AN EXAMPLE OF GRAPH LIMITS OF GROWING SEQUENCES OF RANDOM GRAPHS

SVANTE JANSON AND SIMONE SEVERINI

ABSTRACT. We consider a class of growing random graphs obtained by creating vertices sequentially one by one: at each step, we choose uniformly the neighbours of the newly created vertex; its degree is a random variable with a fixed but arbitrary distribution, depending on the number of existing vertices. Examples from this class turn out to be the Erdős–Rényi random graph, a natural random threshold graph, *etc.* By working with the notion of graph limits, we define a kernel which, under certain conditions, is the limit of the growing random graph. Moreover, for a subclass of models, the growing graph on any given n vertices has the same distribution as the random graph with n vertices that the kernel defines. The motivation stems from a model of graph growth whose attachment mechanism does not require information about properties of the graph at each iteration.

1. INTRODUCTION

Many models of randomly grown graphs have been studied during the recent years in the attempt of reproducing characteristic properties of natural and engineered networks. For example, it is well-known that the power law (Zipf's law) on the degree distribution observed for many real-world networks can occur as a result of preferential attachment following some local rule (see, e.g. Mitzenmacher [19] and Durrett [11]).

We may initially distinguish between two types of growth depending on whether the random steps require or do not require local knowledge of the graph. Of course, preferential attachment requires local knowledge either available for free or provided by some dynamics that generates it, for example, a random walk. Such a distinction is meaningful because it helps to isolate the type of information needed for the construction of specific network ensembles. Once we have assumed no knowledge, we may further distinguish between rewiring schemes acting on the whole set of vertices and mechanisms concerned only with

the lastly added vertex. This latter scenario is considered in the present note.

We grow graphs by attaching vertices one by one: at each step, the neighbours of the new vertex are chosen uniformly; and the number is a random variable with a fixed but arbitrary distribution depending on the number of vertices already present. This mechanism reflects the idea that the graph is constructed by an agent without any kind of knowledge of the graph, apart from the labels of the vertices. The role of the agent is to attach vertices according to the chosen distribution. Related results for some other models of growing random graphs can be found in [8].

We study examples of growing sequences of these random graphs within the framework of graph limits. (See Lovász [16]; for additional references and basic definitions see Section 4 below.) Every convergent sequence of growing graphs, where "convergent" means Cauchy in a specific metric, has a limit which can be represented in the form of a symmetric measurable function in two variables also called a *graphon*. The notion of graph limits has been central to a general theory of parameter testing as developed by Borgs *et al.* [7]. The wider perspective of graph limits is to propose an approximation theory of graphs. This would help to study large graphs/networks by looking at the the proportion of copies of any fixed graph as a subgraph.

Section 2 defines our construction and lists some of its natural examples. Section 3 recasts a special case of the construction in terms of a certain infinite random graph. Section 4 gives the necessary definitions concerned with graph limits and kernels. Section 5 contains the main result. Section 6 states further remarks and formulates several open problems.

2. PRELIMINARIES

Consider a growing sequence of random graphs $(G_n)_{n=1}^{\infty}$ defined by the following Markov process:

Construction 2.1. For each $n \geqslant 1$, let ν_n be a given probability distribution on $\{0, \ldots, n-1\}$. Construct the random graphs $G_1, G_2, \ldots$ as follows.

  (i) $G_1 = K_1$, the graph with a single vertex.
  (ii) For $n \geqslant 2$, let D_n be a random variable with distribution ν_n and construct G_n by adding a new vertex to G_{n-1} and connecting it to D_n of the previously existing vertices; these vertices are chosen randomly and uniformly among all $\binom{n-1}{D_n}$ possibilities. (D_n and the choice of vertices are independent of G_{n-1}.)

We may label the vertices $1, 2, 3, \ldots$ in the order they are added, so G_n has vertex set $[n] := \{1, \ldots, n\}$. Since edges are added only incident to the new vertex, and edges are never removed, we can define the infinite random graph $G_\infty := \bigcup_{n=1}^\infty G_n$ with vertex set $[\infty] := \{1, 2, \ldots\}$; then $G_n = G_\infty|_{[n]}$, the restriction of G_∞ to the vertex set $[n]$.

