Analytic representations of large graphs*

Andrzej Grzesik† Daniel Král'‡

Abstract

The recently emerged theory of graph limits provides analytic tools to represent and analyse large graphs, which appear in various scenarios in mathematics and computer science. We survey basic concepts concerning dense graph limits and then focus on recent results on finitely forcible graph limits. We conclude with presenting some of the existing notions concerning sparse graph limits and discussing their mutual relation.

1 Introduction

Large graphs appear as representations of huge networks in many different areas of life. One should mention in particular the internet network of hyperlinks, acquaintance graphs of social networks, etc. Since such graphs are often too huge to be examined by standard graph theoretic or algorithmic approaches, there has been a need for developing tools specifically for large graphs that can be used to gain some information from local sampling, studying global properties, or observing the behaviour of various processes on the graph through a longer time interval. In this short survey, we present analytic tools for representing and analysing large graphs provided by the theory of graph limits as a response to these new challenges.

*The work of both authors has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 648509). This publication reflects only its authors' view; the European Research Council Executive Agency is not responsible for any use that may be made of the information it contains.

†Faculty of Mathematics and Computer Science, Jagiellonian University, Łojasiewicza 6, 30-348 Kraków, Poland. E-mail: Andrzej.Grzesik@uj.edu.pl.

‡Faculty of Informatics, Masaryk University, Botanická 68A, 602 00 Brno, Czech Republic, and Mathematics Institute, DIMAP and Department of Computer Science, University of Warwick, Coventry CV4 7AL, UK. E-mail: dkral@fi.muni.cz.

The theory of graph limits has highlighted 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 graphs and sparse graphs, which is also reflected in the way that this survey is structured. For many applications, the concept of a convergence of a sequence of graphs, without explicitly defining an analytic object representing its limit, could be sufficient. However, a better understanding can often be gained if an analytic object that properly captures the interplay of local and global parameters is available. In our exposition, we will be concerned with the convergence and limit representations of graphs. However, many of the results presented further can be translated to other discrete objects, e.g., permutations [37, 53, 54, 63] or partial orders [51, 55]. We also refer the reader to a recent monograph by Lovász [67], where the theory of graph limits is treated in a more detailed and thorough way.

In this survey, we are primarily concerned with results on limits of dense graphs, i.e., graphs where the number of edges is quadratic in its number of vertices. The foundations of the theory of dense graph limits were laid in a series of papers by Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi [15–17,69,70]. In Section 3, we survey basic concepts concerning limits of dense graphs, and we then focus on the uniqueness of the limit structures in Section 4. Limits of dense graphs turned out to be very useful with respect to applications in extremal combinatorics. In particular, the closely related flag algebra method, which was introduced by Razborov [84], enables the use semidefinite programming to search for bounds on problems studied in extremal graph theory. Using this method, Razborov [85] solved the famous problem, which dates back to the work of Rademacher in the 1940's and Erdős in the 1950's, on the minimum possible density of triangles in a graph with a given edge density. This result was later generalized by Nikiforov [81] and by Reiher [87] using similar but finer techniques from triangles to larger complete graphs. It should be emphasized that the flag algebra method can also be used in relation to other combinatorial objects such as directed graphs, hypergraphs, permutations, etc. The method has seen many profound applications and resulted in substantial progress on many long standing open problems in extremal combinatorics, e.g. [5–8, 45, 48, 49, 59, 61, 82, 83, 86].

Another application of graph limits that we would like to mention here belongs to computer science. A property testing algorithm is an algorithm that determines with high probability a property or approximates a parameter of a large input based on a constant size sample; such algorithms started to be systematically studied in the 1990's [41–43, 88], also see, e.g., the surveys in [40]. The theory of graph limits led to an analytic characterization of properties and parameters that can be computed in this way [67, 70]. In particular, it is possible to define a notion of distance, which is called cut distance, between large graphs of not necessarily the same order; this notion extends to the setting of graphons representing graph limits. Inputs that are close in the cut distance cannot be distinguished using property testing algorithms. The converse, which is also true, can be exploited to provide a characterization of properties and parameters amenable to such algorithms [67, 70].

The area of limits of sparse graphs, such as graphs of bounded degree, is less developed than the area of limits of dense graphs. Several notions of convergence of such graphs were proposed and 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 neighbourhood 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 concerning the sparse graph convergence, including the conjecture of Aldous and Lyons, in Section 5.

2 Preliminaries

In this section, we introduce basic notation used throughout the paper. The set of all positive integers is denoted by $\mathbb{N}$, the set of all non-negative integers by $\mathbb{N}_0$, and the set of integers between 1 and k (inclusive) by $[k]$. All measures considered in this paper are Borel measures on $\mathbb{R}^d$, $d \in \mathbb{N}$. If a set $X \subseteq \mathbb{R}^d$ is measurable, then we write $|X|$ for its measure, and if X and Y are two measurable sets, then we write $X \sqsubseteq Y$ if $|X \setminus Y| = 0$.

All graphs considered in this paper are simple graphs without loops. If G is a graph, we write $|G|$ for its order, i.e., the number of its vertices, and $||G||$ for its size, i.e., the number of its edges.

For completeness, we next give a brief overview of results from the probability theory that we particularly need in our exposition; we refer the reader to, e.g., [3] for further details. We start with the Borel-Cantelli lemma.

Lemma 1 (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 E_n occur is zero.

We next define a notion of a martingale. 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. For a martingale $(X_n)_{n\in\mathbb{N}}$, we can bound the probability of a large deviation of X_n from its expected value.

Theorem 2 (Azuma-Hoeffding inequality)**.** *Let $(X_n)_{n\in\mathbb{N}}$ be a martingale with $\mathbb{E}X_n = X_0$ for all $n \in \mathbb{N}$, and let $(c_n)_{n\in\mathbb{N}}$ be a sequence of reals. If it 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 $t \in \mathbb{R}$.

Finally, we will need the following corollary of Doob's Martingale Convergence Theorem.

Corollary 3. *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 $\mathbb{E}|X_n| < K$ for every $n \in \mathbb{N}$, then there exists a random variable X on Ω such that*

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

for μ-almost all $\omega \in \Omega$.

3 Dense graph limits

In this section, we are primarily concerned with limits of dense graphs, i.e., graphs where the number of edges is quadratic in the number of vertices. 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. 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 sequences of graphs include the sequence of complete graphs K_n, the sequence of complete bipartite graphs $K_{n,n}$ with parts of equal size and the sequence of complete bipartite graphs $K_{\lfloor \alpha n \rfloor,n}$ for $\alpha \in (0,1)$. 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 the Borel-Cantelli lemma (Lemma 1) and the Azuma-Hoeffding inequality (Theorem 2).

Assume now that $(G_n)_{n\in\mathbb{N}}$ is a sequence of sparse graphs, which means that the number of edges of G_n is $o(|G_n|^2)$, i.e.,

$$\lim_{n\to\infty} \frac{||G_n||}{|G_n|^2} = 0.$$

Consequently, 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 in the sense that we have defined earlier. Consequently, it holds that any sequence of sparse graphs is convergent regardless of its structure. This is the reason why the notion of convergence that we have just defined is of interest for dense graphs. Notions of convergence appropriate for sparse graphs will be described in Section 5.

Another way of defining convergent sequences of graphs is to consider homomorphic densities of graphs. If G and H are two graphs, the *homomorphic density* of H in G, which is denoted by $t(H,G)$, is the probability that a random map from the vertex set of H to the vertex set of G is a homomorphism from H to G. A simple application of the Principle of Inclusion and Exclusion shows that $d(H,G)$ is determined by the values of $t(H',G)$ for all spanning subgraphs H' of H, and $t(H,G)$ is determined by the values of $d(H',G)$ for all supergraphs H' of H with the same number of vertices. Hence, a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is convergent if and only if if the sequence of homomorphic densities $t(H,G_n)$ converges for every graph H.

We next introduce an analytic object that is used to represent a convergent sequence of graphs. This object is 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 as we will see in the following. The

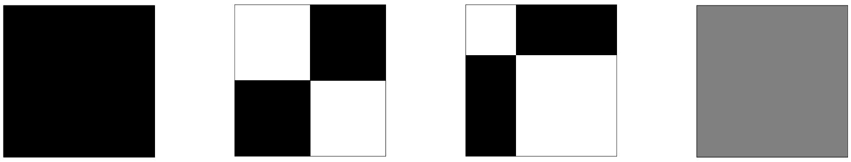

Figure 1: Graphons that are limits of the sequences $(K_n)_{n\in\mathbb{N}}$, $(K_{n,n})_{n\in\mathbb{N}}$, $(K_{n,2n})_{n\in\mathbb{N}}$ and $(G(n,1/2))_{n\in\mathbb{N}}$.

analogy with adjacency matrices also motivates some of the definitions that follow.

A graphon can be viewed as a recipe for creating a random graph as we know present. 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 $G(n,p)$. Graphons are usually depicted in the unit square with values being different shades of gray, where white represents zero and black represents one. The origin of the coordinate system is usually in the top left corner to follow the analogy with adjacency matrices. An example of such visualization can be found in Figure 1.

We now relate graphons to convergent sequences of graphs. Let the *density* of a graph H in a graphon W be the probability that the W-random graph of order $|H|$ is isomorphic to H; this probability is denoted by $d(H,W)$. It can be shown that 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. Examples of graphons that are limits of some simple convergent sequences of graphs are given in Figure 1. In what follows, we will also consider a special type of graphons called step graphons: a graphon W is a *step graphon* if there exist an integer k and a partition of $[0,1]$ into k measurable sets $A_1,\dots,A_k$ such that the graphon W is constant on $A_i\times A_j$ for all $i,j\in[k]$. Examples of step graphons can be found in Figure 2.

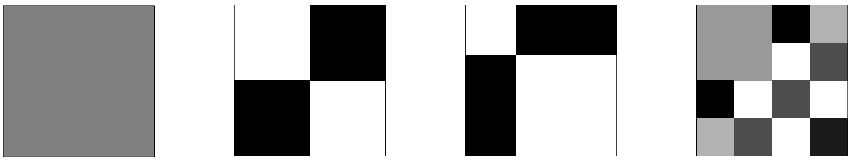

Figure 2: Examples of step graphons.

