European Journal of Combinatorics 32 (2011) 985–999

Contents lists available at ScienceDirect

European Journal of Combinatorics

journal homepage: www.elsevier.com/locate/ejc

Limits of randomly grown graph sequences

Christian Borgs [a], Jennifer Chayes [a], László Lovász [b], Vera Sós [c], Katalin Vesztergombi [b]

[a] *Microsoft Research, Cambridge, MA, United States*
[b] *Institute of Mathematics, Eötvös Loránd University, Budapest, Hungary*
[c] *A. Rényi Institute of Mathematics, Budapest, Hungary*

ARTICLE INFO

Article history:
Available online 6 May 2011

ABSTRACT

Motivated in part by various sequences of graphs growing under random rules (such as Internet models), Borgs, Chayes, Lovász, Sós, Szegedy and Vesztergombi introduced convergent sequences of dense graphs and their limits. In this paper we use this framework to study one of the motivating classes of examples, namely randomly growing graphs. We prove the (almost sure) convergence of several such randomly growing graph sequences, and determine their limit. The analysis is not always straightforward: in some cases the cut-distance from a limit object can be directly estimated, while in other cases densities of subgraphs can be shown to converge.

1. Introduction

Convergent graph sequences and their limits have been studied in connection with Internet models, statistical physics, extremal graph theory, and more. In the context of dense graphs, a rather complete theory has emerged. One can define a notion of convergence based on the convergence of densities of subgraphs. An appropriate notion of distance between two graphs, called their *cut-distance*, can be defined, such that convergent sequences are Cauchy in this metric and vice versa. The completion of the metric space of graphs relative to this metric can be described, and its elements, i.e., limit objects for convergent graph sequences, can be characterized in various ways. To mention one of these, limit objects can be described using two-variable symmetric measurable functions $[0, 1]^2 \to [0, 1]$.

The goal of this paper is study in this framework one of the motivating classes of examples, namely randomly growing graphs. Typically, such a sequence of graphs grows by every now and then adding

E-mail address: lovasz@cs.elte.hu (L. Lovász).

0195-6698/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejc.2011.03.015

a new node, and then creating new edges (between the new node and the old ones, or between two old nodes) randomly, from some simple distribution determined by local conditions.

We will prove the (almost sure) convergence of several such randomly growing graph sequences, and determine their limit. This analysis is not always straightforward: in some cases the cut-distance from a limit object can be directly estimated, while in other cases densities of subgraphs can be shown to converge.

2. Preliminaries

In this section we summarize those notions and results concerning convergent graph sequences and their limits which are relevant for the rest of the paper.

2.1. Convergent graph sequences

For two simple graphs F and G, $\hom(F, G)$ denotes the number of homomorphisms (adjacency-preserving maps) from $V(F)$ to $V(G)$. We also consider the *homomorphism densities*

$$t(F, G) = \frac{\hom(F, G)}{|V(G)|^{|V(F)|}}. \tag{1}$$

(Thus $t(F, G)$ is the probability that a random map $V(F) \to V(G)$ is a homomorphism.)

A sequence (G_n) of graphs is *convergent* if the sequence $t(F, G_n)$ has a limit for every simple graph F. This notion was defined in [3,4].

Convergent graph sequences have a limit object, which can be represented as a measurable function [8]. Let $\mathcal{W}$ denote the space of all bounded measurable functions $W : [0, 1]^2 \to \mathbb{R}$ such that $W(x, y) = W(y, x)$ for all $x, y \in [0, 1]$. We also define $\mathcal{W}_0 = \{W \in \mathcal{W} : 0 \le W \le 1\}$. For every simple graph F and $W \in \mathcal{W}$, we define

$$t(F, W) = \int_{[0,1]^{V(F)}} \prod_{ij \in E(F)} W(x_i, x_j)\, dx.$$

Every finite simple graph G can be represented by a function $W_G \in \mathcal{W}_0$: let $V(G) = \{1, \dots, n\}$. Split the interval $[0, 1]$ into n equal intervals $J_1, \dots, J_n$, and for $x \in J_i$, $y \in J_j$ define

$$W_G(x, y) = \begin{cases} 1, & \text{if } ij \in E(G), \\ 0, & \text{otherwise.} \end{cases}$$

Informally, we replace the (i, j) entry in the adjacency matrix of G by a square of size $(1/n) \times (1/n)$, and define the value of the function W_G on this square as the corresponding entry of the adjacency matrix.

Such functions W represent limits of convergent graph sequences in the following sense.

Theorem 2.1. (a) *For every convergent graph sequence (G_n) there is a $W \in \mathcal{W}$ such that $t(F, G_n) \to t(F, W)$ for every simple graph F.*

(b) *This function W is uniquely determined up to measure-preserving transformations in the following sense: for every other limit function W' there are measure-preserving maps $\phi, \psi : [0, 1] \to [0, 1]$ such that $W(\phi(x), \phi(y)) = W'(\psi(x), \psi(y))$.*

(c) *Every function $W \in \mathcal{W}_0$ arises as the limit of a convergent graph sequence.*

Parts (a) and (c) of the theorem were proved in [8], and part (b), in [2]. The proof of (c) in [8] depends on W-random graphs, to be discussed in the next section.

We could consider any probability space $(\Omega, \mathcal{A}, \pi)$ instead of $[0, 1]$, with a symmetric measurable function $W : \Omega \times \Omega \to [0, 1]$. These structures are called *graphons*. The densities $t(F, W)$ in a graphon could be defined by a similar integral. Considering graphons would not give greater generality, since we could always replace $(\Omega, \mathcal{A}, \pi)$ by the uniform measure on $[0, 1]$. Still, it is sometimes useful to represent the limit object by other probability spaces, as we shall see.

2.2. The distance of graphs

The *cut-norm* introduced in [7] is defined for $W \in \mathcal{W}$ by

$$\|W\|_\square = \sup_{S,T\subset[0,1]} \left| \int_{S\times T} W(x,y)dx\,dy \right|,$$

where the supremum is over measurable subsets of $[0, 1]$. We define the *cut-distance* of two functions in $\mathcal{W}$ by

$$\delta_\square(U,W) = \inf_{\phi:[0,1]\to[0,1]} \|U - W^\phi\|_\square \tag{2}$$

where the infimum is over all invertible maps $\phi : [0, 1] \to [0, 1]$ such that both ϕ and its inverse are measure preserving, and W^ϕ is defined by $W^\phi(x, y) = W(\phi(x), \phi(y))$. For two graphs G and G', this yields a distance

$$\delta_\square(G,G') = \delta_\square(W_G, W_{G'}).$$

Remark 2.2. (a) We call this a "distance" rather than a "metric" since two different graphs can have distance 0. This is the case when one graph can be obtained from the other by replacing each node by the same number of twins, or more generally, when both can be obtained from a third graph in this way. To get a metric, we should identify such pairs of graphs. Similarly, to get a metric on $\mathcal{W}_0$, we have to identify functions U, W for which $\delta_\square(U, W) = 0$. Several characterizations of such pairs are given in [2].