We may regard G_n as a directed graph by directing each edge towards the endpoint with largest label. Then D_k is the indegree of vertex k in G_n, for any $n \geqslant k$. The outdegree of k is 0 in G_k, and increases (weakly) as n grows.

Example 2.2. Fix $p \in [0, 1]$ and let $\nu_n = \mathrm{Bi}(n - 1, p)$, $n \geqslant 1$. Then Construction 2.1 yields the same result as connecting the new vertex n to each previous vertex i with probability p, with these events independent for $i = 1, \ldots, n - 1$. Hence, $G_n = G(n, p)$, the Erdős–Rényi random graph where all edges appear independently and with probability p each. This random graph has been extensively studied, see e.g. [2] and [14].

Example 2.3. Fix $p \in [0, 1]$ and let ν_n be concentrated on $\{0, n - 1\}$ with $\nu_n\{n - 1\} = \mathbb{P}(D_n = n - 1) = p$ and $\nu_n\{0\} = \mathbb{P}(D_n = 0) = 1 - p$. Thus each new vertex is with probability p joined to all previous vertices, and with probability p to none. This is an example of a random threshold graph, see [9, Section 6.3], where this G_n is denoted $T_{n,p}$.

Note that each pair of vertices in G_n is joined by an edge with probability p, just as in Example 2.2. However, in the present example these events are not always independent for different pairs.

Example 2.4. Let ν_n be the uniform distribution on $\{0, \ldots, n - 1\}$. In this case, the degree of vertex n in G_n is then chosen uniformly at random among all possibilities. Thus, if we only consider the number of added edges, this example uses the "highest possible amount of randomness" for the construction of the n-th iteration graph, in the sense that the *entropy* of this number is maximal. Hence, of all graph ensembles obtained with Construction 2.1, G_n is in some sense the less predictable one. Note also that the neighbours of n are also chosen at random once the degree has been determined, again maximising the entropy of this step. Nevertheless, as is well-known, the total entropy of the growing random graph is not maximised by this procedure but by Example 2.2 with $p = 1/2$.

The purpose of the present note is to find the limit of the sequence G_n in the sense of graph limits, see Section 4.

All graphs are undirected and finite except when we explicitly say otherwise. All unspecified limits are as $n \to \infty$.

3. A RELATED CONSTRUCTION

A class of examples, including the three examples above, can be obtained as follows.

Construction 3.1. Let ν be a given probability measure on $[0, 1]$. Let $\theta_1, \theta_2, \ldots,$ be an i.i.d. sequence of random variables with distribution ν. Then, conditionally given this sequence, let G_∞ be the infinite random graph on $[\infty]$ where the edge $\{i, j\}$ appears with probability $\theta_{\max\{i,j\}}$, and all edges appear independently (conditionally on $(\theta_j)_{j=1}^\infty$). Further, let $G_n := G_\infty|_{[n]}$.

If $D_n := |\{i < n : in \in E(G_n)\}|$, i.e. the indegree of n if we orient the edges as above, then D_n conditioned on $(\theta_j)_j$ has the distribution $\mathrm{Bi}(n - 1, \theta_n)$. Hence, the distribution of D_n is a mixture of binomial distributions:

$$\mathbb{P}(D_n = k) = \mathbb{E}\,\mathrm{Bi}(n - 1, \theta_n)\{k\} = \mathbb{E}\binom{n-1}{k}\theta_n^k(1 - \theta_n)^{n-1-k}$$

$$= \binom{n-1}{k}\int_0^1 \theta^k(1 - \theta)^{n-1-k}\,\mathrm{d}\nu(\theta), \qquad 0 \leqslant k \leqslant n - 1.$$

$$(3.1)$$

It is obvious that Construction 3.1 is a special case of Construction 2.1, with $\nu_n := \mathcal{L}(D_n)$ given by (3.1).

Example 3.2. Let $\nu = \delta_p$, a point mass at $p \in [0, 1]$. Then $\theta_n = p$ and $D_n \sim \mathrm{Bi}(n - 1, p)$, in other words, $\nu_n = \mathrm{Bi}(n - 1, p)$; hence Construction 3.1 with this ν yields the Erdős–Rényi random graph $G(n, p)$ in Example 2.2.