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 latter of these questions, which is simpler to answer, and we discuss the former of the questions later in this section.

Theorem 4. *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, $i = 0, \ldots, n$, 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 the W-random graph of order n. 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 the Azuma-Hoeffding inequality (Theorem 2) with $c_i \leq n^{|H|-1}$ and get that

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

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

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

which yields that

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

if $n \geq 2|H|$. The 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. □

Proving that there exists a limit graphon for every convergent sequence of graphs is harder. We will present here the proof by Lovász and Szegedy from [69]. The proof uses weak regularity of graphs introduced by Frieze and Kannan in [35]; this notion is weaker than the more well-known notion of Szemerédi regularity. However, it is simpler and sufficient for our purposes. If G is a graph and S and T are two subsets of its vertices, then $e(S,T)$ denotes the number of pairs of vertices $s \in S$ and $t \in T$ joined by an edge and $d(S,T)$ denotes the corresponding density, i.e., $d(S,T) = \frac{e(S,T)}{|S|\cdot|T|}$. 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 $i, j \in [k]$, and it is *weak 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 [35] proved the following theorem.

Theorem 5. *For every $\varepsilon \in (0,1)$, there exists $K = 2^{O(\varepsilon^{-2})}$ such that every graph G has a weak ε-regular partition with at most K parts.*

We will need a strengthening of Theorem 5, whose proof follows the same lines as the proof of Theorem 5. We say that 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 $V'_j \subseteq V_i$.

Theorem 6. *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 weak ε-regular partition with at most $K \cdot k$ parts.*

Finally, weak regular partitions are related to subgraph densities as follows [35].

Theorem 7. *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 $V_1, \ldots, V_k$ is a weak ε-regular partition of its vertex set, then*

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

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

We are now ready to prove that every convergent sequence of graphs can be represented by a graphon.

Theorem 8 (Lovász and Szegedy [69])**.** *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 weak 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 5. Suppose that we have already fixed a weak 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 6, there exists a weak $\varepsilon_{\ell+1}$-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 5 and 6 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 the following holds for every $\ell \in \mathbb{N}$:

- 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 $k_{n,\ell}$ for the 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 7 implies that the following holds for every graph H:

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

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 3 implies that there exists a measurable function W 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). □

The proof that we have presented here is not the only proof of the existence of a limit graphon of a convergent sequence of graphs that is known. The existence of the limit graphon can be derived from the representation theorem on symmetrically exchangeable random variables due to Aldous [2] and Hoover [52] and further developed by Kallenberg [56]; see [4, 27] for further details. Another way of proving the existence of a limit graphon is using the arguments concerning a suitable measure space defined using the ultraproduct of graphs in the sequence as presented by Elek and Szegedy in [32]. More recently, another approach was given by Doležal, Grebík, Hladký, Rocha and Rozhoň [28–30]: in a certain sense, they consider all weak* accumulation points of zero-one step graphons associated with the graphs in the sequence and define a certain "structuredness" order on them such that the most structured points are limit graphons.

While graphons were originally developed to represent large graphs, there are various mathematical properties of graphons that are of their own interest to study. Among many such properties, we would like to mention the notion of weakly norming graphs and relate it to one of the most important open problems in extremal graph theory—Sidorenko's Conjecture. This beautiful conjecture of Erdős and Simonovits [90] and of Sidorenko [89] asserts, in the language of graphons, that $t(K_2, W)^{\|H\|} \le t(H, W)$ for every bipartite graph H and every graphon W and every graphon W, i.e., a quasirandom graph minimizes the density of H among all graphs with the same edge density. Sidorenko [89] confirmed the conjecture for trees, cycles and bipartite graphs with one of the sides having at most three vertices; it is interesting that the case of paths is equivalent to the Blakley-Roy inequality for matrices, which was proven in [10]. Additional graphs were added to the list of graphs satisfying the conjecture by Conlon, Fox and Sudakov [21], by Hatami [47], and by Szegedy [93]. More general results concerning recursively described classes of bipartite graphs were obtained by Conlon, Kim, Lee and Lee [22], by Kim, Lee and Lee [58], by Li and Szegedy [64] and by Szegedy [92]. In particular, Szegedy [92] has described a class of graphs called thick graphs that satisfy the conjecture. More recently, Conlon and Lee [24] showed that the conjecture is satisfied by bipartite graphs such that one of the parts has many vertices of maximum degree. Sidorenko's Conjecture is also known to hold in the local sense [67, Proposition 16.27], i.e., it holds for graphons W close to the constant graphon; a stronger statement with uniform quantitative bounds has recently been proven by Fox and Wei [34].

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. A stronger notion of norming graphs concerns a generalization of graphons to functions on $[0, 1]^2$ that do not need to be non-negative on $[0, 1]^2$, however, we prefer not deviating from the main topic of our survey and we avoid giving further details here. It is easy to show that every weakly norming graph satisfies Sidorenko's conjecture, and some results on Sidorenko's conjecture actually deal with this stronger property of graphs. Hatami [47] characterized weakly norming graphs as those satisfying a certain Hölder-type inequality involving graphs edge-decorated by graphons; also see [23, 62] for additional results on weakly norming graphs. However, it is interesting that the property of being weakly norming is equivalent to a generalization of the property concerned in Sidorenko's conjecture. To state this link precisely, we need several definitions. Let $\mathcal{P} = \{J_1, \ldots, J_k\}$ be a partition of the interval $[0,1]$ into non-null measurable sets. If W is a graphon, then the graphon $W^{\mathcal{P}}$ is defined as the average on the parts in $\mathcal{P}$, i.e.,

$$W^{\mathcal{P}}(x,y) = \frac{1}{|J_i| \cdot |J_j|} \int_{J_i \times J_j} W(s,t) \mathrm{d}s \mathrm{d}t$$

where J_i and J_j are the unique parts from $\mathbb{P}$ such that $x \in J_i$ and $y \in J_j$. We say that a graph H has the *step Sidorenko property* if $t(H, W^{\mathcal{P}}) \leq t(H, W)$ for every graphon W and every finite partition $\mathcal{P}$. Considering the partition $\mathcal{P}$ with a single part implies that every graph that has the step Sidorenko property satisfies Sidorenko's conjecture. The converse is not true; the graph obtained from C_4 by adding a new vertex adjacent to one of the vertices of the cycle is known to satisfy Sidorenko's conjecture but does not have the step Sidorenko property [62]. However, a graph H is weakly norming if and only if H has the step Sidorenko property. The proof of one of the implications can be found in [67, Proposition 14.13] and the other implication has recently been proven by Doležal et al. in [28].

We conclude this section by describing an analytic object representing k-uniform hypergraphs. While it may be natural to expect this object to be a function from $[0,1]^k$ to $[0,1]$, the situation is more complex for the same reasons why graph regularity does not straightforwardly generalize to the setting of hypergraphs. A *k-hypergraph* is a hypergraph where every edge contains exactly k vertices. In the analogy to graphs, the *density* of an ℓ-vertex hypergraph H in a hypergraph G is the probability that a randomly chosen subset of ℓ vertices of G induces a subhypergraph isomorphic to H. A sequence $(G_n)_{n \in \mathbb{N}}$ of hypergraphs is *convergent* if the density of every hypergraph H in the hypergraphs G_n converges.

We next define an analytic object, which we call k-hypergraphon. A *k-hypergraphon* is a measurable function W from $[0,1]^{2^k-2}$ to $[0,1]$ such that the $2^k - 2$ variables of W are associated with the $2^k - 2$ proper subsets of $[k]$ and satisfy that $W(\vec{x}) = W(\pi(\vec{x}))$ for every $\vec{x} \in [0,1]^{2^k-2}$ and every permutation $\pi \in S_k$, where $\pi(J) = \{\pi(j), j \in J\}$ for $J \subseteq [k]$. Observe that the definition of a 2-hypergraphon coincide with the definition of a graphon. Given a k-hypergraphon W, we may define a W-random k-hypergraph of order n as follows. Fix n vertices and assign to every ℓ-element subset of vertices, $\ell \in [k-1]$, independently and uniformly a number from the unit interval $[0,1]$. The vertices $v_1, \ldots, v_k$ form an edge with probability $W(\vec{x})$ where the coordinate of x associated with a $J \subseteq [k]$, $J \notin \{\emptyset, [k]\}$, is equal to the number assigned to the $|J|$-tuple of vertices $\{v_j, j \in J\}$. Again, we define the density of a k-hypergraph H with n vertices as the probability that a W-random k-hypergraph of order n is isomorphic to H, and say that

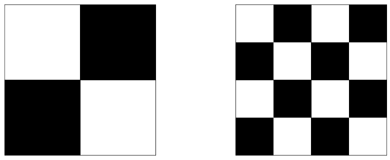

Figure 3: Two weakly isomorphic graphons.

a k-hypergraphon W is a limit of a convergent sequence of k-hypergraphs if the density of every k-hypergraph H in W is the limit density of H in the sequence. The existence of a limit k-hypergraph for every convergent sequence of k-hypergraphs was established by Elek and Szegedy [32] using the ultraproduct argument that we have mentioned earlier in relation to graphons, however, the existence of a limit hypergraphon can also be proven using arguments similar to those that we have presented in the graph setting earlier as shown by Zhao [96].

4 Finite forcibility

In this section, we will discuss in what sense a limit graphon of a convergent sequence of graphs is unique and when its structure is determined by finitely many densities. 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 bipartite graphs with parts of equal sizes.

The following is a general way of constructing weakly isomorphic graphons. Let φ be a measure preserving map from $[0, 1]$ to $[0, 1]$, i.e., $|\varphi^{-1}(A)| = |A|$ for every measurable subset A of $[0, 1]$. If W is a graphon, we define a graphon W^φ by setting $W^\varphi(x, y) = W(\varphi(x), \varphi(y))$. A standard measure theory argument yields that $d(H, W) = d(H, W^\varphi)$, i.e., the graphons W and W^φ are weakly isomorphic. For example, consider the following measure preserving map:

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

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

Borgs, Chayes and Lovász [14], also see [67, Chapter 13] for further discussion, have shown that the above way of constructing weakly isomorphic graphons is in a certain sense the only way of obtaining weakly isomorphic

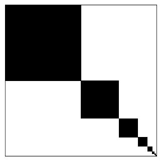

Figure 4: A graphon that is not finitely forcible.

graphons (note that the following two theorems are not obviously equivalent since the maps φ_1 and φ_2 need not be bijective).