(b) There are combinatorial, but somewhat lengthy ways to define this distance between graphs; see [4].

We can define a similar distance function based on other norms. We shall use the L_1-norm

$$\|W\|_1 = \int_{[0,1]^2} |W(x,y)|\,dx\,dy,$$

from which we can define the *edit distance* of two functions in $\mathcal{W}$ by

$$\delta_1(U,W) = \inf_{\phi:[0,1]\to[0,1]} \|U - W^\phi\|_1. \tag{3}$$

The following characterization of convergent graph sequences was proved in [4] (see [5] for other characterizations not used in this paper).

Theorem 2.3. *A sequence of graphs (G_n) is convergent if and only if it is Cauchy in the $\delta_\square$ distance. The sequence (G_n) converges to W if and only if $\delta_\square(W_{G_n}, W) \to 0$. Furthermore, if this is the case, and $|V(G_n)| \to \infty$, then there is a way to label the nodes of the graphs G_n such that $\|W_{G_n} - W\|_\square \to 0$.*

If the graphs G_n are labeled such that $\|W_{G_n} - W\|_\square \to 0$, then

$$\sup_{S,T} \left| \int_{S\times T} (W_{G_n} - W) \right| \to 0 \quad (n \to \infty).$$

In particular, it follows that

$$\int_{S\times T} (W_{G_n} - W) \to 0 \tag{4}$$

for every product set $S \times T$, which implies that $W_{G_n} \to W$ in the weak* topology of $L_\infty([0, 1]^2)$. Convergence in the norm $\|.\|_\square$ is, however, not equivalent to convergence in this weak* topology, as the sequence of prefix attachment graphs shows (Section 3.3).

2.3. W-random graphs and extensions

Let $(\Omega, \mathcal{A}, \pi, W)$ be a graphon. For every finite subset $S \subseteq \Omega$ we define two graphs $\mathbb{G}(S, W)$ and $\mathbb{H}(S, W)$ on $V(\mathbb{G}(S, W)) = V(\mathbb{H}(S, W)) = S$. In $\mathbb{G}(S, W)$, we connect $i, j \in S$, $i \neq j$, with probability $W(i, j)$. In $\mathbb{H}(S, W)$, we connect $i, j \in S$, $i \neq j$, by an edge with weight $W(i, j)$. If W is $\{0, 1\}$ valued, then $\mathbb{G}(S, W) = \mathbb{H}(S, W)$ is deterministic, and can be considered as an "induced subgraph".

Let $\mathbf{S}_n$ be a random n-element subset of Ω (each element of $\mathbf{S}_n$ chosen independently from the distribution π). The graph $\mathbb{G}(n, W) = \mathbb{G}(\mathbf{S}_n, W)$ is called a *W-random* graph. The following fact was shown in [8] (for the case when the underlying probability space is the uniform distribution on [0, 1], but this is no essential restriction of generality).

Lemma 2.4. *With probability* 1*, the sequence $\mathbb{G}(n, W)$ is convergent and its limit is represented by the function W.*

In this paper, we will also need sequences S_n of subsets of Ω that are not random, but for which $\mathbb{G}(S_n, W)$ still converges to W. We prove and use the following sufficient condition for a deterministic sequence S_n. Let (Ω, d) be a metric space, and π, a probability measure on the Borel subsets of (Ω, d). For every $n \geq 1$, let S_n be a finite subset of Ω such that $|S_n| \to \infty$. We say that the sequence (S_n) is *well distributed in a set $X \subseteq \Omega$* if $|S_n \cap X|/|S_n| \to \pi(X)$ as $n \to \infty$. We say that (S_n) is *well distributed in (Ω, d, π)* if for every $\varepsilon > 0$ there exists a partition $\{P_1, \ldots, P_m\}$ of Ω into sets with diameter at most ε such that S_n is well distributed in each P_j.

Lemma 2.5. *Let (Ω, d, π) be a metric space with an atom-free probability measure. Let $W : \Omega \times \Omega \to [0, 1]$ be a symmetric measurable function that is almost everywhere continuous. Let S_n be a sequence of sets that is well distributed in (Ω, d, π).*

(a) *Then $\delta_1(W_{\mathbb{H}(S_n,W)}, W) \to 0$ and with probability* 1*, $\delta_\square(W_{\mathbb{G}(S_n,W)}, W) \to 0$.*
(b) *If W is* 0–1 *valued, then $\delta_1(W_{\mathbb{G}(S_n,W)}, W) \to 0$.*

It is clear that such a conclusion cannot hold without some assumption on W, since a general measurable function could be changed on the sets $S_n \times S_n$ arbitrarily without changing its subgraph densities.

Proof. (a) First we construct a special partition of Ω.

Claim 2.6. *There exists a sequence of partitions $\mathcal{Q}_n$ of Ω into $|S_n|$ sets such that every partition class contains exactly one point of S_n, the maximum diameter of partition classes tends to* 0*, and the maximum of $\left|\pi(Q)|S_n| - 1\right|$ $(Q \in \mathcal{Q}_n)$ tends to* 0*.*

Let $\varepsilon > 0$. Consider a partition $\{P_1, \ldots, P_m\}$ into sets with diameter at most ε such that S_n is well distributed in every P_j. For n large enough, we have $(1 - \varepsilon)\pi(P_j) \leq |S_n \cap P_j|/|S_n| \leq (1 + \varepsilon)\pi(P_j)$ for every j. Let us partition each set P_j into $|S_n \cap P_j|$ sets of equal measure, each containing exactly one point of $S_n \cap P_j$ to get the partition $\mathcal{Q}_n$. It is clear that this sequence of partitions has the properties as required in the claim.