Example 3.3. Let $\nu = p\delta_1 + (1 - p)\delta_0$. (This is the Bernoulli distribution $\mathrm{Be}(p)$.) Then $\theta \in \{0, 1\}$, which implies $D_n = (n - 1)\theta_n$, and

$$\mathbb{P}(D_n = n - 1) = \mathbb{P}(\theta_n = 1) = p,$$
$$\mathbb{P}(D_n = 0) = \mathbb{P}(\theta_n = 0) = 1 - p.$$

Hence, Construction 3.1 yields the random threshold graph in Example 2.3.

Example 3.4. Let ν be the uniform distribution on $[0, 1]$; thus $\nu = \lambda$, the Lebesgue measure on $[0, 1]$. Then each $\theta_n \sim \mathrm{U}(0, 1)$ and (3.1) yields

by the evaluation of a beta integral, as is well-known,

$$
\nu_n\{k\} = \mathbb{P}(D_n = k) = \binom{n-1}{k} \int_0^1 \theta^k (1-\theta)^{n-1-k} \, \mathrm{d}\theta
$$

$$
= \binom{n-1}{k} B(k+1, n-k) = \frac{1}{n}, \qquad 0 \leqslant k \leqslant n-1. \quad (3.2)
$$

Consequently, ν_n is uniform on $\{0, \ldots, n-1\}$, so Construction 3.1 with $\nu = \lambda$ yields the random graphs G_n in Example 2.4.

Example 3.5. The random graph G_n in Examples 2.4 and 3.4 can also be constructed as follows, using some basic results on Pólya–Eggenberger urns.

Recall that a Pólya–Eggenberger urn contains red and black balls; we repeatedly draw a ball at random from the urn, and then replace the ball together with another ball of the same colour. If we start the urn with one ball of each colour, then the sequence of drawn balls has the same distribution as the sequence obtained by first taking a random $\theta \sim \mathrm{U}(0,1)$ and then, conditioned on θ, taking a sequence of i.i.d. balls, each being red with probability θ and black otherwise. This is easily verified by a direct calculation, see [12], [18, Theorem 3.1] and (3.2). Alternatively, it is easily seen (again by direct calculation) that the sequence of drawn balls is exchangeable. By de Finetti's theorem (see e.g. [18, Theorem 1.2] or, in a more general version, [15, Theorem 11.10]), there exists a random variable θ with values in $[0,1]$ such that conditioned on θ, the sequence of drawn balls is i.i.d. with each ball being red with probability θ. The law of large numbers yields $R_n/n \xrightarrow{\text{a.s.}} \theta$, where R_n is the number of red balls drawn in the first n draws. To see the representation above, with $\theta \sim \mathrm{U}(0,1)$, it thus suffices to show that $R_n/n \xrightarrow{\text{d}} \mathrm{U}(0,1)$, see [12], [18, Exercise 3.4].

The sequence of the first $n-1$ drawn balls in this urn thus has the same distribution as the sequences of indicators of edges $\{i, n\}$, $i = 1, \ldots, n-1$ in the random graph G_n in Example 3.4. (We translate *red* = 1 and *black* = 0.) The graph G_n is therefore described by a sequence of (finite) draws from Pólya–Eggenberger urns, independent of each other. This can be formulated as the following, rather curious, construction:

Start with vertices $\{-1, 0, 1, 2, 3, \ldots\}$. Connect 0 to all other vertices (except -1), but do not connect -1 to any vertex. For each $k \geqslant 1$, consider $i = 1, \ldots, k-1$ in order; for each $i < k$ pick a random j in $-1, 0, \ldots, i-1$ (uniformly and independent of everything else), and add an edge $\{i, k\}$ if and only if there already is an edge $\{j, k\}$. The

sequence of edge indicators $\{i, k\}$, $i = 1, \ldots, k-1$, then forms a Pólya–Eggenberger sequence as above, for each k. Consequently, if we discard vertices 0 and -1 at the end, the random graph constructed in this way equals G_∞ in Example 3.4, and we obtain G_n if we do the same construction for $k = 1, \ldots, n$.

We note the following consequence of the law of large numbers.

