goal now is to compute the expected value of $f_W^\bullet(H)$. More precisely, for $a \in \mathcal{A}^\bullet$, we would like to find $a' \in \mathcal{A}$ such that $f_W(a') = \mathbb{E}_x f_W^x(a)$; such a' will be denoted by $[\![a]\!]_\bullet$. By linearity of expectations, it is enough to define the $[\![]\!]_\bullet$ for rooted graphs. Let H be a rooted graph and H' the same graph but with no vertex distinguished as a root. Following the definition of $f_W^x(H)$, the value $\mathbb{E}_x f_W^x(H)$ is the probability that a W-random graph of order $|H|$ with its first vertex distinguished as a root is isomorphic to H. This probability is equal to the probability that a W-random graph of order $|H|$ is H' and it becomes H when a root is chosen randomly. In particular, $[\![H]\!]_\bullet = \alpha H'$ where α is the probability that H' becomes H when one of the vertices of H' is randomly chosen as a root. For example, $[\![e]\!]_\bullet = K_2$ and $[\![c]\!]_\bullet = \frac{1}{3}K_{1,2}$.

The concepts that we have just introduced will now be treated in full generality considering graphs rooted at arbitrary subgraphs. Fix a labelled graph R with r vertices, i.e., a graph with vertices labelled from 1 to r; we will think of the vertices of R as the roots and refer to them as the first root, …, the r-th root. Let $\mathcal{A}^R$ be an algebra of formal linear combinations of R-rooted graphs, i.e., graphs that have R distinguished vertices, which we refer to as the roots, and the subgraph induced by them is R (preserving the order of the distinguished vertices). We next define the probability distribution on mappings from $\mathcal{A}^R$ to $\mathbb{R}$ as follows. Generate a W-random graph H_R of order r and let $x_1, \dots, x_r \in [0,1]$ be the choices for its vertices. If H_R is not R (preserving the order of the vertices, i.e., the vertex corresponding to x_i being the i-th root), then $f_W^R(H) = 0$ for every $H \in \mathcal{A}^R$. Otherwise, $f_W^R(H)$ for an R-rooted graph H with k vertices is the probability that if additional $r-k$ vertices are chosen as in the definition of a W-random graph and the edges between the vertices are chosen as when generating a W-random graph, then the obtained graph with the vertices of H_R being the roots is isomorphic to H; the function f_W^R is extended to the whole $\mathcal{A}^R$ linearly. Note that the definition of f_W^R coincides with that of $f_W^\bullet$ in the case when R is a single vertex. The mapping f_W^R is an algebra homomorphism from $\mathcal{A}^R$ to $\mathbb{R}$ if $f_W^R(R) = 1$. To simplify our presentation, we refer to all random mappings f_W^R as homomorphisms, which is technically incorrect in the case when f_W^R is identically equal to zero.

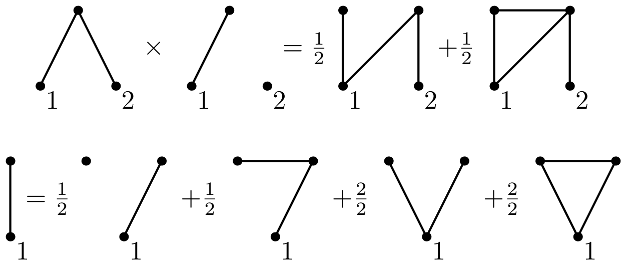

Figure 6: Examples of operations with rooted graphs; the root vertices are labelled with numbers.

We now define a multiplication on $\mathcal{A}^R$. Our goal is that it holds that $f_W^R(H \times H') = f_W^R(H) \cdot f_W^R(H')$. We will define $H \times H'$ for two R-rooted graphs H and H' and linearly extend the definition to the whole $\mathcal{A}^R$. Let k and k' be the number of vertices of H and H', respectively. The product $H \times H'$ is equal to the sum of all R-rooted graphs H'' with $k + k' - r$ vertices where the coefficient at the graph H'' is the probability that a random partition of the non-root vertices to sets A and A' of $k - r$ and $k' - r$ vertices, respectively, has the property that the roots together with the vertices of A induce a subgraph isomorphic to H and the roots together with the vertices of A' induce a subgraph isomorphic to H'.

In a similar way, one may consider the analogue of Lemma 10 and show that if H is an R-rooted graph with k vertices and H' is the sum of all R-rooted graphs H' with $k' > k$ vertices where the coefficient at the graph H' is the probability that H' becomes H after removing $k' - k$ random non-root vertices, then $f_W^R(H) = f_W^R(H')$. We can also define the operator $[\![]\!]_R$ such that $\mathbb{E}_R f_W^R(H) = f_W([\![H]\!]_R)$. For an R-rooted graph H with k vertices, $[\![H]\!]_R$ is equal to the k-vertex graph H' obtained from H by undistinguishing the roots where the coefficient at the graph H' is the probability that choosing r roots randomly in H' yields an R-rooted graph and this graph is isomorphic to H. Some examples of the just introduced definitions can be found in Figure 6.

We will now solve Exercise 5 using the language of flag algebras. Recall that e is the graph obtained from K_2 by choosing one of its vertices as the root; let f be the graph obtained from $\overline{K_2}$ by choosing one of its vertices as the root. For every graphon W, $f_W^\bullet((e-f)^2) = f_W^\bullet(e-f)^2 \geq 0$ with probability one. It follows that

$$0 \leq \mathbb{E}_\bullet f_W^\bullet((e-f)^2) = f_W([\![(e-f)^2]\!]_\bullet) = f_W(H)$$

where

$$H = \overline{K_3} - \frac{1}{3}\overline{K_{1,2}} - \frac{1}{3}K_{1,2} + K_3 \ .$$

Since $f_W(\overline{K_3} + \overline{K_{1,2}} + K_{1,2} + K_3) = 1$, we get that

$$\frac{1}{3} \leq f_W\left(\frac{4}{3}\overline{K_3} + \frac{4}{3}K_3\right) \ .$$

It follows that $f_W(\overline{K_3}+K_3)$ is at least 1/4 as desired.

The just introduced arguments can be practiced while solving the next exercise; a hint how to approach the exercise is given after it.

Exercise 8. *Show that $d(K_3,W) \geq 2d(K_2,W)(d(K_2,W)-1/2)$ for every graphon W.*

To solve Exercise 8, use $(\mathbb{E}f^\bullet_W(e))^2 \leq \mathbb{E}f^\bullet_W(e^2)$ and express K_2 as a linear combination of 3-vertex graphs using Lemma 10.