For each n and $s \in S_n$, let Q_s be the partition class of $\mathcal{Q}$ containing s. Define the function W_n as follows: for $s, s' \in S_n$ and $(x, y) \in Q_s \times Q_{s'}$, let $W_n(x, y) = W(s, s')$. Then $W_n(x, y) \to W(x, y)$ at every point (x, y) where W is continuous; in particular $W_n \to W$ almost everywhere. This implies that

$$\|W_n - W\|_1 \to 0 \quad (n \to \infty). \tag{5}$$

We can view W_n as W_{H_n}, where H_n is a weighted graph with $V(H_n) = S_n$, the weight of node $s \in S_n$ is $\pi(Q_s)$, and the weight of ss' $(s, s' \in S)$ is $W(s, s')$. Note that H_n is almost the same weighted graph as $\mathbb{H}_n = \mathbb{H}(S_n, W)$: they are defined on the same set of nodes, the edges have the same weights, and the nodeweight $\pi(Q_s)$ is asymptotically $1/|S_n|$ by the Claim. Given $\varepsilon > 0$, we have $|\pi(Q_s) - 1/|S_n|| < \varepsilon/|S_n|$ if n is large enough. Hence there is a measure-preserving bijection $\phi : [0, 1] \to [0, 1]$ and a set $R \subseteq [0, 1]$ of measure ε such that

$$W_{H_n}(x, y) = W^{\phi}_{\mathbb{H}_n}(x, y) \quad (x, y \notin R).$$

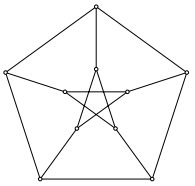
This implies that

$$\delta_1(H_n, \mathbb{H}_n) \to 0 \quad (n \to \infty). \tag{6}$$

By Lemma 4.3 from [4], it follows that with probability 1,

$$\delta_\square(\mathbb{H}(S_n, W), \mathbb{G}(S_n, W)) \to 0 \quad (n \to \infty). \tag{7}$$

Eqs. (5)–(7) imply that $\mathbb{G}(S_n, W) \to W$ with probability 1.



$$\begin{pmatrix} 0&1&0&0&1&1&0&0&0&0\\ 1&0&1&0&0&0&1&0&0&0\\ 0&1&0&1&0&0&0&1&0&0\\ 0&0&1&0&1&0&0&0&1&0\\ 1&0&0&1&0&0&0&0&0&1\\ 1&0&0&0&0&0&0&1&1&0\\ 0&1&0&0&0&0&0&0&1&1\\ 0&0&1&0&0&1&0&0&0&1\\ 0&0&0&1&0&1&1&0&0&0\\ 0&0&0&0&1&0&1&1&0&0 \end{pmatrix}$$

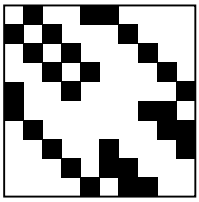
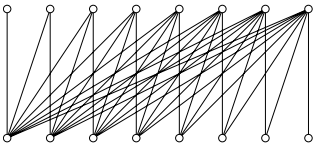
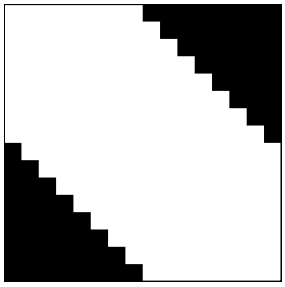


Fig. 1. The Petersen graph, its adjacency matrix, and its pixel picture.





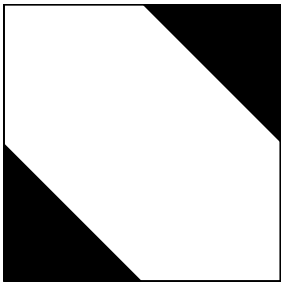


Fig. 2. A half-graph, its pixel picture, and the limit function.

(b) follows trivially, since in this case $\mathbb{H}(S_n, W) = \mathbb{G}(S_n, W)$. □

We note that (b) would also follow from the result of Pikhurko [9] that if a graph sequence tends to a 0–1 valued function W in the $\delta_\square$ distance, then it also tends to W in the δ_1 distance.

2.4. The pixel picture

We have seen that every finite simple graph G can be represented by a function $W_G \in \mathcal{W}_0$. In fact, this representation is very useful for creating figures representing graphs.

Every function $W \in \mathcal{W}_0$ can be represented by a grayscale picture on the unit square: the point (x, y) is black if $W(x, y) = 1$, it is white if $W(x, y) = 0$, and it is appropriately dark grey if $0 < W(x, y) < 1$. For a graph, this picture gives a black-and-white picture consisting of a finite number of "pixels". The origin is in the upper left corner (as for a matrix). Fig. 1 illustrates this construction. Note that the function associated with a graph depends on the ordering of the nodes.

Example 1 (*Half-Graphs*)**.** Consider the *half-graphs* $H_{n,n}$: they are bipartite graphs on $2n$ nodes $\{1, \dots, n, 1', \dots, n'\}$, where i is connected to j' if and only if $i \le j'$. It is easy to see that this sequence is convergent, and to guess the limit function (Fig. 2).

Example 2 (*Erdős–Rényi Random Graphs*)**.** The pixel picture of a random graph is essentially grey (Fig. 3).

The following simple example illustrates the importance of the ordering of the nodes:

Example 3 (*Chessboard*)**.** The 100×100 chessboard in Fig. 4 is the pixel picture of a complete bipartite graph. It is also uniformly grey, so one might assume that it represents a graph that is close to random. But rearranging the rows and columns so that odd indexed columns come first, we see that it is isomorphic to the graph represented by the 2×2 chessboard.

This example also shows that different graphs may be represented by the same pixel picture: all complete bipartite graphs with equal color classes have the same pixel picture. If we restrict our attention to graphs with no twin nodes, the pixel picture will determine the graph.

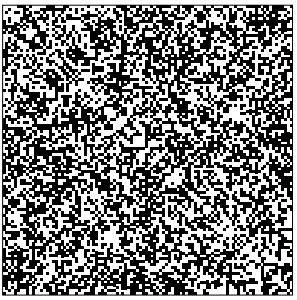


Fig. 3. A random graph with 100 nodes and with edge density 1/2.

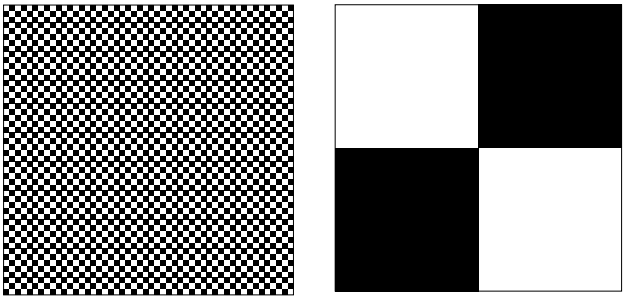


Fig. 4. A chessboard and the pixel picture obtained by rearranging the rows and columns.

The pixel picture of a random graph remains uniformly grey, no matter how you reorder the nodes. It is easy to verify that

$$t(F, G) = t(F, W_G)$$

for every finite simple graph G.

3. Convergent graph sequences and their limits

3.1. Growing uniform attachment graphs

We generate a randomly growing graph sequence G_n^{ua} as follows. We start with a single node. At the nth iteration, a new node is born, and then every pair of nonadjacent nodes is connected with probability $1/n$. We call this graph sequence a *randomly grown uniform attachment graph sequence*.

