CONVERGENCE OF GRAPHS
WITH INTERMEDIATE DENSITY

PÉTER E. FRENKEL

ABSTRACT. We propose a notion of graph convergence that interpolates between the Benjamini–Schramm convergence of bounded degree graphs and the dense graph convergence developed by László Lovász and his coauthors. We prove that spectra of graphs, and also some important graph parameters such as numbers of colorings or matchings, behave well in convergent graph sequences. Special attention is given to graph sequences of large essential girth, for which asymptotics of coloring numbers are explicitly calculated. We also treat numbers of matchings in approximately regular graphs.

We introduce tentative limit objects that we call graphonings because they are common generalizations of graphons and graphings. Special forms of these, called Hausdorff and Euclidean graphonings, involve geometric measure theory. We construct Euclidean graphonings that provide limits of hypercubes and of finite projective planes, and, more generally, of a wide class of regular sequences of large essential girth. For any convergent sequence of large essential girth, we construct weaker limit objects: an involution invariant probability measure on the sub-Markov space of consistent measure sequences (this is unique), or an acyclic reversible sub-Markov kernel on a probability space (non-unique). We also pose some open problems.

CONTENTS

Notation and terminology. Graphs are finite, simple and undirected, unless otherwise specified. On k nodes, the complete graph, cycle, path, and path with a fork at the end is denoted by K_k, C_k, P_k, and D_k, respectively. For a graph $G = (V(G), E(G))$, we write $v(G) = |V(G)|$ and $e(G) = |E(G)|$. A graph F has

$c(F)$ connected components, out of which $c_{\geq 2}(F)$ have at least two nodes. The neighborhood (i.e., set of neighbors) of a node o is written $N(o)$.

The number of homomorphisms and injective homomorphisms from F to G is denoted by $\hom(F, G)$ and $\operatorname{inj}(F, G)$, respectively. The number of automorphisms of F is $\operatorname{aut} F$. The symbols $\times$ and $\square$ stand for the categorical (or weak) direct product and the Cartesian sum of graphs, respectively.

The product of σ-algebras is denoted by $\otimes$. We write a.e. for "almost every(where)" and a.s. for "almost surely", i.e., "with probability 1". The indicator of an event A is $\mathbb{1}_A$.

1. Homomorphism densities and graph convergence

The two most developed graph limit theories are the Benjamini–Schramm limit theory of bounded degree graphs and the dense graph limit theory developed by Borgs, Chayes, Lovász, T. Sós, Szegedy, and Vesztergombi. The convergence of dense graphs is defined in terms of homomorphism densities. The convergence of bounded degree graphs is defined in terms of neighborhood statistics, but this easily translates into convergence of homomorphism frequencies. We now propose a common generalization that works for both cases and also for intermediate density.

Definition 1.1. An *admissible pair* is a pair (G, d), where $d \geq 1$ and G is a graph with all degrees $\leq d$. For a connected graph F and an admissible pair (G, d), we define the *homomorphism density*

$$t(F, G, d) = \frac{\hom(F, G)}{\operatorname{v}(G) d^{\operatorname{v}(F)-1}} \in [0, 1].$$

We extend this to arbitrary F by making it multiplicative:

$$t(F, G, d) = \frac{\hom(F, G)}{\operatorname{v}(G)^{\operatorname{c}(F)} d^{(\operatorname{v} - \operatorname{c})(F)}} \in [0, 1].$$

An *admissible sequence* is a sequence of admissible pairs. An admissible sequence (G_n, d_n) is *convergent* if the number sequence $t(F, G_n, d_n)$ converges for any (or, equivalently, any connected) graph F.

Remark 1.2. Note that

$$t(F, G, \operatorname{v}(G)) = \frac{\hom(F, G)}{\operatorname{v}(G)^{\operatorname{v}(F)}} = t(F, G)$$

is the usual homomorphism density. Thus, a sequence of the form $(G_n, \operatorname{v}(G_n))$ — which is always admissible — is convergent precisely if (G_n) is a convergent dense graph sequence.

Note also that for connected F we have

$$t(F, G, d) = t^*(F, G)/d^{\operatorname{v}(F)-1},$$

where $t^*(F, G) = \hom(F, G)/\operatorname{v}(G)$ is the usual homomorphism frequency, so if $d_n = d$ does not depend on n, then an admissible sequence (G_n, d) is convergent precisely if (G_n) is a Benjamini–Schramm convergent graph sequence (alternatively called a locally convergent graph sequence).

When F is a forest, the normalization used in Definiton 1.1 is similar to the one used by Bollobás and Riordan [6] and by Borgs, Chayes, Cohn, and Zhao [7, 8]. However, for general F, our normalization is quite different. The goal in those papers was to generalize dense graph convergence to the sparse case, but no attempt

was made to also include Benjamini–Schramm convergence in a unified treatment. In the present approach, both extremes are included as special cases. This is also reflected in the limit objects — generalized graphons —, which are L_p graphons in [7, 8] but graphonings in Section 4 of the present paper. Admittedly, the results presented in this paper are less conclusive.

Remark 1.3. Let (G, d) be an admissible pair. Removing an edge from a connected graph F without destroying connectivity cannot decrease $t(F, G, d)$. Removing a vertex of degree 1 from a connected graph F cannot either. Thus, we have $t(F, G, d) \leq t(F', G, d)$ if $F' \subseteq F$ are connected graphs.

Proposition 1.4. *Let (G, d) be an admissible pair. Let F be a graph. Then we have $t(F, G, d) = 1$ if and only if at least one of the following holds:*

(a) *F is an empty graph, or*
(b) *F is a forest and G is d-regular, or*
(c) *F is bipartite and G is a disjoint union of complete bipartite graphs $K_{d,d}$.*

Proof. We may assume that F and G are connected.

If any of (a), (b), (c) holds, then an easy induction on $v(F)$ shows that $\hom(F, G) = v(G) d^{v(F)-1}$ and the claim follows.

For the converse, assume that $t(F, G, d) = 1$.

If (a) does not hold, then F contains K_2 as a subgraph, thus $2 e(G)/(v(G)d) = t(K_2, G, d) \geq t(F, G, d) = 1$ and therefore G is d-regular.

If F contains an odd cycle C_{2k+1}, then consider the path $P_{2k+1} = C_{2k+1} - e$ for an edge $e \in E(C_{2k+1})$. We have

$$t(C_{2k+1}, G, d) = 1 = t(P_{2k+1}, G, d),$$

thus

$$\hom(C_{2k+1}, G) = \hom(P_{2k+1}, G).$$

But there exists a homomorphism $\phi : P_{2k+1} \to G$ such that images of the two endnodes coincide. Such a ϕ does not extend to C_{2k+1} because G has no loops. This contradiction proves that F is bipartite.

If F contains an even cycle C_{2k} for some $k \geq 2$, then a similar argument shows that in G, the two endnodes of any walk of length $2k - 1$ are joined by an edge. It follows that this holds for 3 in place of $2k - 1$, and thus for any odd length as well. But G has no loops, so it must be bipartite. It is connected, so it is a complete bipartite graph. It is d-regular, so $G \simeq K_{d,d}$. □

Example 1.5. *Let (Γ_i, δ_i) be admissible pairs $(i = 1, 2, \dots)$. Set $G_n = \Gamma_1 \times \cdots \times \Gamma_n$ and $d_n = \delta_1 \cdots \delta_n$. Then the sequence (G_n, d_n) is convergent. The homomorphism density $t(F, G_n, d_n)$ converges to $\prod_{i=1}^{\infty} t(F, \Gamma_i, \delta_i)$.*

Proof. We have $\hom(F, G_n) = \prod_{i=1}^{n} \hom(F, \Gamma_i)$ and $v(G_n) = \prod_{i=1}^{n} v(\Gamma_i)$, whence

$$t(F, G_n, d_n) = \prod_{i=1}^{n} t(F, \Gamma_i, \delta_i).$$

This is decreasing and therefore convergent as $n \to \infty$. □

Corollary 1.6. *Let (Γ, δ) be an admissible pair. Then the sequence $(\Gamma^{\times n}, \delta^n)$ is convergent. The homomorphism density $t(F, \Gamma^{\times n}, \delta^n)$ converges to 1 if $t(F, \Gamma, \delta) = 1$ and to zero otherwise.*

Example 1.7. If G is a disjoint union of graphs G^i $(i = 1, \ldots, \mathrm{v}(G)/d)$ of size d, then

$$t(F, G, d) = \frac{d}{\mathrm{v}(G)} \sum_{i=1}^{\mathrm{v}(G)/d} t(F, G^i)$$

for all connected F. We can think of each G^i as a point in the compact graphon space $\widetilde{\mathcal{W}}_0$ of L. Lovász and B. Szegedy [26, 28], and consider the uniform probability measure on these $\mathrm{v}(G)/d$ points. We can think of $\widetilde{\mathcal{W}}_0$ as sitting in $[0, 1]^\infty$, each graphon W being represented by its profile of homomorphism densities $t(F, W)$ with connected F. A sequence (G_n, d_n), such that G_n is a disjoint union of graphs of size d_n, is convergent if and only if the barycenters of the corresponding probability measures form a convergent sequence. This is strictly weaker than the weak convergence of the probability measures themselves. If (G_n, d_n) converges, then the limit can be represented by the limiting barycenter (which is unique), or any subsequential weak limit measure (which is non-unique in general, but each one has the correct barycenter).

Further examples of convergent sequences are regular sequences of large essential girth, such as hypercube graphs, large grid graphs, incidence graphs of finite projective spaces, and suitable random nearly regular graphs. See Subsections 1.3 and 1.4.

1.1. **Injective homomorphism densities.** It is sometimes useful to count injective, rather than arbitrary, homomorphisms. We introduce injective homomorphism densities. Even in the dense case, our normalization deviates slightly from the standard one in Lovász's monograph [26].

Definition 1.8. Let (G, d) be an admissible pair. For a connected graph F, we define the *injective homomorphism density*

$$t_{\mathrm{inj}}(F, G, d) = \frac{\mathrm{inj}(F, G)}{\mathrm{v}(G)d(d-1)^{\mathrm{v}(F)-2}} \in [0, 1]$$

unless F is a single point, in which case $t_{\mathrm{inj}}(F, G, d) = 1$. We extend this to arbitrary F by making the denominator multiplicative:

$$t_{\mathrm{inj}}(F, G, d) = \frac{\mathrm{inj}(F, G)}{\mathrm{v}(G)^{\mathrm{c}(F)} d^{\mathrm{c}_{\geq 2}(F)} (d-1)^{(\mathrm{v} - \mathrm{c} - \mathrm{c}_{\geq 2})(F)}} \leq \prod_i t_{\mathrm{inj}}(F_i, G, d),$$

where the F_i are the connected components of F.

Remark 1.9. Let $d > 1$ and let (G, d) be an admissible pair. Removing an edge from a connected graph F without destroying connectivity cannot decrease $t_{\mathrm{inj}}(F, G, d)$. Removing a vertex of degree 1 from a connected graph F cannot either. Thus, $t_{\mathrm{inj}}(F, G, d) \leq t_{\mathrm{inj}}(F', G, d)$ if $F' \subseteq F$ are connected graphs.

Proposition 1.10. *For any fixed connected graph F, we have*

$$t(F, G, d) - t_{\mathrm{inj}}(F, G, d) = O(1/d),$$

where the constant in the O depends only on F.

Proof. For $F = K_1$, both densities are 1 and the claim is trivial. For $\mathrm{v}(F) \geq 2$, we have

$$t(F, G, d) \geq \frac{\mathrm{inj}(F, G)}{\mathrm{v}(G)d^{\mathrm{v}(F)-1}} = t_{\mathrm{inj}}(F, G, d) \left(1 - \frac{1}{d}\right)^{\mathrm{v}(F)-2},$$

whence

$$t_{\text{inj}}(F, G, d) - t(F, G, d) \leq t_{\text{inj}}(F, G, d) \left(1 - \left(1 - \frac{1}{d}\right)^{\text{v}(F)-2}\right) \leq \frac{\text{v}(F) - 2}{d}.$$

On the other hand, we have the well-known formula

$$\hom(F, G) = \sum_{F'} \text{inj}(F', G),$$

where F' runs over the quotients of F. Note that quotients of connected graphs are connected, and proper quotients have fewer vertices than the original graph. Thus,

$$t(F, G, d) = \sum_{F'} \frac{\text{inj}(F', G)}{\text{v}(G)d^{\text{v}(F)-1}} \leq \sum_{F'} \frac{t_{\text{inj}}(F', G, d)}{d^{\text{v}(F)-\text{v}(F')}} = t_{\text{inj}}(F, G, d) + O\left(\frac{1}{d}\right).$$

$\square$

Corollary 1.11. *An admissible sequence (G_n, d_n) with $d_n \to \infty$ is convergent precisely if the injective homomorphism density $t_{\text{inj}}(F, G_n, d_n)$ converges for all connected graphs F. If this is the case, then*

$$\lim_{n \to \infty} t_{\text{inj}}(F, G_n, d_n) = \lim_{n \to \infty} t(F, G_n, d_n)$$

for any connected F.

This is well known in the dense case: homomorphism and injective homomorphism densities are almost the same.

It will be useful to also compare injective and componentwise injective homomorphisms.

Proposition 1.12. *Let F have connected components F_i. Then we have*

$$\left(\prod_i t_{\text{inj}}(F_i, G, d)\right) - t_{\text{inj}}(F, G, d)$$

$$\leq \frac{1}{\text{v}(G)} \sum_{F'} t_{\text{inj}}(F', G, d) + O\left(\frac{1}{\text{v}(G)^2}\right) = O\left(\frac{1}{\text{v}(G)}\right),$$

where F' runs over the quotients of F such that each F_i maps injectively to F' and $\text{c}(F') = \text{c}(F) - 1$. The constant in the O depends only on F.

Proof. We have

$$\prod_i \text{inj}(F_i, G) = \sum_{F'} \text{inj}(F', G),$$

where F' runs over the quotients of F that arise by only identifying nodes from distinct components. We always have $(\text{v} - \text{c})(F') \leq (\text{v} - \text{c})(F)$ and $\text{c}_{\geq 2}(F') \leq \text{c}_{\geq 2}(F)$, hence

$$\prod_i t_{\text{inj}}(F_i, G, d) = \sum_{F'} \frac{\text{inj}(F', G)}{\text{v}(G)^{\text{c}(F)}d^{\text{c}_{\geq 2}(F)}(d-1)^{(\text{v}-\text{c}-c_{\geq 2})(F)}} \leq \sum_{F'} \frac{t_{\text{inj}}(F', G, d)}{\text{v}(G)^{\text{c}(F)-\text{c}(F')}}$$

and the claim follows. $\square$

1.2. **Rooted homomorphism densities.**

Definition 1.13. Let (F, o) and (G, p) be rooted graphs, where F is connected. Let $\hom((F, o), (G, p))$ be the number of homomorphisms of F into G that map o to p. If (G, d) is admissible, we define the *rooted homomorphism density*

$$t((F, o), (G, p), d) = \frac{\hom((F, o), (G, p))}{d^{\mathrm{v}(F)-1}} \in [0, 1].$$

Remark 1.14. For any connected rooted graph (F, o) and any admissible pair (G, d), we have

$$t(F, G, d) = \mathbb{E}t((F, o), (G, p), d),$$

where p is a uniform random node of G.

1.3. **Regular sequences.**

Definition 1.15. Let $0 \le \alpha \le 1$. The admissible sequence (G_n, d_n) is α-*regular* if the degree of a uniform random vertex of G_n, divided by d_n, tends stochastically to α.

If the graph G_n is αd_n-regular for every n, then of course the sequence (G_n, d_n) is α-regular. Let us look at less trivial examples.

Example 1.16. Let G_n be the d_n-dimensional grid graph with $n \times \cdots \times n$ points. Then $\mathrm{v}(G_n) = n^{d_n}$ and

$$\mathrm{e}(G_n) = d_n n^{d_n - 1}(n - 1),$$

so

$$t(K_2, G_n, 2d_n) = \mathrm{e}(G_n)/(\mathrm{v}(G_n)d_n) = (n - 1)/n \to 1,$$

i.e., the sequence $(G_n, 2d_n)$ is 1-regular; cf. Proposition 1.20 below. If the sequence d_n either stabilizes to some d or tends to ∞, then $(G_n, 2d_n)$ is convergent; cf. Subsection 1.4.

The case when $d_n \to \infty$ can be generalized as follows.

Example 1.17. Consider a triangular array $(\Gamma_{ni}, \delta_{ni})$ $(1 \le i \le n)$ of admissible pairs with normalized average degree $\alpha_{ni} = t(K_2, \Gamma_{ni}, \delta_{ni})$. Set $G_n = \Gamma_{n1} \square \ldots \square \Gamma_{nn}$ and $d_n = \delta_{n1} + \cdots + \delta_{nn}$. Assume that

$$(1.1) \qquad\qquad\qquad \max_{1 \le i \le n} \delta_{ni}/d_n \to 0$$

and the weighted average

$$\frac{1}{d_n} \sum_{i=1}^{n} \delta_{ni}\alpha_{ni} \to \alpha.$$

Then the sequence (G_n, d_n) is α-regular.

Proof. Let X_{ni} be the degree of a uniform random node in Γ_{ni}, divided by δ_{ni}. Then X_{ni} is a random variable with range in $[0, 1]$, and $\mathbb{E}X_{ni} = \alpha_{ni}$. The degree of a uniform random node in G_n, divided by d_n, is

$$X_n = (\delta_{n1}X_{n1} + \cdots + \delta_{nn}X_{nn})/d_n,$$

where the X_{ni} are independent. We have $\mathbb{E}X_n \to \alpha$ and

$$\mathbb{D}^2 X_n = (\delta_{n1}^2 \mathbb{D}^2 X_{n1} + \cdots + \delta_{nn}^2 \mathbb{D}^2 X_{nn})/d_n^2 \le \frac{1}{d_n^2} \sum_{i=1}^{n} \delta_{ni}^2 \to 0.$$

Thus $X_n \to \alpha$ stochastically as claimed, by Chebyshev's inequality. $\square$

Again we refer to Subsection 1.4 where it will be proved that such a sequence (G_n, d_n) of Cartesian sums is always convergent.