The flag algebra method can be used in a computer assisted way to find proofs of statements in extremal combinatorics. We will now sketch the main ideas behind this approach. In the example above, we used that $f_W([\![H^2]\!]_\bullet)$ is always non-negative for every $H \in \mathcal{A}^\bullet$. This can be generalized as follows. If $H_1,\dots,H_m \in \mathcal{A}^R$, A is an $m \times m$ positive semidefinite matrix and

$$H = (H_1 \cdots H_m) A \begin{pmatrix} H_1 \\ \vdots \\ H_m \end{pmatrix}, \tag{8}$$

then $f_W([\![H]\!]_\bullet)$ is non-negative.

Let $G_0,\dots,G_k \in \mathcal{A}$ and $\alpha_1,\dots,\alpha_k \in \mathbb{R}$. Suppose that we would like to prove for every graphon W satisfying $f_W(G_i) \geq \alpha_i$ for every $i \in [k]$ that that $f_W(G_0) \geq \alpha_0$ for α_0 as large as possible. For example, if $G_0 = K_3$, $G_1 = K_2$ and $\alpha_1 = 3/4$, we are looking for a maximum value α_0 such that every graphon with edge density at least 3/4 has triangle density at least α_0. If we had $G_0 = K_3$, $G_1 = K_2$, $G_2 = -K_2$, $\alpha_1 = 3/4$ and $\alpha_2 = 3/4$, then we would be looking for a maximum value α_0 such that every graphon with edge exactly density 3/4 has triangle density at least α_0.

Let $G'_1,\dots,G'_\ell$ be all graphs with a certain fixed number N of vertices. A possible way of proving this is to find $\gamma_1,\dots,\gamma_k \geq 0$, $\delta_0 \in \mathbb{R}$, $\delta_1,\dots,\delta_\ell \geq 0$, and an $m \times m$ positive semidefinite matrix A such that

$$G_0 = \sum_{i=1}^{k} \gamma_i G_i + \sum_{i=1}^{\ell} (\delta_0+\delta_i) G'_i + [\![H]\!]_R \text{ and}$$

$$\alpha_0 = \delta_0 + \sum_{i=1}^{k} \gamma_i \alpha_i$$

where H is as in (8). For simplicity, we consider here that we use (8) for a single choice of the root R but it is possible to work with matrices for several different choices of the root. If such coefficients γ_i and δ_i exists, then $f_W(G_0) \geq \alpha_0$ for every graphon W satisfying $f_W(G_i) \geq \alpha_i$, $i \in [k]$. Indeed, note that

$$f_W(G_1) \;\geq\; \alpha_1$$

$$\begin{array}{rcl} & \vdots & \\ f_W(G_k) & \geq & \alpha_k \\ \sum_{i=1}^{\ell} f_W(G'_i) & = & 1 \\ f_W(G'_1) & \geq & 0 \\ & \vdots & \\ f_W(G'_\ell) & \geq & 0 \\ f_W([\![H]\!]_R) & \geq & 0 \end{array}$$

where H is as in (8). Summing these inequalities and the equality with the coefficients $\gamma_1, \ldots, \gamma_k, \delta_0, \ldots, \delta_\ell, 1$, respectively, and using that f_W is an algebra homomorphism from $\mathcal{A}$ to $\mathbb{R}$, we get that

$$f_W(G_0) \geq \alpha_0; .$$

The search for the coefficients γ_i and δ_i can be formulated as a semidefinite program: one fixes a number of vertices N and expresses all the sides as combinations of N-vertex graphs using Lemma 10. For each such graph, one gets a single equation that becomes part of the semidefinite program.

Let us demonstrate this approach on a specific example. Suppose that we aim at proving $f_W(\overline{K_3} + K_3) \geq \alpha_0$ for α_0 as large as possible, i.e., $G_0 = \overline{K_3} + K_3$. We fix $N = 3$, $G'_1 = \overline{K_3}$, $G'_2 = \overline{K_{1,2}}$, $G'_3 = K_{1,2}$ and $G'_4 = K_3$. Further, we set $R = K_1^\bullet$ and $(H_1, H_2) = (\overline{K_2^\bullet}, K_2^\bullet)$. This setting leads to the following semidefinite program: maximize $\langle C, X\rangle$ subject to $\langle A_i, X\rangle = b_i$, $i = 1, 2, 3, 4$, and $X \succeq 0$, where the matrices $A_1, \ldots, A_4$ and C are as follows.

$$A_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \qquad b_1 = 1$$

$$A_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \end{pmatrix} \qquad b_2 = 0$$

$$A_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/3 \end{pmatrix} \qquad b_3 = 0$$

$$A_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \qquad b_4 = 1$$

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

We remark that $\langle M_1, M_2 \rangle$ is equal to trace of $M_1^T M_2$, i.e., the sum of the products of the corresponding entries of M_1 and M_2, and $X \succeq 0$ means that X is a symmetric positive semidefinite matrix. An optimal solution of the program is formed by the following matrix.

$$X = \begin{pmatrix} 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3/4 & -3/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & -3/4 & 3/4 \end{pmatrix}$$

This yields the following flag algebra proof of $f_W(\overline{K_3} + K_3) \geq 1/4$ for every graphon W:

$$\overline{K_3} + K_3 = \frac{1}{4}\left(\overline{K_3} + \overline{K_{1,2}} + K_{1,2} + K_3\right) + \left[\!\!\left[\begin{pmatrix} \overline{K_2}^{\bullet} \\ K_2^{\bullet} \end{pmatrix}^T \begin{pmatrix} 3/4 & -3/4 \\ -3/4 & 3/4 \end{pmatrix} \begin{pmatrix} \overline{K_2}^{\bullet} \\ K_2^{\bullet} \end{pmatrix} \right]\!\!\right]_{\bullet}$$

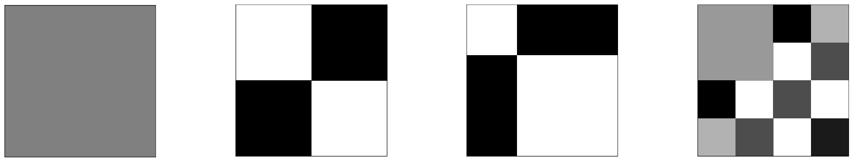

Figure 7: Examples of step graphons.