Lemma 3.6. *Let ν be a probability measure on $[0, 1]$, and let D_n have the mixed binomial distribution in (3.1). Then $D_n/n \xrightarrow{\mathrm{d}} \nu$ as $n \to \infty$.*

Proof. Since D_n conditioned on θ_n has the distribution $\mathrm{Bi}(n-1, \theta_n)$, we have the law of large numbers $D_n/(n-1) - \theta_n \xrightarrow{\mathrm{p}} 0$ as $n \to \infty$. (For example, by computing the variance.) The result follows since $\theta_n \sim \nu$. $\qquad\square$

4. Graph limits and kernels

We assume that the reader is familiar with the theory of graph limits developed in Lovász and Szegedy [17] and Borgs, Chayes, Lovász, Sós and Vesztergombi [5, 6], see also e.g. Austin [1], Bollobás and Riordan [3], Borgs, Chayes and Lovász [4], Lovász [16], Diaconis and Janson [10], Janson [13]. We recall only a few definitions; these will help to fix our notation.

If F and G are finite graphs, let $t(F, G)$ be the probability that a random mapping $\varphi : V(F) \to V(G)$ is a *graph homomorphism*, i.e., satisfies $\varphi(i) \sim \varphi(j)$ in G whenever $i \sim j$ in F. We say that a sequence (G_n) of graphs with $|G_n| \to \infty$ *converges* if $\lim_{n \to \infty} t(F, G_n)$ exists for every graph F.

Graph limits. The *graph limits* are objects in a suitable space defined such that each convergent sequence of graphs has a graph limit as its limit. If Γ is a graph limit, then $t(F, \Gamma)$ is defined for every graph F, and a sequence of graphs G_n with $|G_n| \to \infty$ converges to Γ if and only if $t(F, G_n) \to t(F, \Gamma)$ for every F. Hence a graph limit Γ is determined by the numbers $t(F, \Gamma) \in [0, 1]$ for graphs F. Formally, the graph limits may be defined as equivalence classes of convergent sequences of graphs, or as suitable families $(t_F)_{F \in \mathcal{U}}$ of numbers, where $\mathcal{U}$ is the set of graphs. The graph limits can be equivalently defined as classes of kernels, as we do below. This distinction is immaterial. We tacitly refer to unlabelled graphs.

It is important that the set of all graphs together with all graph limits is a compact metric space.

Kernels. Let $(\mathcal{S}, \mathcal{F}, \mu)$ be a probability space. (We usually denote this space simply by $\mathcal{S}$ or $(\mathcal{S}, \mu)$, with $\mathcal{F}$ and perhaps μ being clear from the context.) A *kernel* or *graphon* on $(\mathcal{S}, \mu)$ is a measurable symmetric function $W : \mathcal{S}^2 \to [0, 1]$. We will consider graphons with domain $[0, 1]^2$. For this specific setting see e.g. Borgs, Chayes, Lovász, Sós and Vesztergombi [5].

The basic fact is that every kernel W on a probability space $(\mathcal{S}, \mu)$ defines a graph limit Γ_W. Conversely, every graph limit equals Γ_W for some kernel W. We say that *the graph limit is represented by the kernel W*. Note that Γ_W implicitly depends on $\mathcal{S}$ and μ as well as on W. However, such representations of graph limits are not unique. We say that two kernels W_1 and W_2, possibly on different probability spaces, are *equivalent* if they represent the same graph limit, i.e., if $\Gamma_{W_1} = \Gamma_{W_2}$. Since every kernel is equivalent to some kernel on $[0, 1]$, every graph limit may be represented by a kernel W on $[0, 1]$, equipped with Lebesgue measure λ, but even then W is not unique. Detailed results are in Borgs, Chayes and Lovász [4], Bollobás and Riordan [3] and Janson [13].

If G_n is a sequence of graphs with $G_n \to \Gamma_W$, for some kernel W, we also write $G_n \to W$.

Random graphs. Let W be a kernel, defined on a probability space $(\mathcal{S}, \mu)$. We define a random graph $G(n, W)$ with vertex set $[n]$, for $1 \leqslant n \leqslant \infty$, by first taking an i.i.d. sequence $\{X_i\}_{i=1}^n$ of random points in $\mathcal{S}$ with the distribution μ, and then, given this sequence, letting $\{i, j\}$ be an edge in $G(n, W)$ with probability $W(X_i, X_j)$. For a given sequence $(X_i)_i$, this is done independently for all pairs $(i, j) \in [n]^2$ with $i < j$. Note that we may construct $G(n, W)$ for all n by first constructing $G(\infty, W)$ and then taking the subgraph induced by the first n vertices. A fundamental result is that for every kernel W, $G(n, W) \to W$ a.s.