Corollary 1.18. *Let (Γ, δ) be admissible with normalized average degree $t(K_2, \Gamma, \delta) = \alpha$. Then the sequence $(\Gamma^{\square n}, n\delta)$ is α-regular.*

Regular sequences can be characterized in terms of homomorphism densities.

Proposition 1.19. *For an admissible sequence (G_n, d_n), the following are equivalent:*

(a) *The sequence (G_n, d_n) is α-regular.*
(b) *We have $t(K_2, G_n, d_n) \to \alpha$ and $t(P_3, G_n, d_n) \to \alpha^2$ as $n \to \infty$.*
(c) *For all forests F, we have $t(F, G_n, d_n) \to \alpha^{e(F)}$ as $n \to \infty$.*
(d) *For all rooted trees (F, o), we have $t((F, o), (G_n, p_n), d_n) \to \alpha^{e(F)}$ stochastically, as $n \to \infty$. Here p_n is a uniform random node of G_n.*

Note that the statement (d) for $F = K_2$ is exactly the same as (a).

Proof. Let X_n be the degree of a uniform random node of G_n, divided by d_n. Then X_n is a random variable with values in $[0, 1]$. The equivalence of (a) and (b) is clear since we have $t(K_2, G_n, d_n) = \mathbb{E}X_n$ and $t(P_3, G_n, d_n) = \mathbb{E}X_n^2$.

As (d) $\Rightarrow$ (c) $\Rightarrow$ (b) is trivial, it suffices to show that (a) implies (d). We use induction on $v(F)$. The case $v(F) = 1$ is trivial. Let $v(F) \geq 2$. Let o_i $(i = 1, \ldots, k)$ be the neighbors of o in F, i.e., $N(o) = \{o_1, \ldots, o_k\}$. Let F_i be the connected component of $F - o$ containing o_i.

Let $\epsilon > 0$. By the induction hypothesis, for $n \geq n_0(\epsilon)$ there exists an $S_n \subset V(G_n)$, with $|S_n| < \epsilon^2 v(G_n)$, such that for all $q \in V(G_n) - S_n$ and all i, we have

$$\left| \sqrt[e(F_i)]{t((F_i, o_i), (G_n, q), d_n)} - \alpha \right| < \epsilon.$$

Let T_n be the set of nodes in G_n that have at least ϵd_n neighbors in S_n. Since all nodes in S_n have at most d_n neighbors, we have $|T_n| \leq |S_n|/\epsilon < \epsilon v(G_n)$. For all $p \in V(G_n) - T_n$, we have

$$0 \leq \hom((F, o), (G_n, p)) - \hom((F, o, N(o)), (G_n, p, V(G_n) - S_n)) \leq k\epsilon d_n^{v(F)-1}.$$

Let U_n be the set of nodes in G_n whose degree divided by d_n is not in $(\alpha - \epsilon, \alpha + \epsilon)$. For $n \geq n_0(\epsilon)$, we have $|U_n| < \epsilon v(G_n)$ by (a). For all $p \in V(G_n) - T_n - U_n$, we have

$$((\alpha - 2\epsilon)d_n)^{e(F)} \leq \hom((F, o, N(o)), (G_n, p, V(G_n) - S_n)) \leq ((\alpha + \epsilon)d_n)^{e(F)}$$

and therefore

$$(\alpha - 2\epsilon)^{e(F)} \leq t((F, o), (G_n, p), d_n) \leq (\alpha + \epsilon)^{e(F)} + k\epsilon.$$

This is true for all $\epsilon > 0$, $n \geq n_0(\epsilon)$, and $p \in V(G_n) - T_n - U_n$, where $|T_n| + |U_n| < 2\epsilon v(G_n)$. Statement (d) follows. $\square$

Proposition 1.20. *For an admissible sequence (G_n, d_n), the following are equivalent:*

(a) *The sequence (G_n, d_n) is 1-regular.*
(b) *The average degree in G_n is asymptotically d_n.*
(c) *We have $t(K_2, G_n, d_n) \to 1$ as $n \to \infty$.*
(d) *For all forests F, we have $t(F, G_n, d_n) \to 1$ as $n \to \infty$.*

Proof. The equivalence of (a) and (b) is clear since (G_n, d_n) is admissible.

Observe that $t(K_2, G, d)$ is the average degree in G, divided by d. This shows the equivalence of (b) and (c).

Since trivially (d) $\Rightarrow$ (c), it suffices to prove (a) $\Rightarrow$ (d). This follows from Proposition 1.19 and Remark 1.14. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

1.4. Sequences with large essential girth.

Definition 1.21. The graph sequence (G_n) *has large girth* if, for any $k \geq 3$, we have $\mathrm{inj}(C_k, G_n) = 0$ for $n \geq n_0(k)$. The admissible sequence (G_n, d_n) *has large essential girth* if, for all $k \geq 3$, we have $t_{\mathrm{inj}}(C_k, G_n, d_n) \to 0$ as $n \to \infty$.

Remark 1.22. If (G_n, d_n) has large essential girth and G'_n is a spanning subgraph of G_n, then (G'_n, d_n) has large essential girth.

Remark 1.23. The injective homomorphism density of a cycle in a graph G satisfies

$$(1.2) \qquad\qquad t_{\mathrm{inj}}(C_k, G, d) \leq \frac{N(G)}{\mathrm{v}(G)d(d-1)^{k-2}},$$

where $N(G)$ is the number of non-backtracking walks of length $k-1$ in G whose starting point and endpoint are adjacent. The fraction on the right hand side has the following interpretation in terms of random walks. Choose $v_0 \in V(G)$ uniformly at random. With probability $\deg(v_0)/d$, let v_1 be a neighbor of v_0 chosen uniformly at random. With probability $1 - \deg(v_0)/d$, do not define v_1. If $i \geq 2$ and v_{i-1} is defined, then with probability $(\deg(v_{i-1}) - 1)/(d-1)$, let v_i be a neighbor of v_{i-1}, distinct from v_{i-2}, chosen uniformly at random. With probability $1 - (\deg(v_{i-1}) - 1)/(d-1)$, do not define v_i. Then the right hand side of (1.2) is the probability that v_{k-1} is defined and adjacent to v_0.

When G is d-regular, v_i is almost surely defined for every i.

For example, we look at two classical examples of regular graphs with intermediate density: hypercubes and projective planes (and their generalizations below). Let $Q_n = \{0, 1\}^n$ be the hypercube graph.

Proposition 1.24.

(a) *Let $q_n \to \infty$ and let G_n be the (bipartite) incidence graph of points and hyperplanes in a projective space of order q_n and dimension r_n. Let*

$$d_n = (q_n^{r_n} - 1)/(q_n - 1).$$

Then G_n is d_n-regular and (G_n, d_n) has large essential girth.

(b) *Consider a triangular array $(\Gamma_{ni}, \delta_{ni})$ $(1 \leq i \leq n)$ of admissible pairs. Let G_n be a subgraph of $\Gamma_{n1} \square \ldots \square \Gamma_{nn}$, and let $d_n = \delta_{n1} + \cdots + \delta_{nn}$. As in Example 1.17, assume that (1.1) holds. Then the sequence (G_n, d_n) has large essential girth.*

In particular, if (G, d) is admissible, then the sequence $(G^{\square n}, nd)$ has large essential girth. For example, the sequence (Q_n, n) has large essential girth. More generally, if G_n is any finite subgraph of the n-dimensional grid $\mathbb{Z}^n$, then the sequence $(G_n, 2n)$ has large essential girth.

Proof. (b) We make use of Remark 1.23. It suffices to prove that if we do in G_n the random walk defined there, then for any fixed $k \geq 3$, the probability that v_{k-1} exists and is adjacent to v_0 tends to 0 as $n \to \infty$. Clearly, the Hamming distance

of v_0 and v_{k-1} will be $k - 1 \geq 2$ with probability tending to 1 as $n \to \infty$. Indeed, when doing (at most) $k - 1$ steps, the probability that there will be two steps in the same coordinate goes to zero as $n \to \infty$, because of (1.1).

(a) We omit the subscript n for easier reading. The regularity claim is clear since any hyperplane has d points and any point is on d hyperplanes.

We have

$$\text{inj}(C_k, G) \leq \left(\frac{q^{r+1} - 1}{q - 1} \right)^{k/2} \left(\frac{q^{r-1} - 1}{q - 1} \right)^{k/2},$$

whence

$$t_{\text{inj}}(C_k, G, d) \leq \frac{(q^{r+1} - 1)^{k/2 - 1}(q^{r-1} - 1)^{k/2}}{2(q^r - 1)^{k-1}} \leq \frac{q^{r-1} - 1}{2(q^r - 1)} < \frac{1}{2q} \to 0$$

as $n \to \infty$. $\qquad\square$

Statement (a) is maybe a bit surprising since in a large girth 1-regular sequence (G_n, d_n), the number $v(G_n)$ of nodes would have to be superpolynomial in d_n, whereas for projective spaces we have $d = (q^r - 1)/(q - 1)$ and

$$v(G) = 2(q^{r+1} - 1)/(q - 1) = 2(qd + 1) < 2 \left(d^{r/(r-1)} + 1 \right).$$

This means in particular that we cannot delete $o(v(G_n)d_n)$ edges from G_n to make the sequence have large girth (if $r_n \geq 2$ for all n). This is in contrast to the bounded degree case. Thus, the word 'essential' is essential. Another instance of this will be Proposition 3.7.

In other words, $v(G_n)/d_n$ can go to ∞ arbitrarily slowly (compared to $v(G_n)$ and d_n) in a 1-regular sequence of large essential girth: the dimension r and thus the cardinality of the projective space can grow arbitrarily fast compared to the order q, while we have $v(G)/d \sim 2q$. A sequence of large essential girth can thus be almost dense. It it easy to see, however, that it cannot be dense:

Proposition 1.25. *If (G_n, d_n) is an admissible sequence with large essential girth, where $v(G_n) \to \infty$ and $d_n = O(v(G_n))$, then*

$$\frac{d_n}{v(G_n)} t(K_2, G_n, d_n) = \frac{2 \, e(G_n)}{v^2(G_n)} \to 0$$

as $n \to \infty$.

Proof. We have

$$\frac{2 \, e(G_n)}{v^2(G_n)} = t(K_2, G_n) \leq \sqrt[4]{t(C_4, G_n)} = O \left(\sqrt[4]{t(C_4, G_n, d_n)} \right),$$

where

$$t(C_4, G_n, d_n) = t_{\text{inj}}(C_4, G_n, d_n) + O(1/d_n)$$

and $t_{\text{inj}}(C_4, G_n, d_n) \to 0$ as $n \to \infty$. Also,

$$\frac{d_n}{v(G_n)} t(K_2, G_n, d_n) \leq \frac{d_n}{v(G_n)}.$$

The proposition follows. $\qquad\square$

A further important example of regular sequences of large essential girth is given by random graphs. Let $G(n, d)$ be the random (almost) d-regular multigraph generated by the configuration model: we take n nodes with d legs emanating from

each node, and take a uniform random perfect matching on the dn legs (if dn is odd, leave out a leg). Let $G(n, d)^{\mathrm{simp}}$ be the underlying simple graph.

Proposition 1.26. *Fix* $\epsilon > 0$. *Let* $d_n = O(n^{1-\epsilon})$ *and let the random graph* G_n *have the distribution of* $G(n, d_n)^{\mathrm{simp}}$. *Then the sequence* (G_n, d_n) *is a.s. 1-regular and has large essential girth.*

Proof. Let L be the proportion of loops among the edges of $G(n, d_n)$. It is easy to see that $\mathbb{E}L^2 = O(1/n^2)$, whence $L \to 0$ a.s.

Let r be so large that $\sum (d_n/n)^r < \infty$.

For easier reading, we omit the subscript n from d_n and G_n.

Let M be the proportion of edges in $G(n, d)$ that have a parallel edge. We have $\mathbb{E}M^r = O((d/n)^r)$, whence $M \to 0$ a.s. Thus, the sequence is a.s. 1-regular.

Let $k \geq 3$. Using Proposition 1.12, we have

$$\mathbb{E}t_{\mathrm{inj}}^r(C_k, G, d) \leq \mathbb{E}t_{\mathrm{inj}}(C_k^r, G, d) + \frac{1}{n} \sum_{F'} \mathbb{E}t_{\mathrm{inj}}(F', G, d) + O\left(\frac{1}{n^2}\right),$$

where F' runs over quotients of C_k^r into which each C_k maps injectively, such that F' has $r - 1$ components. It is easy to see that $\mathbb{E}t_{\mathrm{inj}}(C_k^r, G, d) = O((d/n)^r)$ and $\mathbb{E}t_{\mathrm{inj}}(F', G, d) = O((d/n)^{r-1})$ for each F', whence $\sum \mathbb{E}t_{\mathrm{inj}}^r(C_k, G, d) < \infty$ and therefore $t_{\mathrm{inj}}(C_k, G, d) \to 0$ a.s. $\square$

This concludes our set of examples of sequences with large essential girth. Putting together Remark 1.9 and Propositions 1.10 and 1.19, we obtain

Proposition 1.27. *If* $d_n \to \infty$ *and* (G_n, d_n) *is* α-*regular and has large essential girth, then*

(a) *the homomorphism density* $t(F, G_n, d_n)$ *and the injective homomorphism density* $t_{\mathrm{inj}}(F, G_n, d_n)$ *converge to* $\alpha^{\mathrm{e}(F)}$ *for any tree* F *and to* 0 *for any other connected* F;

(b) *the sequence* (G_n, d_n) *is convergent.*

In particular, hypercubes, or — more generally — Cartesian powers of a fixed graph, or grids of size $n \times \cdots \times n$ where both n and the dimension tend to ∞, or point-hyperplane incidence graphs of projective spaces whose order tends to ∞, or the random graphs of Proposition 1.26, form convergent sequences.

2. CONVERGENCE OF SPECTRA

Let $\sigma_{G,d}$ be the uniform probability measure on the $\mathrm{v}(G)$ numbers λ/d, where λ runs over the eigenvalues of G. If (G, d) is admissible, then all eigenvalues of G are in $[-d, d]$, so $\sigma_{G,d}$ is supported on $[-1, 1]$. We have

$$\epsilon^2 \sigma_{G,d}(\{x : |x| \geq \epsilon\}) \leq \int_{-1}^{1} x^2 \mathrm{d}\sigma_{G,d}(x) = \frac{2\,\mathrm{e}(G)}{\mathrm{v}(G)d^2} = \frac{t(K_2, G, d)}{d} \leq \frac{1}{d}.$$

This proves

Proposition 2.1. *Let* (G_n, d_n) *be an admissible sequence with* $d_n \to \infty$. *Then the measure* $\sigma_n = \sigma_{G_n, d_n}$ *converges weakly to the Dirac measure at* 0. *More precisely,*

$$\sigma_n((-\epsilon, \epsilon)) \geq 1 - 1/(\epsilon^2 d_n)$$

for all n *and all* $\epsilon > 0$.

This is probably well known but we couldn't find a reference.

Up to now, we used only the second moment of $\sigma_{G,d}$, but to get more precise results for convergent sequences, we shall need the other moments as well. The 0-th moment is 1, the first moment is 0, and we have

$$\int x^k \mathrm{d}\sigma_{G,d}(x) = \frac{\mathrm{hom}(C_k, G)}{\mathrm{v}(G)d^k} = \frac{t(C_k, G, d)}{d}$$

for $k \geq 3$. In fact, this formula holds for $k \geq 1$ if we agree that $C_2 = K_2$ and C_1 is a node with a loop. We infer

Lemma 2.2. *Let (G_n, d_n) be convergent and $g : [-1, 1] \to \mathbb{R}$ be continuous. Then*

$$d_n \int_{-1}^{1} x^2 g(x)\mathrm{d}\sigma_{G_n,d_n}(x)$$

converges as $n \to \infty$.

Proof. For $g(x) = x^k$, $k \geq 0$, the statement is clear from the preceding discussion. The general case follows by the Weierstrass approximation theorem. □

For Benjamini–Schramm convergent graph sequences (d_n independent of n), it is well known that the spectral measure converges weakly, to a non-trivial measure in general.

For convergent dense graph sequences, C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi (see [9, Subsection 6.3] and [26, Section 11.6]) have given a much more precise description of the limiting behavior of the spectrum than the one in Proposition 2.1. Namely, the k-th largest (resp. k-th smallest) eigenvalue, divided by the number of nodes, converges to the k-th largest (smallest) eigenvalue of the limiting graphon, which is non-negative (non-positive).

We shall now show that these two results (bounded degree and dense) carry over to intermediate density — at least partially: we don't (yet) have limit objects; cf. Section 4.

Let G have all degrees $\leq d$. Let $1 \leq r \leq \mathrm{v}(G)$ be an integer. Let $\sigma_{G,d,r}$ and $\sigma'_{G,d,r}$ be the uniform probability measures on the numbers λ/d, where λ runs over the r largest and r smallest eigenvalues of G, respectively. These measures are supported on $[-1, 1]$. Note that $\sigma_{G,d,\mathrm{v}(G)} = \sigma'_{G,d,\mathrm{v}(G)} = \sigma_{G,d}$. For any r, the probability measures $\sigma_{G,d,r}$ and $\sigma'_{G,d,r}$ are the restrictions of the measure $(\mathrm{v}(G)/r)\sigma_{G,d}$ to the intervals $[\lambda_r/d, 1]$ and $[-1, \lambda_{\mathrm{v}(G)-r+1}]$, respectively.

Theorem 2.3. *Let (G_n, d_n) be a convergent sequence with $d_n \to \infty$. Let $1 \leq r_n \leq \mathrm{v}(G_n)$ ($n = 1, 2, \ldots$) be integers such that $r_n d_n / \mathrm{v}(G_n)$ converges to a positive limit α. Then the measures $\sigma_n = \sigma_{G_n,d_n,r_n}$ and $\sigma'_n = \sigma'_{G_n,d_n,r_n}$ converge weakly to probability measures σ supported on $[0, 1]$ and σ' supported on $[-1, 0]$, respectively.*

Proof. We only treat σ_n since everything works the same way for σ'_n.