$$\geq \frac{1}{4} + 0 = \frac{1}{4} \, .$$

We now return back to the general setting. Suppose that the value α_0 produced by the semidefinite program solver is optimal. However, since the solution found by the solver is numerical, it may be needed to round it to a rational solution that will provide the actual proof of the desired inequality $f_W(G_0) \geq \alpha_0$. The most tricky part of this is to find the right matrix A for (8). Consider the eigenvectors of the matrix A found by the semidefinite program solver; each such eigenvector gives a linear combination of graphs $H_1, \ldots, H_m$, and let $h \in \mathcal{A}^R$ be this linear combination. If W is a graphon such that $f_W(G_0) = \alpha_0$ and $f_W(G_i) \geq \alpha_i$ for every $i \in [k]$, then $f_W^R(h) = 0$ with probability one for every such $h \in \mathcal{A}^R$ corresponding to a non-zero eigenvalue. Hence, we know that the eigenvectors corresponding to the non-zero eigenvalues of A must lie in the hyperplane determined by $h \in \mathcal{A}^R$ such that $f_W^R(h) = 0$ with probability one for any graphon W satisfying $f_W(G_0) = \alpha_0$ and $f_W(G_i) \geq \alpha_i$. After projecting the eigenvectors of the matrix A found by the solver to this hyperplane, we may obtain a proof of the optimal inequality $f_W(H_0) \geq \alpha_0$.

5 Finitely forcible graphons

In this section, we study finitely forcible graphons, which are of particular interest in relation to problems from extremal graph theory. A graphon W is *finitely forcible* if there exist finitely many graphs $H_1, \ldots, H_k$ such that every graphon W' satisfying $d(H_i, W') = d(H_i, W)$ for every $i \in [k]$ is weakly isomorphic to W. We will say that the graphon W is *forced* by the graphs $H_1, \ldots, H_k$. Since $h(H, W)$ can be expressed as a linear combination of $d(H', W)$, where H' ranges through all supergraphs of H with the same number of vertices as H, Theorem 8 implies that every constant graphon is finitely forcible; the theorem also implies that this graphon is forced by 4-vertex graphs. Lovász and Sós [37] generalized Theorem 8 to step graphons. A *step graphon* is a graphon such that there exist disjoint measurable sets $J_1, \ldots, J_k$ satisfying $J_1 \cup \cdots \cup J_k = [0, 1]$ and a symmetric matrix $D \in [0, 1]^{k \times k}$ such that $W(x, y) = D_{ij}$ for every $x \in J_i$ and $y \in J_j$. Examples of step graphons can be found in Figure 7.

Theorem 11 (Lovász and Sós [37]). *Every step graphon is finitely forcible.*

We will prove two special cases of Theorem 11 as an illustration of the application of the flag algebra method in the setting of finitely forcible graphons. Before we do so, we would like to exhibit one of the links between finitely forcible graphons and extremal graph theory.

Proposition 12. *Every finitely forcible graphon W_0 is the unique (up to a weak isomorphism) minimizer of a linear combination of subgraph densities, i.e., there exist $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ and graphs $H_1, \ldots, H_k$ such that the graphon W_0 minimizes the expression*

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

and any graphon minimizing this expression is weakly isomorphic to W_0.

Proof. Let $H'_1, \ldots, H'_\ell$ be the graphs forcing W_0 and let $\beta_j = d(H'_j, W_0)$, $j \in [\ell]$. Further, let

$$\sum_{i=1}^{\ell} (H'_j - \beta_j)^2 = \sum_{i=1}^{k} \alpha_i H_i \in \mathcal{A} .$$

Note that

$$\sum_{i=1}^{k} \alpha_i d(H_i, W) = \sum_{i=1}^{\ell} (d(H'_j, W) - \beta_j)^2$$

for every graphon W, in particular, the sum

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

is always non-negative. If $W = W_0$, then the sum

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

is zero, i.e., W_0 is a minimizer. Moreover, if

$$\sum_{i=1}^{k} \alpha_i d(H_i, W) = 0$$

for a graphon W, then $d(H'_j, W) = d(H'_j, W_0)$ for every $j \in [\ell]$, which implies that W is weakly isomorphic to W_0 since W_0 is forced by $H'_1, \ldots, H'_\ell$. ∎

It was conjectured that the converse of Proposition 12 is also true.

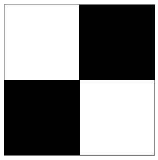

Figure 8: The graphon shown to be finitely forcible in Proposition 14.

Conjecture 13 (Lovász and Szegedy [38])**.** *Let $\alpha_1, \ldots, \alpha_k$ be reals and $H_1, \ldots, H_k$ graphs. There exists a finitely forcible graphon that minimizes*

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

We now prove two special kinds of step graphons are finitely forcible.

Proposition 14. *The graphon in Figure 8 is finitely forcible.*

Proof. Let W_0 be the graphon from Figure 8. We will show that W_0 is forced by 5-vertex graphs. Let W be another graphon such that the densities of all 5-vertex graphs are the same in W and W_0, and define $W_x(y) = W(x, y)$ for $x \in [0, 1]$. Further, let T be the set of all functions $f \in L_2[0, 1]$ such that the set

$$N_\varepsilon(f) = \left\{ x \in [0,1] \text{ such that } \int_{[0,1]} |W_x(y) - f(y)| \mathrm{d}y \le \varepsilon \right\}$$

has positive measure for every $\varepsilon > 0$.

We first observe that $\mathbb{E}_\bullet f_W^\bullet(e - 1/2)^2 = 0$: indeed, the left hand side can be expressed in terms of 4-vertex graph densities, which can then be expressed in terms of 5-vertex graph densities by Lemma 10. This implies that $f_W^\bullet(e) = 1/2$ with probability one. It follows that

$$\int_{[0,1]} f(y) \mathrm{d}y = 1/2 \tag{9}$$

for every $f \in T$.