Theorem 9. *If W_1 and W_2 are weakly isomorphic graphons, then there exist a graphon W and measure preserving maps φ_1 and φ_2 such that the graphons W^{φ_1} and W_1 are equal almost everywhere, and W^{φ_2} and W_2 are equal almost everywhere.*

Theorem 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.*

In general, it is necessary to know the densities $d(G, W)$ of all graphs G in a graphon W to know the structure of W. For example, see [37] for a more detailed discussion, if W is the graphon depicted in Figure 4, then for every finite set $\mathcal{G}$ of graphs, there exists a graphon W' such that $d(H, W) = d(H, W')$ for every $H \in \mathcal{G}$ but W and W' are not weakly isomorphic, i.e., there exists a graph H' such that $d(H, W) \neq d(H', W)$. On the other hand, the classical results on quasirandom graphs due to Thomasson [94] and Chung, Graham and Wilson [19] yield that if a graphon W satisfies that $t(K_2, W) = p$ and $t(C_4, W) = p^4$ for some $p \in [0, 1]$, then W is equal to p almost everywhere. In particular, there are graphons such that their structure is determined by finitely many densities. In the rest of the section, we will be interested in such graphons.

The ideas presented in the previous paragraph leads to the following definition: a graphon W is *finitely forcible* if there exists a finite set $\mathcal{G}$ of graphs such that any graphon W' satisfying $d(H, W) = d(H, W')$ for every graph $H \in \mathcal{G}$ is weakly isomorphic to W; such a set $\mathcal{G}$ is called a *forcing family* of W. In particular, the constant graphon is finitely forcible and its forcing family is $\{K_2, C_4, K_4 \setminus e, K_4\}$ (note that $t(C_4, W)$ is determined by $d(C_4, W)$, $d(K_4 \setminus e, W)$ and $d(K_4, W)$), and the graphon given in Figure 4 is not finitely forcible. In fact, finitely forcible graphons are rather rare in the sense that the set of finitely forcible graphons is of the first category in the space $L^2([0, 1]^2)$ as shown by Lovász and Szegedy [71].

The study of finitely forcible graphons is motivated by the link to extremal combinatorics captured in the following (folklore) proposition.

Proposition 11. *Let W_0 be a finitely forcible graphon. There exists a linear combination of subgraph densities such that W_0 is its unique (up to a weak isomorphism) minimizer, 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.

Examples of finitely forcible graphons include many graphons that appear as optimal solutions of problems in extremal graph theory. For example, Lovász and Sós [68], also see [91], showed that every step graphon is finitely forcible. A more systematic study of finitely forcible graph limits was initiated by Lovász and Szegedy in [71]. In particular, they showed that if p is a polynomial in x and y such that the function $W : [0,1]^2 \to [0,1]$ defined as

$$W(x,y) = \begin{cases} 1 & \text{if } p(x,y) \geq 0, \text{ and} \\ 0 & \text{otherwise,} \end{cases}$$

is symmetric, then W is a finitely forcible graphon.

Inspired by the known examples of finitely forcible graphons, Lovász and Szegedy [71] conjectured that all finitely forcible graphons posses a simple structure in the sense that we now describe. To state this precisely, we need the following definition. For a graphon W and $x \in [0,1]$, define a function $f_x^W : [0,1] \to [0,1]$ to be

$$f_x^W(y) := W(x,y).$$

Since the function f_x^W belongs to $L^1([0,1])$ for almost every $x \in [0,1]$, the graphon W naturally defines a probability measure μ on $L^1([0,1])$ [71]. The space $T(W)$ is formed by the support of the measure μ equipped with the topology inherited from $L^1([0,1])$, and is referred to as the *space of typical vertices* of W. A vertex x of the graphon W is called *typical* if $f_x^W \in T(W)$. Lovász and Szegedy [71, Conjectures 9 and 10] conjectured the following; we cite both conjectures verbatim.

Conjecture 1. *If W is a finitely forcible graphon, then $T(W)$ is a compact space. (We can't even prove that $T(W)$ is locally compact.)*

Conjecture 2. *If W is a finitely forcible graphon, then $T(W)$ is finite dimensional. (We intentionally do not specify which notion of dimension is meant here—a result concerning any variant would be interesting.)*

The interest in Conjecture 2 comes from the following link to weak regularity partitions of graphons. It is possible to define a different notion of the space of typical vertices of a graphon as follows. If f and g are two functions from $L^1([0,1])$, we define

$$d_W(f,g) := \int_{[0,1]} \left| \int_{[0,1]} W(x,y)(f(y)-g(y))\mathrm{d}y \right| \mathrm{d}x \,,$$

and refer to $d_W(f,g)$ as the *similarity distance* of the functions f and g. Note that the similarity distance d_W depends on the graphon W. The space $\overline{T}(W)$ is formed by the closure (with respect to d_W) of the support of the measure μ, which we have defined earlier, equipped with the topology given by the metric d_W. The structure of the space $\overline{T}(W)$ is related to weak regularity partitions of W as follows [67, Chapter 13]: if the Minkowski dimension of $\overline{T}(W)$ (with respect to the metric d_W) is d, then W has a weak ε-regular partition with $O(\varepsilon^{-d})$ parts for every $\varepsilon > 0$. Note that the number of parts of a weak ε-regular partition may need to be $2^{\Theta(\varepsilon^{-2})}$, and this is the best possible as shown by Conlon and Fox [20].

Conjectures 1 and 2 were disproved in [39] and [38], respectively. More specifically, a construction of a finitely forcible graphon W such that $T(W)$ fails to be locally compact was given in [39] (the graphon can be found in Figure 5) and a construction of a finitely forcible graphon W such that $T(W)$ contains a space homeomorphic to $[0,1]^{\mathbb{N}}$ in [38] (this graphon is depicted in Figure 6). A stronger counterexample to Conjecture 2 was given in [25], where the authors constructed a finitely forcible graphon W such that any weak ε-regular partition must have a number of parts almost exponential in ε^{-2} for infinitely many $\varepsilon > 0$, which is close to the general lower bound. This graphon can be found in Figure 7. This line of research culminated with the following general result of Cooper et al. [26] (the graphon is visualized in Figure 8), which we state as Theorem 12. To state the result, we need the following definition: if W_1 and W_2 are two graphons and $X \subseteq [0,1]$ a non-null measurable set, then we say that W_1 is a *subgraphon* of W_2 induced by X if there exist measure-preserving maps $\varphi_1 \colon X \to [0,|X|)$ and $\varphi_2 \colon X \to X$ such that

$$W_1\left(|X|^{-1}\cdot\varphi_1(x), |X|^{-1}\cdot\varphi_1(y)\right) = W_2\left(\varphi_2(x),\varphi_2(y)\right)$$

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

Theorem 12. *For every graphon W_F, there exists a finitely forcible graphon W_0 such that W_F is a subgraphon of W_0 induced by a $1/14$ fraction of the vertices of W_0.*

Theorem 12 provides a universal framework for constructing finitely forcible graphons with very complex structure, including counterexamples to Conjectures 1 and 2. In view of Proposition 11, Theorem 12 says that problems on minimizing a linear combination of subgraph densities, which are among the problems of the simplest kind in extremal graph theory, may have unique optimal solutions with highly complex structure. Given the general nature of Theorem 12, it is surprising that the forcing family for the graphon W_0 in Theorem 12 is the same for all choices of W_F, i.e., the structure of W_0 is controlled by choosing the densities of the graphs in the forcing family only.

It is natural to ask whether the fraction $1/14$ given in Theorem 12 can be improved. The techniques presented in [26] would easily yield that the fraction $1/14$ can be replaced by $1/2-\varepsilon$ for any $\varepsilon > 0$. A recent result given in [60] shows that it is possible to improve this fraction to be arbitrarily close to 1.

Theorem 13. *For every $\varepsilon > 0$ and every graphon W_F, there exists a finitely forcible graphon W_0 such that W_F is a subgraphon of W_0 induced by a $1-\varepsilon$ fraction of the vertices of W_0.*

Recall that the forcing family in Theorem 12 was the same for all choices of W_F. However, the forcing family in Theorem 13 depends on ε and this dependance is necessary as shown in [60].