Let us do some simple calculations. After n steps, let $\{0, 1, \ldots, n-1\}$ be the nodes (born in this order). The probability that nodes $i < j$ are not connected is $\frac{j}{j+1} \cdot \frac{j+1}{j+2} \cdots \frac{n-1}{n} = \frac{j}{n}$. These events are independent for all pairs (i, j). The expected degree of j is (see Fig. 5)

$$\sum_{i=0}^{j-1} \frac{n-j}{n} + \sum_{i=j+1}^{n-1} \frac{n-i}{n} = \frac{n-1}{2} - \frac{j(j-1)}{2n}.$$

The expected number of edges is

$$\frac{1}{2} \sum_{j=0}^{n-1} \left(\frac{n-1}{2} - \frac{j(j-1)}{2n} \right) = \frac{n^2-1}{6}.$$

To figure out the limit function, note that the probability that nodes i and j are connected is $1 - \max(i, j)/n$. If $i = xn$ and $j = yn$, then this is $1 - \max(x, y)$. This motivates the following:

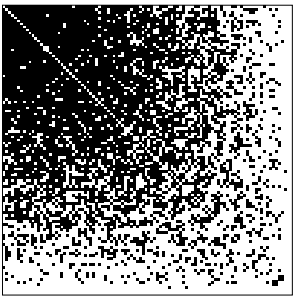


Fig. 5. A randomly grown uniform attachment graph with 100 nodes.

Theorem 3.1. *The sequence G_n^{ua} tends to the limit function $1-\max(x,y)$ with probability* 1.

Proof. For a fixed n, the events that nodes i and j are connected are independent for different i, j, and so by the computation above, G_n^{ua} has the same distribution as $\mathbb{G}(S_n, 1-\max(x,y))$, where $S_n = \{0, 1/n, \dots, (n-1)/n\}$. It is easy to see that this sequence is well distributed in the metric space $[0,1]$ with the uniform measure, and so the theorem follows by Lemma 2.5. □

One can get a good explicit bound on the convergence rate by estimating the cut-distance of $W_{G_n^{\mathrm{ua}}}$ and $1-\max(x,y)$, using the Chernoff–Hoeffding bound.

3.2. Growing ranked attachment graphs

This randomly growing graph sequence G_n^{ra} is generated somewhat similarly. We start with a single node. At the nth iteration, a new node is born, and it is connected to node i with probability $1-i/n$. Then every pair of nonadjacent nodes is connected with probability $2/n$. We call this graph sequence a *randomly grown ranked attachment graph sequence*.

Theorem 3.2. *The sequence G_n^{ra} tends to the limit function $1-xy$ with probability* 1.

Proof. The probability that nodes i and j are *not* connected after the nth step is

$$p_{ij} = \frac{i}{j}\cdot\left(1-\frac{2}{j}\right)\cdot\left(1-\frac{2}{j+1}\right)\cdots\left(1-\frac{2}{n}\right) = \frac{i(j-2)(j-1)}{j(n-1)n}$$
$$= \frac{ij}{n^2} - \frac{(3n-j)ij-2ni}{jn(n-1)} = \frac{ij}{n^2} - q_{ij},$$

where $0 < q_{ij} < \min\{\frac{3}{n}, ij/n^2\}$. Furthermore, these events are independent for different pairs i, j. Therefore, we can generate the graph G_n^{ra} as follows: we generate $\mathbb{G}(S_n, 1-xy)$, where $S_n = \{0, 1/n, \dots, (n-1)/n\}$, and then connect each nonadjacent i and j with probability $1-p_{ij}$. Since $\mathbb{G}(S_n, 1-xy)$ tends to the function $1-xy$ by Lemma 2.5 and the added edges change $\mathbb{G}(S_n, 1-xy)$ negligibly in $\delta_\square$ distance, the theorem follows. □

3.3. Growing prefix attachment graphs

In this construction, it will be more convenient to label the nodes starting with 1. At the nth iteration, a new node n is born, a node z is selected at random, and node n is connected to nodes $1, \dots, z-1$. We denote the nth graph in the sequence by G_n^{pfx}, and call this graph sequence a *randomly grown prefix attachment graph sequence* (Fig. 6).

Remark 3.3. In a recent paper [6], Diaconis, Holmes and Janson study a related graph sequence, where at each step, the new node is connected either to all previous nodes or to none of them, with probability (in the simplest case) 1/2. These graphs are random threshold graphs, and they tend to the limit graphon $\mathbb{1}(x+y \geq 1)$ with probability 1.

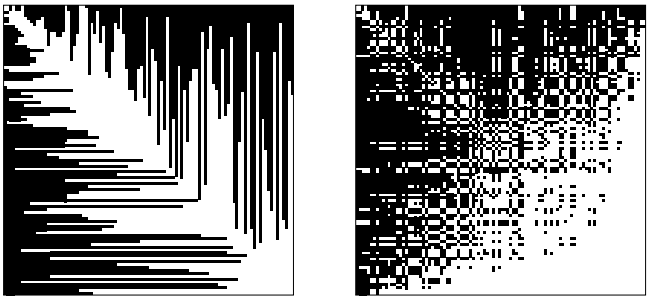


Fig. 6. A randomly grown prefix attachment graph with 100 nodes, and the same graph with nodes ordered by their degrees.

Again we start with some simple calculations. The probability that nodes $i < j$ are connected is $\frac{j-i}{j}$ (but these events are not independent in this case!). The expected degree of j is therefore

$$\sum_{i=1}^{j-1} \frac{j-i}{j} + \sum_{i=j+1}^{n} \frac{i-j}{i} = n - \frac{j}{2} + j \ln \frac{n}{j} + o(n).$$

The expected number of edges is $n(n-1)/4$.