Let s and t be the graphs $K_{1,2}$ and K_3, respectively, with two root vertices such that the roots are non-adjacent in the case of s; let σ and τ be the subgraph induced by the roots of s and t, respectively. Since $d(K_3, W_0) = 0$, we do not have to consider the graph t here but we prefer presenting the arguments in a full generality. Observe that

$$\mathbb{E}_\sigma f_W^\sigma(s(s - 1/4)^2) = 0 \text{ and } \mathbb{E}_\tau f_W^\tau(t(t - 1/4)^2) = 0 \ .$$

Since both the expected values can be expressed in terms of 5-vertex graph densities, it follows that for any two functions f and g in T, it holds that

$$\int_{[0,1]} f(y)g(y)\mathrm{d}y = 0 \text{ or } \int_{[0,1]} f(y)g(y)\mathrm{d}y = 1/4\ . \tag{10}$$

Following the lines of the proof of Lemma 9, we derive using (9) and (10) that

$$\int_{[0,1]} f^2(y)\mathrm{d}y = 1/4$$

for every $f \in T$.

Since W is not a constant graphon (by Theorem 8), the set T contains two functions f and g that differ on a set of non-zero measure. Let f and g be two such functions. If

$$\int_{[0,1]} f(y)g(y)\mathrm{d}y = 1/4,$$

we use the Cauchy-Schwartz inequality together with the equalities

$$\int_{[0,1]} f^2(y)\mathrm{d}y = 1/4 \text{ and } \int_{[0,1]} g^2(y)\mathrm{d}y = 1/4$$

to derive that f and g are multiple of each other almost everywhere, which yields that they are equal almost everywhere, contradicting their choice. Hence, it holds that

$$\int_{[0,1]} f(y)g(y)\mathrm{d}y = 0.$$

Consequently, the supports S_f and S_g of the functions f and g intersect at a set of measure zero. Since the integral of f on $[0,1]$ is at most the measure of S_f and the integral of g on $[0,1]$ is at most the measure of S_g, it follows from (9) that both S_f and S_g has measure $1/2$. Since the choice of f and g in T was arbitrary, it follows that there exist disjoint subsets A and B of $[0,1]$, each of measure $1/2$, such that every function in T differ from the characteristic function χ_A of A or from from the characteristic function χ_B of B on a set of measure zero.

Let A' be the set of all $x \in [0,1]$ such that W_x differs from χ_A on a set of measure zero, and let B' be the set of all $x \in [0,1]$ such that W_x differs from χ_B on a set of measure zero. Observe that A' and B' are disjoint and their union has measure one. If $A' \cap A$ had positive measure, then it would hold that $d(K_3, W) > 0$. Likewise, if $B' \cap B$ had positive, then $d(K_3, W) > 0$. We conclude that A and B' differ on a set measure zero, and A' and B differ on a set of measure zero. Hence, the graphon W is equal to one almost everywhere on $(A \times B) \cup (B \times A)$ and equal to zero almost everywhere else. It follows that W is weakly isomorphic to W_0. □

In Proposition 14, we have shown that the graphon from Figure 8 is forced by 5-vertex graphs. Using classical results from extremal graph theory, it would be possible to argue that this graphon is forced by setting the density of K_2 to $1/2$ and the density of K_3 to 0. This statment can also be established using the graph limit language. However, we have intentionally decided to prove a weaker statement to let ourselves to be less careful with some of the arguments.

Proposition 15. *Let $D \in [0,1]^{k\times k}$ be a symmetric matrix, $k \geq 2$, such that no two rows of D have the same sum. Let W_0 be the graphon such that*

$$W_0(x,y) = D_{\lfloor k\cdot x\rfloor+1,\lfloor k\cdot y\rfloor+1}$$

for $x, y \in [0,1)$, and set $W_0(x,y) = 0$ if $x = 1$ or $y = 1$. The graphon W_0 is finitely forcible.

Proof. We show that W_0 is forced by densities of $(8k-4)$-vertex graphs. Let W be the a graphon such that $d(H,W) = d(H,W_0)$ for every $(8k-4)$-vertex graph, and let $W_x(z) = W(x,z)$ for $x \in [0,1]$. Further, set

$$\delta_i = \frac{1}{k}\sum_{j=1}^{k} D_{ij}$$

for $i \in [k]$. Recall that e is the rooted graph obtained from K_2 by choosing one of its vertices to be the root. Since it holds that (note that the left hand can be expressed in terms of $2k+1$-vertex graphs)

$$\mathbb{E}_\bullet f_W^\bullet\left(\prod_{i=1}^{k}(e-\delta_i)^2\right) = 0$$

it follows that

$$f_W^x(e) = \int\limits_{[0,1]} W_x(z)\mathrm{d}z \in \{\delta_1,\ldots,\delta_k\}$$

for almost every $x \in [0,1]$. Let A_j, $j \in [k]$, be the set of all $x \in [0,1]$ such that $f_W^x(e) = \delta_j$. To simplify our presentation, we will assume that $A_1\cup\cdots\cup A_k = [0,1]$; since $[0,1]\setminus(A_1\cup\cdots\cup A_k)$ has measure zero, this assumption does not affect the generality of our arguments. Note that it holds

$$\mathbb{E}_\bullet f_W^\bullet\left(\prod_{i=1,i\neq j}^{k}(e-\delta_i)^2\right) = \frac{1}{k}\prod_{i=1,i\neq j}^{k}(\delta_j-\delta_i)^2$$

for every $j \in [k]$. Consequently, it holds

$$f_W^\bullet\left(\prod_{i=1,i\neq j}^{k}(e-\delta_i)^2\right) \neq 0$$

only for $f^{\bullet}_W = \delta_j$, we obtain that the measure of the set A_j is $1/k$ for every $j \in [k]$. We will think of A_j as corresponding to the interval $[(j-1)/k, j/k)$ in the definition of W_0 and refer to both A_j and to this interval as to the *j-th part*.

A graph H is *decorated* if each vertex of H is labelled with one of the numbers between $1, \dots, k$ (different vertices may be assigned the same number). For a decorated graph H, we define $d(H,W)$ to be the probability that a W-random graph is isomorphic to H and the vertices belong to the parts corresponding to their labels. We will show that $d(H,W) = d(H,W_0)$ for every 4-vertex decorated graph. Fix a 4-vertex decorated graph H and let H_0 be the graph obtained from H by making each of the four vertices to be a root. Further, let $\ell_1, \dots, \ell_4$ be the labels of the roots of H, and let H_i, $i = 1, \dots, 4$, be the sum of all 5-vertex H_0-rooted graphs where the non-root vertex is adjacent to the i-th root (there are exactly $2^3 = 8$ such H_0-rooted graphs). Observe that

$$f^{H_0}_W \left(\prod_{i=1}^{4} \prod_{j=1, j \neq \ell_i}^{k} (H_i - \delta_j)^2 \right) \neq 0$$

if and only if the i-th root is chosen from the i-th part of graphon, and if it is non-zero, then it is equal to

$$\prod_{i=1}^{4} \prod_{j=1, j \neq \ell_i}^{k} (\delta_{\ell_i} - \delta_j)^2 \; .$$

It follows that

$$\mathbb{E}_{H_0} f^{H_0}_W \left(\prod_{i=1}^{4} \prod_{j=1, j \neq \ell_i}^{k} (H_i - \delta_j)^2 \right) = d(H,W) \cdot \prod_{i=1}^{4} \prod_{j=1, j \neq \ell_i}^{k} (\delta_{\ell_i} - \delta_j)^2 \; .$$

Since the left hand side of this expression can be expressed in term of densities of $(8k-4)$-vertex graphs in W, it follows that $d(H,W) = d(H,W_0)$ for every 4-vertex decorated graph H.