Let λ_r be the r-th largest eigenvalue of the graph G with all degrees $\leq d$. We have

$$0 = \sum_{i=1}^{\mathrm{v}(G)} \lambda_i \leq (\mathrm{v}(G) - r)\lambda_r + rd,$$

whence

$$\frac{\lambda_r}{d} \geq -\frac{r}{\mathrm{v}(G) - r}.$$

Thus, the measure σ_n is supported on the halfline with left endpoint

$$-\frac{r_n}{\mathrm{v}(G_n) - r_n} \to 0$$

since $r_n/\mathrm{v}(G_n) \sim \alpha/d_n \to 0$. Thus, it suffices to show that for any $0 < a < b < 1$, we have

(2.1) $$\liminf \sigma_n([a, 1]) \geq \limsup \sigma_n([b, 1]).$$

Let $g : [-1, 1] \to [0, 1]$ be continuous, non-decreasing, $g(a) = 0$, $g(b) = 1$.

For all n, either $\sigma_n([a, 1]) = 1$ or

$$\sigma_n([a, 1]) = (\mathrm{v}(G_n)/r_n)\sigma_{G_n, d_n}([a, 1]).$$

Hence, $\liminf \sigma_n([a, 1]) = 1$ or

$$\liminf \sigma_n([a, 1]) \geq \liminf \left(\frac{d_n}{\alpha} \int g \mathrm{d}\sigma_{G_n, d_n}\right) = \frac{1}{\alpha} \lim \left(d_n \int g \mathrm{d}\sigma_{G_n, d_n}\right).$$

On the other hand, for all n, we have $\sigma_n([b, 1]) \leq 1$ and

$$\sigma_n([b, 1]) \leq \int g \mathrm{d}\sigma_n \leq \frac{\mathrm{v}(G_n)}{r_n} \int g \mathrm{d}\sigma_{G_n, d_n}.$$

Hence, $\limsup \sigma_n([b, 1]) \leq 1$ and

$$\limsup \sigma_n([b, 1]) \leq \limsup \left(\frac{d_n}{\alpha} \int g \mathrm{d}\sigma_{G_n, d_n}\right) = \frac{1}{\alpha} \lim \left(d_n \int g \mathrm{d}\sigma_{G_n, d_n}\right),$$

and the inequality (2.1) follows. $\square$

3. Graph polynomials

The convergence of a sequence (G_n, d_n) was defined in Section 1 by the convergence of certain graph parameters, the homomorphism densities. This forces certain further parameters to converge (sometimes only under further conditions); such parameters are called *estimable* (some parameters are only estimable for a certain class of convergent sequences). Theorem 2.3 can be thought of as an estimability statement. In this section, we present some more estimable parameters.

Following the paper [11] by P. Csikvári and the present author, let f be an isomorphism-invariant monic multiplicative graph polynomial of linearly bounded exponential type. I.e.,

- for every graph G, a monic polynomial $f(G, x) \in \mathbb{C}[x]$ of degree $\mathrm{v}(G)$ is given,
- $f(G_1, x) = f(G_2, x)$ if $G_1 \simeq G_2$,
- $f(G_1 \cup G_2, x) = f(G_1, x)f(G_2, x)$ for any disjoint union,
-
$$f(G, x + y) = \sum_{S \subseteq V(G)} f(G[S], x)f(G[V(G) - S], y)$$

  for all G, and finally

-

(3.1) $$\sum\{|f'(G[S], 0)| : v \in S \subseteq V(G), |S| = t\} \leq (cd)^{t-1}$$

  for all G with maximal degree $\leq d$, all $v \in V(G)$, and all $t \geq 1$, with a constant c depending only on f.

Examples include the chromatic, adjoint, and Laplacian characteristic polynomials, and also the modified matching polynomial defined as

$$M(G, x) = \sum_{k=0}^{\lfloor \mathrm{v}(G)/2 \rfloor} (-1)^k m_k(G) x^{\mathrm{v}(G)-k},$$

where $m_k(G)$ is the number of matchings in G that consist of k edges.

The characteristic polynomial $f(G, x) = \det(xI - A_G)$ of the adjacency matrix of G is not a valid example because it is not of exponential type. Nevertheless everything that follows, including Theorem 3.1 below, applies to this case in a trivial way; in fact, much more is true, even without assuming graph convergence, as we have seen in Proposition 2.1.

We wish to study the distribution of roots of $f(G, x)$. By [11, Theorem 1.6], we can choose a constant C depending only on f such that for any G, all roots have absolute value $\leq Cd$. It is shown there that $C = 7.04 \cdot c$ is an appropriate choice if c is the constant in (3.1). For some of the specific graph polynomials mentioned above, smaller appropriate values of C are known.

Let $p_k(G)$ be the k-th power sum of the roots of $f(G, x)$. By [11, Theorem 5.6.(b)], for each $k \geq 1$, there exist constants $c_k(F)$ such that

$$(3.2) \qquad p_k(G) = \sum_{2 \leq \mathrm{v}(F) \leq k+1} c_k(F) \operatorname{inj}(F, G)$$

for all G, where F runs over the isomorphism classes of connected graphs. We also have

$$p_0(G) = \mathrm{v}(G) = c_0(K_1) \operatorname{inj}(K_1, G),$$

where $c_0(K_1) = 1$.

Let $\nu_{G,d}$ be the uniform probability measure on the points λ/d, where λ runs over the roots of $f(G, x)$. This measure is supported on the disc of radius C and has k-th holomorphic moment

$$(3.3) \qquad \begin{aligned} \int z^k d\nu_{G,d}(z) &= \frac{1}{\mathrm{v}(G)} \sum_{f(G,\lambda)=0} \frac{\lambda^k}{d^k} = \frac{p_k(G)}{\mathrm{v}(G)d^k} = \sum_{2 \leq \mathrm{v}(F) \leq k+1} c_k(F) \frac{\operatorname{inj}(F, G)}{\mathrm{v}(G)d^k} \\ &= \sum_{2 \leq \mathrm{v}(F) \leq k+1} c_k(F) \frac{(d-1)^{\mathrm{v}(F)-2}}{d^{k-1}} t_{\operatorname{inj}}(F, G, d) \end{aligned}$$

for $k \geq 1$.

Theorem 3.1. *Let $d_n \to \infty$. Let (G_n, d_n) be a convergent sequence, or, more generally, an admissible sequence such that $t(F, G_n, d_n)$ converges whenever $c_{\mathrm{v}(F)-1}(F) \neq 0$. Write*

$$t(F) = \lim_{n \to \infty} t(F, G_n, d_n)$$

for the limiting homomorphism density. Set $\nu_n = \nu_{G_n, d_n}$.

(1) *For all $k \geq 0$, we have*

$$\int z^k d\nu_n(z) \to \sum_{\mathrm{v}(F)=k+1} c_k(F) t(F)$$

as $n \to \infty$.

(2) *For any function $g(z)$ that is continuous for $|z| \leq C$ and harmonic for $|z| < C$, the integral $\int g(z) d\nu_n(z)$ converges as $n \to \infty$.*

(3) *For any $|\xi| > C$, the normalized absolute value*

$$\frac{\sqrt[\mathrm{v}(G_n)]{|f(G_n, \xi d_n)|}}{d_n}$$

of f converges to a positive limit.

(4) *If $f(G_n, x)$ has only real roots for all n, then ν_n converges weakly.*

For the bounded degree case, the analogous theorem is [11, Theorem 1.10], which, in turn, was a generalization (with a simpler proof) of the result of M. Abért and T. Hubai [2, Theorems 1.1, 1.2], who first discovered this phenomenon in the case of the chromatic polynomial. For the dense case, essentially the same was proved by P. Csikvári, J. Hladký, T. Hubai and the author in [12, Theorems 1.4, 1.5, 4.3], using the approach of [11]. The proof carries over to intermediate density almost unchanged.

Proof.   (1) The 0-th moment is always 1. Let $k \geq 1$. From (3.3), we have

$$\int z^k d\nu_n(z)$$

$$= \sum_{2 \leq \mathrm{v}(F) \leq k+1} c_k(F) \frac{(d_n - 1)^{\mathrm{v}(F)-2}}{d_n^{k-1}} t_{\mathrm{inj}}(F, G_n, d_n) \to \sum_{\mathrm{v}(F)=k+1} c_k(F) t(F)$$

as $n \to \infty$.

(2) The claim follows from (1) because $g(z)$ can be uniformly approximated by real parts of polynomials.

(3) For any G, d, and ξ, we have

$$\log \frac{\sqrt[\mathrm{v}(G)]{|f(G, \xi d)|}}{d} = \frac{1}{\mathrm{v}(G)} \log \frac{|f(G, \xi d)|}{d^{\mathrm{v}(G)}} = \frac{1}{\mathrm{v}(G)} \log \prod_{i=1}^{\mathrm{v}(G)} \left| \xi - \frac{\lambda_i}{d} \right|,$$

where the λ_i are the roots of $f(G, x)$. The last expression can be rewritten as

$$\frac{1}{\mathrm{v}(G)} \sum_{i=1}^{\mathrm{v}(G)} \log \left| \xi - \frac{\lambda_i}{d} \right| = \int g(z) \mathrm{d}\nu_{G,d},$$

where $g(z) = \log |\xi - z|$. The claim now follows from the previous statement (2).

(4) The claim follows from (1) because each ν_n is supported on the interval $[-C, C]$.   □

3.1. **Number of proper colorings (large essential girth case).** We now wish to prove, for intermediate density graph sequences of large essential girth, a qualitative variant of Abért and Hubai's [2, Theorem 1.4] about the asymptotic number of proper colorings. They only treated the large girth case, but gave an explicit bound on the error in their formula.

Let $\mathrm{ch}(G, x)$ be the chromatic polynomial of the graph G. I.e., for integral $q \geq 0$, $\mathrm{ch}(G, q)$ is the number of proper q-colorings of G.

Theorem 3.2. *Let (G_n, d_n) be a sequence of large essential girth, such that $d_n \to \infty$ and*

$$t(K_2, G_n, d_n) = \frac{2\,\mathrm{e}(G_n)}{\mathrm{v}(G_n) d_n} \to t(K_2)$$

as $n \to \infty$. Let $|\xi| \geq 8$. Then

$$(3.4) \qquad \frac{\sqrt[\mathrm{v}(G_n)]{|\,\mathrm{ch}(G_n, \xi d_n)|}}{|\xi| d_n} \to \exp(-t(K_2)\Re(1/2\xi)).$$

Proof. We have

$$\log \frac{\sqrt[\mathrm{v}(G)]{|\,\mathrm{ch}(G, \xi d)|}}{|\xi| d} = \int_{|z| \leq C} \log \left| 1 - \frac{z}{\xi} \right| d\nu_{G,d}(z)$$

$$(3.5) \qquad = -\sum_{k=1}^{\infty} \frac{1}{k} \Re \int_{|z| \leq C} \left(\frac{z}{\xi} \right)^k d\nu_{G,d}(z),$$

where $\nu_{G,d}$ is the uniform probability measure on the $\mathrm{v}(G)$ points λ/d for which $\mathrm{ch}(G, \lambda) = 0$, and $C < 8$ is Sokal's constant such that $|\lambda| \leq Cd$ for all λ. The series on the right hand side of (3.5) converges uniformly in G and d.

By [11, Theorem 6.6], in the formula (3.2) for the power sum $p_k(G)$ of the roots of $\mathrm{ch}(G, x)$, the coefficient $c_k(F)$ is 0 unless F is 2-connected. On the other hand, $t_{\mathrm{inj}}(F, G_n, d_n) \to 0$ if F contains a cycle. Thus, $c_k(F)t_{\mathrm{inj}}(F, G_n, d_n) \to 0$ unless F is a 2-connected tree, i.e., $F = K_2$. Note also that $c_1(K_2) = 1/2$ because $p_1(G) = \mathrm{e}(G) = \mathrm{inj}(K_2, G)/2$. From formula (3.3), we see that $\int z^k d\nu_n(z)$ tends to $t(K_2)/2$ for $k = 1$ and to 0 for $k \geq 2$.

Putting all this together, the logarithm of the left hand side of (3.4) tends to $-\Re(t(K_2)/2\xi)$, as claimed. $\qquad \square$

3.2. **Matching measure and graph convergence.** In this subsection, we prove intermediate degree analogs of some results of the recent paper [1] by Abért, Csikvári, Kun and the author. Contrary to the bounded degree case treated there, large girth will not play any role in relation to matchings.

Definition 3.3. Let G be a graph and let $m_k(G)$ denote the number of matchings of size k. Then the *matching polynomial* $\mu(G, x)$ is defined as follows:

$$\mu(G, x) = \sum_{k=0}^{\lfloor \mathrm{v}(G)/2 \rfloor} (-1)^k m_k(G) x^{\mathrm{v}(G)-2k}.$$

Note that $m_0(G) = 1$. Let $d > 0$ be an upper bound on all degrees in G. The *matching measure* $\rho_{G,d}$ is defined to be the uniform probability distribution on the points $\lambda/\sqrt{d}$, where λ runs over the roots of $\mu(G, x)$ (with multiplicity).

The fundamental theorem for the matching polynomial is the following.

Theorem 3.4 (Heilmann and Lieb [23]). (a) *The roots of the matching polynomial $\mu(G, x)$ are real.*
    (b) *If $d \geq 2$ is an upper bound for all degrees in G, then all roots of $\mu(G, x)$ have absolute value $\leq 2\sqrt{d-1}$.*

Many graph parameters related to matchings can be read off from the matching measure, for example, the number

$$\mathbb{M}(G) = \sum_{k=0}^{\lfloor \mathrm{v}(G)/2 \rfloor} m_k(G)$$

of all matchings and the number $\mathrm{pm}(G) = m_{\mathrm{v}(G)/2}$ of perfect matchings. The latter is zero if $\mathrm{v}(G)$ is odd.

Proposition 3.5.

(3.6) $$\log \frac{\mathbb{M}(G)^{2/\,\mathrm{v}(G)}}{d} = \int_{-2}^{2} \log\left(\frac{1}{d} + x^2\right) \mathrm{d}\rho_{G,d}(x).$$

(3.7) $$\log \frac{\mathrm{pm}(G)^{2/\,\mathrm{v}(G)}}{d} = 2\int_{-2}^{2} \log|x| \mathrm{d}\rho_{G,d}(x).$$

Proof. (3.6) The number of matchings in G is

$$\mathbb{M}(G) = \sum_{k=0}^{\lfloor \mathrm{v}(G)/2 \rfloor} m_k(G) = |\mu(G, \sqrt{-1})|.$$

Thus,

$$\frac{\log \mathbb{M}(G)}{\mathrm{v}(G)} - \frac{1}{2}\log d = \frac{\log|\mu(G, \sqrt{-1})|}{\mathrm{v}(G)} - \frac{1}{2}\log d$$

$$= \int_{-2}^{2} \log\left|\frac{\sqrt{-1}}{\sqrt{d}} - x\right| \mathrm{d}\rho_{G,d}(x) = \frac{1}{2}\int_{-2}^{2} \log\left(\frac{1}{d} + x^2\right) \mathrm{d}\rho_{G,d}(x).$$

(3.7) The number of perfect matchings in G is

$$\mathrm{pm}(G) = |\mu(G, 0)|.$$

Thus,

$$\frac{\log \mathrm{pm}(G)}{\mathrm{v}(G)} - \frac{1}{2}\log d = \frac{\log|\mu(G, 0)|}{\mathrm{v}(G)} - \frac{1}{2}\log d = \int_{-2}^{2} \log|x| \mathrm{d}\rho_{G,d}(x).$$

$\square$

Let

$$w(x) = \frac{\sqrt{4 - x^2}}{2\pi} \qquad (-2 \le x \le 2)$$

denote Wigner's semicircle density function. The *semicircle distribution on the interval* $[-2\beta, 2\beta]$ is the distribution of βX, where X is a random variable with density w.

Theorem 3.6. *Let $d_n \to \infty$. Let (G_n, d_n) be an admissible sequence with matching measures $\rho_n = \rho_{G_n, d_n}$.*

(a) *If $t(F, G_n, d_n)$ is convergent for any tree F, then the sequence of matching measures ρ_n converges weakly to a probability measure ρ on $[-2, 2]$. Moreover, we have*

$$\limsup_{n \to \infty} \log \frac{\mathbb{M}(G_n)^{2/\,\mathrm{v}(G_n)}}{d_n} \le 2\int_{-2}^{2} \log|x| \mathrm{d}\rho(x).$$

(b) *If the sequence (G_n, d_n) is α-regular, then ρ is the semicircle distribution on the interval $[-2\sqrt{\alpha}, 2\sqrt{\alpha}]$, and we have*

(3.8) $$\limsup_{n \to \infty} \frac{\mathbb{M}(G_n)^{2/\,\mathrm{v}(G_n)}}{d_n} \le \frac{\alpha}{e}.$$

For example, the matching polynomial of the complete graph K_n is the n-th Hermite polynomial, so (b) recovers the ancient fact that root distributions of Hermite polynomials converge to the semicircle law [19, 20, 24, 33]. Similarly, complete bipartite graphs $K_{n,n}$ yield Laguerre polynomials.

When each graph G_n is d_n-regular, the first statement in (b) has been also independently obtained by Abért, Csikvári and Hubai with a different proof (unpublished), and the inequality in (b) follows from the much stronger result of Davies, Jenssen, Perkins and Roberts [13, Theorem 4].

When each graph G_n is d_n-regular and bipartite, $\mathrm{pm}(G_n)^{2/\,\mathrm{v}(G_n)} \sim d_n/e$, which is well known to follow from classical results of Brègman ($\leq$) and Schrijver ($\geq$); see [27, pp. 311–312]. From the inequality in (b), we see that $\mathbb{M}(G_n)^{2/\,\mathrm{v}(G_n)} \sim d_n/e$ as well, and we only need Schrijver's lower bound

$$\mathrm{pm}(G)^{2/\mathrm{v}(G)} \geq \frac{(d-1)^{d-1}}{d^{d-2}} \sim \frac{d}{e} \qquad (d \to \infty)$$

on the number of perfect matchings to get this. Note that Propp's 1999 survey on the enumeration of matchings cites [10] for the asymptotic formula for the number of perfect matchings of the hypercube, and asks for a formula for the number of all matchings [30, Problem 19].