We now briefly outline the ideas used in the proofs that the graphons depicted in Figures 5–8 are finitely forcible. The arguments are based on the method of decorated constraints, which was developed in [39] and formalized in [38], and which builds on the flag algebra method of Razborov. Each of the graphons depicted in Figures 5–8 have the property that the interval $[0,1]$ is split into finitely many sets $X_1,\dots,X_k$ such that the integral

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

is the same for all x from the same set X_i, i.e., the degrees of the vertices in each of the parts are the same. We will refer to these sets as *parts* and to graphons with this structure as *partitioned* graphons. The flag algebra arguments can be used to show that there is a polynomial combination of densities that is zero if and only if a graphon has a given number of parts

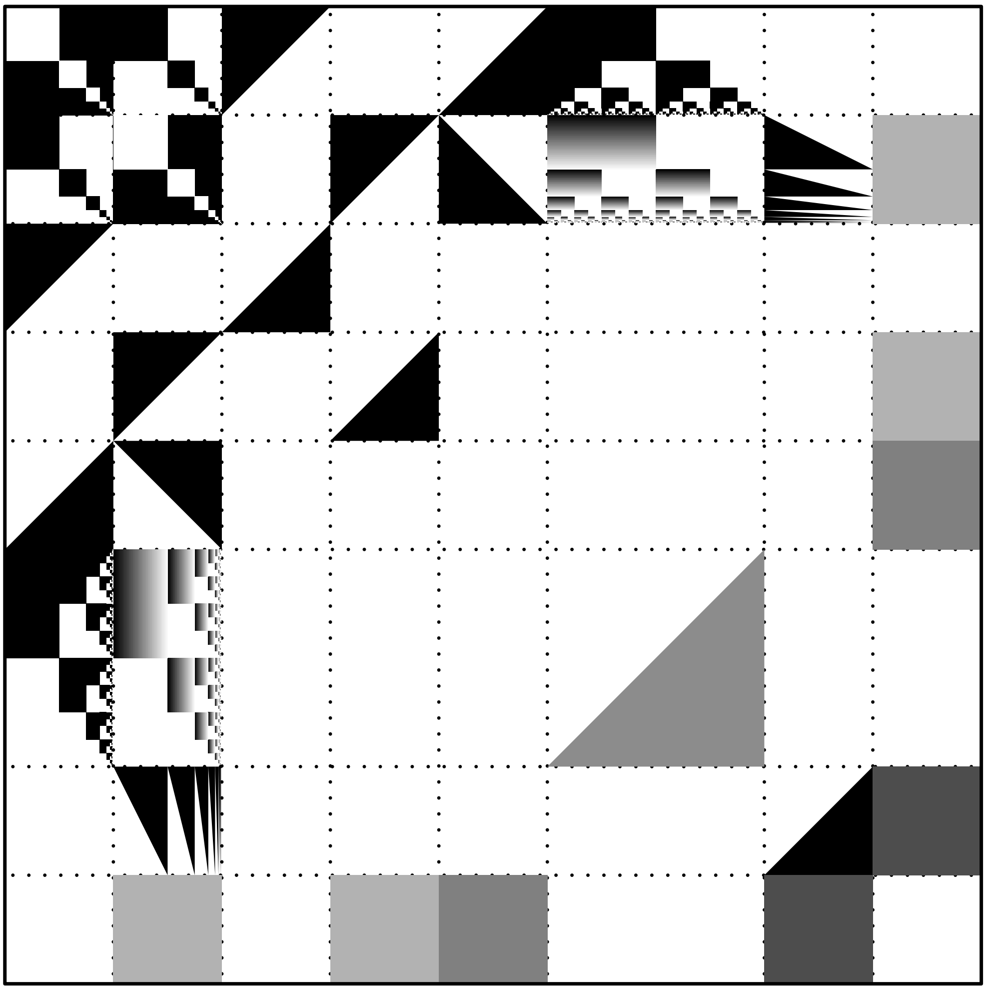

Figure 5: The finitely forcible graphon W with the space $T(W)$ of typical vertices that is not compact constructed in [39].

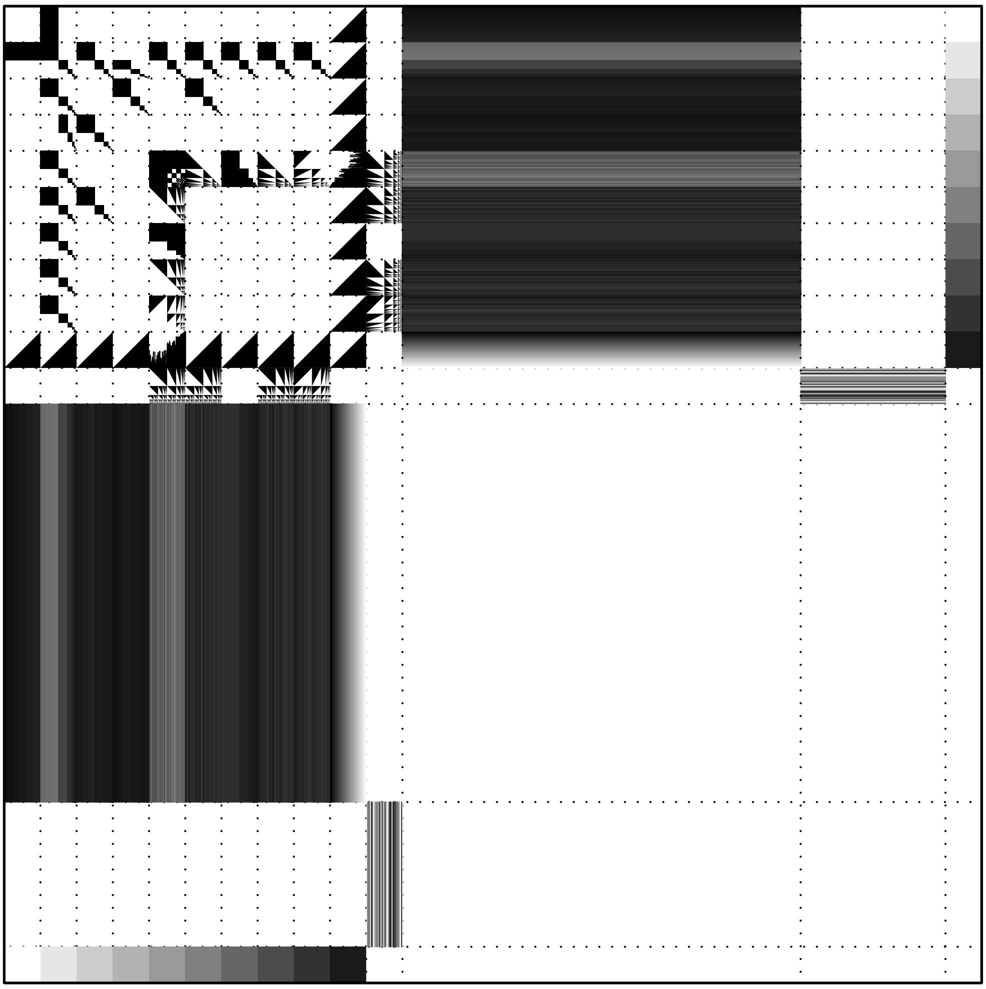

Figure 6: The finitely forcible graphon W with the space $T(W)$ of typical vertices containing a subspace homeomorphic to $[0,1]^{\mathbb{N}}$ that was constructed in [38].

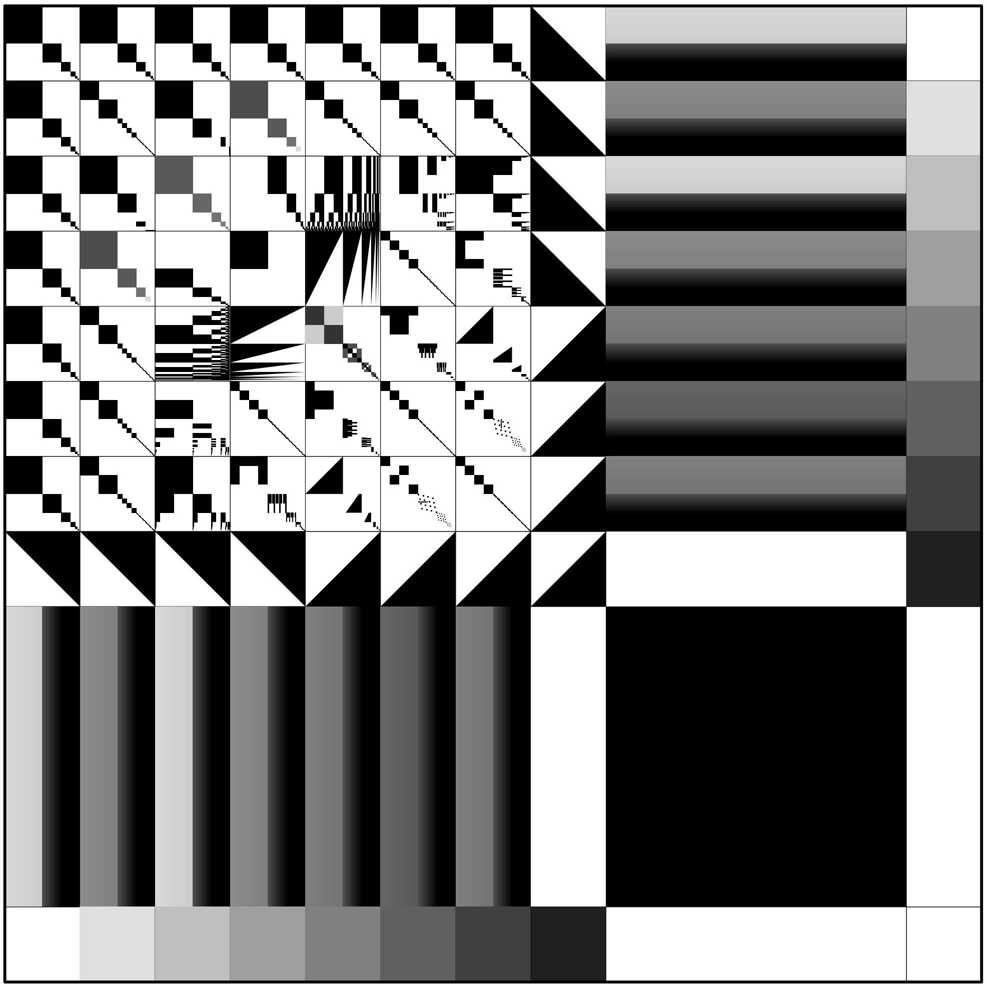

Figure 7: The finitely forcible graphon W constructed in [25]. Any weak ε-regular partition of W must have a number of parts almost exponential in ε^{-2} for infinitely many $\varepsilon > 0$.