We now follow the lines of the proof of Theorem 8. Since the densities of 4-vertex decorated graphs are the same in W and W_0, it holds for every $i, j \in [k]$ that

$$\int\limits_{A_i \times A_j} W(x,z) \mathrm{d}xz = \frac{D_{ij}}{k^2} \quad \text{and} \quad \int\limits_{A_i^2} \left(\int\limits_{A_j} W(x,y) W(x,z) \mathrm{d}z - \frac{D_{ij}^2}{k} \right)^2 \mathrm{d}xy = 0 \; .$$

The arguments given in the proof of Theorem 8 imply that

$$\int\limits_{A_j} W_x(z) \mathrm{d}z = \frac{D_{ij}}{k} \quad \text{and} \quad \int\limits_{A_j} W_x(z)^2 \mathrm{d}z = \frac{D_{ij}^2}{k}$$

for almost every $x \in A_i$. We conclude that $W_x(z) = D_{ij}$ for almost every $(x,y) \in A_i \times A_j$, which implies that W is weakly isomorphic to W_0. $\square$

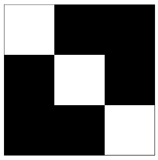

Figure 9: The graphon from Exercise 9.

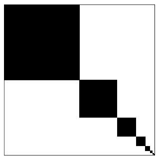

Figure 10: A graphon that is not finitely forcible.

The methods used in the proofs of Propositions 14 and 15 can be used to solve the following three exercises.

Exercise 9. *Show that the graphon in Figure 9 is finitely forcible.*

Exercise 10. *Let* $J_1, \dots, J_k$ *be disjoint measurable sets such that* $J_1 \cup \dots \cup J_k = [0,1]$*, and define* W *to be the graphon such that* $W(x,y) = 1$ *if* x *and* y *belongs to the same set* J_j*, and* $W(x,y) = 0$*, otherwise. Show that* W *is finitely forcible.*

Exercise 11. *Let* W *be a step graphon and let* $J_1, \dots, J_k$ *be the measurable sets from its definition. Suppose that*

$$\int\limits_{[0,1]} W(x,z)\ dz \neq \int\limits_{[0,1]} W(y,z)\ dz$$

for any $x \in J_i$ *and* $y \in J_j$ *such that* $i \neq j$*. Show that* W *is finitely forcible.*

Lovász and Szegedy carried a systematic study of finitely forcible graphons in [38]. They conjectured that all finitely forcible graphons have simple structure, which was disproved using the flag algebra method in [16, 17, 22, 23]. The most general of these results asserts that every computable graphon is a subgraphon of a finitely forcible graphon [17]. However, even some simple graphons need not be finitely forcible. For example, the graphon in Figure 10 is not finitely forcible [21].

6 Sparse graph convergence

We finish with giving a brief overview of the main notions of convergence of sparse graphs. We restrict our attention to graphs with bounded maximum degree though many of the presented concepts can be extended to more general settings. We will also be less technical than in the previous sections, primarily focusing on presenting the main ideas behind the relevant concepts.

The most used notion of convergence in relation to graphs with bounded degrees is the one defined by Benjamini and Schramm [6], known as the *Benjamini-Schramm convergence*, shortly BS-convergence, and also as the *left convergence*. Suppose that $(G_n)_{n\in\mathbb{N}}$ is a sequence of graphs with maximum degree at most Δ. For every $d \in \mathbb{N}$, let $\mathcal{G}^\bullet(d,\Delta)$ be the set of all rooted graphs with maximum degree Δ where all vertices has distance at most d from the root. Note that $\mathcal{G}^\bullet(d,\Delta)$ is finite for every pair d and Δ. By choosing a root in G_n randomly and restricting the graph G_n to the d-neighborhood of the root, i.e., the vertices at distance at most d from the root, we get a probability distribution on graphs from $\mathcal{G}^\bullet(d,\Delta)$. Let $p_{n,d} \in [0,1]^{\mathcal{G}^\bullet(d,\Delta)}$ be the corresponding vector of probabilities. We say that the sequence $(G_n)_{n\in\mathbb{N}}$ is *BS-convergent* if the sequence $(p_{n,d})_{n\in\mathbb{N}}$ converges for every d. Benjamini-Schramm convergent sequences of graphs can be associated with an analytic representation called a graphing [18], however, we omit further details here.

Every BS-convergent sequence yields a probability measure on the space $\mathcal{G}^\bullet(\Delta)$ of (not necessarily finite) rooted graphs with maximum degree Δ. The topology on $\mathcal{G}^\bullet(\Delta)$ is generated by clopen sets of rooted graphs with the same d-neighborhood of the root for some d, and the limit probabilities from the definition of the Benjamini-Schramm convergence give a probability measure on the corresponding σ-algebra on $\mathcal{G}^\bullet(\Delta)$ by Carathéodory's Extension Theorem. In what follows, we will just write $\mathcal{G}^\bullet$ instead of $\mathcal{G}^\bullet(\Delta)$ when Δ is clear from the context.

It is not true that every probability measures μ on $\mathcal{G}^\bullet$ corresponds to a BS-convergent sequence of graphs. Let T be the infinite rooted tree where the vertices at even levels (including the root) have degree three and the vertices at odd levels have degree two. If $\Delta = 3$ and $\mu(\{T\}) = 1$, then there is no BS-convergent sequence of graphs corresponding to μ. Indeed, graphs in such a sequence would have almost all vertices of degree three but almost every vertex of degree three would have neighbors of degree two only, which is impossible.