Looking at the picture, it seems that it tends to some function, which we can try to figure out similarly as we did in the case of uniform attachment graphs. The probability that i and j are connected can be written in a symmetric form as

$$\frac{|j-i|}{\max(i,j)}.$$

If $i = xn$ and $j = yn$, then this is

$$\frac{|x-y|}{\max(x,y)}.$$

Does this mean that the function $U(x,y) = |x-y|/\max(x,y)$ is the limit? Somewhat surprisingly, the answer is negative, which we can see by computing triangle densities. The probability that nodes $i < j < k$ form a triangle is $\left(1-\frac{j}{k}\right)\left(1-\frac{i}{j}\right)$ (since if k is connected to j, then it is also connected to i). Hence the expected number of triangles is

$$\sum_{i<j<k} \left(1-\frac{j}{k}\right)\left(1-\frac{i}{j}\right) = \frac{1}{6}\binom{n}{3}.$$

Hence

$$t(K_3, G_n) = \frac{1}{n^3}\binom{n}{3} \to \frac{1}{6}.$$

On the other hand,

$$t(K_3, U) = \int_{[0,1]^3} \frac{|x-y|}{\max(x,y)} \cdot \frac{|x-z|}{\max(x,z)} \cdot \frac{|y-z|}{\max(y,z)}\, dx\, dy\, dz.$$

Since the integrand is independent of the order of the variables, we can compute this easily:

$$t(K_3, U) = 6 \int_{0\le x<y<z\le 1} \left(1-\frac{x}{y}\right)\left(1-\frac{x}{z}\right)\left(1-\frac{y}{z}\right) dx\, dy\, dz = \frac{5}{36}.$$

So U is not the limit of the sequence G_n^{pfx}. On the other hand, it is not hard to verify that

$$\int_{S\times T} (W_{G_n^{\mathrm{pfx}}} - W) \to 0 \tag{8}$$

for every $S, T \subseteq [0,1]$. Indeed, it is enough to prove this for sets S, T from a generating set of the σ-algebra of Borel sets, e.g. rational intervals. Since there are only a countable number of these intervals, it suffices to prove that (8) holds with probability 1 for each fixed S and T. It is also easy to see that it suffices to consider the case $S = T$. For a node j with $j/n \in S$, let $X_{n,j}$ denote the number of edges ij ($i < j$) in G_n^{pfx} with $i/n \in S$, and let $X_n = \sum_{j/n \in S} X_{n,j}$. Then direct computation shows that

$$\frac{1}{n^2}\mathsf{E}(X_n) \to \int_{S\times S} U.$$

Furthermore, the variables $X_{n,j}$ are independent for fixed n; hence the Chernoff–Hoeffding Inequality implies that $\mathsf{P}(|X_n - \mathsf{E}(X_n)| > \varepsilon n^2)$ drops exponentially with n. Hence it follows that $X_n/n^2 \to \int_{S\times S} U$ with probability 1.

So $W_{G_n^{\mathrm{pfx}}} \to W$ in the weak-star topology of $L_\infty[0,1]^2$, but not in our sense. This example also shows that had we defined convergence of a graph sequence through this convergence in weak-star topology (after appropriate relabeling), the limit would not be unique.

Perhaps ordering the nodes by degrees helps? The second pixel picture in Fig. 6 suggests that after this reordering, the functions $W_{G_n^{\mathrm{pfx}}}$ converge to some other continuous function. But again this convergence is only in the weak-star topology, not in the $\delta_\square$ distance. We will see that no continuous function can represent the "right" limit: the limit graphon is 0–1 valued, and it is uniquely determined up to measure-preserving transformations by Theorem 2.1, which do not change this property.

Is this graph sequence convergent at all? Our computation of the triangle densities above can be extended to computing the density of any subgraph, and it follows that the sequence of densities $t(F, G_n^{\mathrm{pfx}})$ is convergent for every n. How do we figure out the limit?

Let us label a node born in step k, connected to $\{1, \dots, m\}$, by $(k/n, m/k) \in [0,1] \times [0,1]$. Then we can observe that *nodes with label (x_1, y_1) and (x_2, y_2) are connected if and only if either $x_1 < x_2 y_2$ or $x_2 < x_1 y_1$.*

Consider the function $W : [0,1]^2 \times [0,1]^2 \to [0,1]$, given by

$$W^{\mathrm{pfx}}((x_1, y_1), (x_2, y_2)) = \begin{cases} 1, & \text{if } x_1 < x_2 y_2 \text{ or } x_2 < x_1 y_1, \\ 0, & \text{otherwise.} \end{cases}$$

Proposition 3.4. *The prefix attachment graphs G_n^{pfx} tend to W^{pfx} with probability* 1.

Proof. Let S_n be the (random) set of points in $[0,1]^2$ of the form $(i/n, z_i/i)$ where $i = 1, \dots, n$ and z_i is a uniformly chosen random integer in $\{1, \dots, i\}$. Then $G_n^{\mathrm{pfx}} = \mathbb{G}(S_n, W^{\mathrm{pfx}}) = \mathbb{H}(S_n, W^{\mathrm{pfx}})$.

Furthermore, with probability 1, the sets S_n are well distributed in $[0,1]^2$. Indeed, for $m \geq 1$, let $J_{m,k}$ denote the interval $(k/m, (k+1)/m]$, and let $\mathcal{P}_m$ denote the partition of $[0,1]^2$ into the sets $J_{m,k} \times J_{m,l}$ $(k, l = 0, \dots, m-1)$. We want to prove that for every fixed m and $0 \leq k, l \leq m-1$, $|S_n \cap (J_{m,k} \times J_{m,l})|/n \to 1/m^2$ as $n \to \infty$ with probability 1. Let

$$X_i = \begin{cases} 1, & \text{if } (i, z_i) \in J_{m,k} \times J_{m,l}, \\ 0, & \text{otherwise.} \end{cases}$$

Then

$$|S_n \cap (J_{m,k} \times J_{m,l})| = \sum_{i=1}^{n} X_i.$$

We have

$$\mathsf{E}(X_i) = \begin{cases} \dfrac{1}{i}\left(\left\lfloor \dfrac{(l+1)i}{m} \right\rfloor - \left\lfloor \dfrac{li}{m} \right\rfloor\right), & \text{if } \dfrac{k}{m} \leq \dfrac{i}{n} \leq \dfrac{k+1}{m}, \\ 0, & \text{otherwise,} \end{cases}$$

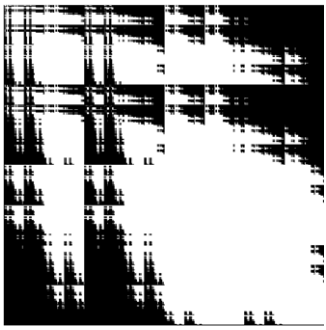


Fig. 7. The limit of randomly grown prefix attachment graphs (as a function on $[0,1]^2$).

and hence

$$\mathsf{E}|S_n \cap (J_{m,k} \times J_{m,l})| = \sum_{i \in nJ_{m,k}} \frac{1}{i}\left(\left\lfloor \frac{(l+1)i}{m} \right\rfloor - \left\lfloor \frac{li}{m} \right\rfloor\right) = \sum_{i \in nJ_{m,k}} \frac{1}{m} + O(\log n)$$
$$= \frac{1}{m}\left(\left\lfloor \frac{(k+1)n}{m} \right\rfloor - \left\lfloor \frac{kn}{m} \right\rfloor\right) + O(\log n) = \frac{n}{m^2} + O(\log n).$$

Thus

$$\mathsf{E}\left(\frac{1}{n}|S_n \cap (J_{m,k} \times J_{m,l})|\right) \to \frac{1}{m^2} \quad (n \to \infty).$$

The fact that $|S_n \cap (J_{m,k} \times J_{m,l})|/n \to 1/m^2$ with probability 1 (not just in expectation) follows by the Law of Large Numbers, since the X_i are independent.