Furthermore, two kernels W_1 and W_2 are equivalent, i.e. $\Gamma_{W_1} = \Gamma_{W_2}$, if and only if $G(n, W_1) \stackrel{\mathrm{d}}{=} G(n, W_2)$ for every finite n, and then also for $n = \infty$.

5. Main results

Given a probability measure ν on $[0, 1]$, let $\mu = \mu_\nu := \nu \times \lambda$ be a measure on the product space $\mathcal{S} := [0, 1]^2$. Define the kernel $W : \mathcal{S}^2 \to [0, 1]$ by

$$W\big((s_1, t_1), (s_2, t_2)\big) := \begin{cases} s_2, & \text{if } t_1 < t_2; \\ s_1, & \text{if } t_1 > t_2. \end{cases} \tag{5.1}$$

We may define $W\big((s_1, t_1), (s_2, t_2)\big) := 0$ if $t_1 = t_2$; this is not important since it really is sufficient to have W defined μ-almost everywhere.

Theorem 5.1. *Let ν be a probability measure on $[0, 1]$, and let $1 \leqslant n < \infty$. The random graph G_n defined by Construction 3.1 and the random graph $G(n, W)$ defined by the kernel W in (5.1) on the probability space $(\mathcal{S}, \mu_\nu)$ are, regarded as unlabelled graphs, equal in the sense that they have the same distribution.*

Remark 5.2. We have to regard the graphs as unlabelled here, since the vertices in G_n are (in general) not equivalent, while they are in $G(n, W)$. For example, in Example 3.3, the edges in G_n incident to vertex 1 appear independently of each other, so the degree of 1 has distribution $\mathrm{Bi}(n - 1, p)$, while the degree of n is D_n, which is 0 or $n - 1$.

If we prefer to consider labelled graphs, the correct conclusion is that G_n with a (uniform) random relabelling of the vertices has the same distribution as $G(n, W)$, for any finite n.

Remark 5.3. Similarly, the conclusion of Theorem 5.1 fails for $n = \infty$. Consider again Example 3.3. It is easily verified that in G_∞ there is a pair of vertices i and j with the same closed neighbourhoods $\bar{N}(i)$ and $\bar{N}(j)$ (for example, vertices 1 and 2; in fact, there are a.s. infinitely many such pairs), while there is a.s. no such pair in $G(\infty, W)$.

Proof of Theorem 5.1. Let $X_i = (\xi_i, \eta_i)$, $i = 1, 2, 3, \ldots$, be i.i.d. points in $\mathcal{S} = [0, 1]^2$ with distribution μ_ν; thus each ξ_i has distribution ν and $\eta_i \sim \mathrm{U}(0, 1)$, and all ξ_i, η_i are independent.

The numbers $\eta_1, \ldots, \eta_n$ are a.s. distinct. Order them in increasing order as $\eta_{i_1} < \eta_{i_2} < \cdots < \eta_{i_n}$, and let $\theta_k := \xi_{i_k}$. Then $\theta_1, \ldots, \theta_n$ are i.i.d. with distribution ν, and $(\theta_i)_{i=1}^n$ is independent of the random permutation $(i_1, \ldots, i_n)$.

Conditioned on $(X_1, \ldots, X_n)$, the edges in $G(n, W)$ appear independently, and the probability of an edge between i_j and i_k, with $j < k$, is $W(X_{i_j}, X_{i_k}) = \xi_{i_k} = \theta_k$. Thus, given $(i_1, \ldots, i_n)$, $G(n, W)$ has the same distribution as G_n in Construction 3.1 after the relabelling $k \mapsto i_k$. Hence, $G(n, W)$ has the same distribution as G_n with a uniform random relabelling. Consequently, $G(n, W) \overset{\mathrm{d}}{=} G_n$ as unlabelled graphs. $\qquad \square$

Theorem 5.4. *If G_n is defined by Construction 3.1 for some probability measure ν on $[0, 1]$, then $G_n \overset{\mathrm{P}}{\longrightarrow} \Gamma_\nu$ as $n \to \infty$, where Γ_ν is the graph limit defined by the kernel W in (5.1) on the probability space $([0, 1]^2, \mu_\nu)$.*

Proof. An immediate consequence of Theorem 5.1 and $G(n, W) \xrightarrow{\mathrm{p}}$ $\Gamma_W = \Gamma_\nu$. $\qquad\square$

We have a similar result for the more general construction Construction 2.1, provided the distributions ν_n converge to ν after rescaling by n (or $n - 1$).