Leaving regular graphs, note that statement (a) applies in particular to the special case when (G_n, d_n) is convergent. The first claim in (a), for the special case of convergent dense graph sequences, is [12, Theorem 4.3] of Csikvári, Hladký, Hubai and the author.

We prove Theorem 3.6.

Proof. By the Heilmann–Lieb Theorem, the measures ρ_n are all supported on $[-2, 2]$. We shall exploit the relation between the modified and the ordinary matching polynomial: $M(G, x^2) = x^{\mathrm{v}(G)}\mu(G, x)$. Let $\nu_{G,d}$ be the uniform probability measure on the points λ/d, where λ runs over the roots of the modified matching polynomial $M(G, x)$. This measure is supported on the interval $[0, 4]$.

There is a very nice interpretation of the $2k$-th power sum of the roots of the matching polynomial $\mu(G, x)$. It counts the number of closed tree-like walks of length $2k$ in the graph G [21, Chapter 6]. Note that for $k \geq 1$, this is twice the k-th power sum of the roots of the modified matching polynomial $M(G, x)$. Thus, in the formula (3.2) written for the graph polynomial $M(G, x)$, the coefficient $c_k(F)$ is half the number of tree-like walks of length $2k$ in F that use all edges of F, divided by aut F. Thus, $c_{\mathrm{v}(F)-1} = 0$ unless F is a tree.

(a) Let $\nu_n = \nu_{G_n, d_n}$. By Theorem 3.1(4), ν_n converges weakly as $n \to \infty$. But from $\nu_{G,d}$ we get $\rho_{G,d}$ by decreasing the mass at 0 by $1/2$ and then relocating the mass of any point x to both points $\pm\sqrt{x}$, so as to get a probability measure again. This operation clearly preserves weak convergence. Thus, ρ_n also converges weakly to a measure ρ.

Let $u(x) = 2\log|x|$ and

$$u_k(x) = \log\left(\frac{1}{k} + x^2\right)$$

for $k = 1, 2, \ldots$. Then

$$\log \frac{\mathbb{M}(G)^{2/\,\mathrm{v}(G)}}{d} \leq \int_{-2}^{2} u_k \mathrm{d}\rho_{G,d}$$

if $d \geq k$. Thus, for any k,

$$\limsup_{n\to\infty} \log \frac{\mathbb{M}(G_n)^{2/\operatorname{v}(G_n)}}{d_n} \leq \lim_{n\to\infty} \int_{-2}^{2} u_k \mathrm{d}\rho_n = \int_{-2}^{2} u_k \mathrm{d}\rho,$$

since the measures ρ_n are supported on the compact interval $[-2, 2]$ not depending on n, and u_k is continuous and bounded on $[-2, 2]$.

Since $u_k \geq u_{k+1}$ and $u_k \to u$ pointwise, the claim follows using the Monotone Convergence Theorem.

(b) Matching measures are symmetric about 0, and so is the semicircle measure, so it suffices to show convergence of even moments of ρ_n to those of the semicircle law. Let $k \geq 1$. By Theorem 3.1, we have

$$(3.9) \qquad \int_{-2}^{2} x^{2k} \mathrm{d}\rho_n(x) = 2 \int_{0}^{4} x^k \mathrm{d}\nu_n(x) \to \sum_{\operatorname{v}(F)=k+1} 2c_k(F)t(F),$$

where

$$t(F) = \lim_{n\to\infty} t(F, G_n, d_n) = \alpha^k$$

by Proposition 1.19. So the limit in (3.9) is α^k times the number of non-isomorphic pairs (F, γ), where F is a tree with $k + 1$ nodes and γ is an Eulerian trail in the graph $\bar{F}$ which is F with all edges doubled. These pairs (F, γ) correspond to Dyck words of length $2k$, so their number is the Catalan number

$$\frac{1}{k+1} \binom{2k}{k} = \int_{-2}^{2} x^{2k} w(x) dx = \mathbb{E}X^{2k},$$

where X has density w. Therefore

$$\int_{-2}^{2} x^{2k} \mathrm{d}\rho(x) = \lim_{n\to\infty} \int_{-2}^{2} x^{2k} \mathrm{d}\rho_n(x) = \alpha^k \mathbb{E}X^{2k} = \mathbb{E}\left(\sqrt{\alpha}X\right)^{2k}$$

as claimed.

The inequality (3.8) is immediate from statement (a) and the fact that

$$\int_{-2}^{2} w(x) \log |x| \mathrm{d}x = -\frac{1}{2};$$

cf. [22, integral 4.241.9]. $\qquad\qquad\square$

3.3. **Spectral measure rescaled (regular, large girth case).** We know from Proposition 2.1 that scaling down the spectrum by the degree bound $d \to \infty$ leads to trivial behavior in terms of weak convergence. What happens if we only scale down by $\sqrt{d}$?

Let $\Sigma_{G,d}$ be the uniform probability measure on the $\operatorname{v}(G)$ points $\lambda/\sqrt{d}$, where λ runs over the eigenvalues of G. If all degrees in G are $\leq d$, then $\Sigma_{G,d}$ is supported on the interval $\left[-\sqrt{d}, \sqrt{d}\right]$.

Proposition 3.7. *Let (G_n, d_n) be an α-regular sequence of large girth, such that $d_n \to \infty$. Let $\Sigma_n = \Sigma_{G_n, d_n}$. Then, for each $k \geq 0$,*

$$(3.10) \qquad \int_{-\sqrt{d}}^{\sqrt{d}} x^k \mathrm{d}\Sigma_n \to \int_{-2}^{2} \left(\sqrt{\alpha}x\right)^k w(x) \mathrm{d}x$$

as $n \to \infty$. Thus, Σ_n converges weakly to the semicircle distribution on the interval $[-2\sqrt{\alpha}, 2\sqrt{\alpha}]$.

Note that the limit in (3.10) is 0 for k odd and is $\alpha^{k/2}$ times the Catalan number $\frac{1}{k/2+1}\binom{k}{k/2}$ for k even.

Proof. We have

$$\int_{-\sqrt{d}}^{\sqrt{d}} x^k \mathrm{d}\Sigma_n = \frac{1}{\mathrm{v}(G_n)} \sum \left(\frac{\lambda}{\sqrt{d_n}}\right)^k = \frac{\hom(C_k, G_n)}{\mathrm{v}(G_n)d_n^{k/2}}.$$

For $n \geq n_0(k)$, all walks in G_n of length k are tree-like, whence

$$\int_{-\sqrt{d}}^{\sqrt{d}} x^k \mathrm{d}\Sigma_n(x) = \int_{-2}^{2} x^k \mathrm{d}\rho_n(x) \to \int_{-2}^{2} \left(\sqrt{\alpha}x\right)^k w(x)\mathrm{d}x$$

as $n \to \infty$, by Theorem 3.6. To deduce the weak convergence, we use that the semicircle measure is compactly supported. □

A different proof is possible based on the fact that Kesten–McKay measures converge to the semicircle law.

For random graphs, results similar to Proposition 3.7 have been proved by Dumitriu and Pal [15] and by Tran, Vu and Wang [34]. Those results are of course much deeper than Proposition 3.7.

Proposition 3.7 fails for large *essential* girth, even if each G_n is exactly d_n-regular. Indeed, for the hypercube sequence (Q_d, d), the measure Σ_d is the (binomial) distribution of

$$(X_1 + \cdots + X_d)/\sqrt{d},$$

where the X_i are i.i.d. random variables with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$; see [25, Exercise 11.9]. Thus, Σ_d converges weakly to the standard Gaussian distribution and not to the semicircle distribution, therefore its moments do not all converge to those of the semicircle law.

4. GRAPHONINGS

We propose a common generalization of graphons and graphings.

Definition 4.1. A *graphoning* is a tuple $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$, where $(X, \mathcal{B}, \lambda)$ is a probability space, $\mu : \mathcal{B} \to [0, \infty]$ is a measure, and $W : X^2 \to [0, 1]$ is a symmetric $(\mathcal{B} \otimes \mathcal{B})$-measurable function such that

- (degree bound)

$$\deg_X(x) \stackrel{\text{def}}{=} \int_X W(x, y)\mathrm{d}\mu(y) \leq 1$$

  for all $x \in X$,
- (degree measurability)

$$(4.1) \qquad \deg_A(x) \stackrel{\text{def}}{=} \int_A W(x, y)\mathrm{d}\mu(y)$$

  is a measurable function of $x \in X$ for all $A \in \mathcal{B}$, and
- (measure preserving property)

$$(4.2) \qquad \int_A \deg_B \mathrm{d}\lambda = \int_B \deg_A \mathrm{d}\lambda$$

  for all $A, B \in \mathcal{B}$.

A graphoning with $\lambda = \mu$ is the same thing as a graphon, except that graphons are measurable only w.r.t. the completion of $\mathcal{B} \otimes \mathcal{B}$, and thus their degrees are only almost measurable. A graphoning on a Borel probability space $(X, \mathcal{B}, \lambda)$, such that μ is the counting measure divided by d, and W only takes values in $\{0, 1\}$, is the same thing as a graphing.

For these two special cases of graphonings, it is known that degree measurability follows from the degree bound condition. It is unclear to the author whether this holds for general graphonings, maybe under the assumption that the σ-algebra $\mathcal{B}$ is Borel.

Note that μ is not in general σ-finite, so the Fubini Theorem is not applicable to the iterated integrals in (4.2).

4.1. Sub-Markov kernels and rooted homomorphism densities.

We wish to generalize the homomorphism densities of graphons that play a fundamental role in the limit theory of dense graphs developed by László Lovász and his coauthors [9, 26, 28]. Technical difficulties are caused by the lack of the Fubini Theorem, but these can be dealt with. We treat rooted homomorphism densities first. Even this requires some preparation. It will save work later on if we introduce structures even more general than graphonings. For this, let us recall a basic concept from the theory of Markov chains.

Definition 4.2. A *sub-Markov kernel* on a measurable space $(X, \mathcal{B})$ is a function

$$\deg : X \times \mathcal{B} \to [0, 1]$$
$$(x, A) \mapsto \deg_A(x)$$

such that the function $\deg_A : x \mapsto \deg_A(x)$ is measurable for all $A \in \mathcal{B}$ and the set function $\deg(x) : A \mapsto \deg_A(x)$ is a measure for all $x \in X$.

Clearly, the degree function of a graphoning is a sub-Markov kernel. The measurability of $\deg_A$ implies its seemingly stronger form below. This is probably well known but we prove it to be self-contained.

Lemma 4.3. *Consider a sub-Markov kernel* $\deg$ *on the measurable space* $(X, \mathcal{B})$. *If* $(Z, \mathcal{C})$ *is a measurable space and* $f : X \times X \times Z \to [0, 1]$ *is measurable, then*

$$\deg_f(x, z) \overset{\text{def}}{=} \int_X f(x, y, z) \mathrm{d} \deg(x)(y)$$

is a measurable function of $(x, z) \in X \times Z$, *and takes values only in* $[0, 1]$.

Proof. We have

$$0 \leq \deg_f(x, z) \leq \deg_X(x) \leq 1$$

for all x and z. The function $\deg_f$ is measurable, by the definition of a sub-Markov kernel, when f is the indicator of a direct product of measurable sets. By linearity and the Monotone Convergence Theorem, it follows that $\deg_f$ is measurable when f is any measurable stepfunction.

Any measurable function to $[0, 1]$ can be uniformly approximated by measurable stepfunctions. An error with uniform upper bound ϵ in f leads to an error with uniform upper bound ϵ in $\deg_f(x, z)$. This proves the lemma because a uniform limit of measurable functions is measurable. $\qquad \square$

Corollary 4.4. *If* deg *is a sub-Markov kernel on* $(X, \mathcal{B})$ *and* $f : X \to [0, 1]$ *is measurable, then the function* $\deg_f : X \to [0, 1]$ *defined by*

$$\deg_f(x) \stackrel{\text{def}}{=} \int_X f \, \mathrm{d} \deg(x)$$

is measurable and takes values only in $[0, 1]$.

Proof. Use Lemma 4.3 for $F(x, y, z) = f(y)$, with $Z = \{z\}$ being a single point. $\qquad \square$

Definition 4.5. A sub-Markov kernel deg on a measurable space $(X, \mathcal{B})$ is *compatible* with a $(\mathcal{B} \otimes \mathcal{B})$-measurable function $W : X^2 \to [0, 1]$ if

$$(4.3) \qquad \int_A W(x_2, y) \mathrm{d} \deg(x_1)(y) = \int_A W(x_1, y) \mathrm{d} \deg(x_2)(y)$$

for all $x_1, x_2 \in X$ and all $A \in \mathcal{B}$.

Lemma 4.6. *In a graphoning,* deg *is compatible with* W.

Proof. Both sides of (4.3) equal

$$\int_A W(x_1, y) W(x_2, y) \mathrm{d} \mu(y).$$

Indeed, this is a special case of the well-known formula

$$\int_A f \frac{\mathrm{d} \nu}{\mathrm{d} \mu} \mathrm{d} \mu = \int_A f \mathrm{d} \nu$$

involving a Radon-Nikodym derivative. $\qquad \square$

To define rooted homomorphism densities, we will have to introduce labelings on the test graphs F. The compatibility discussed above will ensure that the density is independent of the labeling chosen.

Definition 4.7. An *admissible labeling* of a connected graph F is a bijection

$$\phi : V(F) \to \{0, 1, \ldots, \mathrm{v}(F) - 1\}$$

such that for all $1 \le i \le \mathrm{v}(F)$, the nodes with labels less than i span a connected subgraph. Two admissible labelings are *adjacent* if a transposition $(i-1, i)$ of labels takes one to the other. This turns the set of admissible labelings of F into a graph.

Lemma 4.8.
    (a) *For any connected graph F, the graph of admissible labelings is connected.*
    (b) *The admissible labelings such that a fixed node o gets label 0 span a connected subgraph.*

Proof. (b) Consider two admissible labelings ϕ and ψ such that $\phi(o) = \psi(o) = 0$. We prove that they are connected by a path. We use induction on the number of inversions between them, i.e., the number of pairs $x, y \in V(F)$ such that

$$(\phi(x) - \phi(y))(\psi(x) - \psi(y)) < 0.$$

If there are no inversions, then $\phi = \psi$. If there are inversions, then there are nodes x and y with $\phi(x) = i$, $\phi(y) = i+1$, $\psi(x) > \psi(y)$. Choose such x and y so that i is largest possible. Since ψ is admissible, there is an edge in F from y to a node z with $\psi(z) < \psi(y)$ and therefore $\psi(z) < i$. Thus, composing ϕ with the transposition $(i, i+1)$ yields an admissible labeling ϕ that has less inversions when compared to ψ than ϕ does.

(a) We may assume that F has at least two nodes. It suffices to show that for any two adjacent nodes x and y in F, there exist adjacent admissible labelings ϕ and ψ such that $\phi(x) = \psi(y) = 0$. Let ϕ be an admissible labeling with $\phi(x) = 0$ and $\phi(y) = 1$. Let ψ arise from ϕ by swapping the labels of x and y. Then ψ is admissible and adjacent to ϕ. □

Definition 4.9. Let $\mathbf{G} = (X, \mathcal{B}, W, \deg)$ be a measurable space endowed with a symmetric measurable function $W : X^2 \to [0,1]$ and a sub-Markov kernel $\deg$ that is compatible with W. Let $x_0 \in X$. Let (F, o) be a connected rooted graph. Fix any admissible labeling of $V(F)$ such that o gets label 0. For any label $i = 1, \ldots, \mathrm{v}(F) - 1$, let $j(i)$ be a label such that $j(i) < i$ and j is adjacent to i in F. Note that $j(1) = 0$. Let T be the spanning tree consisting of the edges $(i, j(i))$. We define the *rooted homomorphism density*

$$t((F,o),(X,x_0))$$
$$= \int_X \cdots \int_X \prod_{kl \in E(F) - E(T)} W(x_k, x_l) \mathrm{d}\deg(x_{j(\mathrm{v}(F)-1)})(x_{\mathrm{v}(F)-1})$$
$$\cdots \mathrm{d}\deg(x_{j(1)})(x_1).$$

Proposition 4.10. *The rooted homomorphism density*
   (a) *is well defined, is in $[0,1]$, is measurable as a function of x_0, and*
   (b) *is independent of the admissible labeling and the function j chosen.*
   (c) *If F is a tree, then it is also independent of the function W.*

Proof. (a) By repeated application of Lemma 4.3, we see that each integration yields a measurable function of the remaining variables, with values in $[0,1]$.

(b) For a given admissible labeling, the rooted homomorphism density does not depend on the function j because of the condition (4.3).

Let us assume that $V(F) = \{0, 1, \ldots, \mathrm{v}(F) - 1\}$, and the identity as well as the transposition $(i - 1, i)$ are admissible labelings, where $i \geq 2$ is fixed. Then $j(i) < i - 1$ and we may apply the Fubini Theorem to swap the two factors

$$\mathrm{d}\deg(x_{j(i)})(x_i)\mathrm{d}\deg(x_{j(i-1)})(x_{i-1}),$$

showing that the two admissible labelings in consideration define the same value of the rooted homomorphism density.

An application of Lemma 4.8(b) finishes the proof.

(c) The product in Definition 4.9 is empty if $F = T$. □

Definition 4.9 may be frightening, but it becomes much nicer for graphonings. From now on, we abbreviate $\mathrm{d}\mu(x_i)$ to $\mathrm{d}x_i$.

Remark 4.11. Consider a graphoning $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$ with a specified point $x_0 \in X$. Let (F, o) be a connected rooted graph. Then we have

$$t((F,o),(\mathbf{G},x_0)) = \int_X \cdots \int_X \prod_{ij \in E(F)} W(x_i, x_j)\mathrm{d}x_{\mathrm{v}(F)-1}\cdots\mathrm{d}x_1$$

if $V(F)$ is admissibly labeled by 0, 1, ..., $\mathrm{v}(F) - 1$ so that o gets label 0. Note that the Fubini Theorem is not directly applicable to the right hand side of this formula because μ is not in general σ-finite. Note also that λ plays no role here.