Thus Lemma 2.5 applies and proves the proposition. □

Lemma 2.5 in fact implies (since W^{pfx} is 0–1 valued) that $W_{G_n^{\mathrm{pfx}}}$ tend to W^{pfx} with probability 1 in the edit distance, not just in the cut-distance. This means that while the graphs G_n^{pfx} are random, they are very highly concentrated: two instances of G_n^{pfx} only differ in $o(n^2)$ edges if overlaid properly (not in the original ordering of the nodes!). Informally, they have a relatively small amount of randomness in them, which disappears as $n \to \infty$. Indeed, G_n^{pfx} is generated using only $O(n \log n)$ bits, as opposed to, say, $\mathbb{G}(n, 1/2)$, which is generated using $\binom{n}{2}$ bits. It would be interesting to explore this phenomenon.

Proposition 3.4 gives a nice and simple representation of the limit object with the underlying probability space $[0,1]^2$ (with the uniform measure). If we want a representation on $[0,1]$, we can map $[0,1]$ into $[0,1]^2$ by a measure-preserving map ϕ; then $W^{\phi}_{\mathrm{pfx}}(x,y) = W^{\mathrm{pfx}}(\phi(x), \phi(y))$ gives a representation of the same graphon as a two-variable function. For example, using the map ϕ that separates even and odd bits of x, we get the fractal-like picture in Fig. 7.

It is interesting to note that the graphs $\mathbb{G}(n, W)$ form another (different) sequence of random graphs tending to the same limit W with probability 1.

3.4. Growing preferential attachment graphs

This randomly growing graph sequence G_n^{pf} is generated as follows. We start with a single node. At the nth step (when we already have a graph with n nodes), a new node labeled $n+1$ is created, and is connected to each old node i with probability $(d_n(i)+1)/(n+1)$, independently for different nodes i (here $d_n(i)$ is the current degree of node i). (Adding 1 to the degree in the numerator is needed in order to generate anything other than empty graphs.)

The behavior of the graph sequence G_n^{pf} is somewhat unexpected: it is convergent with probability 1, but the limit is not determined. More precisely:

Proposition 3.5. *With probability* 1*, the sequence* G_n^{pf} *is quasirandom, i.e., it converges to a constant function. The constant is not determined.*

Proof. Set $X_n = |E(G_n^{\mathrm{pf}})|$. Then

$$\mathsf{E}(X_n \mid G_{n-1}^{\mathrm{pf}}) = X_{n-1} + \sum_{i=1}^{n-1} \frac{d_{n-1}(i)+1}{n} = X_{n-1} + \frac{2}{n}X_{n-1} + \frac{n-1}{n}.$$

Hence

$$\frac{1}{(n+2)(n+1)}\mathsf{E}(2X_n + 2n + 1 \mid X_{n-1}) = \frac{1}{(n+1)n}(2X_{n-1} + 2n - 1),$$

which shows that the values

$$Y_n = \frac{2X_n + 2n + 1}{(n+2)(n+1)}$$

form a martingale. Since they are obviously bounded, the Martingale Convergence Theorem implies that with probability 1 there is a value a such that $Y_n \to a$. Clearly, $Y_n \sim t(K_2, G_n^{\mathrm{pf}})$, and so $t(K_2, G_n^{\mathrm{pf}}) \to a$.

Given G_{n-1}^{pf}, the degree of node n when it is born is $\sum_{i=1}^{n-1} X_i$, where the X_i are independent 0–1 random variables with $\mathsf{E}(X_i) = (d_{n-1}(i)+1)/(n+1)$. Hence

$$\mathsf{E}(d_n(n) \mid G_{n-1}^{\mathrm{pf}}) = \sum_{i=1}^{n-1} \frac{d_{n-1}(i)+1}{n} = \frac{2}{n}|E(G_{n-1}^{\mathrm{pf}})| + \frac{n-1}{n},$$

and hence $(d_n(n)+1)/(n+1)$ will be heavily concentrated around a. In particular, $(d_n(n)+1)/(n+1) \to a$ as $n \to \infty$.

Next, observe that the development of $d_n(i)$, for a fixed i, follows a Pólya Urn model with $d_i(i)+1$ red and $i - d_i(i)$ green balls, whence $(d_n(i)+1)/(n+1)$ is a martingale converging to the beta distribution with parameters $d_i(i)+1$ and $i - d_i(i)$. So for large i, $(d_n(i)+1)/(n+1)$ will be heavily concentrated around its expectation $(d_i(i)+1)/(i+1)$, which in turn is heavily concentrated around a. So for large n, most nodes will have degree around an.

It follows that the process is almost the same as $\mathbb{G}(n, a)$, where we can also think of the nodes created one by one and joined to each previous node with probability a. We can couple the two processes to show that with probability 1, they converge to the same limit, which is clearly the function that is identically a.

Note that from the basic properties of martingales,

$$\mathsf{E}(a \mid G_n^{\mathrm{pf}}) = Y_n = \frac{2|E(G_n^{\mathrm{pf}})| + 2n + 1}{(n+2)(n+1)}.$$

Since G_n^{pf} can be any simple graph on n nodes with positive probability, it follows that a is not determined, and with a more careful computation one can see that a falls into any interval with positive probability. It would be interesting to determine the distribution of a. □

3.5. A preferential attachment graph on n fixed nodes

A *preferential attachment graph with n fixed nodes and m edges* PAG(n, m) is the random graph obtained by the following procedure. Let $v_1 \dots v_n$ be a set of nodes. We extend this sequence, one by one, by picking an element of the current sequence randomly and uniformly, and append a copy of it at the end. We repeat this until $2m$ further elements have been added. So we get a sequence $v_1 \dots v_n v_{n+1} \dots v_{n+2m}$.

Now we connect nodes v_{n+2k-1} and v_{n+2k} for $k = 1, 2, \dots, m$, to get PAG(n, m). (Note that PAG(n, m) may have multiple edges and loops, which we have to live with for the time being.)