Theorem 5.5. *Let G_n be defined by Construction 2.1 for some probability measures ν_n, and suppose that $D_n/n \xrightarrow{\mathrm{d}} \nu$ as $n \to \infty$ for some probability measure ν on $[0, 1]$, where $D_n \sim \nu_n$. Then $G_n \xrightarrow{\mathrm{p}} \Gamma_\nu$ as $n \to \infty$, where Γ_ν is the graph limit defined by the kernel W in (5.1) on the probability space $([0, 1]^2, \mu_\nu)$.*

Proof. If F and G are labelled graphs, let $n(F, G)$ be the number of graph homomorphisms $\varphi : F \to G$; thus $t(F, G) = n(F, G)/|G|^{|F|}$. Further, let $n_<(F, G)$ be the number of graph homomorphisms $\varphi : F \to G$ that are increasing, i.e., $\varphi(i) < \varphi(j)$ when $i < j$, and let $n_0(F, G)$ be the number of graph homomorphisms $F \to G$ that are not injective.

Let F be a fixed graph with vertices labelled $1, \dots, m = |F|$. If σ is a permutation of $[m]$, let F_σ be F relabelled by $i \mapsto \sigma(i)$. For any labelled graph G,

$$n(F, G) = \sum_\sigma n_<(F_\sigma, G) + n_0(F, G), \tag{5.2}$$

since an injective map $V(F) \to V(G)$ is increasing as a map $F_\sigma \to G$ for exactly one permutation σ.

Fix a permutation σ and consider $n_<(F_\sigma, G_n)$, with G_n as in Construction 2.1. We regard F_σ as a directed graph by directing each edge towards the endpoint with the largest label. Let $d_j^- := |\{i < j : \{i, j\} \in E(F_\sigma)\}|$ be the indegree in F_σ of $j \in [m]$.

Let $\varphi : [m] \to [n]$ be an increasing map. Then φ is a graph homomorphism $F_\sigma \to G_n$ if and only if, for each $j = 1, \dots, m$, G_n contains the d_j^- edges $\{\varphi(i), \varphi(j)\}$ for $i < j$ with $\{i, j\} \in E(F_\sigma)$. Conditioned on the indegrees $D_1, \dots, D_n$ in G_n, this happens with probability

$$\prod_{j=1}^m \frac{\binom{D_{\varphi(j)}}{d_j^-}}{\binom{\varphi(j)-1}{d_j^-}} = \prod_{j=1}^m \frac{\left(D_{\varphi(j)}\right)_{d_j^-}}{(\varphi(j) - 1)_{d_j^-}}. \tag{5.3}$$

Hence, taking the expectation, summing over all φ, and using the independence of $D_1, \dots, D_n$,

$$\mathbb{E}\, n_<(F_\sigma, G_n) = \sum_{1 \leqslant \varphi(1) < \cdots < \varphi(m) \leqslant n} \prod_{j=1}^m \frac{\mathbb{E}\left(D_{\varphi(j)}\right)_{d_j^-}}{(\varphi(j) - 1)_{d_j^-}}. \tag{5.4}$$

By assumption, $D_k/k \xrightarrow{\text{d}} \nu$ as $k \to \infty$. By dominated convergence, since $0 \leqslant D_k/k \leqslant 1$, we have

$$\frac{\mathbb{E} \, D_k^d}{k^d} \to M_d := \int_0^1 x^d \, \mathrm{d}\nu(x), \qquad k \to \infty, \tag{5.5}$$

for every $d \geqslant 0$. Hence also

$$\frac{\mathbb{E} \, (D_k)_d}{(k-1)_d} = \frac{\mathbb{E} \, D_k^d + O(k^{d-1})}{k^d + O(k^{d-1})} \to M_d, \qquad k \to \infty. \tag{5.6}$$