4.2. **Reversible kernels and unrooted homomorphism densities.** To define unrooted homomorphism densities, we will need the measure preserving property (4.2). Again it is worthwhile to generalize this first. We recall another basic concept from Markov chain theory.

Definition 4.12. A sub-Markov kernel deg on a probability space $(X, \mathcal{B}, \lambda)$ is *reversible* w.r.t. λ if the measure preserving condition (4.2) holds.

In particular, the degree function of a graphoning is reversible.

On a measurable space $(X, \mathcal{B})$, there can be many probability measures that make a given sub-Markov kernel deg reversible. We call such measures λ *involution-invariant* w.r.t. deg because if we choose a λ-random point $a \in X$ and then a point $b \in X$ with conditional (sub-probability) distribution $\deg(a)$, then the pairs (a, b) and (b, a) have the same (sub-probability) distribution. Indeed, (4.2) precisely means the equality of these two measures on measurable product sets $A \times B \subseteq X^2$, and this implies equality on the entire σ-algebra $\mathcal{B} \otimes \mathcal{B}$. This implies the following well-known, crucial fact.

Lemma 4.13. *If* deg *is a reversible sub-Markov kernel on a probability space* $(X, \mathcal{B}, \lambda)$, *and* $f : X^2 \to [0, 1]$ *is measurable, then*

$$\int_X \int_X (f(x, y) - f(y, x)) \mathrm{d} \deg(x)(y) \mathrm{d}\lambda(x) = 0.$$

Corollary 4.14. *If* deg *is a reversible sub-Markov kernel on* $(X, \mathcal{B}, \lambda)$ *and* $f, g : X \to [0, 1]$ *are measurable functions, then*

$$(4.4) \qquad \int_X f \cdot \deg_g \mathrm{d}\lambda = \int_X g \cdot \deg_f \mathrm{d}\lambda.$$

Proof. Use Lemma 4.13 for $F(x, y) = f(x)g(y)$. $\qquad \square$

Corollary 4.15. *If* deg *is a reversible sub-Markov kernel on* $(X, \mathcal{B}, \lambda)$ *and* $f : X^2 \to [0, 1]$ *is a symmetric* $(\mathcal{B} \otimes \mathcal{B})$-*measurable function, then the sub-Markov kernel* $f \deg$ *defined by*

$$(f \deg)_A(x) = \int_A f(x, y) \mathrm{d} \deg(x)(y)$$

is again reversible.

Proof. The fact that $f \deg$ is a sub-Markov kernel follows from Lemma 4.3. For reversibility, we need to show that the value

$$\int_A (f \deg)_B \mathrm{d}\lambda = \int_A \int_B f(x, y) \mathrm{d} \deg(x)(y) \mathrm{d}\lambda(x)$$

is symmetric w.r.t. A and B. This is Lemma 4.13 for $F(x, y) = \mathbb{1}_A(x) f(x, y) \mathbb{1}_B(y)$.
$\qquad \square$

Definition 4.16. A *pseudo-graphoning* is a probability space $(X, \mathcal{B}, \lambda)$ endowed with a symmetric $(\mathcal{B} \otimes \mathcal{B})$-measurable function $W : X^2 \to [0, 1]$ and a reversible sub-Markov kernel deg that is compatible with W.

Every graphoning is also a pseudo-graphoning. A pseudo-graphoning is a graphoning if and only if there exists a measure $\mu : \mathcal{B} \to [0, \infty]$ such that the equality (4.1) holds for all $x \in X$ and $A \in \mathcal{B}$.

Proposition 4.17. *If* $\mathbf{G} = (X, \mathcal{B}, \lambda, W, \deg)$ *is a pseudo-graphoning and* $f : X^2 \to [0,1]$ *is symmetric and* $(\mathcal{B} \otimes \mathcal{B})$-*measurable, then* $f\mathbf{G} = (X, \mathcal{B}, \lambda, fW, f \deg)$ *is also a pseudo-graphoning.*

Proof. The function fW is symmetric and measurable. By Corollary 4.15, $f \deg$ is a reversible sub-Markov kernel. It remains to check that $f \deg$ is compatible with fW, which is trivial. □

Corollary 4.18. *If* $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$ *is a graphoning and* $f : X^2 \to [0,1]$ *is symmetric and* $(\mathcal{B} \otimes \mathcal{B})$-*measurable, then* $f\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, fW)$ *is also a graphoning.*

Proof. We have
$$\int_A (fW)(x, y)\mathrm{d}\mu(y) = (f \deg)_A(x),$$
so the claim follows from the previous proposition. □

This is a generalization of [26, Lemma 18.19] from László Lovász's monograph: a Borel subgraph of a graphing is a graphing.

Using reversibility, we can define unrooted homomorphism densities.

Definition 4.19. Consider a pseudo-graphoning $\mathbf{G} = (X, \mathcal{B}, \lambda, W, \deg)$. Let F be a connected graph. Choose a root o in F. Choose $x_0 \in X$ randomly with distribution λ. We define the *homomorphism density*

(4.5) $$t(F, \mathbf{G}) = \mathbb{E}t((F, o), (\mathbf{G}, x_0)) = \int_X t((F, o), (\mathbf{G}, x_0))\mathrm{d}\lambda(x_0).$$

Since the rooted homomorphism density is a measurable function of x_0 and takes values in $[0, 1]$ only, the expectation above exists and is in $[0, 1]$.

Proposition 4.20.
  (a) *The homomorphism density* $t(F, \mathbf{G})$ *is independent of the root* o.
  (b) *If* F *is a tree, then it is also independent of the function* W.

Proof. (a) Given two adjacent nodes o_0 and o_1 in F, consider an admissible labeling such that o_0 and o_1 get labels 0 and 1 respectively. For each $i \geq 2$, let $j(i) < i$ be such that the nodes with labels i and $j(i)$ are adjacent, and let T be the spanning tree given by the edges $(i, j(i))$ and (01). Consider the birooted homomorphism density

$$f(x_0, x_1) = \int_X \cdots \int_X \prod_{kl \in E(F) - E(T)} W(x_k, x_l)\mathrm{d}\deg(x_{j(\mathrm{v}(F)-1)})(x_{\mathrm{v}(F)-1})$$

$$\cdots \mathrm{d}\deg(x_{j(2)})(x_2).$$

This does not depend on the function j chosen because deg is compatible with W. We have
$$t((F, o_0), (\mathbf{G}, x_0)) = \int_X f(x_0, x_1)\mathrm{d}\deg(x_0)(x_1)$$
and
$$t((F, o_1), (\mathbf{G}, x_0)) = \int_X f(x_1, x_0)\mathrm{d}\deg(x_0)(x_1)$$
— note that the labeling that arises by swapping the labels 0 and 1 is also admissible. These two rooted densities have the same expectation by Lemma 4.13.

  (b) Immediate from Proposition 4.10(c). □

Remark 4.21. Let F be a connected graph.

For a graphoning $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$, we have

$$t(F, \mathbf{G}) = \int_X \int_X \cdots \int_X \prod_{kl \in E(F)} W(x_k, x_l) \mathrm{d}x_{\mathrm{v}(F)-1} \cdots \mathrm{d}x_1 \mathrm{d}\lambda(x_0)$$

if $V(F)$ is admissibly labeled by 0, 1, ..., $\mathrm{v}(F) - 1$. Note again that the Fubini Theorem is not directly applicable to the right hand side of this formula because μ is not in general σ-finite.

For a graphon $\mathbf{G}$ — which is a graphoning with $\mu = \lambda$ — we recover the well-known homomorphism density

$$t(F, \mathbf{G}) = \int_{X^{V(F)}} \prod_{kl \in E(F)} W(x_k, x_l) \prod_{i \in V(F)} \mathrm{d}\lambda(x_i).$$

For a graphing $\mathbf{G}$ — which is a graphoning with μ being $(1/d)$ times the counting measure — we recover a normalized version of the well-known homomorphism frequency:

$$t(F, \mathbf{G}) = t^*(F, \mathbf{G})/d^{\mathrm{v}(F)-1},$$

where

$$t^*(F, \mathbf{G}) = \int_X \hom((F, o), (\mathbf{G}, x)) \mathrm{d}\lambda(x).$$

For a graph G with all degrees $\leq d$, we can define a graphoning as follows. Let $X = V(G)$ and $\mathcal{B} = \mathcal{P}(X)$. Let λ be the uniform probability measure on X. Let $\mu = (\mathrm{v}(G)/d)\lambda$. Let $W : X^2 \to \{0, 1\}$ be the adjacency matrix of G. This graphoning has the same (rooted and unrooted) homomorphism densities as (G, d).

4.3. **Graph limits.**

Definition 4.22. A *limit* for a convergent sequence (G_n, d_n) is a pseudo-graphoning $\mathbf{G}$ such that $t(F, G_n, d_n) \to t(F, \mathbf{G})$ for all connected F. In this case, we write $(G_n, d_n) \to \mathbf{G}$. A *true limit* is a limit which is a graphoning.

In the rest of this paper, our main interest is in the existence of limits. Very little is known. We start with a very special example.

Proposition 4.23. *Let (G_n, d_n) be a convergent sequence such that G_n is the disjoint union of graphs with d_n nodes each. Then the sequence has a true limit.*

Proof. As explained in Example 1.7, there exists a Borel probability measure γ on the compact graphon space $\widetilde{\mathcal{W}}_0$, such that

$$t(F, G_n, d_n) \to \int_{\widetilde{\mathcal{W}}_0} t(F, U) \mathrm{d}\gamma(U)$$

for all connected graphs F.

Let $\mathcal{W}_0$ be the space of labeled graphons endowed with the 1-norm — not the cut norm, which is used to define the topology in $\widetilde{\mathcal{W}}_0$. I.e., $\mathcal{W}_0$ is the subset of the Banach space $\mathcal{L}^1([0, 1]^2)$ that consists of all symmetric functions with values in $[0, 1]$. By [32, Theorem 1] of Orbanz and Szegedy, there exists a measurable map $\xi : \widetilde{\mathcal{W}}_0 \to \mathcal{W}_0$ which is a section (one-sided inverse) of the canonical quotient map $\mathcal{W}_0 \to \widetilde{\mathcal{W}}_0$. Note that for each $U \in \widetilde{\mathcal{W}}_0$, the function $\xi(U) \in \mathcal{W}_0$ is defined only

almost everywhere, but for each $f \in \mathcal{W}_0$, we may use a variant of [3, Definition 2.2] to choose a canonical representative which is defined everywhere:

$$\bar{f}(x, y) = \limsup_{\epsilon \to 0} \frac{1}{4\epsilon^2} \int_{x-\epsilon}^{x+\epsilon} \int_{y-\epsilon}^{y+\epsilon} f \mathrm{d}\lambda_2,$$

where λ_2 stands for 2-dimensional Lebesgue measure, and undefined values of f are taken to be zero. It is easy to see that the function

$$\mathcal{W}_0 \times [0,1]^2 \to [0,1], \qquad (f, x, y) \mapsto \bar{f}(x, y)$$

is Borel measurable; this was observed by Viktor Kiss (unpublished). It follows that the function

$$\widetilde{\mathcal{W}}_0 \times [0,1]^2 \to [0,1], \qquad (W, x, y) \mapsto \overline{\xi(U)}(x, y)$$

is also Borel measurable.

Let $X = \widetilde{\mathcal{W}}_0 \times [0,1]$ and define $W : X^2 \to [0,1]$ by putting

$$W((U, x), (V, y)) = \mathbb{1}_{U=V} \overline{\xi(U)}(x, y).$$

The function W is clearly symmetric and Borel measurable.

For all $A \subseteq X$, let

$$A_U = \{x \in [0,1] : (U, x) \in A\} \qquad \left(U \in \widetilde{\mathcal{W}}_0\right).$$

Let $\mathcal{B} \subset \mathcal{P}(X)$ be the σ-algebra of Borel sets. For all $A \in \mathcal{B}$, define

$$\mu(A) = \sum_{U \in \widetilde{\mathcal{W}}_0} \lambda_1(A_U).$$

Let $\lambda = \gamma \times \lambda_1$, where λ_1 stands for 1-dimensional Lebesgue measure.

Let $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$. It is straightforward to check that $\mathbf{G}$ is a graphoning and

$$t(F, \mathbf{G}) = \int_{\widetilde{\mathcal{W}}_0} t(F, U) \mathrm{d}\gamma(U) = \lim_{n \to \infty} t(F, G_n, d_n)$$

for all connected graphs F. $\hspace{10cm} \square$

4.4. Hausdorff limits.
We now introduce special graphonings that involve geometric measure theory.

Definition 4.24. A *Hausdorff graphoning* is a graphoning of the form

$$\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W),$$

where X is a metric space, $\mathcal{B}$ is the σ-algebra of Borel sets, λ is 1-dimensional Hausdorff measure, and μ is a Hausdorff measure with some gauge function h. I.e., $h \geq 0$ is a right-continuous non-decreasing function on a right neighborhood of 0 and

$$\mu(B) = \lim_{\delta \to 0} \inf \left\{ \sum_{i=0}^{\infty} h(\mathrm{diam}(I_i)) : \mathrm{diam}\, I_i < \delta \text{ for all } i, \text{ and } B \subseteq \bigcup_{i=0}^{\infty} I_i \right\}$$

for any Borel set B.

A *Euclidean graphoning* is a Hausdorff graphoning where $X = [0,1]$ with the Euclidean metric.

Note that if gauge functions h_1 and h_2 satisfy $(1-\epsilon)h_1(x) \leq h_2(x) \leq (1+\epsilon)h_1(x)$ for $0 \leq x < \delta(\epsilon)$, then they define the same Hausdorff measure.

The gauge function $h(x) = x$ gives rise to the 1-dimensional Hausdorff measure. For $X = [0, 1]$, this is Lebesgue measure; the corresponding Euclidean graphonings are Borel measurable graphons. The constant gauge function $h(x) = 1/d$ gives rise to the counting measure divided by d; in this case $\{0, 1\}$-valued Hausdorff graphonings are graphings.

Definition 4.25. A *Hausdorff* (resp. *Euclidean*) *limit* for a convergent sequence (G_n, d_n) is a limit which is a Hausdorff (resp. Euclidean) graphoning with a gauge function h such that $h(1/\operatorname{v}(G_n)) \sim 1/d_n$ as $n \to \infty$.

Recall from Definition 1.1 that the homomorphism density $t(F, G, d)$ involved a normalization by an appropriate power of d in order to be in $[0, 1]$. The role of the gauge function h is to encode in the limit object not only the limiting homomorphism densities, but also the growth rate of the degree bound d_n.

For a convergent sequence (G_n) of dense graphs with $\operatorname{v}(G_n) = n$, a Euclidean limit for the convergent sequence (G_n, n) is the same thing as a limiting (Borel measurable) graphon on $[0, 1]$. For a Benjamini–Schramm convergent sequence (G_n) with degree bound d and with $\operatorname{v}(G_n) \to \infty$, a $\{0, 1\}$-valued Euclidean limit for the convergent sequence (G_n, d) is the same thing as a limiting graphing on $[0, 1]$.

Example 4.26. The sequence (G_n, d_n) of Example 1.5, provided that $\operatorname{v}(\Gamma_i) \geq 2$ for all i, always has a Hausdorff limit such that in the underlying metric space, all non-zero distances are of the form $1/\operatorname{v}(G_n)$, and W is $\{0, 1\}$-valued.

Proof. Let $X = \prod_{i=1}^{\infty} V(\Gamma_i)$. The distance of two points in $x, y \in X$ is defined to be $1/\operatorname{v}(G_n)$ if $n + 1 = \inf\{i : x_i \neq y_i\}$. The corresponding 1-dimensional Hausdorff measure λ will be the product of the uniform probability measures λ_i on $V(\Gamma_i)$. Set $h(1/\operatorname{v}(G_n)) = 1/d_n$. This is well defined since $\operatorname{v}(G_n) < \operatorname{v}(G_{n+1})$ for all n. The corresponding Hausdorff measure μ will be the product of the measures $\mu_i = (\operatorname{v}(\Gamma_i)/\delta_i)\lambda_i$. Let $\mathbf{G} = (X, \mathcal{B}, \lambda, \mu, W)$, where $W(x, y) = 1$ if x_i and y_i are adjacent in Γ_i for all i, and $W(x, y) = 0$ otherwise. This $\mathbf{G}$ is the direct product of the graphonings that correspond to the (Γ_i, δ_i) by the end of Remark 4.21. It is easy to see that $\mathbf{G}$ is a Hausdorff limit of (G_n, d_n). $\square$

The author is unable to answer the fundamental

Problem 4.27.

    (a) Which convergent sequences have (true, Hausdorff, Euclidean) limits?

    (b) Which pseudo-graphonings arise as (Hausdorff) limits?

In the dense case, the Euclidean (i.e., graphon) versions of both questions have been answered by L. Lovász and B. Szegedy [26, 28]; the answer is "all". In the bounded degree case, the graphing version of (a) was solved by D. Aldous and R. Lyons [4] and by G. Elek [16], see also [26, Theorem 18.37]; the answer is "all"; while the answer "all" for the graphing version of (b) is the Aldous–Lyons Conjecture. (In our setting, we should say "all simple graphings" because we are only allowing simple graphs.)

4.5. **Acyclicity and regularity.** In the remaining part of this paper, our main focus is on constructing limit objects for convergent sequences of large essential girth. First, we characterize when the cycle densities of a graphoning vanish.

A sub-Markov kernel generates a sub-Markov chain in the usual way:

Definition 4.28. Let deg be a sub-Markov kernel on the measurable space $(X, \mathcal{B})$. For $x \in X$, let $\deg^0(x) : \mathcal{B} \to \{0, 1\}$ be the Dirac measure at x. Let $\deg^1 = \deg$. If i and j are positive integers summing to k, then define

$$(4.6) \qquad \deg_A^k(x) = \int_X \deg_A^i \, \mathrm{d} \deg^j(x).$$

This yields a well defined sub-Markov kernel $\deg^k$ on $(X, \mathcal{B})$.