The following condition on a probability measure μ is necessary in order that μ corresponds to a BS-convergent sequence of graphs. First, let μ' be a probability measure on $\mathcal{G}^\bullet$ defined as

$$\mu'(S) = \frac{\int\limits_S \delta(G)\mathrm{d}G}{\int\limits_{\mathcal{G}^\bullet} \delta(G)\mathrm{d}G}$$

where $\delta(G)$ for $G \in \mathcal{G}^\bullet$ is the degree of the root of G (we may assume that $\delta(G) > 0$ with non-zero probability since otherwise μ clearly corresponds to a BS-convergent sequence of graphs). We next define a probability measure μ_e on rooted graphs $\mathcal{G}^e$ with one distinguished edge at the root. Choose a rooted graph

$G \in \mathcal{G}^\bullet$ according to μ' and make randomly one of the edges incident with the root distinguished. This defines the probability measure μ_e on rooted graphs $\mathcal{G}^e$. Another probability measure μ'_e on $\mathcal{G}^e$ can be obtained from μ_e by choosing a random graph $G \in \mathcal{G}^e$ and making the other end of the distinguished edge to be the root. If μ corresponds to a BS-convergent sequence of graphs, then the probability measures μ_e and μ'_e are the same. The conjecture of Aldous and Lyons [1] asserts that this necessary condition is also sufficient for a probability measure μ on $\mathcal{G}^\bullet$ to correspond to a BS-convergent sequence of graphs.

The Benjamini-Schramm convergence have the following drawback. Let us consider a setting of graphs with maximum degree three. A sequence of random cubic graphs of increasing orders is BS-convergent with probability one since the probability that a randomly chosen vertex is contained in a cycle of length k tends to 0 for any fixed integer k. In particular, the corresponding probability measure μ on $\mathcal{G}^\bullet$ satisfies that $\mu(\{T\}) = 1$ for the infinite rooted cubic tree T. By the same argument, a sequence of random cubic bipartite graphs of increasing orders is BS-convergent with probability one. Consequently, any sequence obtained by mixing these two is BS-convergent with probability one. However, the independence number of a random n-vertex cubic graph is bounded away from $n/2$. Hence, the BS-convergence is not robust enough to distinguish bipartite graphs from graphs that are far from being bipartite. Another example of the same phenomenon is the following: consider a BS-convergent sequence of cubic expanders and consider the same sequence where each graph is the union of two copies of the corresponding graph in the former sequence. Any sequence obtained by mixing these two sequences is BS-convergent despite of one of these two sequences being formed by well-connected graphs and the other by disconnected graphs.

To overcome this, a finer notion of convergence called local-global convergence was proposed in [7] and further studied in [27]. This notion of convergence takes into account possible partitions of vertices of graphs in a sequence. Formally, let $\mathcal{G}^\bullet(d, k, \Delta)$ be the set of all rooted k-vertex-colored graphs with maximum degree Δ such that every vertex is at distance at most d from the root (the vertex coloring need not be proper). For a graph G with maximum degree Δ, let $P_{d,k}(G)$ be the set of all vectors from $[0, 1]^{\mathcal{G}^\bullet(d,k,\Delta)}$ that corresponds to the probability distribution on d-neighborhoods for at least one k-vertex-coloring of G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with maximum degree Δ is *local-global convergent* if for every $d \in \mathbb{N}$, $k \in \mathbb{N}$ and $\varepsilon > 0$, there exists n_0 such that the sets $P_{d,k}(G_i)$ and $P_{d,k}(G_j)$ are ε-close for every $i, j \geq n_0$; two sets $X, Y \subseteq \mathbb{R}^m$ are *ε-close* if for every $x \in X$, there exists $y \in Y$ such that $||x-y||_1 \leq \varepsilon$, and for every $y \in Y$, there exists $x \in X$ such that $||x - y||_1 \leq \varepsilon$. If a sequence of graphs is local-global convergent, it is also BS-convergent (set $k = 1$ in the definition) but the converse is not necessarily true because of the examples given in the previous paragraph.

Another notion of convergence related to the BS-convergence is that of the right convergence. Let H be a complete graph with a loop at each vertex such that all its vertices and edges are assigned positive weights. We refer to such a graph H as to a *target*. The number of weighted homomorphisms from a graph G to H, denoted by $\hom(G,H)$, is

$$\sum_{f:V(G)\to V(H)} \prod_{v\in V(G)} w(f(v)) \prod_{vv'\in E(G)} w(f(v)f(v'))$$

where w is the weight function of H. A sequence of graphs $(G_n)_{n\in\mathbb{N}}$ is *right convergent* if the fraction

$$\frac{\log \hom(G_n,H)}{|G_n|}$$

convergences for every target H. It can be shown [9], also see [41], that if a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with bounded maximum degree is right convergent, then it is also BS-convergent. The right convergence can distinguish graphs close to being bipartite and those far from being bipartite but it cannot distinguish connected and disconnected graphs. Let us also remark that the way we define the right convergence is weaker than the original definition, which allows targets with zero weights, but every sequence of graphs that is right convergent in the definition we gave can be modified by changing sublinear number of edges to a sequence of graphs convergent in the original (stronger) definition [8].

Another notion of convergence of sparse graphs, which is fully based on possible partitions, was proposed by Bollobás and Riordan [7]. A *k-partition* of a graph G is a partition of its vertex set into k subsets. The *statistic* of a k-partition $\mathbf{P}=(P_1,\dots,P_k)$ is a vector $s(\mathbf{P})\in\mathbb{R}^{k+\binom{k+1}{2}}$ whose coordinates are the relative sizes $p_i=\frac{|P_i|}{|G}$ of the parts and the edge densities $e_{ij}=\frac{e(P_i,P_j)}{|G|}$ between them, where $e(P_i,P_j)$ stands for the number of edges between parts P_i and P_j. Note that the normalization here is different than the one used in Section 2 when dealing with dense graphs. Let $P_k(G)\subseteq\mathbb{R}^{k+\binom{k+1}{2}}$ be the set of statistics $s(\mathbf{P})$ of at least one k-partition $\mathbf{P}$ of a graph G. A sequence of graphs $(G_n)_{n\in\mathbb{N}}$ with bounded maximum degree is *partition convergent* if for every $k\in\mathbb{N}$ and $\varepsilon>0$, there exists n_0 such that the sets $P_k(G_i)$ and $P_k(G_j)$ are ε-close for all $i,j\ge n_0$.