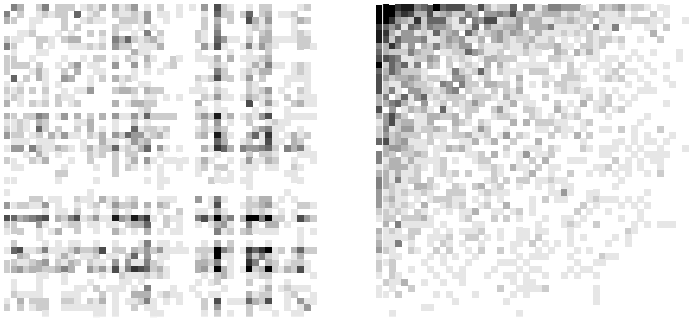


Fig. 8. (a) A preferential attachment graph PAG(50, 1000). Darkness of a pixel indicates multiplicity of the edge. (b) The same graph with the nodes ordered by decreasing degrees.

Another way of describing this construction is to view it as adding edges one by one, where the probability of adding an edge connecting u and v is proportional to the product of their degrees. To be more precise, the probability that the $(k+1)$st edge connects u and v is

$$\begin{cases} \dfrac{2(d_k(u)+1)(d_k(v)+1)}{(n+2k)(n+2k+1)} & \text{if } u \neq v, \\ \dfrac{(d_k(u)+1)(d_k(u)+2)}{(n+2k)(n+2k+1)} & \text{if } u = v, \end{cases}$$

where $d_k(u)$ is the current degree of the node (adding 1 to the degree is needed to start the procedure at all; adding 2 to the second factor in the case when $u = v$ is a minor trick that makes everything come out more nicely; Fig. 8).

Preferential attachment graphs are motivated by the (sparse) Albert–Barabási graphs [1], and they have been studied in detail by Pittel [10].

The somewhat awkward definition of preferential attachment graphs is justified by the following nice properties. First, let us compute the probability that this process yields a multigraph G on $V(G) = [n]$, with degrees $d_1, \ldots, d_n$, with m edges and m' non-loop edges. Fix any order of the edges, and for the non-loop edges fix an order in which their endpoints were inserted (i.e., an orientation of G). Then the probability that G arises this way is

$$\frac{d_1! \ldots d_n!}{n(n+1)\cdots(n+2m-1)}. \tag{9}$$

Summing over all orientations and orderings of the edges, we get that the probability that $\mathrm{PAG}(n, m) = G$ is

$$m! 2^{m'} \frac{d_1! \ldots d_n!}{n(n+1)\cdots(n+2m-1)}. \tag{10}$$

An important observation that we can make from this computation is the following:

Lemma 3.6. *Conditioning on the graph $G(n, m)$, all the $2^{m'} m!$ possibilities in which the edges could have been inserted have the same probability.*

We can use this lemma to determine the expected subgraph densities in $\mathrm{PAG}(n, m)$. For two multigraphs F and G, let $\mathrm{inj}(F, G)$ denote the number of embeddings of f into G, i.e., the number of pairs (ϕ, ψ) of injective maps $\phi : V(F) \to V(G)$ and $\psi : E(F) \to E(G)$ that preserve incidence. Let

$$t_{\mathrm{inj}}(F, G) = \frac{\mathrm{inj}(F, G)}{(n)_k},$$

where $k = |V(F)|$ and $n = |V(G)|$.

Let F be a multigraph on $V(G) = [k]$, with degrees $r_1, \dots, r_k$, with l edges and l' non-loop edges. Fix an order of the edges of F and also an orientation σ of the non-loop edges of F as above. Let $\overrightarrow{e}_1, \dots, \overrightarrow{e}_m$ be the order and orientation in which PAG(n, m) arises. Let $p(\sigma, v_1, \dots, v_k, j_1, \dots, j_l)$ denote the probability that edges $\overrightarrow{e}_{j_1}, \dots, \overrightarrow{e}_{j_l}$ form a copy of F on nodes $v_1, \dots, v_k$ (with the given labeling of the nodes, the given order of the edges, and the given orientation). By Lemma 3.6, this number is the same for any l-tuple $(j_1, \dots, j_l)$, and trivially, it is the same for every k-tuple $(v_1, \dots, v_k)$. Hence

$$\begin{aligned}\mathsf{E}\big(\mathrm{inj}(F, \mathrm{PAG}(n, m))\big) &= \sum_{v_1,\dots,v_k} \sum_{j_1,\dots,j_l} \sum_{\sigma} p(\sigma, v_1, \dots, v_k, j_1, \dots, j_l) \\ &= (n)_k (m)_l 2^{l'} p(\sigma_0, 1, \dots, k, 1, \dots, l),\end{aligned}$$

where σ_0 is any fixed orientation of F. By (9), we have

$$p(\sigma_0, 1, \dots, k, 1, \dots, l) = \frac{r_1! \dots r_k!}{(n + 2l - 1)_{2l}},$$

and so

$$\mathsf{E}\big(t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m))\big) = \frac{1}{(n)_k} (n)_k (m)_l 2^{l'} \frac{r_1! \dots r_k!}{(n + 2l - 1)_{2l}} = 2^{l'} r_1! \dots r_k! \frac{(m)_l}{(n)_{2l}}. \tag{11}$$

Suppose that $n, m \to \infty$ so that $m \sim cn^2/2$. Then

$$\mathsf{E}\big(t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m))\big) \sim 2^{l'} r_1! \dots r_k! \frac{m^l}{n^{2l}} \longrightarrow 2^{l'-l} c^l r_1! \dots r_k!.$$

If we assume that F has no loops, then

$$\mathsf{E}\big(t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m))\big) \longrightarrow c^l r_1! \dots r_k!.$$

Using high concentration results, one can show not only that this convergence holds in expectation, but also that with probability 1,

$$t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m)) \longrightarrow c^l r_1! \dots r_k!.$$

Note that the relation $t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m)) \sim t(F, \mathrm{PAG}(n, m))$ does not hold in general if F has multiple edges. In fact, it is easy to see that

$$t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m)) \sim \sum_{F'} t(F, \mathrm{PAG}(n, m)) \prod_{i,j \in V(F)} m'_i j! \left\{ \begin{matrix} m_{ij} \\ m'_{ij} \end{matrix} \right\},$$

where F' ranges through all multigraphs obtained from F by reducing the edge multiplicities (not strictly, but keeping at least one copy of each edge), m_{ij} and m'_{ij} denote the multiplicities of the edge ij in F and F', respectively, and $\left\{ \begin{matrix} a \\ b \end{matrix} \right\}$ denotes the Stirling number of the second kind. For example, if $K_2^{(2)}$ denotes the graph on two nodes having two parallel edges, then

$$t(K_2^{(2)}, \mathrm{PAG}(n, m)) \sim t_{\mathrm{inj}}(K_2^{(2)}, \mathrm{PAG}(n, m)) + t_{\mathrm{inj}}(K_2, \mathrm{PAG}(n, m)).$$

Let $L_c(x, y) = c(\ln x)(\ln y)$. Then for a multigraph F without loops we have

$$t(F, L_c) = \int_{[0,1]^k} \prod_{ij \in E(F)} W(x_i, x_j)\, dx = \int_{[0,1]^k} c^l \prod_{i=1}^{k} (\ln x_i)^{r_i}\, dx = c^l r_1! \dots r_k!.$$

This implies that the limit of preferential attachment graphs PAG(n, cn^2), with probability 1, is described by the function L_c. To be precise, the graphs PAG(n, cn^2) have multiple edges, and hence the theory of convergent graph sequences developed in [4,5] does not apply, but the computations above show that the convergence does occur in at least one possible sense.