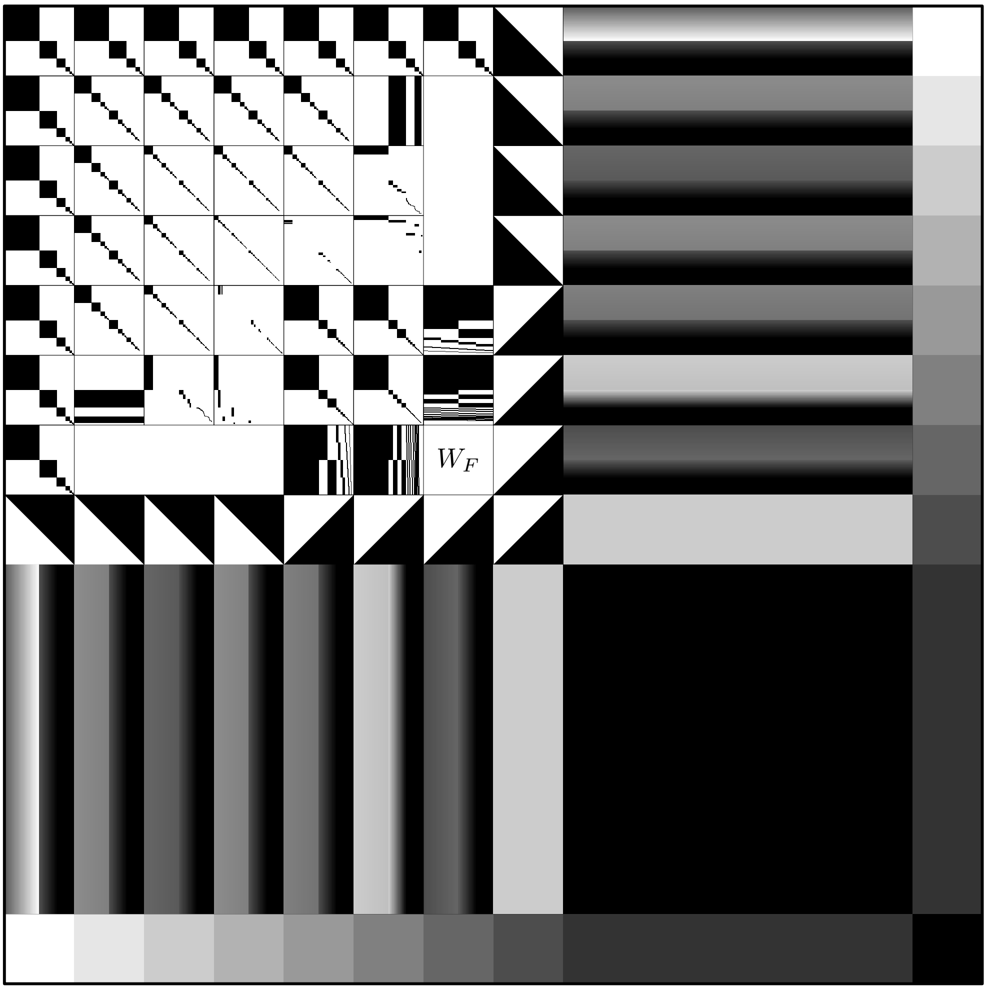

Figure 8: Visualization of the universal construction of complex finitely forcible graphon given in [26].

with given sizes and vertices of given degrees. In particular, the structure of a partitioned graphon can be forced by finitely many densities. The method of decorated constraints uses the power of the flag algebra method to restrict the structure inside and between the parts of a partitioned graphon by constraints that are simple to analyse even for complex graphons such as those in Figures 5–8.

We would like to conclude this section with a recent result concerning the relation of finitely forcible graphons and optimal solutions of problems in extremal graph theory. As a motivation, let us have a look at several classical results in extremal graph theory. One of the oldest results in extremal graph theory is the theorem of Mantel [73], which says that the maximum number of edges of an n-vertex triangle-free graph is $\lfloor n/2 \rfloor \cdot \lceil n/2 \rceil$ and the maximum is attained only by the balanced complete bipartite graph, i.e., the graph $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$. In the language of graph limits, Mantel's theorem says that the maximum value of $d(K_2, W)$ among all graphons W with $d(K_3, W) = 0$ is $1/2$ and every graphon achieving this maximum is weakly isomorphic to the graphon representing (large) complete bipartite graphs with parts of equal sizes. Mantel's theorem was extended by Turán [95] to graphs avoiding complete graphs of arbitrary sizes and by Erdős and Stone [33] to all graphs. In the language of graph limits, we obtain that, for every graph H, the maximum value of $d(K_2, W)$ among all graphons W with $t(H, W) = 0$ is equal to $\frac{\chi(H)-2}{\chi(H)-1}$ and every graphon achieving this maximum is weakly isomorphic to the graphon representing (large) complete $(\chi(H) - 1)$-partite graphs with parts of equal sizes.

Since the graphon representing (large) complete $(\chi(H)-1)$-partite graphs with parts of equal sizes is finitely forcible, it may be tempting to think that the converse of Proposition 11 could hold, i.e., the optimal configurations for every extremal graph theory problem are asymptotically unique. However, the following shows that this is not true. Let us consider the problem of minimizing the sum $d(K_3, W) + d(\overline{K_3}, W)$, i.e., the sum of the induced densities of K_3 and its complement. A classical result of Goodman [44] implies that this sum is minimized by any graphon such that

$$\int_{[0,1]} W(x, y) \mathrm{d}y = \frac{1}{2}$$

for almost every $x \in [0, 1]$, i.e., by any graphon representing graphs where almost every vertex has degree close to the number of vertices divided by two. However, the structure of an optimal solution can be made unique by adding additional constraints. For example, any graphon W that minimizes the sum and that satisfies $d(K_3, W) = 0$ corresponds to (large) complete bipartite graphs with parts of equal sizes, any graphon W that minimizes the sum and that satisfies $d(\overline{K_3}, W) = 0$ corresponds to (large) graphs that are the union of two complete graphs of equal sizes, or any graphon W that minimizes the sum and that satisfies $t(C_4, W) = 1/16$ is equal to $1/2$ almost everywhere, i.e., it corresponds to Erdős-Rényi random graphs $G_{n,1/2}$.

A conjecture of Lovász asserts that the phenomenon that we have just described is a more general one. The conjecture has been the most frequently quoted conjecture concerning dense graph limits, it also sometimes appeared as a question, and we include only some of the many references to its statement.

Conjecture 3 (Lovász [65, Conjecture 3], [66, Conjecture 9.12], [67, Conjecture 16.45], and [71, Conjecture 7])**.** *Let $H_1, \ldots, H_\ell$ be graphs and $d_1, \ldots, d_\ell$ reals. If there exists a convergent sequence of graphs with the limit density of H_i equal to d_i, $i = 1, \ldots, \ell$, then there exists such a sequence that its limit graphon is finitely forcible.*

Informally speaking, the conjecture says that "every extremal problem has a finitely forcible optimum", see [67, p. 308]. The conjecture has been recently disproved in [46] using the universal construction of complex finitely forcible graph limits from [26]. More precisely, the authors proved the following theorem in [46].

Theorem 14. *There exists a family of graphons $\mathcal{W}$, graphs $H_1, \ldots, H_\ell$ and reals $d_1, \ldots, d_\ell$ such that*

- *a graphon W is weakly isomorphic to a graphon contained in $\mathcal{W}$ if and only if $d(H_i, W) = d_i$ for every $i \in [\ell]$, and*
- *no graphon in $\mathcal{W}$ is finitely forcible, i.e., for all graphs $H'_1, \ldots, H'_r$ and reals $d'_1, \ldots, d'_r$, the family $\mathcal{W}$ contains either zero or infinitely many graphons W with $d(H'_i, W) = d'_i$, $i \in [r]$.*

The family $\mathcal{W}$ of graphons from Theorem 14 is visualized in Figure 9. Unlike in the results that we have mentioned earlier, the graphons in the family $\mathcal{W}$ have a part that depends on a countable vector $z \in [0, 1]^{\mathbb{N}}$ and analytic tools are applied to understand and to restrict the behaviour of graphons in the family $\mathcal{W}$. We remark that Theorem 14 can be further generalized [46] in the way that all graphons in the family $\mathcal{W}$ have the same value of a given graphon parameter that behaves nicely on the space of graphons. An example of such a parameter may be the graphon entropy, i.e., informally speaking, there are problems in extremal graph theory with no single "typical" graphon at the exponential scale.

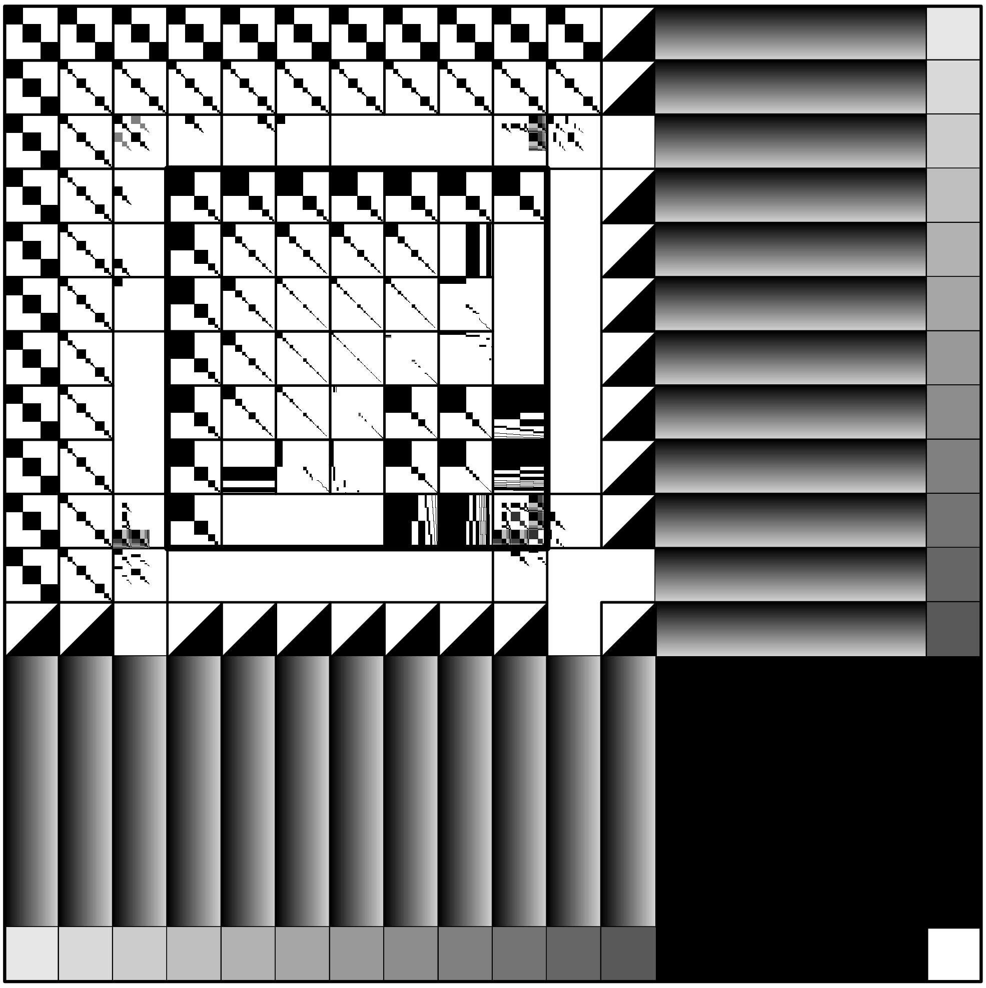

Figure 9: Visualization of the graphons forming the family $\mathcal{W}$ in Theorem 14. The family is obtained by varying densities in the sqaure in the third row and the third column in a controlled way.