Let $U_2\subseteq\mathbb{R}^5$ be the set of all non-negative real vectors $(p_1,p_2,e_{11},e_{12},e_{22})$ such that $p_1+p_2=1$, $p_1=e_{11}+e_{12}/2$ and $p_2=e_{22}+e_{12}/2$. Observe that $P_2(G)\subseteq U_2$ for every 2-regular graph G.

Exercise 12. *Show that for every $\varepsilon>0$, there exists n_0 such that the sets $P_2(nC_4)$ and $P_2(nC_6)$ are ε-close to U_2 for every $n\ge n_0$.*

Exercise 12 can be generalized to show that for every $k\in\mathbb{N}$ and every $\varepsilon>0$, there exists n_0 such that the sets $P_k(nC_4)$ and $P_k(nC_6)$ are ε-close [8]. In particular, the sequence of graphs containing nC_4 and nC_6 in a mixed way is partition convergent. Consequently, the partition convergence does not imply any of the earlier mentioned notions of convergence. On the other hand, the local

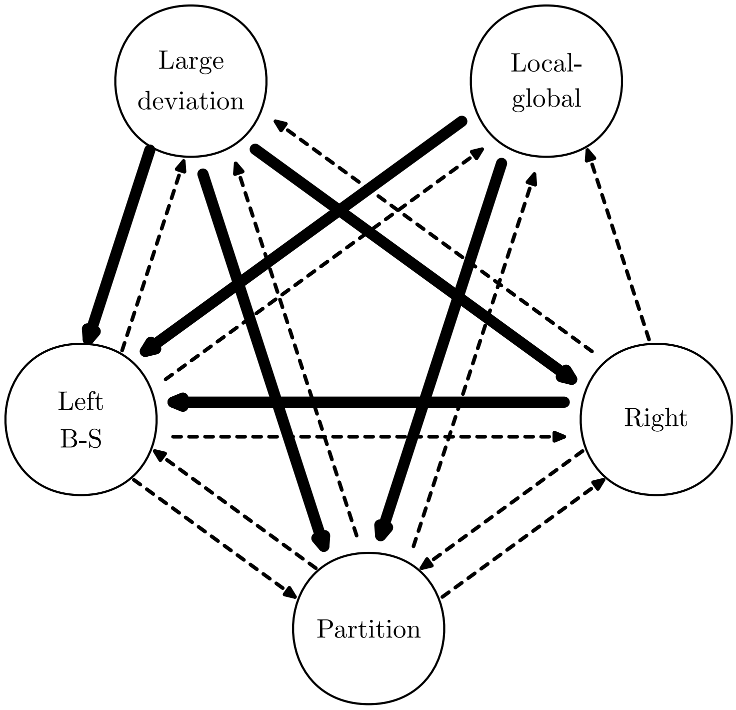

Figure 11: The relation between the presented notions of convergence of bounded degree graphs. The bold arrows represent that the notion implies the other and the dashed arrows that this is not the case in general. When an arrow is missing, the relation between the notions is not known.

global convergence trivially implies the partition convergence but the connected vs. disconnected example yield that neither the BS-convergence nor the right convergence implies the partition convergence. We refer the reader to Figure 11 for the relation between the notions of convergence of sparse that we have already introduced and the notion of large deviation convergence that we introduce next.

The recent notion of large deviation convergence, which was introduced in [8], is a common refinement of the right convergence and partition convergence. A sequence of graphs $(G_n)_{n\in\mathbb{N}}$ with bounded maximum degree is *LD-convergent* if the following limit exists (while possibly being infinite)

$$I_k(x) = \lim_{\varepsilon\to 0}\lim_{n\to\infty} -\frac{\log \frac{|\{\mathbf{P} \text{ such that } ||s(\mathbf{P})-x||_1\leq\varepsilon\}|}{k^{|G_n|}}}{|G_n|}$$

for every k and $x \in \mathbb{R}^{k+\binom{k+1}{2}}$. Note that $I_k(x) \in [0, \log k] \cup \{\infty\}$. On the intuitive level, one can think that the number of k-partitions of G_n, if n is large, with statistic close to x is approximately $k^{|G_n|} \cdot e^{-I_k(x)|G_n|}$. If a sequence $(G_n)_{n\in\mathbb{N}}$ is LD-convergent, then it is also partition convergent. Indeed, $P_k(G_n)$ is ε-close to the set $\{x \mid I_k(x) < \infty\}$ when n is sufficiently large. A more involved argument shows that every LD-convergent sequence of graphs is right convergence [8], which implies that it is also BS-convergent. However, there exist LD-convergent sequences of graphs that are not right convergent when targets with zero weights are allowed (e.g., the sequence of cycles of increasing lengths where the parities of the cycle lengths alternate).

The last notion that we would like to mention is the notion of the first order convergence introduced in [42, 43] and further studied in [14, 20, 32, 44]. This notion is an attempt to provide a universal notion of graph convergence that can be applied both in the sparse and in the dense settings. If ψ is a first order formula with k free variables and G is a (finite) graph, then the *Stone pairing* $\langle\psi, G\rangle$ is the probability that a uniformly chosen k-tuple of vertices of G satisfies ψ. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *first order convergent* if the limit $\lim\limits_{n\to\infty} \langle\psi, G_n\rangle$ exists for every first order formula ψ. It is not hard to show that every first order convergent sequence of dense graphs is convergent in the sense defined in Section 2 and every first order convergent sequence of graphs with bounded maximum degree is Benjamini-Schramm convergent. Neither of the opposite implications is true. Some of first order convergent sequence graphs can be represented by an analytic object called a *modeling* but not every first order convergent sequence of graphs has such a representation [43]. A conjecture of Nešetřil and Ossona de Mendez [43] asserts that if $\mathcal{G}$ is a nowhere-dense class of graphs, then any first order convergent sequence of graphs from $\mathcal{G}$ can be represented by a modeling.

Appendix: Probability theory tools

In this section, we provide a brief overview of the most important results from the probability theory that are used in the lecture notes. We start with Borel-Cantelli lemma.

Lemma 16 (Borel-Cantelli lemma)**.** *Let $(E_n)_{n\in\mathbb{N}}$ be a sequence of probability events. If the sum of probabilities of E_n, $n \in \mathbb{N}$, is finite, i.e.,*

$$\sum_{n\in\mathbb{N}} \mathbb{P}(E_n) < \infty \ ,$$

then the probability that infinitely many of the events occur is zero.