Let $\varepsilon > 0$, it follows from (5.6) that there exists n_ε such that if $\varphi(1) \geqslant n_\varepsilon$, then the product in (5.4) differs by at most ε from $\prod_{j=1}^m M_{d_j^-}$. For smaller $\varphi(1)$ we use the fact that the product is bounded by 1. The total number of terms in the sum in (5.4) is $\binom{n}{m}$, of which $O(n^{m-1})$ have $\varphi(1) < n_\varepsilon$, and thus we obtain

$$\left| \mathbb{E} \, n_<(F_\sigma, G_n) - \binom{n}{m} \prod_{j=1}^m M_{d_j^-} \right| \leqslant \varepsilon \binom{n}{m} + O(n^{m-1}), \tag{5.7}$$

which implies,

$$\mathbb{E} \, n_<(F_\sigma, G_n) = \binom{n}{m} \prod_{j=1}^m M_{d_j^-} + o(n^m) = \frac{n^m}{m!} \prod_{j=1}^m M_{d_j^-} + o(n^m), \tag{5.8}$$

since $\varepsilon > 0$ is arbitrary.

We have so far considered a fixed σ, but we now sum (5.8) over all σ and use (5.2). Since $n_0(F, G_n) = O(n^{m-1})$,

$$\mathbb{E} \, n(F, G_n) = t_F n^m + o(n^m) \tag{5.9}$$

for some constant t_F depending on F and ν. We have $m! \, t_F = \sum_\sigma \prod_{j=1}^m M_{d_j^-}$, where d_j^- depends on F and σ.

Since $t(F, G_n) = n(F, G_n)/n^m$, (5.9) is the same as

$$\mathbb{E} \, t(F, G_n) \to t_F. \tag{5.10}$$

We have proved this for any graph F, and it follows by [10, Corollary 3.2] that $G_n \xrightarrow{\text{p}} \Gamma$ for some graph limit Γ.

It remains to identify the limit Γ as Γ_ν. We have proved that t_F for each graph F, and thus the limit Γ, depends on ν but not otherwise on the distributions ν_n. For a given distribution ν, we consider Construction 3.1, which is a special case of Construction 2.1 with ν_n the mixture of binomial distributions given by (3.1). By Lemma 3.6, we have $D_n/n \xrightarrow{\text{d}} \nu$. We are then in the setting of the present theorem and the proof above shows $G_n \xrightarrow{\text{p}} \Gamma$. On the other hand, Theorem 5.4 shows $G_n \xrightarrow{\text{p}} \Gamma_\nu$. Hence, $\Gamma = \Gamma_\nu$. $\qquad\square$

6. Further comments and open problems

We have found the limit of the random sequence G_n as a graph limit defined by a kernel on $([0,1]^2, \nu \times \lambda)$. It is easy to find an equivalent kernel on $([0,1]^2, \lambda \times \lambda)$: Let $\psi : [0,1] \to [0,1]$ be the right-continuous inverse of the distribution function of ν. If $U \sim \mathrm{U}(0,1)$, then $\psi(U) \sim \nu$. We define W_ν as the pullback of W via the map $(s,t) \mapsto (\psi(s), t)$, i.e.,

$$W_\nu\big((s_1,t_1),(s_2,t_2)\big) := W\big((\psi(s_1),t_1),(\psi(s_2),t_2)\big) = \begin{cases} \psi(s_2), & \text{if } t_1 < t_2; \\ \psi(s_1), & \text{if } t_1 > t_2. \end{cases}$$
$$(6.1)$$

Then W_ν is a kernel on $([0,1]^2, \lambda^2)$ which is equivalent to W on $([0,1]^2, \nu \times \lambda)$; thus we also have $G_n \xrightarrow{\mathrm{p}} W_\nu$ under the conditions of Theorem 5.4 or Theorem 5.4.

However, it is at least sometimes possible to find simpler representations.

Example 6.1. In Example 2.2 and 3.2, $\nu = \delta_p$ and $\psi(s) = p$ for all s; thus $W_\nu = p$ is constant. (Similarly, $W = p$ a.e. with respect to μ_ν.) In fact, as is well known, the graph limit of $G(n,p)$ is represented by the constant kernel p on any probability space. (Conversely, any representing kernel equals p a.s., see [13, Corollary 8.12].)

Example 6.2. In Example 2.3 and 3.3, ν is concentrated on $\{0,1\}$, so μ_ν is concentrated on $\{0,1\} \times [0,1]$. In particular, the kernel W is a.e. 0/1-valued. (This is a general property of kernels representing limits of threshold graphs; see [9] and [13, Section 9].)