5 Sparse graph limits

In this section, we give a brief overview of the main notions of convergence for 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.

As we said earlier, the theory of limits of sparse graphs is developed in a less satisfactory way than the theory of limits of dense graphs. While this can be caused by the lack of our understanding of the structure of sparse graphs, many believe that there is no perfect notion of convergence because of the nature of sparse graphs. Such a perfect notion of convergence should be able to distinguish graphs with different local and global structures, i.e., sequences of graphs such that their local or global properties differ substantially should not be convergent. The notion should also be robust enough that sublinear modifications of graphs in the sequence do not affect the convergence, i.e., a convergent sequence should stay convergent if a sublinear number of edges is added or removed. Finally, the notion should ideally allow representing convergent sequences of sparse graphs with an analytic object that captures the interplay between local and global properties, similarly to the way that graphons in the dense setting capture the interplay between subgraph densities (a local property) and regularity partitions (a global property). In what follows, we present several notions of convergence for sparse graphs that have been studied and we will discuss their mutual relation and demonstrate their power on examples of particular sequences of graphs that do or do not converge with respect to these notions.

The most widely used notion of convergence in relation to graphs with bounded degrees is the one defined by Benjamini and Schramm [9], known as *Benjamini-Schramm convergence*, shortly BS-convergence, and also as *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 $\mathcal{G}^v(d,\Delta)$ be the set of all rooted graphs with maximum degree Δ where all vertices have distance at most d from the root. Note that the set $\mathcal{G}^v(d,\Delta)$ 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-neighbourhood of the root, i.e., the vertices at distance at most d from the root, we get a (finite) probability distribution on graphs from $\mathcal{G}^v(d,\Delta)$. Let $p_{n,d} \in [0,1]^{\mathcal{G}^v(d,\Delta)}$ 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 [31], however, we omit further details concerning this representation here and explore the view on limits of BS-convergent sequences in terms of distribution on rooted neighbourhoods of vertices.

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

It is not true that every probability measure μ on $\mathcal{G}^v$ corresponds to a BS-convergent sequence of graphs. Let us fix $\Delta = 3$, i.e., we restrict our attention to graphs with maximum degree three in the following exposition. 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 $\mu(\{T\}) = 1$, 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 neighbours of degree two only—this is clearly impossible.

We now describe a condition on a probability measure μ that is necessary in order that μ corresponds to a BS-convergent sequence of graphs. We start with defining a different probability measure μ' on $\mathcal{G}^v$ as

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

where the integration is with respect to the measure μ and $\delta(G)$ for $G \in \mathcal{G}^v$ is the degree of the root of G (we may assume that $\delta(G) > 0$ with non-zero probability, i.e., μ' is well-defined, 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 $G \in \mathcal{G}^v$ 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 $G \in \mathcal{G}^e$ according to μ_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 that is known as the conjecture of Aldous and Lyons [1] asserts that this necessary condition is also sufficient for a probability measure μ on $\mathcal{G}^v$ to correspond

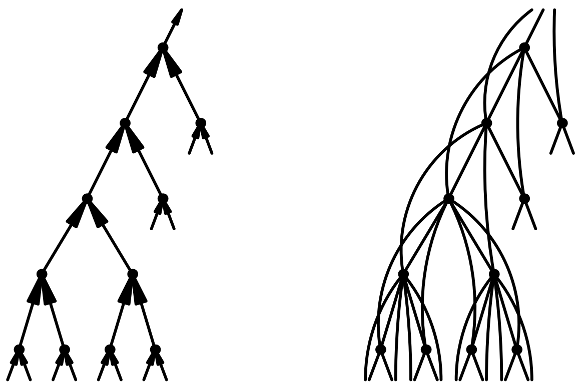

Figure 10: The construction of an infinite graph presented in relation to Benjamini-Schramm convergence: a part the original directed infinite tree and the corresponding part of the obtained undirected graph.

to a BS-convergent sequence of graphs. We remark that this conjecture is closely related to a question of Gromov whether all countable discrete groups are sofic, see [67, Chapter 19].

It is tempting to think that the following condition, which is weaker than the one presented in the previous paragraph, can also be sufficient for a probability measure μ to correspond to a BS-convergent sequence of graphs. Let μ_v be the probability distribution on $\mathcal{G}^v$ obtained as follows: sample a rooted graph with a distinguished edge based on μ_e and keep the root, i.e., forget that any edge of the sampled graph is distinguished. Note that μ_v differs from μ if $\mu(\{T\}) > 0$ where T is the single vertex graph with its only vertex being the root, and μ_v and μ are the same if $\mu(\{T\}) = 0$. We define μ'_v based on μ'_e in the analogous way. Informally speaking, μ'_v is the distribution obtained from μ_v by rerooting to a random neighbour of the root (in an appropriately weighted way). Clearly, if μ corresponds to a BS-convergent sequence of graphs, then the probability measures μ_v and μ'_v are the same. It may be tempting to think that if μ_v and μ'_v are the same for a measure μ on $\mathcal{G}^v$, then μ corresponds to a BS-convergent sequence of graphs. However, this is not true as we explain in the next paragraph.

Consider an infinite tree T_0 where every vertex has degree three and each edge is directed in such a way that each vertex has out-degree exactly one (note that this determines the tree T_0 completely) and let T be the infinite (undirected) graph obtained from T_0 by joining two vertices by an edge if they are joined by a directed path of length one or two; see Figure 10 for an illustration of the construction. Since T_0 is vertex-transitive, T is also vertex-transitive. In particular, if $\mu(\{T\}) = 1$, then the corresponding measures μ_v

and μ'_v are the same. However, there is no sequence $(G_n)_{n\in\mathbb{N}}$ of graphs such that μ is the resulting measure on $\mathcal{G}^v$. To see this, we proceed as follows. Assume that a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with maximum degree eight is BS-convergent and μ is the resulting measure on $\mathcal{G}^v$. A vertex of G_n is *typical* if its 2-neighbourhood is the same as the 2-neighbourhood of vertices in T. For $\varepsilon > 0$, consider $n \in \mathbb{N}$ such that the 2-neighbourhood of all but $\varepsilon|G_n|$ vertices of G_n are typical. We consider each typical vertex v of G_n and orient some of the edges incident with v as follows. The vertex v is incident with exactly three edges e_1, e_2 and e_3 contained in three triangles and all but a single pair of these three edges are contained in a common triangle, i.e., we can assume by symmetry that e_1 and e_2 are contained in a common triangle and e_1 and e_3 are contained in a common triangle. We now orient the edge e_1 from the vertex v and the edges e_2 and e_3 towards v. Because the 2-neighbourhood of v is the same as the 2-neighbourhood of the vertices in T, no edge is oriented in two different ways. Observe that the sum of in-degrees of the vertices of G_n is at least $2(1-\varepsilon)|G_n|$ (each typical vertex has two incoming edges) but the sum of out-degrees is at most $(1+8\varepsilon)|G_n|$ (each typical vertex has a single outgoing edge and each vertex that is not typical can have at most eight such edges). However, this is impossible if $\varepsilon < 1/10$. We conclude that there is no BS-convergent sequence of graphs such that μ is the resulting measure on $\mathcal{G}^v$.

Benjamini-Schramm convergence has the drawback that we next describe. Let us consider a setting of graphs with maximum degree three.

Example 1. *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that G_n is a random $(2n)$-vertex cubic graph when n is odd, and G_n is a random $(2n)$-vertex cubic bipartite graph when n is even.*

We claim that the sequence from Example 1 is BS-convergent with probability one. Indeed, the probability that a randomly chosen vertex of a random cubic graph is contained in a cycle of length k tends to 0 for any fixed integer k. The same is true for random cubic bipartite graphs. Hence, the sequence from Example 1 is BS-convergent and the corresponding probability measure μ on $\mathcal{G}^v$ satisfies that $\mu(\{T\}) = 1$ for the infinite rooted cubic tree T. However, the independence number of a random n-vertex cubic graph is at most $0.455n$ with probability tending to one [74], i.e., it is bounded away from $n/2$. In other words, Example 1 shows that BS-convergence is not robust enough to distinguish bipartite graphs from graphs that are far from being bipartite. We consider one more example.

Example 2. *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that G_n is H_n when n is odd, and G_n is the union of two copies of H_n when n is even, where $(H_n)_{n\in\mathbb{N}}$*

is a BS-convergent sequence of cubic expanders (to obtain $(H_n)_{n\in\mathbb{N}}$, consider a sequence of cubic expanders and one of its convergent subsequences).

Since $(H_n)_{n\in\mathbb{N}}$ is BS-convergent, the sequence $(G_n)_{n\in\mathbb{N}}$ is also BS-convergent. This example shows that BS-convergence is not robust to distinguish well-connected graphs, which appear in the sequence $(G_n)_{n\in\mathbb{N}}$ on even positions, from disconnected graphs, which appear in the sequence on odd positions.