Proposition 4.29. *Let $k \geq 2$. For a graphoning $\mathbf{G}$, the following are equivalent:*
  (a) $t(C_{k+1}, \mathbf{G}) = 0$;
  (b) *the neighborhood $N(x) = \{y \in X : W(x, y) > 0\}$ has $\deg^k(x)$-measure zero for λ-a.e. x;*
  (c) $\deg^k(x) \perp \deg(x)$ *(singular measures) for λ-a.e. $x \in X$.*

Proof. (a) $\Leftrightarrow$ (b): We have

$$t(C_{k+1}, \mathbf{G}) = \int_X \int_X W(x, y) \mathrm{d} \deg^k(x)(y) \mathrm{d}\lambda(x).$$

Statement (a) holds if and only if this is zero, i.e.,

$$\int_X W(x, y) \mathrm{d} \deg^k(x)(y) = 0$$

for λ-a.e. x, which is equivalent to (b).

(b) $\Rightarrow$ (c): The measure $\deg(x)$ is concentrated on the set $N(x)$.

(c) $\Rightarrow$ (b): The formula (4.6) for $i = 1$, together with the definition (4.1) of deg in a graphoning, show that $\deg^k(x)$ is absolutely continuous with respect to μ, for all x, if $k \geq 1$. Assume that $\deg^k(x) \perp \deg(x)$ for a fixed x; we prove that the set $N(x)$ has $\deg^k(x)$-measure zero. The set $N(x)$ can be written as a union $A \cup B$, where $\deg_A^k(x) = \deg_B(x) = 0$, because X can be written as such a union, by the definition of singular measures. By the definition of the measure $\deg(x)$, we have $\mu(B) = 0$, whence $\deg_B^k(x) = 0$ and therefore $\deg_{N(x)}^k(x) = \deg_{A \cup B}^k(x) = 0$ as claimed. □

Definition 4.30. Consider a probability space endowed with a sub-Markov kernel: $\mathbf{G} = (X, \mathcal{B}, \lambda, \deg)$. The space $\mathbf{G}$ (or the kernel deg) is *acyclic* if $\deg^k(x) \perp \deg(x)$ for λ-a.e. $x \in X$ and all $0 \leq k \neq 1$.

In particular, a graphoning is acyclic if and only if all cycle densities are zero.

In the next subsection, we will be interested in limits of *regular* sequences (of large essential girth). We now introduce the corresponding limit objects.

Definition 4.31. Let $0 \leq \alpha \leq 1$. Consider a probability space endowed with a sub-Markov kernel: $\mathbf{G} = (X, \mathcal{B}, \lambda, \deg)$. The space $\mathbf{G}$ (or the kernel deg) is *α-regular* if for all $k \geq 0$, and for λ-a.e. $x \in X$, we have $\deg_X(y) = \alpha$ for $\deg^k(x)$-a.e. $y \in X$.

In particular, a Markov kernel is 1-regular.

Regular kernels can be characterized in terms of homomorphism densities of rooted trees. Note that rooted tree densities as in Definition 4.9 depend neither on the function W — cf. Proposition 4.10(c) —, nor on the probability measure λ, therefore rooted tree densities of a measurable space endowed with a sub-Markov kernel make sense.

Proposition 4.32. *Consider a probability space endowed with a sub-Markov kernel:* $\mathbf{G} = (X, \mathcal{B}, \lambda, \deg)$. *The following are equivalent:*

(a) *The space* $\mathbf{G}$ *is* α-*regular.*

(b) *For* λ-*a.e.* $x \in X$, *we have*

$$t((P_{k+2}, o), (\mathbf{G}, x)) = \alpha^{k+1}$$

and

$$t((D_{k+3}, o), (\mathbf{G}, x)) = \alpha^{k+2}$$

for all $k \geq 0$, *where* o *is a leaf (farthest from the trivalent node in the case of* D_{k+3}, $k \geq 1$), *except in* D_3, *where* o *is the non-leaf.*

(c) *For all rooted trees* (F, o), *we have* $t((F, o), (\mathbf{G}, x)) = \alpha^{\mathrm{e}(F)}$ *for* λ-*a.e.* $x \in X$.

Proof. Assuming (a), we easily get (c) by induction on $\mathrm{v}(F)$. The implication (c) $\Rightarrow$ (b) is trivial. Assuming (b), we prove (a). We have

$$\alpha^k = t((P_{k+1}, o), (\mathbf{G}, x)) = \int_X \mathrm{d} \deg^k(x)(y),$$

$$\alpha^{k+1} = t((P_{k+2}, o), (\mathbf{G}, x)) = \int_X \deg_X(y) \mathrm{d} \deg^k(x)(y),$$

and

$$\alpha^{k+2} = t((D_{k+3}, o), (\mathbf{G}, x)) = \int_X (\deg_X(y))^2 \mathrm{d} \deg^k(x)(y)$$

for all $k \geq 0$ and λ-a.e. $x \in X$; note that $P_1 \simeq K_1$. From the condition of equality in the Cauchy–Schwarz inequality, we see that for all k, there exists an α_k such that for λ-a.e. $x \in X$, we have $\deg_X(y) = \alpha_k$ for $\deg^k(x)$-a.e. $y \in X$. Then

$$\alpha_0 \cdots \alpha_k = t((P_{k+2}, o), (\mathbf{G}, x)) = \alpha^{k+1}$$

for all $k \geq 0$ and λ-a.e. $x \in X$. If $\alpha > 0$, then this implies that $\alpha_k = \alpha$ for all k, and $\mathbf{G}$ is α-regular. If $\alpha = 0$, then we get $\alpha_0 = 0$, i.e., $\deg_X(x) = 0$ for λ-a.e. $x \in X$. But then $\deg^k(x) = 0$ for λ-a.e. $x \in X$ and all $k \geq 1$, and therefore $\mathbf{G}$ is 0-regular. $\square$

For reversible kernels, the characterization of regularity becomes much nicer.

Lemma 4.33. *Let* $0 \leq \alpha \leq 1$. *Consider a probability space endowed with a reversible sub-Markov kernel:* $\mathbf{G} = (X, \mathcal{B}, \lambda, \deg)$. *The space* $\mathbf{G}$ *is* α-*regular if and only if for* λ-*a.e.* $x \in X$, *we have* $\deg_X(x) = \alpha$.

Proof. If $\mathbf{G}$ is α-regular, then for λ-a.e. $x \in X$, we have $\deg_X(y) = \alpha$ for $\deg^0(x)$-a.e. $y \in X$. But $\deg^0$ is Dirac measure at x, so we have $\deg_X(x) = \alpha$ for λ-a.e. $x \in X$, as claimed.

Conversely, assume that the set $A = \{y \in X : \deg_X(y) \neq \alpha\}$ has $\lambda(A) = 0$. What we want to prove is that $\deg_A^k(x) = 0$ for all $k \geq 0$ and λ-a.e. $x \in X$. Let

$A_k = \{x \in X : \deg_A^k(x) > 0\}$. We use induction on k to show that $\lambda(A_k) = 0$. For $k = 0$, this holds because $A_0 = A$. If it holds for $k - 1$, then it also holds for k because

$$\deg_A^k(x) = \int_X \deg_A^{k-1} \mathrm{d} \deg(x) = \int_{A_{k-1}} \deg_A^{k-1} \mathrm{d} \deg(x) = 0$$

for λ-a.e. $x \in X$. Indeed, $\deg_{A_{k-1}}(x) = 0$ for λ-a.e. $x \in X$ because

$$\int_X \deg_{A_{k-1}} \mathrm{d}\lambda = \int_{A_{k-1}} \deg_X \mathrm{d}\lambda = 0$$

by reversibility of the kernel deg and by the induction hypothesis. $\qquad\square$

Regular reversible kernels can be characterized in terms of homomorphism densities of trees. Recall from Proposition 4.20(b) that tree densities of a pseudographoning do not depend on the function W, therefore tree densities of a probability space endowed with a reversible sub-Markov kernel make sense.

Proposition 4.34. *Consider a probability space endowed with a reversible sub-Markov kernel:* $\mathbf{G} = (X, \mathcal{B}, \lambda, \deg)$. *The following are equivalent:*

(a) *The space* $\mathbf{G}$ *is* α-*regular.*
(b) *We have* $t(K_2, \mathbf{G}) = \alpha$ *and* $t(P_3, \mathbf{G}) = \alpha^2$.
(c) *For all trees* F, *we have* $t(F, \mathbf{G}) = \alpha^{\mathrm{e}(F)}$.

Proof. Assuming (a), we get (c) from Proposition 4.32(c). The implication (c) $\Rightarrow$ (b) is trivial. Assuming (b), we prove (a). The degree of a λ-random point $x \in X$ has expectation $t(K_2, \mathbf{G}) = \alpha$ and variance $t(P_3, \mathbf{G}) - t(K_2, \mathbf{G})^2 = \alpha^2 - \alpha^2 = 0$, therefore it is a.s. α. $\qquad\square$

4.6. Hausdorff limits of regular sequences of large essential girth.

Lemma 4.35. *If* $0 \leq \alpha \leq 1$, *and* $h \geq 0$ *is a continuous non-decreasing function on a right neighborhood of* 0, *such that* $h(0) = 0$ *but* $h(x)/x \to \infty$ *as* $x \to 0$, *then*

(a) *there exists an* α-*regular acyclic Hausdorff graphoning* $\mathbf{G}$ *with gauge function* h, *such that* W *is* $\{0, 1\}$-*valued.*
(b) *If, in addition, the gauge function* h *is concave, then* $\mathbf{G}$ *can be chosen to be Euclidean.*

Proof. If $\alpha = 0$, let $X = [0, 1]$ with the Euclidean metric, and let $W = 0$ everywhere. This is a 0-regular acyclic Euclidean graphoning with gauge function h.

If $\alpha > 0$, then we may, and do, assume that $\alpha = 1$, since we may replace h by h/α.

(a) For $i = 1, 2, \ldots$, choose positive integers γ_i and δ_i such that δ_i is even and $\gamma_i > \delta_i$ for all i, δ_i and γ_i/δ_i both tend to ∞ as $i \to \infty$, and $h(1/(\gamma_1 \cdots \gamma_n)) \sim 1/(\delta_1 \cdots \delta_n)$ as $n \to \infty$. Let $V(\Gamma_i) = \mathbb{Z}/\gamma_i\mathbb{Z}$, and join two nodes by an edge if their distance is $\leq \delta_i/2$ to get a δ_i-regular graph Γ_i. For all $k \geq 3$, we have

$$\limsup_{i \to \infty} t(C_k, \Gamma_i, \delta_i) \leq 3/4,$$

so the 1-regular sequence (G_n, d_n), where $G_n = \Gamma_1 \times \cdots \times \Gamma_n$ and $d_n = \delta_1 \cdots \delta_n$, has large essential girth. The claim now follows from Example 4.26.

(b) For $i = 1, 2, \ldots$, choose integers $\gamma_i \geq \delta_i > 1$ such that $\delta_i - 1 | \gamma_i - 1$ for all i, δ_i and γ_i/δ_i both tend to ∞ as $i \to \infty$, and

$$|\delta_1 \cdots \delta_n h(1/(\gamma_1 \cdots \gamma_n)) - 1| < 1/2^n$$

for all n. Let

$$S = \left\{ \sum_{n=1}^{\infty} \frac{a_n}{\gamma_1 \cdots \gamma_n} \ : \ 0 \le a_n < \gamma_n, \ a_n \equiv 0 \mod (\gamma_n - 1)/(\delta_n - 1) \right\}.$$

I.e., $S \subset [0,1]$ is the set of numbers which, in the mixed radix system with base γ_1, $\gamma_2, \ldots$, have a representation such that the n-th digit is divisible by $(\gamma_n-1)/(\delta_n-1)$ for all n. In other words, $S = \bigcap_{n=0}^{\infty} S_n$, where

$$S_n = \bigcup_{r \in \mathcal{R}_n} I_r,$$

$\mathcal{R}_n$ is the set of integer sequences $(r_1, \ldots, r_n)$ such that $0 \le r_i \le \delta_i - 1$ for all i, $I_\emptyset = [0,1]$, and I_r is a compact interval of length $1/(\gamma_1 \cdots \gamma_n)$, such that the two intervals $I_{r_1,\ldots,r_{n-1},0}$ and $I_{r_1,\ldots,r_{n-1},\delta_n-1}$ share a left, resp. right endpoint with $I_{r_1,\ldots,r_{n-1}}$, and the midpoints of the δ_n intervals $I_{r_1,\ldots,r_{n-1},0}, \ldots, I_{r_1,\ldots,r_{n-1},\delta_n-1}$ form an arithmetic progression.

Since each I_r is compact, so is S_n, and therefore so is S.

Let μ be the Hausdorff measure with gauge function h. We shall now prove that $\mu(S) = 1$. This is closely related to [31, Theorem 1]. The basic idea is found already in [18, pp. 14–15].

We have $S \subset S_n = \bigcup I_r$, where

$$\sum_r h(\operatorname{diam} I_r) = \delta_1 \cdots \delta_n \cdot h(1/(\gamma_1 \cdots \gamma_n)) \to 1$$

and

$$\max_r \operatorname{diam} I_r = 1/(\gamma_1 \cdots \gamma_n) \to 0$$

as $n \to \infty$, whence $\mu(S) \le 1$.

For the converse inequality, assume that $S \subseteq \bigcup_{J \in \mathcal{J}} J$, where $\mathcal{J}$ is countable and each $\operatorname{diam} J$ is smaller than the length of a shortest component of $[0,1] - S_n$ for a given n. We show that

$$\sum_{J \in \mathcal{J}} h(\operatorname{diam} J) > 1 - \frac{1}{2^{n-2}}.$$

We may assume (by taking convex hull and fattening a bit) that the sets J are open intervals with endpoints not in S. By compactness, we may assume that there are only finitely many of them. Now we may change our mind and assume (by cutting off superfluous bits) that each J is the convex hull of two intervals I_r, where $r \in R_N$ with a fixed N, while each J is contained in some I_r with $r \in R_n$. We may also assume that the intervals $J \in \mathcal{J}$ are pairwise disjoint.

Let $\mathcal{J}'$ be the set of non-empty intervals arising by intersecting each $J \in \mathcal{J}$ with each connected component of S_{n+1}.

It suffices to show that

$$\sum_{J' \in \mathcal{J}'} h(\operatorname{diam} J') - \sum_{J \in \mathcal{J}} h(\operatorname{diam} J) \le \sum_{r' \in \mathcal{R}_{n+1}} h(\operatorname{diam} I_{r'}) - \sum_{r \in \mathcal{R}_n} h(\operatorname{diam} I_r),$$

because the right hand side is

$$\delta_1 \cdots \delta_n \delta_{n+1} h(1/(\gamma_1 \cdots \gamma_n \gamma_{n+1})) - \delta_1 \cdots \delta_n h(1/(\gamma_1 \cdots \gamma_n)) < 3/2^{n+1}.$$

Let $r \in \mathcal{R}_n$ be fixed. It suffices to show that

$$\sum_{J' \in \mathcal{J}', J' \subset I_r} h(\operatorname{diam} J') - \sum_{J \in \mathcal{J}, J \subseteq I_r} h(\operatorname{diam} J) \leq \sum_{r_{n+1}=0}^{\delta_{n+1}-1} h(\operatorname{diam} I_{r,r_{n+1}}) - h(\operatorname{diam} I_r),$$

because summation upon r gives our previous claim. The last inequality follows from the concavity of h. Indeed, if an interval $J \in \mathcal{J}$ with $J \subseteq I_r$ contains exactly s of the $\delta_{n+1} - 1$ connected components of $I_r \setminus S_{n+1}$, then

$$\sum_{J' \in \mathcal{J}', J' \subseteq J} h(\operatorname{diam} J') - h(\operatorname{diam} J)$$

$$\leq \frac{s}{\delta_{n+1}-1} \left(\sum_{r_{n+1}=0}^{\delta_{n+1}-1} h(\operatorname{diam} I_{r,r_{n+1}}) - h(\operatorname{diam} I_r) \right),$$

and summation w.r.t. J yields our previous inequality. This proves that $\mu(S) = 1$.

For $x, y \in [0,1]$, put $W(x, y) = 1$ if $|x - y| \in S$ and $W(x, y) = 0$ otherwise. Let λ be a Lebesgue measure on $\mathcal{B} = \mathcal{B}[0,1]$. We must prove that the tuple

$$\mathbf{G} = ([0,1], \mathcal{B}, \lambda, \mu, W)$$

is a 1-regular Euclidean graphoning with gauge function h. First, the function W is semicontinuous and therefore Borel measurable. We have $\deg_{[0,1]}(x) = 1$ for all $x \in [0,1]$. When A is an interval, the function $\deg_A$ is continuous and therefore Borel measurable. When A is an open set, the function $\deg_A$ is still Borel measurable because A is a countable disjoint union of intervals, thus $\deg_A$ is Baire 1 (i.e., a pointwise limit of continuous functions). The class of Borel subsets A of $[0,1]$ such that $\deg_A$ is Borel measurable is closed under monotone sequential limits and contains all open sets, therefore contains all Borel sets; cf. [14, Section II.6].

It remains to check the measure preserving property (4.2). Observe that W is the indicator of a set that is a union of lines with slope $45°$ intersected with the unit square. Any union of such lines is symmetric w.r.t. any line of slope $-45°$. We deduce (4.2) for intervals $A, B \subseteq [0,1]$ of equal length. Since any rectangle can be exhausted by squares, (4.2) holds for any intervals A and B by the Monotone Convergence Theorem. For a fixed interval A, both sides of (4.2), as functions of the Borel set B, are finite measures that coincide on intervals, so coincide on all Borel sets. For a fixed Borel set B, the two sides coincide on all intervals A, so on all Borel sets A. This proves that $\mathbf{G}$ is indeed a 1-regular Euclidean graphoning with gauge function h.