The representation theorem in [9] for general limits of threshold graphs yields a kernel on [0,1]. (This kernel is monotone, and this property makes it uniquely determined a.e.) In the present case, the kernel is the indicator function of the quadrilateral S_p having vertices in $(0,1)$, $(1-p, 1-p)$, $(1,0)$ and $(1,1)$, see [9, Section 6]. Denote this kernel by W'.

It is easy to find a relation between the two representations. Let $\varphi : [0,1] \to \{0,1\} \times [0,1]$ be defined by $\varphi(x) = (0, 1 - x/(1-p))$ for $0 \leqslant x \leqslant 1 - p$ and $\varphi(x) = (1, (x - 1 + p)/p)$ for $1 - p < x \leqslant 1$. Then φ is measure preserving $([0,1], \lambda) \to ([0,1]^2, \mu_\nu)$ and $W'(x,y)$ is the pullback $W(\varphi(x), \varphi(y))$ of W.

As said in Section 4, it is always possible to find an equivalent kernel on $[0,1]$. In the two examples above, there are simple and natural choices of such kernels. However, in Example 2.4 and 3.4 we do not know any natural kernel on $[0,1]$ representing the limit.

Problem 6.3. Find a natural kernel on $[0, 1]$ representing the limit in Example 2.4, i.e., a natural kernel on $[0, 1]$ that is equivalent to W in (5.1) on $([0, 1]^2, \lambda^2)$.

More generally, find a natural representing kernel on $[0, 1]$ for any ν.

We close with two different problems inspired by the results above.

Problem 6.4. We have stated Theorems 5.4 and 5.5 with convergence in probability. Are these results true also a.s.?

Problem 6.5. In Theorem 5.5, we assume that D_n/n converges in distribution, i.e., that the distributions ν_n converge after rescaling. What happens for more general sequences ν_n? Is it possible to characterize the sequences ν_n that give convergence of G_n to some graph limit?

References

[1] T. Austin, On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probability Surveys* **5** (2008), 80–145.

[2] B. Bollobás, *Random Graphs.* 2nd ed., Cambridge Univ. Press, Cambridge, 2001.

[3] B. Bollobás, and O. Riordan, Metrics for sparse graphs. *Surveys in Combinatorics 2009*, LMS Lecture Notes Series 365, Cambridge Univ. Press, 2009, pp. 211–287.

[4] C. Borgs, J. T. Chayes, and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** (2010), no. 6, 1597–1619.

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

[6] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, Convergent sequences of dense graphs II: Multiway cuts and statistical physics. Preprint, 2007. `http://research.microsoft.com/~borgs/`

[7] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, Graph limits and parameter testing. In *STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, 261–270. ACM, New York, 2006.

[8] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, Limits of randomly grown graph sequences, *Eur. J. Comb.* **32** (2011), no. 7, 985–999.

[9] P. Diaconis, S. Holmes, and S. Janson, Threshold graph limits and random threshold graphs. *Internet Mathematics* **5** (2009), no. 3, 267–318.

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

[11] R. Durrett, *Random Graph Dynamics.* Cambridge University Press, Cambridge, 2007.

[12] F. Eggenberger and G. Pólya, Über die Statistik verketteter Vorgänge. *Zeitschrift Angew. Math. Mech.* **3** (1923), 279–289.

[13] S. Janson, Graphons, cut norm and distance, couplings and rearrangements. Preprint, 2010. `arXiv:1009.2376`.

[14] S. Janson, T. Łuczak, and A. Ruciński, *Random Graphs.* Wiley, New York, 2000.

[15] O. Kallenberg, *Foundations of Modern Probability.* 2nd ed., Springer, New York, 2002.

[16] L. Lovász, *Large Networks and Graph Limits.* American Mathematical Society, Providence, RI, 2012.

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

[18] H. M. Mahmoud, *Pólya Urn Models.* CRC Press, Boca Raton, FL, 2009.

[19] M. Mitzenmacher, A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, **1** (2003), no. 2, 226–251.

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: `svante.janson@math.uu.se`

Department of Computer Science and Department of Physics & Astronomy, University College London, Gower Street, WC1E 6BT London, UK
E-mail address: `s.severini@ucl.ac.uk`