To overcome the phenomenon demonstrated by Examples 1 and 2, a finer notion of convergence called local-global convergence was proposed in [11] and further studied in [50]. This notion of convergence takes into account possible partitions of vertex sets of graphs in a sequence. Formally, let $\mathcal{G}^v(d,k,\Delta)$ be the set of all rooted k-vertex-coloured graphs with maximum degree Δ (the vertex colouring need not be proper) such that every vertex is at distance at most d from the root. For a graph G with maximum degree Δ, let $P_{d,k}(G)$ be the set of all vectors from $[0,1]^{\mathcal{G}^v(d,k,\Delta)}$ that corresponds to the probability distribution on d-neighbourhoods for all k-vertex-colourings of G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with maximum degree Δ is *local-global convergent* if the sets $\left(P_{d,k}(G_n)\right)_{n\in\mathbb{N}}$ converge in the Hausdorff metric for every $d \in \mathbb{N}$ and $k \in \mathbb{N}$, i.e., for every $\varepsilon > 0$, there exists n_0 such that the Hausdorff distance of $P_{d,k}(G_i)$ and $P_{d,k}(G_j)$ is at most ε for every $i, j \geq n_0$. Recall, that the *Hausdorff distance* of two subsets A and B of $\mathbb{R}^D$ is

$$\max\{\sup_{x\in A}\inf_{y\in B} d(x,y), \sup_{x\in B}\inf_{y\in A} d(x,y)\},$$

where $d(x,y)$ is the distance between points x and y (in this definition, it does not matter which of the standard metrics on $\mathbb{R}^D$ we use, so, we can use the L_1-metric for example). Informally speaking, the definition says that the sequence $(G_n)_{n\in\mathbb{N}}$ is local-global convergent if and only if for every k-colouring of G_i, there exists a k-colouring of G_j with a close statistic of d-neighbourhoods assuming that both i and j are sufficiently large.

Observe that if a sequence of graphs is local-global convergent, it is also BS-convergent (set $k = 1$ in the definition). However, the converse is not necessarily true: neither of the sequences given in Examples 1 and 2 is local-global convergent. In Example 1, an $(2n)$-vertex random cubic bipartite graphs has a vertex-colouring with two colours, say red and blue, such that the number of red vertices is n and there are no red-red edges. However, a 2-vertex-colouring with a neighbourhood statistic close to this 2-vertex-colouring does not exist for $(2n)$-vertex random cubic graphs with high probability since the size of their largest independent set is at most $0.91n$ with high probability as we have mentioned earlier. In Example 2, the union of two

n-vertex (cubic) expanders has a 2-vertex-colouring such that each colour is used on half of the vertices and all edges are monochromatic but no n-vertex cubic expander has a 2-vertex-colouring with a neighbourhood statistic close to this 2-vertex-colouring.

Another notion of convergence related to BS-convergence is that of 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*. We remark that such graphs are also often called soft cores, while graphs, where non-negative weights are allowed are called hard cores. The definition that we use here is weaker than the original definition, which was using hard cores instead of soft cores, however, every sequence of graphs that is convergent in the definition that we use can be modified by changing a sublinear number of edges to a sequence of graphs convergent in the original (stronger) definition, see [12] for further details.

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. For a homomorphism f, the corresponding summand in the expression above is referred as the *weight* of the homomorphism f. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *right convergent* if the fraction

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

convergences for every target H. It can be shown [13], also see [72], 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. However, the converse is not true since the sequence given in Example 1 is not right convergent, i.e., informally speaking, right convergence can distinguish graphs close to being bipartite and those far from being bipartite. To see that the sequence given in Example 1 is not right convergent, consider the target graph H_K, $K \in \mathbb{N}$, with two vertices v and w such that the weight of the vertex v is K, the weight of the loop at v is $1/K$, the weights of w, the loop at w and the edge vw are one. Observe that G_n has a homomorphism to H of weight K^m if and only if the independence number of G_n is m. This can be used to show that

$$\lim_{K\to\infty}\lim_{n\to\infty}\frac{\log\hom(G_n, H)}{|G_n|} = \lim_{n\to\infty}\frac{\alpha(G_n)}{|G_n|}.$$

Hence, the sequence given in Example 1 is not right convergent.

We next consider the following modification of Example 2.

Example 3. *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that G_n is H_n when n is odd, and G_n is the union of two copies of H_n when n is even, where $(H_n)_{n\in\mathbb{N}}$ is a right convergent sequence of cubic expanders (to obtain $(H_n)_{n\in\mathbb{N}}$, consider a sequence of cubic expanders and one of its convergent subsequences).*

Observe that it holds that

$$\frac{\log \hom(G,H)}{|G|} = \frac{\log \hom(G\cup G,H)}{|G\cup G|}$$

for every graph G and every target H, where $G\cup G$ stands for the union of two disjoint copies of G. Consequently, Example 3 shows that right convergence does not imply local-global convergence.

Another notion of convergence of sparse graphs, which is entirely based on possible vertex partitions, was proposed by Bollobás and Riordan in [11]. 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 first k coordinates are the relative sizes $p_i = \frac{|P_i|}{|G|}$ of the parts and the remaining $\binom{k+1}{2}$ coordinates are the edge densities $e_{ij} = \frac{e(P_i,P_j)}{|G|}$ between the parts (including the cases when $i = j$), 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 3 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 all k-partitions $\mathbf{P}$ of a graph G. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with bounded maximum degree is *partition convergent* if the sequence $(P_k(G_n))_{n\in\mathbb{N}}$ converges in the Hausdorff metric for every $k \in \mathbb{N}$. Observe that local-global convergence of a sequence of graphs trivially implies partition convergence but Example 2 and its modification considered in the previous paragraph yield that neither BS-convergence nor right convergence implies partition convergence.

We will now show that there exists a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs that is partition convergent but is not BS-convergent (and so is neither right convergent nor local-global convergent). Consider the following sequence of 2-regular graphs.

Example 4. *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that G_n is the union of n cycles of length four, i.e., the graph $n\,C_4$, when n is odd, and it is the union of n cycles of length six, i.e., the graph $n\,C_6$, when n is even.*

The sequence from Example 4 is clearly not BS-convergent, however, the sequence $(P_k(G_n))_{n\in\mathbb{N}}$ converges in the the Hausdorff metric for every $k \in \mathbb{N}$, i.e., the sequence from Example 4 is partition convergent. We sketch the argument for $k = 2$. Let $U_2 \subseteq \mathbb{R}^5$ be the set of all non-negative real vectors

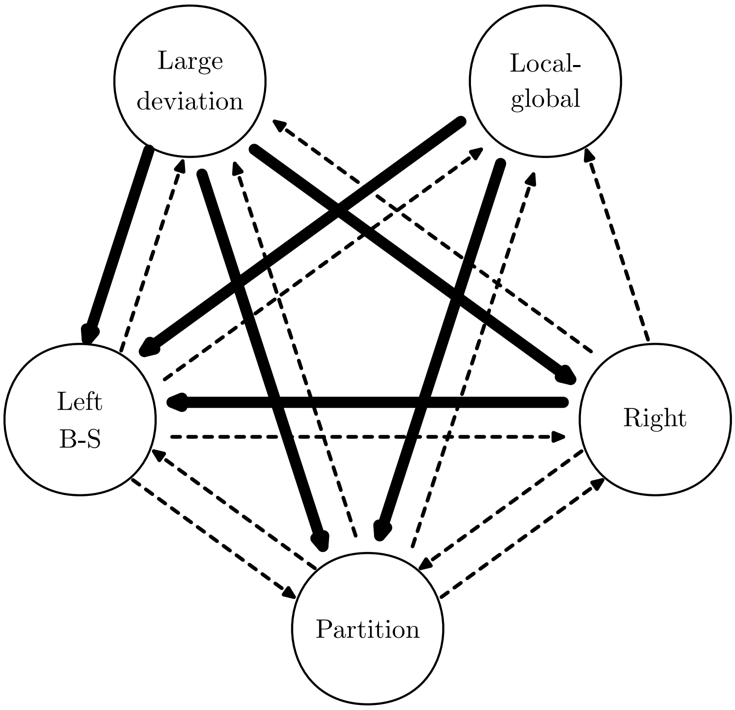

Figure 11: The relation between the presented notions of convergence of bounded degree graphs. The bold arrows represent that the notion at the tail of an arrow implies the other and the dashed arrows that this is not the case in general. When an arrow is missing, the relation between the notions is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

$(p_1, p_2, e_{11}, e_{12}, e_{22})$ 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. It can be shown that both the sequence $(P_2(n\ C_4))_{n\in\mathbb{N}}$ and the sequence $(P_2(n\ C_6))_{n\in\mathbb{N}}$ converge to U_2 in the Hausdorff metric, i.e., the statistics of the partitions into two parts of the vertices of the graphs in these sequences converge to the set of all possible statistics of the partitions into two parts of the vertices of 2-regular graphs.

We refer the reader to Figure 11 for the relation between the notions of convergence of sparse graphs that we have already discussed and the notion of large deviation convergence that we introduce next. The recent notion of large deviation convergence, which was introduced in [12], is a common refinement of right convergence and partition convergence. A sequence $(G_n)_{n\in\mathbb{N}}$ of graphs with bounded maximum degree is *LD-convergent* if the following limit exists (while possibly being infinite)

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

for every k and $x \in \mathbb{R}^{k+\binom{k+1}{2}}$, where $s(\mathbf{P})$ is the statistic of a k-partition $\mathbf{P}$ as defined earlier. 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. In fact, the sequence $(P_k(G_n))_{n\in\mathbb{N}}$ converges to the set $\{x \mid I_k(x) < \infty\}$ in the Hausdorff metric. A more involved argument shows that every LD-convergent sequence of graphs is right convergent [12], which implies that it is also BS-convergent. We would like to emphasize here that it is important here that we consider targets with positive weights only (soft cores). If the definition of right convergence uses targets with non-negative weights (hard cores), when LD-convergence does not imply right convergence. An example showing this is the sequence $(C_n)_{n\in\mathbb{N}}$ of cycles with alternating parities that can be shown to be LD-convergent but it is not right convergent when targets are allowed to have elements with zero weight (a cycle can be homomorphically mapped to K_2 if and only if its length is even).