We show that $\mathbf{G}$ is acyclic. Since S is symmetric w.r.t. $1/2$, this amounts to saying that for any $k \geq 2$, the modulo 1 sum of k independent μ-random elements of S is a.s. not in S. But μ-random means that for each n, we choose a value from $\{0, 1, \ldots, \delta_n - 1\}$ uniformly and multiply it by $(\gamma_n - 1)/(\delta_n - 1)$ to get the n-th digit in the mixed radix expansion; and we do this independently for all n. There will a.s. be infinitely many indices n such that there is carrying from the n-th digit to the previous digit when we perform the k-fold addition, but there is no carrying from the $(n+1)$-th digit to the n-th. If such an n is large enough, then in the k-fold modulo 1 sum the n-th digit a_n is not divisible by the corresponding $(\gamma_n - 1)/(\delta_n - 1)$. $\qquad\square$

Theorem 4.36. *If $0 \leq \alpha \leq 1$, and (G_n, d_n) is an α-regular sequence of large essential girth, such that $\mathrm{v}(G_n) < \mathrm{v}(G_{n+1})$ and $d_n \leq d_{n+1}$ for all n, then*

  (a) *(G_n, d_n) has a Hausdorff limit such that W is $\{0,1\}$-valued.*
  (b) *If, in addition, the function $1/\mathrm{v}(G_n) \mapsto 1/d_n$ is concave, then (G_n, d_n) has a Euclidean limit such that W is $\{0,1\}$-valued.*

Proof. Let $h(1/\mathrm{v}(G_n)) = 1/d_n$, and let h be linear on each of the intervals

$$[1/\mathrm{v}(G_{n+1}), 1/\mathrm{v}(G_n)].$$

Put $h(0) = \lim(1/d_n)$ to make h right-continuous. If $h(0) > 0$, then d_n stabilizes to a value d. Then αd must be an integer, and G_n is Benjamini–Schramm convergent to the αd-regular tree, which can be represented by a graphing on $[0,1]$. From now on, we assume that $d_n \to \infty$, i.e., $h(0) = 0$.

If $\alpha = 0$, let $X = [0,1]$ with the Euclidean metric, and let $W = 0$ everywhere. This is a Euclidean limit for (G_n, d_n).

If $\alpha > 0$, then, by Proposition 1.25, we have $\mathrm{v}(G_n)/d_n \to \infty$ as $n \to \infty$, and therefore $h(x)/x \to \infty$ as $x \to 0$. The theorem now follows from Lemma 4.35. $\square$

The rest of this subsection is not logically necessary, it is only to illustrate Lemma 4.35 and Theorem 4.36. We work out two examples: we explicitly construct Euclidean limits of the sequence of hypercubes and the sequence of projective planes. Let μ_{cube} and μ_{proj} be the Hausdorff measures on $[0,1]$ corresponding to the gauge functions

$$h_{\mathrm{cube}}(x) = 1/\log_2(1/x) \qquad \text{and} \qquad h_{\mathrm{proj}} = \sqrt{2x},$$

respectively. Note that these gauge functions have the right growth rate:

$$(4.7) \qquad h_{\mathrm{cube}}(1/\mathrm{v}(Q_d)) = 1/d \qquad \text{and} \qquad h_{\mathrm{proj}}(1/\mathrm{v}(G)) \sim 1/(q+1)$$

if G is the incidence graph of a projective plane of order $q \to \infty$. Observe also that h_{cube} is concave on $[0, 1/e]$ and h_{proj} is concave on $[0, +\infty)$. This will help us to calculate the Hausdorff measures of carefully constructed sets. The following construction relies on a rather special property of these two functions h: the numbers $1/h^{-1}(1/2)$ and $h^{-1}(1/2^n)/h^{-1}\left(1/2^{n+1}\right)$ $(n = 1, 2, \ldots)$ are integral powers of 2. Thus, we can get away with binary expansions instead of the mixed radix expansions above, and the inequalities involved in the proof of Lemma 4.35(b) become much simpler.

Let
(4.8)
$$S_{\mathrm{cube}} = \left\{ \sum_{j=1}^{\infty} a_j 2^{-j} : a_j \in \{0,1\}, a_1 = a_2, a_3 = a_4, a_5 = \cdots = a_8, a_9 = \cdots = a_{16}, \ldots \right\}$$

and
(4.9)
$$S_{\mathrm{proj}} = \left\{ \sum_{j=1}^{\infty} a_j 2^{-j} : a_j \in \{0,1\}, a_1 = a_2 = a_3, a_4 = a_5, a_6 = a_7, a_8 = a_9, \ldots \right\}.$$

Proposition 4.37.   (a) *The sets S_{cube} and S_{proj} are compact.*
  (b) *$\mu_{\mathrm{cube}}(S_{\mathrm{cube}}) = \mu_{\mathrm{proj}}(S_{\mathrm{proj}}) = 1$.*

For the 'proj' case, this is well known, [18, p. 15], it is also a special case of [31, Theorem 1]. The 'cube' case can be proved by the same technique, or a proof can be extracted from that of Lemma 4.35(b). We include a proof for the convenience of the reader.

Proof. In both cases, we have $S = \bigcap_{n=0}^{\infty} S_n$, where

$$S_n = \bigcup_{i \in \{0,1\}^n} I_i,$$

and I_i is a compact interval with $h(\operatorname{diam} I_i) = 1/2^n$ for each $i \in \{0,1\}^n$, such that $I_\emptyset = [0,1]$, and $I_{i,0}$ and $I_{i,1}$ share a left, resp. right endpoint with I_i.

(a) Since each I_i is compact, so is S_n, and therefore S.

(b) We have $S \subset S_n = \bigcup I_i$, where

$$(4.10) \qquad \sum_{i \in \{0,1\}^n} h(\operatorname{diam} I_i) = 2^n \cdot (1/2^n) = 1$$

and $\max_i \operatorname{diam} I_i = h^{-1}(1/2^n) \to 0$ as $n \to \infty$, whence $\mu(S) \leq 1$.

For the converse inequality, assume that $S \subseteq \bigcup_{J \in \mathcal{J}} J$, where $\mathcal{J}$ is countable. We show that $\sum h(\operatorname{diam} J) \geq 1$. We may assume (by taking convex hull and fattening a bit) that the sets J are open intervals with endpoints not in S. By compactness, we may assume that there are only finitely many of them. Now we may change our mind and assume (by cutting off superfluous bits) that each J is the convex hull of two intervals I_i, where $i \in \{0,1\}^n$ with a fixed n.

In view of (4.10), it suffices to show that

$$(4.11) \qquad h(\operatorname{diam} J) \geq \sum_{i \in \{0,1\}^n, I_i \subseteq J} h(\operatorname{diam} I_i) = \frac{1}{2^n} \sum_{i \in \{0,1\}^n, I_i \subseteq J} 1$$

for any such J. We use induction on n. For $n = 1$, we have $J = I_0$ or $J = I_1$, when (4.11) holds with equality, or $J = [0,1]$, when it holds with strict inequality. Let $n \geq 2$ and assume that (4.11) holds for $n-1$ in place of n, whenever J is the convex hull of two intervals I_i, $i \in \{0,1\}^{n-1}$. Let us prove the same for n.

If the rightmost I_i contained in J has an index $i \in \{0,1\}^n$ that ends on 0, then let i' be the same i with the last digit modified to 1. Let J' be the convex hull of J and $I_{i'}$. By concavity of h, we have

$$h(\operatorname{diam} J') - h(\operatorname{diam} J) \leq 1/2^n,$$

except, in the hypercube case, if $\operatorname{diam} J' > 1/e$, but then we have $\operatorname{diam} J > 1/2$ and $h(\operatorname{diam} J) > 1$ trivially.

Thus, it suffices to prove (4.11) for J' in place of J. We may perform a similar trick at the left end of J. After all, we may assume that the leftmost interval I_i in J has i ending on 0 and the rightmost one has i ending on 1. But then J is the convex hull of two intervals I_i with $i \in \{0,1\}^{n-1}$ and we are done by the induction hypothesis. $\qquad \square$

We continue to treat the hypercube and the projective plane simultaneously. We omit the subscripts cube and proj. As in the proof of Lemma 4.35(b), we use S to construct a graphoning. For $x, y \in [0,1]$, put $W(x,y) = 1$ if $|x - y| \in S$ and $W(x,y) = 0$ otherwise. Let λ be Lebesgue measure on $\mathcal{B} = \mathcal{B}[0,1]$. The tuple $\mathbf{G} = ([0,1], \mathcal{B}, \lambda, \mu, W)$ is an acyclic 1-regular Euclidean graphoning with gauge function h. Acyclicity means that for any $k \geq 2$, the modulo 1 sum of k

independent μ-random elements of S is a.s. not in S. Here μ-random means that for each block of binary digits in (4.8) or (4.9), we choose the common value 0 or 1 with equal probability, and we do this independently for all blocks. There will a.s. be a block where we choose 1 exactly twice for the common value, but for the following k blocks, we choose 0 all k^2 times. In the k-fold modulo 1 sum this block will not consist of equal digits.

Proposition 4.38.   (a) $(Q_d, d) \to \mathbf{G}_{\mathrm{cube}}$ *as a Euclidean limit as $d \to \infty$.*
   (b) *If G_n is the incidence graph of a projective plane of order q_n, and $q_n \to \infty$, then*

$$(G_n, q_n + 1) \to \mathbf{G}_{\mathrm{proj}}$$

   as a Euclidean limit as $n \to \infty$.

Proof. From Propositions 4.34 and 4.37, we have $t(F, \mathbf{G}) = 1$ for any tree F. Since $\mathbf{G}$ is acyclic, we have $t(F, \mathbf{G}) = 0$ if F contains a cycle. By Propositions 1.24 and 1.27, the convergence claims in the proposition hold. By (4.7), the limiting graphonings are Hausdorff, and therefore, Euclidean limits. $\square$

5. Sub-Markov spaces

In the previous subsection, we dealt with regular sequences of large essential girth. In this section, we shall construct limit objects for arbitrary sequences of large essential girth. Sadly, these limit objects will not be graphonings, they will be weaker structures: probability spaces with a reversible sub-Markov kernel — in other words, pseudo-graphonings with $W = 0$.

5.1. **Tree densities and kernel preserving maps.** In this subsection, we do the easy part of the preparations.

We only need to care about tree densities. Recall again that the rooted tree densities, as in Definition 4.9, depend only on the sub-Markov kernel deg, not on the function W — cf. Proposition 4.10(c) — or on the probability measure λ. They satisfy a simple recursion whose proof is trivial from Definition 4.9:

Lemma 5.1. *Let* deg *be a sub-Markov kernel on $(X, \mathcal{B})$, and let $x_0 \in X$. Let (F, o) be a rooted tree. We have*

$$t((F, o), (X, x_0)) = \prod_{i=1}^{k} \int_X t((F_i, o_i), (X, x_i)) \mathrm{d} \deg(x_0)(x_i)$$

if $F - o$ is the disjoint union of trees F_1, ..., F_k whose nodes adjacent to o in F are o_1, ..., o_k respectively.

Definition 5.2. Let $(X, \mathcal{A}, \deg)$ and $(Y, \mathcal{B}, \deg)$ be spaces with sub-Markov kernels (by abuse of notation, both kernels are denoted by deg). A measurable map $\phi : X \to Y$ is *kernel preserving* if $\phi_*(\deg(x)) = \deg(\phi(x))$ for all $x \in X$.

Proposition 5.3. *If $\phi : X \to Y$ is kernel preserving, then*

$$t((F, o), (X, x_0)) = t((F, o), (Y, \phi(x_0)))$$

for all rooted trees (F, o) and all $x_0 \in X$.

Proof. We use induction on $\mathrm{v}(F)$. Assume that the proposition is true for all rooted trees with less than $\mathrm{v}(F)$ nodes. Let $F - o$ be the disjoint union of trees $F_1, \ldots, F_k$ whose nodes adjacent to o in F are $o_1, \ldots, o_k$ respectively. We have

$$t((F, o), (Y, \phi(x_0))) = \prod_{i=1}^{k} \int_Y t((F_i, o_i), (Y, y_i)) \mathrm{d} \deg(\phi(x_0))(y_i)$$

by Lemma 5.1. Here we may replace $\deg(\phi(x_0))$ by $\phi_*(\deg(x_0))$ because ϕ is kernel preserving. But

$$\int_Y t((F_i, o_i), (Y, y_i)) \mathrm{d}(\phi_*(\deg(x_0)))(y_i) = \int_X t((F_i, o_i), (Y, \phi(x_i))) \mathrm{d} \deg(x_0)(x_i)$$

for all i by the definition of ϕ_*. By the induction hypothesis, we may replace $t((F_i, o_i), (Y, \phi(x_i)))$ by $t((F_i, o_i), (X, x_i))$. The proposition follows by using Lemma 5.1 again.   □

The unrooted tree densities of a probability space with a reversible sub-Markov kernel are defined by formula (4.5). By Proposition 4.20, they are well defined. That proposition is about pseudo-graphonings, but we can always put $W = 0$ to get a pseudo-graphoning.

Simultaneously kernel preserving and measure preserving maps also preserve reversibility and unrooted tree densities:

Proposition 5.4. *If* $\mathbf{G} = (X, \mathcal{A}, \kappa, \deg)$ *and* $\mathbf{H} = (Y, \mathcal{B}, \lambda, \deg)$ *are probability spaces with sub-Markov kernels on each,* $\phi : X \to Y$ *is measurable, kernel preserving and measure preserving, and the kernel on X is reversible, then*

(a) *the kernel on Y is reversible, and*
(b) *we have* $t(F, \mathbf{G}) = t(F, \mathbf{H})$ *for all trees F.*

Proof. (a) For all $A, B \in \mathcal{B}$, we have

$$\int_A \deg_B \mathrm{d}\lambda = \int_{\phi^{-1}(A)} (\deg_B \circ \phi) \mathrm{d}\kappa = \int_{\phi^{-1}(A)} \deg_{\phi^{-1}(B)} \mathrm{d}\kappa,$$

which is symmetric w.r.t. A and B because the kernel on X is reversible.

(b) We have

$$t(F, \mathbf{G}) = \int_X t((F, o), (\mathbf{G}, x)) \mathrm{d}\kappa(x) = \int_Y t((F, o), (\mathbf{H}, y)) \mathrm{d}\lambda(y) = t(F, \mathbf{H})$$

for all rooted trees (F, o).   □

5.2. The space of consistent measure sequences.

We now wish to construct a compact metrizable space that, for sequences of large essential girth, will play a role analogous to that of the space of bounded degree rooted graphs in the Benjamini–Schramm limit theory [26, Subsection 18.3].

Given a compact metric space K, let $\mathcal{M}(K)$ be the space of Borel measures on K whose total mass is ≤ 1 (i.e., sub-probability measures). This, endowed with the Lévy–Prokhorov metric, is again a compact metric space, where convergence is the weak convergence of measures. A continuous map $f : K \to L$ of compact metric spaces induces a continuous map $f_* : \mathcal{M}(K) \to \mathcal{M}(L)$.

Let M_0 be a point and let $M_r = \mathcal{M}(M_{r-1})$. E.g., $M_1 \simeq [0, 1]$. Let $f_0 : M_1 \to M_0$ be the unique map, and let $f_r = (f_{r-1})_* : M_{r+1} \to M_r$. A *consistent sequence* is a

sequence

$$\sigma = (\sigma_r)_{r=0}^{\infty} \in \prod_{r=0}^{\infty} M_r$$

such that $f_r(\sigma_{r+1}) = \sigma_r$ for all r. Let M be the set of consistent sequences. This is the inverse limit of the spaces M_r with respect to the maps f_r. It is closed in the above product space, and therefore compact. Let $\mathcal{B}$ be the σ-algebra of Borel sets in M.

There is a canonical sub-Markov kernel on $(M, \mathcal{B})$. Let

$$\tilde{A} = \{\sigma \in M : \sigma_r \in A\}$$

whenever $A \subseteq M_r$ is Borel. Let

$$\mathcal{A} = \{\tilde{A} : A \text{ Borel in } M_r, r = 0, 1, \dots\}.$$

This is an algebra of sets, and it generates $\mathcal{B}$ as a σ-algebra. Define $\deg_{\tilde{A}}(\sigma) = \sigma_{r+1}(A)$ whenever $A \subseteq M_r$ is a Borel set and $\sigma \in M$. Then $\deg(\sigma)$ is a finite measure on $\mathcal{A}$, therefore it extends to a unique measure on $\mathcal{B}$ by the Hahn–Kolmogorov Theorem [14, Section IV.4]. This defines $\deg : M \times \mathcal{B} \to [0, 1]$. The class of sets $A \in \mathcal{B}$ such that $\deg_A : M \to [0, 1]$ is measurable contains $\mathcal{A}$ and is closed under monotone sequential limits, therefore equals $\mathcal{B}$. Thus, $\deg$ is a sub-Markov kernel.

This sub-Markov kernel, when viewed as a map $\deg : M \to \mathcal{M}(M)$, is a homeomorphism. Indeed, its inverse is given by projecting a measure $\sigma \in \mathcal{M}(M)$ to each M_r to get a consistent sequence of measures $\sigma_{r+1} \in \mathcal{M}(M_r) = M_{r+1}$ which we complete by the unique element σ_0 of M_0. This two-sided inverse map of $\deg$ is continuous and $\mathcal{M}(M)$ is compact, so $\deg$ is a homeomorphism.

A useful consequence of this is

Lemma 5.5. *If $f : M \to [0, 1]$ is continuous, then so is $\deg_f : M \to [0, 1]$.*

Proof. Let $x_n \to x$ in M. Since $\deg$ is continuous, $\deg(x_n) \to \deg(x)$ weakly. Since f is continuous,

$$\deg_f(x_n) = \int_M f \, \mathrm{d}\deg(x_n) \to \int_M f \, \mathrm{d}\deg(x) = \deg_f(x).$$

$\square$

We now have a sub-Markov kernel on M, so rooted tree homomorphism densities of M are defined.