Several results in the lecture notes use the notion of martingales. Fix a probability space Ω and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random real variables on Ω. The sequence $(X_n)_{n\in\mathbb{N}}$ forms a *martingale* if the expected value of each X_n is equal to a real number X_0 and

$$\mathbb{E}(X_{n+1}|X_1, \dots, X_n) = X_n \text{ for every } n \in \mathbb{N},$$

i.e., the expected value of X_{n+1} conditioned on the values of $X_1, \dots, X_n$ is the value of X_n. With a slight abuse of notation, X_0 can be understood to be the random variable on Ω equal to X_0 everywhere. Martingales can be used to bound the probability of a large deviation of X_n from its expected value.

Theorem 17 (Azuma-Hoeffding inequality)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a martingale with $\mathbb{E}X_n = X_0$, and let $(c_n)_{n\in\mathbb{N}}$ be a sequence of reals. If holds for every $n \in \mathbb{N}$ that $|X_n - X_{n-1}| \le c_n$ with probability one, then*

$$\mathbb{P}\left(|X_n - X_0| \ge t\right) \le 2e^{\frac{-t^2}{2\sum_{k=1}^{n} c_k^2}}$$

for every $n \in \mathbb{N}$ and every $t \in \mathbb{R}$.

We also need the following corollary of Doob's martingale convergence theorem.

Corollary 18. *Let $(X_n)_{n\in\mathbb{N}}$ be a martingale on a probability space Ω with probability μ. If there exists $K \in \mathbb{R}$ such that $\mathbb{E}|X_n| < K$, then there exists a random variable X on Ω such that*

$$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$

for μ-almost all $\omega \in \Omega$.

Acknowledgment

The work of the author on the topics covered in these lecture notes has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No 648509). This publication reflects only its author's view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains.

References

[1] D. Aldous and R. Lyons: *Processes on unimodular random networks*, Electron. J. Probab. **12** (2007), no. 54, 1454–1508.

[2] R. Baber: *Turán densities of hypercubes*, preprint available as arXiv 1201.3587.

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

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

[5] 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.

[6] I. Benjamni and O. Schramm: *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001), no. 23, 1–13.

[7] B. Bollobás and O. Riordan: *Sparse graphs: Metrics and random models*, Random Structures Algorithms **39** (2011), 1–38.

[8] C. Borgs, J.T. Chayes and D. Gamarnik: *Convergent sequences of sparse graphs: A large deviations approach*, available as arXiv:1302.4615.

[9] C. Borgs, J.T. Chayes, J. Kahn and L. Lovász: *Left and right convergence of graphs with bounded degree*, Random Structures Algorithms **42** (2013), 1–28.

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

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

[13] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, B. Szegedy and K. Vesztergombi: *Graph limits and parameter testing*, in: Proc. 38rd Annual ACM Symposium on the Theory of Computing (STOC), ACM, New York, 2006, 261–270.

[14] D. Christofides and D. Král': *First order convergence and roots*, Combin. Probab. Comput. **25** (2016), 213-221.

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

[16] J. W. Cooper, T. Kaiser and D. Král' and J. A. Noel: *Weak regularity and finitely forcible graph limits*, available as arXiv:1507.00067.

[17] J. W. Cooper, D. Král' and T. Martins: *Finite forcibility and computability of graphons*, in preparation.

[18] G. Elek: *On limits of finite graphs*, Combinatorica **27** (2007), 503–507.

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

[20] J. Gajarský, P. Hliněný, T. Kaiser, D. Král', M. Kupec, J. Obdržálek, S. Ordyniak and V. Tůma: *First order limits of sparse graphs: Plane trees and path-width*, to appear in Random Structures Algorithms.

[21] R. Glebov, A. Grzesik, T. Klimošová and D. Král': *Finitely forcible graphons and permutons*, J. Combin. Theory Ser. B **110** (2015), 112-135.

[22] R. Glebov, T. Klimošová and D. Král': *Infinite dimensional finitely forcible graphon*, available as arXiv:1404.2743.

[23] R. Glebov, D. Král' and J. Volec: *Compactness and finite forcibility of graphon*, available as arXiv:1309.6695.

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

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

[26] 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.

[27] H. Hatami, L. Lovász and B. Szegedy: *Limits of local-global convergent graph sequences*, to appear in Geom. Funct. Anal..

[28] H. Hladký, A. Mathé, V. Patel and O. Pikhurko: *Poset limits can be totally ordered*, to appear in Trans. Amer. Math. Soc.

[29] C. Hoppen, Y. Kohayakawa, C.G. Moreira, B. Ráth and R.M. Sampaio: *Limits of permutation sequences*, J. Combin. Theory Ser. B **103** (2013), 93–113.

[30] C. Hoppen, Y. Kohayakawa, C.G. Moreira and R.M. Sampaio: *Limits of permutation sequences through permutation regularity*, preprint available as arXiv:1106.1663.

[31] S. Janson: *Poset limits and exchangeable random posets*, Combinatorica **31** (2011), 529–563.

[32] F. Kardoš, D. Král', A. Liebenau and L. Mach: *First order convergence of matroids*, available as arXiv:1501.06518.

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

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

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

[36] L. Lovász: *Large networks and graph limits*, AMS, Providence, RI, 2012.

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

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

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

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

[41] L. M. Lovász: *A short proof of the equivalence of left and right convergence for sparse graphs*, European J. Combin. **53** (2016), 1–7.

[42] J. Nešetřil and P. Ossona de Mendez: *A model theory approach to structural limits*, Comment. Math. Univ. Carolin. 53 (2012), 581–603.

[43] J. Nešetřil and P. Ossona de Mendez: *A unified approach to structural limits, and limits of graphs with bounded tree-depth*, available as arXiv:1303.6471.

[44] J. Nešetřil and P. Ossona de Mendez: *Modeling limits in hereditary classes: reduction and application to trees*, available as arXiv:1312.0441.

[45] O. Pikhurko and A. Razborov: *Asymptotic structure of graphs with the minimum number of triangles*, preprint available as arXiv:1204.2846.

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

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

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

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

[50] A. Thomason: *Pseudo-random graphs*, North-Holland Math. Stud. **144** (1987), 307–311.

[51] A. Thomason: *Random graphs, strongly regular graphs and pseudorandom graphs*, London Math. Soc. Lecture Note Ser. **123** (1987), 173–195.