The final notion of convergence of graphs that we would like to mention is the notion of first order convergence introduced in [76, 77, 80] and further studied in [18,36,57,78,79]. 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 $\lim_{n\to\infty} \langle\psi, G_n\rangle$ 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 3 and every first order convergent sequence of graphs with bounded maximum degree is Benjamini-Schramm convergent. Neither of the opposite implications is true. We present an argument in the case of Benjamini-Schramm convergence. Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that G_n is the union of n copies of K_2, and let $(G'_n)_{n\in\mathbb{N}}$ be a sequence of graphs such that $G'_n = G_n$ if n is even and $G'_n = G_n \cup K_1$ if n is odd. The sequence $(G'_n)_{n\in\mathbb{N}}$ is Benjamini-Schramm convergent but not first order convergent: if ψ is a first order formula that is true if and only if a graph contains an isolated vertex, then the values $\langle\psi, G'_n\rangle$ alternate between zero and one. This example also shows that first order convergence is not preserved by constant size modifications of graphs in the sequence: the sequence $(G_n)_{n\in\mathbb{N}}$ is first order convergent unlike $(G'_n)_{n\in\mathbb{N}}$.

Some 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 [76, 77]; an interesting example of a sequence of a first order convergent sequence of graphs with no modeling is the sequence of Erdős-Rényi random graphs $G_{n,1/2}$ that has no modeling with probability one. In general, a sequence of dense graphs converging to a graphon W has a modeling if and only if the graphon W is random-free [76, 77], i.e., $W(x,y) \in \{0,1\}$ for almost every $(x,y) \in [0,1]^2$. A nice conjecture of Nešetřil and Ossona de Mendez [76, 77] asserted the following: if $\mathcal{G}$ is a nowhere-dense class of graphs (see [75] for the definition and further exposition), then any first order convergent sequence of graphs from $\mathcal{G}$ can be represented by a modeling. Another conjecture of Nešetřil and Ossona de Mendez [79] asserted that every residual first order convergent sequence of graphs has a limit modeling; a sequence $(G_n)_{n\in\mathbb{N}}$ of graphs is *residual* if for every $d \in \mathbb{N}$ and $\varepsilon > 0$, there exists n_0 such that the number of vertices at distance at most d from any vertex in G_n, $n \geq n_0$, is at most $\varepsilon|G_n|$. Both conjectures were proven in [78], however, their stronger forms asserting that every first order convergent sequence of graphs from a nowhere-dense class of graphs $\mathcal{G}$ has a limit modeling obeying a property called the finitary mass transport principle (see [77] for the definition of this property) and that every residual first order convergent sequence of graphs has a limit modeling obeying the finitary mass transport principle remain open and present very interesting problems.

Acknowledgement

The authors would like to thank the anonymous reviewer for the many detailed comments, which helped to improve the survey significantly.

References

[1] D. Aldous and R. Lyons: *Processes on unimodular random networks*, Electron. J. Probab. **12** (2007), 1454–1508.

[2] D. J. Aldous: *Representations for partially exchangeable arrays of random variables*, J. Multivariate Anal. **11** (1981), 581–598.

[3] N. Alon and J. H. Spencer: The probabilistic method, 2004.

[4] T. Austin: *On exchangeable random variables and the statistics of large graphs and hypergraphs*, Probab. Surv. **5** (2008), 80–145.

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

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

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

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

[9] I. Benjamini and O. Schramm: *Recurrence of distributional limits of finite planar graphs*, Electron. J. Probab. **6** (2001).

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

[11] B. Bollobás and O. Riordan: *Sparse graphs: metrics and random models*, Random Struct. Algor. **39** (2011), 1–38.

[12] C. Borgs, J. Chayes and D. Gamarnik: *Convergent sequences of sparse graphs: A large deviations approach*, Random Struct. Algor. **51** (2017), 52–89.

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

[14] 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.

[15] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, B. Szegedy and K. Vesztergombi: *Graph limits and parameter testing*, Proc. 38th annual ACM Symposium on Theory of computing (STOC) (2006), 261–270.

[16] 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.

[17] 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. Math. (2012), 151–219.

[18] D. Christofides and D. Kráľ: *First-order convergence and roots*, Combin. Probab. Comput. **25** (2016), 213–221.

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

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

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

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

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

[24] D. Conlon and J. Lee: *Sidorenko's conjecture for blow-ups*, preprint arXiv:1809.01259 (2018).

[25] J. W. Cooper, T. Kaiser, D. Kráľ and J. A. Noel: *Weak regularity and finitely forcible graph limits*, Trans. Amer. Math. Soc. **370** (2018), 3833–3864.

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

[27] P. Diaconis and S. Janson: *Graph limits and exchangeable random graphs*, Rend. Mat. Appl. **28** (2008), 33–61.

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

[29] M. Doležal, J. Grebík, J. Hladký, I. Rocha and V. Rozhoň: *Relating the cut distance and the weak* topology for graphons*, preprint arXiv:1806.07368 (2018).

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

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

[32] G. Elek and B. Szegedy: *Limits of hypergraphs, removal and regularity lemmas. A non-standard approach*, preprint arXiv:1705.2179 (2007).

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

[34] J. Fox and F. Wei: *On the local approach to Sidorenko's conjecture*, Electron. Notes Discrete Math. **61** (2017), 459–465.

[35] A. Frieze and R. Kannan: *Quick approximation to matrices and applications*, Combinatorica **19** (1999), 175–200.

[36] 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*, Random Struct. Algor. **50** (2017), 612–635.

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

[38] R. Glebov, T. Klimošová and D. Král': *Infinite dimensional finitely forcible graphon*, to appear in Proc. Lond. Math. Soc.

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

[40] O. Goldreich: Property testing: current research and surveys, *Lecture Notes in Computer Science*, volume 6390, 2010.

[41] O. Goldreich, S. Goldwasser and D. Ron: *Property testing and its connection to learning and approximation*, Proc. 37th annual Symposium on Foundations of Computer Science (FOCS) (1996), 339–348.

[42] O. Goldreich, S. Goldwasser and D. Ron: *Property testing and its connection to learning and approximation*, Journal of the ACM (JACM) **45** (1998), 653–750.

[43] O. Goldreich and L. Trevisan: *Three theorems regarding testing graph properties*, Random Struct. Algor. **23** (2003), 23–57.

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

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

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

[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] H. Hatami, L. Lovász and B. Szegedy: *Limits of locally–globally convergent graph sequences*, Geom. Funct. Anal. **24** (2014), 269–296.

[51] J. Hladký, A. Máthé, V. Patel and O. Pikhurko: *Poset limits can be totally ordered*, Trans. Amer. Math. Soc. **367** (2015), 4319–4337.

[52] D. N. Hoover: *Relations on probability spaces and arrays of random variables*, Institute for Advances Study, Princeton, NJ, 1979 (1979).

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

[54] C. Hoppen, Y. Kohayakawa, C. G. T. de A. Moreira and R. M. Sampaio: *Testing permutation properties through subpermutations*, Theor. Comput. Sci. **412** (2011), 3555–3567.

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

[56] O. Kallenberg: *Symmetries on random arrays and set-indexed processes*, J. Theoret. Probab. **5** (1992), 727–765.

[57] F. Kardoš, D. Král', A. Liebenau and L. Mach: *First order convergence of matroids*, Eur. J. Combin. **59** (2017), 150–168.

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

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

[60] D. Král', L. M. Lovász, J. A. Noel and J. Sosnovec: *Finitely forcible graphons with an almost arbitrary structure*, preprint arXiv:1809.05973 (2018).

[61] 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.

[62] D. Král', T. Martins, P. P. Pach and M. Wrochna: *The step Sidorenko property and non-norming edge-transitive graphs*, J. Combin. Theory Ser. A **162** (2019), 34–54.

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

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

[65] L. Lovász: *Graph homomorphisms: Open problems*, manuscript available as `http://www.cs.elte.hu/~lovasz/problems.pdf` (2008).

[66] L. Lovász: *Very large graphs*, Curr. Dev. Math. **2008** (2009), 67–128.

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

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

[73] W. Mantel: *Problem 28*, Wiskundige Opgaven **10** (1907), 60–61.

[74] D. B. McKay: *Independent sets in regular graphs of high girth*, Ars Combin. **23A** (1987), 179–185.

[75] J. Nešetril and P. O. de Mendez: Sparsity: Graphs, Structures, and Algorithms, *Algorithms and Combinatorics*, volume 28, 2012.

[76] J. Nešetřil and P. Ossona de Mendez: *A unified approach to structural limits (with applications to the study of graphs limits with bounded tree-depth)*, to appear in Mem. Amer. Math. Soc.

[77] J. Nešetřil and P. Ossona de Mendez: *A unified approach to structural limits, and limits of graphs with bounded tree-depth*, arXiv preprint arXiv:1303.6471 (2013).

[78] J. Nešetřil and P. Ossona de Mendez: *Existence of modeling limits for sequences of sparse structures*, arXiv preprint arXiv:1608.00146 (2016).

[79] J. Nešetřil and P. Ossona de Mendez: *Modeling limits in hereditary classes: Reduction and application to trees*, Electron. J. Comb. **23** (2016), 2–52.

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

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

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

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

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

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

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

[87] C. Reiher: *The clique density theorem*, Ann. Math. (2016), 683–707.

[88] R. Rubinfeld and M. Sudan: *Robust characterizations of polynomials with applications to program testing*, SIAM J. Comput. **25** (1996), 252–271.

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

[90] M. Simonovits: *Extremal graph problems, degenerate extremal problems, and supersaturated graphs*, in: Progress in graph theory (Waterloo, Ont., 1982) (1984), 419–437.

[91] J. Spencer: *Quasirandom multitype graphs*, in: An Irregular Mind, *Bolyai Soc. Math. Stud.*, volume 21 (2010), 607–617.

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

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

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

[95] P. Turán: *On an extremal problem in graph theory*, Matematikai és Fizikai Lapok **48** (1941), 436–452.

[96] Y. Zhao: *Hypergraph limits: A regularity approach*, Random Struct. Algor. **47** (2015), 205–226.