Lemma 5.6. *For a fixed rooted tree (F, o) of radius $\leq r$, the rooted homomorphism density*

$$t((F, o), (M, \sigma)) = t((F, o), (M_r, \sigma_r))$$

depends only on σ_r, and this dependence is continuous.

Proof. Induction on r, using Lemmas 5.1 and 5.5. $\square$

Let $\mathcal{T}^{\bullet}$ be the set of rooted trees such that the root has exactly one neighbor.

Proposition 5.7. *The map*

$$t : M \to [0, 1]^{\mathcal{T}^{\bullet}}$$
$$\sigma \mapsto (t((F, o), (M, \sigma)))_{(F, o) \in \mathcal{T}^{\bullet}}$$

is a homeomorphism between M and its image $t(M)$.

Proof. Since M is compact and t is continuous, it suffices to prove that t is injective. Let
$$t_r : M_r \to [0,1]^{\mathcal{T}^\bullet_{\leq r}}$$
$$\sigma \mapsto (t((F,o),(M,\sigma)))_{(F,o)\in\mathcal{T}^\bullet_{\leq r}},$$

where $\mathcal{T}^\bullet_{\leq r}$ is the set of rooted trees in $\mathcal{T}^\bullet$ with radius $\leq r$. It suffices to prove that t_r is injective for all r. For $r = 0$, this holds because M_0 is a point. Assume that it holds for r. Let us prove it for $r + 1$. Suppose that $t_{r+1}(\sigma_{r+1}) = t_{r+1}(\sigma'_{r+1})$ for some $\sigma_{r+1}, \sigma'_{r+1} \in M_{r+1} = \mathcal{M}(M_r)$. We need to show that $\sigma_{r+1} = \sigma'_{r+1}$. We have
$$(t_r)_* \sigma_{r+1} \in \mathcal{M}\left([0,1]^{\mathcal{T}^\bullet_{\leq r}}\right),$$

and similarly for σ'_{r+1}. Since t_r is injective, it suffices to prove that these two measures on the cube coincide, or, equivalently, their moments coincide. But a moment of $(t_r)_* \sigma_{r+1}$ is the same thing as a homomorphism density $t((F,o),(M_{r+1},\sigma_{r+1}))$, where $(F,o) \in \mathcal{T}^\bullet_{\leq r+1}$. Indeed, if we think of $F - o$ as a family of elements of $\mathcal{T}^\bullet_{\leq r}$ that are glued together at their roots (the roots become the node adjacent to o in F), and each rooted tree $(T,p) \in \mathcal{T}^\bullet_{\leq r}$ occurs $m(T,p)$ times in this family, then
$$t((F,o),(M_{r+1},\sigma_{r+1})) = \int_{M_r} \prod_{(T,p)\in\mathcal{T}^\bullet_{\leq r}} t((T,p),(M_r,\sigma_r))^{m(T,p)} \mathrm{d}\sigma_{r+1}(\sigma_r).$$

$\square$

We now show that M is the terminal object in the category of sub-Markov spaces.

Proposition 5.8. *Any space with a sub-Markov kernel admits a unique kernel preserving map to M.*

Proof. Uniqueness is immediate from Propositions 5.3 and 5.7.

To prove existence, let $(X, \mathcal{B}, \deg)$ be a space with a sub-Markov kernel. We construct a kernel preserving map $\sigma : X \to M$. Let $\sigma_0 : X \to M_0$ be the unique map. If $\sigma_r : X \to M_r$ is already defined, then put
$$\sigma_{r+1}(x) = (\sigma_r)_*(\deg(x)) \in \mathcal{M}(M_r) = M_{r+1}$$

for all $x \in X$. Let $\sigma(x) = (\sigma_r(x))_{r=0}^\infty$. This is a consistent sequence, i.e., $f_r(\sigma_{r+1}(x)) = \sigma_r(x)$ for all r. We show this by induction on r. It is true for $r = 0$ because both sides are elements of the singleton M_0. Let us assume it is true for $r - 1$. Then it is true for r because
$$f_r(\sigma_{r+1}(x)) = (f_{r-1} \circ \sigma_r)_*(\deg(x)) = (\sigma_{r-1})_*(\deg(x)) = \sigma_r(x).$$

Thus, we have $\sigma : X \to M$. We must prove that the map σ is kernel preserving, i.e., $\sigma_*(\deg(x)) = \deg(\sigma(x))$ for all x. It suffices to show that these two measures on M coincide when pushed down to M_r, for all r. Using the definition of $\deg(x)$ on the right hand side, this amounts to $(\sigma_r)_*(\deg(x)) = \sigma_{r+1}(x)$. This is true by the very definition of $\sigma_{r+1}(x)$. $\square$

Let us now examine probability measures on $(M, \mathcal{B})$ that make the canonical kernel deg reversible, i.e., involution-invariant measures. These are analogous to a basic concept in the Benjamini–Schramm graph limit theory: involution-invariant probability distributions on the space of rooted graphs with a degree bound.

Proposition 5.9. *The set of involution-invariant probability measures λ on $(M, \mathcal{B})$ is closed under affine combinations that are non-negative measures, and is closed in the weak topology.*

Proof. The measure preserving condition is linear in λ, hence remains true for affine combinations.

To prove closedness under weak limits, let λ_n satisfy the measure preserving equation for $n = 1, 2, \ldots$, and let $\lambda_n \to \lambda$ weakly. We prove the equality (4.4) for λ. Using Lemma 5.5, we get the equality for continuous f and g. For a fixed continuous g, the class of measurable $f : M \to [0, 1]$ for which the equality holds is closed under monotone pointwise limits by the Monotone Convergence Theorem, thus this class contains all measurable f. The same argument for fixed measurable f and varying g finishes the proof. $\square$

We are ready for the main result of this section.

Theorem 5.10. *Let (G_n, d_n) be an admissible sequence such that $t(F, G_n, d_n)$ converges for all trees F. Then there is a unique involution-invariant Borel probability measure λ on M such that $t(F, G_n, d_n) \to t(F, (M, \mathcal{B}, \lambda, \deg))$ for all trees F as $n \to \infty$.*

Proof. To prove uniqueness, observe that if λ and λ' both have the desired property, then $t(F, \mathbf{G}) = t(F, \mathbf{G}')$ for all trees F, whence the measures

$$t_* \lambda, t_* \lambda' \in \mathcal{M}\left([0, 1]^{\mathcal{T}^{\bullet}}\right)$$

have the same moments, so they coincide. Thus, $\lambda = \lambda'$.

To prove existence, consider the graphoning $\mathbf{G}_n = (V(G_n), \mathcal{P}(V(G_n)), \lambda_n, \mu_n, W_n)$ corresponding to (G_n, d_n) by Remark 4.21. Push λ_n forward to M using the unique degree preserving map $\mathbf{G}_n \to M$. Then push it further to $[0, 1]^{\mathcal{T}^{\bullet}}$ using t. The resulting sequence of probability measures converges weakly because all moments converge. The weak limit is a probability measure λ concentrated on $t(M)$ which, when pulled back to M using $t^{-1} : t(M) \to M$, has the desired properties. $\square$

There is a corresponding version of the Aldous–Lyons Conjecture:

Problem 5.11. Is it true that for every involution-invariant Borel probability measure λ on M there exists a convergent sequence (G_n, d_n) of large essential girth such that $t(F, G_n, d_n) \to t(F, (M, \mathcal{B}, \lambda, \deg))$ for all trees F as $n \to \infty$?

In the Benjamini–Schramm case, the affirmative answer was proved by G. Elek [17].

If (G_n, d_n) is a convergent sequence of large essential girth, then the tree densities carry all the information, so the pseudo-graphoning

$$\mathbf{G} = (M, \mathcal{B}, \lambda, W = 0, \deg),$$

where λ is given by Theorem 5.10, is a limit for the sequence. This may be unsatisfactory: we might want large essential girth to be reflected in the acyclicity of the kernel deg rather than only in the fact that $W = 0$ (because an acyclic deg would give us some hope of finding a different W that would turn $\mathbf{G}$ into a true graphoning with unchanged homomorphism densities). This is easy to achieve, as we explain below. The price to pay is that the new probability measure and reversible sub-Markov kernel will not be on the space M, and we lose uniqueness.

If $(X, \mathcal{A})$ and $(Y, \mathcal{B})$ are two measurable spaces with a sub-Markov kernel on each one, then we get a sub-Markov kernel on $(X \times Y, \mathcal{A} \otimes \mathcal{B})$ by defining the measure $\deg(x, y)$ to be the product of the measures $\deg(x)$ and $\deg(y)$. Then $\deg^k(x, y)$ is the product of the measures $\deg^k(x)$ and $\deg^k(y)$ for all k. Thus, if $\deg^k(x) \perp \deg(x)$, then $\deg^k(x, y) \perp \deg(x, y)$.

A product of reversible kernels given on two probability spaces is clearly reversible on the product space. The homomorphism density of any tree in the product will be the product of the densities in the factors.

Thus, we can multiply any probability space endowed with a reversible sub-Markov kernel by either one of the many acyclic 1-regular graphonings constructed in Subsection 4.6 to get an acyclic space with unchanged tree densities. This proves

Theorem 5.12. *Let (G_n, d_n) be an admissible sequence such that $t(F, G_n, d_n)$ converges for all trees F. Then there exists a probability space $\mathbf{G}$ endowed with an acyclic reversible sub-Markov kernel such that $t(F, G_n, d_n) \to t(F, \mathbf{G})$ for all trees F.*

6. Regularity lemma?

To conclude, we briefly speculate on one of the questions involved in Problem 4.27(a): Does every convergent sequence (G_n, d_n) tend to a pseudo-graphoning $\mathbf{G}$? If (G_n, d_n) has large essential girth, the answer is in the affirmative by Theorem 5.10: choose $W = 0$. In general, the proof of an affirmative answer might involve an appropriate version of Szemerédi's Regularity Lemma. The very weak version below is unlikely to suffice.

Proposition 6.1. *For any family $\mathcal{G}$ of admissible pairs (G, d), and for any $\epsilon > 0$, there exists an N such that for every $(G, d) \in \mathcal{G}$ there exists $(G', d') \in \mathcal{G}$ with $\mathrm{v}(G') \leq N$, $d' \leq N$, and*

$$|t(F, G, d) - t(F, G', d')| < \epsilon$$

for all F with $\mathrm{v}(F) \leq 1/\epsilon$.

In the Benjamini–Schramm setting, i.e., when $\mathcal{G}$ is the family of pairs (G, d) such that d is a fixed degree bound, this is equivalent to [26, Proposition 19.10], which answered a question of L. Lovász. The very simple proof by Noga Alon carries over easily to Proposition 6.1.

Acknowledgments

References

[1] Miklós Abért, Péter Csikvári, Péter E. Frenkel, and Gábor Kun, *Matchings in Benjamini-Schramm convergent graph sequences*, Trans. Amer. Math. Soc. **368** (2016), no. 6, 4197–4218, DOI 10.1090/tran/6464. MR3453369

[2] Miklós Abért and Tamás Hubai, *Benjamini-Schramm convergence and the distribution of chromatic roots for sparse graphs*, Combinatorica **35** (2015), no. 2, 127–151, DOI 10.1007/s00493-014-3066-7. MR3347464

[3] J. M. Aldaz and J. Pérez Lázaro, *Functions of bounded variation, the derivative of the one dimensional maximal function, and applications to inequalities*, Trans. Amer. Math. Soc. **359** (2007), no. 5, 2443–2461, DOI 10.1090/S0002-9947-06-04347-9. MR2276629

[4] David Aldous and Russell Lyons, *Processes on unimodular random networks*, Electron. J. Probab. **12** (2007), no. 54, 1454–1508, DOI 10.1214/EJP.v12-463. MR2354165

[5] B. Bollobás and B. D. McKay, *The number of matchings in random regular graphs and bipartite graphs*, J. Combin. Theory Ser. B **41** (1986), no. 1, 80–91, DOI 10.1016/0095-8956(86)90029-8. MR854605

[6] Béla Bollobás and Oliver Riordan, *Metrics for sparse graphs*, Surveys in combinatorics 2009, London Math. Soc. Lecture Note Ser., vol. 365, Cambridge Univ. Press, Cambridge, 2009, pp. 211–287. MR2588543

[7] C. Borgs, J.T. Chayes, H. Cohn, and Y. Zhao: *An L_p theory of sparse graph convergence I: limits, sparse random graph models, and power law distributions,* preprint, 2014. arXiv:1401.2906

[8] C. Borgs, J.T. Chayes, H. Cohn, Y. Zhao: *An L_p theory of sparse graph convergence II: LD convergence, quotients and right convergence* preprint, 2014. arXiv: 1408.0744

[9] 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. (2) **176** (2012), no. 1, 151–219, DOI 10.4007/annals.2012.176.1.2. MR2925382

[10] L. H. Clark, J. C. George, and T. D. Porter, *On the number of 1-factors in the n-cube*, Proceedings of the Twenty-eighth Southeastern International Conference on Combinatorics, Graph Theory and Computing (Boca Raton, FL, 1997), Congr. Numer. **127** (1997), 67–69. MR1604993

[11] Péter Csikvári and Péter E. Frenkel, *Benjamini-Schramm continuity of root moments of graph polynomials. part B*, European J. Combin. **52** (2016), no. part B, 302–320, DOI 10.1016/j.ejc.2015.07.009. MR3425982

[12] Péter Csikvári, Péter E. Frenkel, Jan Hladký, and Tamás Hubai, *Chromatic roots and limits of dense graphs*, Discrete Math. **340** (2017), no. 5, 1129–1135, DOI 10.1016/j.disc.2016.11.009. MR3612454

[13] Ewan Davies, Matthew Jenssen, Will Perkins, and Barnaby Roberts, *Independent Sets, Matchings, and Occupancy Fractions,* arXiv: 1508.04675

[14] J. L. Doob, *Measure theory*, Graduate Texts in Mathematics, vol. 143, Springer-Verlag, New York, 1994. MR1253752

[15] Ioana Dumitriu and Soumik Pal, *Sparse regular random graphs: spectral density and eigenvectors*, Ann. Probab. **40** (2012), no. 5, 2197–2235, DOI 10.1214/11-AOP673. MR3025715

[16] Gábor Elek, *On limits of finite graphs*, Combinatorica **27** (2007), no. 4, 503–507, DOI 10.1007/s00493-007-2214-8. MR2359831

[17] Gábor Elek, *On the limit of large girth graph sequences*, Combinatorica **30** (2010), no. 5, 553–563, DOI 10.1007/s00493-010-2559-2. MR2776719

[18] K. J. Falconer, *The geometry of fractal sets*, Cambridge Tracts in Mathematics, vol. 85, Cambridge University Press, Cambridge, 1986. MR867284

[19] G. Freud, *Orthogonale Polynome,* Akadémiai Kiadó, Budapest, 1969.

[20] Wolfgang Gawronski, *On the asymptotic distribution of the zeros of Hermite, Laguerre, and Jonquière polynomials*, J. Approx. Theory **50** (1987), no. 3, 214–231, DOI 10.1016/0021-9045(87)90020-7. MR892220

[21] C. D. Godsil, *Algebraic combinatorics*, Chapman and Hall Mathematics Series, Chapman & Hall, New York, 1993. MR1220704

[22] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London-Toronto, Ont., 1980. Corrected and enlarged edition edited by Alan Jeffrey; Incorporating the fourth edition edited by Yu. V. Geronimus [Yu. V. Geronimus] and M. Yu. Tseytlin [M. Yu. Tseĭtlin]; Translated from the Russian. MR582453

[23] Ole J. Heilmann and Elliott H. Lieb, *Theory of monomer-dimer systems*, Comm. Math. Phys. **25** (1972), 190–232. MR0297280

[24] Miklós Kornyik and György Michaletzky, *Wigner matrices, the moments of roots of Hermite polynomials and the semicircle law*, J. Approx. Theory **211** (2016), 29–41, DOI 10.1016/j.jat.2016.07.006. MR3547630

[25] László Lovász, *Combinatorial problems and exercises*, 2nd ed., North-Holland Publishing Co., Amsterdam, 1993. MR1265492

[26] László Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012. MR3012035

[27] L. Lovász and M. D. Plummer, *Matching theory*, North-Holland Mathematics Studies, vol. 121, North-Holland Publishing Co., Amsterdam; North-Holland Publishing Co., Amsterdam, 1986. Annals of Discrete Mathematics, 29. MR859549

[28] László Lovász and Balázs Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957, DOI 10.1016/j.jctb.2006.05.002. MR2274085

[29] Brendan D. McKay, *The expected eigenvalue distribution of a large regular graph*, Linear Algebra Appl. **40** (1981), 203–216, DOI 10.1016/0024-3795(81)90150-6. MR629617

[30] James Propp, *Enumeration of matchings: problems and progress*, New perspectives in algebraic combinatorics (Berkeley, CA, 1996), Math. Sci. Res. Inst. Publ., vol. 38, Cambridge Univ. Press, Cambridge, 1999, pp. 255–291. MR1731819

[31] Cheng Qin Qu, Hui Rao, and Wei Yi Su, *Hausdorff measure of homogeneous Cantor set*, Acta Math. Sin. (Engl. Ser.) **17** (2001), no. 1, 15–20, DOI 10.1007/s101140000089. MR1831742

[32] Peter Orbanz and Balazs Szegedy, *Borel liftings of graph limits*, Electron. Commun. Probab. **21** (2016), Paper No. 65, 4, DOI 10.1214/16-ECP14. MR3548777

[33] Gabor Szegö, *Orthogonal Polynomials*, American Mathematical Society, New York, 1939. American Mathematical Society Colloquium Publications, v. 23. MR0000077

[34] Linh V. Tran, Van H. Vu, and Ke Wang, *Sparse random graphs: eigenvalues and eigenvectors*, Random Structures Algorithms **42** (2013), no. 1, 110–134, DOI 10.1002/rsa.20406. MR2999215

Faculty of Science, Institute of Mathematics, ELTE Eötvös Loránd University, 1117 Budapest, Hungary, Pázmány Péter sétány 1/C – and – Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, 1053 Budapest, Hungary Reáltanoda u. 13-15. ORCID iD: 0000-0003-2672-8772

Email address: frenkelp@cs.elte.hu