Proposition 3.7. *If $m(n) = (c + o(1))n^2$, then with probability 1, $t_{\mathrm{inj}}(F, \mathrm{PAG}(n, m)) \to t(F, L_c)$ for every multigraph F without loops.*

Let SPAG(n, cn^2) denote the *simplified preferential attachment graph* obtained from PAG(n, cn^2) by deleting loops and keeping only one copy of parallel edges. Szakács [11] proved that this sequence of graphs is convergent with probability 1, and its limit is the function $1 - \exp(-c \ln x \ln y)$.

4. Convergence to a prescribed function

Lemma 2.4 gives a way to construct a randomly growing graph sequence converging to a given function W. Let $s_1, s_2, \ldots \in \Omega$ be independent random samples from π, and let $S_n = \{s_1, \ldots, s_n\}$. We can construct $\mathbb{G}(S_n, W)$ by taking $G(S_{n-1}, W)$, adding s_n as a new node, and connecting s_n to s_i with probability $W(s_n, s_i)$. Then $\mathbb{G}(S_1, W), \mathbb{G}(S_2, W), \ldots$ is a randomly growing sequence of graphs, and by Lemma 2.4, we have $\mathbb{G}(S_n, W) \to W$ with probability 1.

However, one can have several objections to this method: first, the new edges are not added independently of each other; second, even if $\Omega = [0, 1]$, and the function W is, say, continuous and monotone, the order in which the nodes of $\mathbb{G}(S_n, W)$ are born is random, and has nothing to do with the order of the points $s_i \in [0, 1]$ representing them. In other words, to get a labeling for which $W_{\mathbb{G}(S_n,W)} \to W$ in the $\|.\|_\square$ norm, we have to reorder the nodes.

It may be interesting to consider rules for generating randomly growing graph sequences (G_n) with a prescribed limit function W for which these objections cannot be raised. Given a function $W \in \mathcal{W}_0$, monotone decreasing in each variable, construct a randomly growing simple graph sequence $(G_1, G_2, \ldots)$ as follows. G_1 is a single node labeled 1. For $n > 1$, define

$$p_{n,j} = W\left(\frac{j}{n}, 1\right), \qquad p_{n,ij} = \frac{W(\frac{i}{n}, \frac{j}{n}) - W(\frac{i}{n-1}, \frac{j}{n-1})}{1 - W(\frac{i}{n-1}, \frac{j}{n-1})}.$$

To get G_n from G_{n-1}, we add a new node n, connect it to each node $j < n$ with probability $p_{n,j}$, and connect any two nonadjacent nodes $i, j < n$ with probability $p_{n,ij}$. All of these decisions are independent. The monotonicity of W implies that $0 \le p_{n,ij} \le 1$ is a legal probability.

Proposition 4.1. *The sequence of graphs G_n constructed above has the property that $W_{G_n} \to W$ in the $\|.\|_\square$ norm.*

Proof. The probability that nodes $i < j$ are not connected in G_n is

$$\begin{aligned}
&(1 - p_{j,i})(1 - p_{j+1,ij}) \cdots (1 - p_{n,ij}) \\
&= \left(1 - W\left(\frac{i}{j}, 1\right)\right) \frac{1 - W(\frac{i}{j+1}, \frac{j}{j+1})}{1 - W(\frac{i}{j}, \frac{j}{j})} \cdots \frac{1 - W(\frac{i}{n}, \frac{j}{n})}{1 - W(\frac{i}{n-1}, \frac{j}{n-1})} \\
&= 1 - W\left(\frac{i}{n}, \frac{j}{n}\right),
\end{aligned}$$

and hence the probability that they are adjacent is $W(\frac{i}{n}, \frac{j}{n})$. Thus G_n is the graph $\mathbb{G}(S_n, W)$, where $S_n = \{\frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\}$. It is trivial that this sequence of sets is well distributed in $[0, 1]$, and since W is almost everywhere continuous, it follows by Lemma 2.5 that $G_n \to W$ with probability 1. □

The convergent sequences discussed in Sections 3.1 and 3.2 are special cases of this construction. A more general nice case is when $W = 1 - U$, where U is homogeneous of some degree: $U(\lambda x, \lambda y) = \lambda^c U(x, y)$ with some $c \ge 0$. When a new node n is born we connect it to node $i < n$ with probability $W(\frac{i}{n}, 1)$, and then at each further step, we connect any two nonadjacent nodes with probability $1 - \left(\frac{n-1}{n}\right)^c$.

Acknowledgement

This research was sponsored by OTKA Grant No. 67867.

References

[1] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, Rev. Modern Phys. 74 (2002) 47–97.
[2] C. Borgs, J. Chayes, L. Lovász, Moments of two-variable functions and the uniqueness of graph limits, Geom. Funct. Anal. 19 (2010) 1597–1619.

[3] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Counting graph homomorphisms, in: M. Klazar, J. Kratochvil, M. Loebl, J. Matoušek, R. Thomas, P. Valtr (Eds.), Topics in Discrete Mathematics, Springer, 2006, pp. 315–371.
[4] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Convergent graph sequences I: subgraph frequencies, metric properties, and testing, Adv. Math. 219 (2008) 1801–1851.
[5] C. Borgs, J. Chayes, L. Lovász, V.T. Sós, K. Vesztergombi, Convergent graph sequences II: multiway cuts and statistical physics, (2007). Available from: http://research.microsoft.com/~borgs/Papers/ConRight.pdf.
[6] P. Diaconis, S. Holmes, S. Janson, Threshold graph limits and random threshold graphs, Internet Math. 5 (2008) 267–320.
[7] A. Frieze, R. Kannan, Quick approximation to matrices and applications, Combinatorica 19 (1999) 175–220.
[8] L. Lovász, B. Szegedy, Limits of dense graph sequences, J. Comb. Theory B 96 (2006) 933–957.
[9] O. Pikhurko, An analytic approach to stability, Discrete Math. 310 (2010) 2951–2964.
[10] B. Pittel, On a random graph evolving by degrees, Adv. Math. 223 (2010) 619–671.
[11] L. Szakács, (unpublished).