The diameter of a scale-free random graph

Béla Bollobás[*†‡] Oliver Riordan[†]

August 9, 2002

Abstract

We consider a random graph process in which vertices are added to the graph one at a time and joined to a fixed number m of earlier vertices, where each earlier vertex is chosen with probability proportional to its degree. This process was introduced by Barabási and Albert [3], as a simple model of the growth of real-world graphs such as the world-wide web. Computer experiments presented by Barabási, Albert and Jeong [1, 5] and heuristic arguments given by Newman, Strogatz and Watts [23] suggest that after n steps the resulting graph should have diameter approximately $\log n$. We show that while this holds for $m = 1$, for $m \geq 2$ the diameter is asymptotically $\log n / \log \log n$.

1 Introduction

Recently there has been considerable interest in studying complex real-world networks and attempting to understand their properties using random graphs as models. Attention has been focused particularly on the 'small-world' phenomenon, that graphs with a very large number n of vertices often have diameter around $\log n$. In the context of various standard random graph models, this phenomenon has of course been known for a long time, see for example [13, 14, 6]. That such a small diameter can be expected even when the vertex degrees are constant was shown in [9]. Related results in different contexts include [19, 18, 10]; see also chapter X of [7]. A less standard example showing that when the degrees are constant even a small amount of randomness produces this phenomenon is given in [8]. However, real-world graphs are often not well approximated by these models, as shown by their degree sequences, for example.

Barabási and Albert [3], as well as several other groups (see [5] and the references therein), noticed that in many real-world examples the fraction $P(k)$ of vertices with degree k follows a power law over a large range, with $P(k)$ proportional to $k^{-\gamma}$ for some constant γ independent of the scale of the network.

[*]Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA
[†]Trinity College, Cambridge CB2 1TQ, UK
[‡]Research supported in part by NSF grant no. DSM 9971788

Graphs with this behaviour have been studied in several papers, in particular with respect to how well interconnected they are, as measured by the diameter, or by the average distance between pairs of vertices. Such study has been from one of two points of view. In [1, 5, 23] graphs are constructed at random subject to the degree sequence having the desired distribution, and the diameter is analyzed (by computer experiments in [1, 5], and heuristically in [23]), obtaining an answer of the form $A + B \log n$. This 'small-world phenomenon' is often seen as surprising, as it is in the real world. However, the examples cited in the first paragraph show that in the random graph context if anything is surprising it is that the diameter should be so *large*; one would expect such a skewed degree sequence to lead to a diameter no larger than, and perhaps smaller than, the $\log n$ obtained for all degrees roughly equal. This is discussed briefly in section 10.

The above point of view takes the form of the degree sequence as an experimental fact. Putting only this information into the model, one investigates which other properties of the graph (such as small diameter) follow naturally from such a degree sequence, without considering how the degree distribution arises. A different approach is to try to model the growth of the graph in such a way as to explain the distribution of the degrees. In particular, Barabási and Albert suggested modeling such networks using a random graph process with the following description, taken from [3]:

> ... starting with a small number (m_0) of vertices, at every time step we add a new vertex with $m(\leq m_0)$ edges that link the new vertex to m different vertices already present in the system. To incorporate preferential attachment, we assume that the probability Π that a new vertex will be connected to a vertex i depends on the connectivity k_i of that vertex, so that $\Pi(k_i) = k_i / \sum_j k_j$. After t steps the model leads to a random network with $t + m_0$ vertices and mt edges.

Such a process provides a highly simplified model of the growth of the world-wide web, for example, the vertices representing web sites or web pages, and the edges links from sites to earlier sites, the idea being that a new site is more likely to link to existing sites which are 'popular' (are often linked to) at the time the site is added. It is easy to see heuristically that this process leads to a degree distribution $P(k)$ approximately of the form $P(k) \propto k^{-3}$ [4], so that the number of vertices with degree at least k falls off as ck^{-2} for large k. A precise version of this statement will be proved in a forthcoming paper [12]. (Note added: although written later, [12] has appeared first.) Here we shall study the diameter of the resulting graph, showing that it is asymptotically $\log n / \log \log n$ if $m \geq 2$. In contrast, for $m = 1$ a result of Pittel [24] states essentially that the diameter is $\Theta(\log n)$. The relationship of these results to previous heuristics is discussed briefly in the final section.

Note that a different type of model leading to the 'small-world phenomenon' was introduced earlier by Watts and Strogatz [28], involving regularly arranged 'local' connections together with random 'global' connections. As the graphs produced have all degrees about the same, the study of the diameter in this case is a separate topic to that of this paper.

2 The model

When making the model described in the preceding section precise we have some choice as to how to proceed, since the distribution of a random m-element set is *not* specified by giving the marginal probability that each element is contained in this set. When we add a new vertex to the graph we shall add m new edges one at a time, allowing multiple edges between the same pair of vertices. Also, it will be convenient to allow loops; in terms of the interpretation there is no reason to exclude multiple links from one site to another, or links between different parts of a site or even page.

For precise definitions we start with the case $m = 1$. Consider a fixed sequence of vertices $v_1, v_2, \ldots$. (Most of the time we shall take $v_i = i$ to simplify the notation.) We write $d_G(v)$ for the degree of the vertex v in the graph G. We shall inductively define a random graph process $(G_1^t)_{t\geq 0}$ so that G_1^t is a graph on $\{v_i : 1 \leq i \leq t\}$, as follows. Start with G_1^0, the empty 'graph' with no vertices, or with G_1^1 the graph with one vertex and one loop. Given G_1^{t-1}, we form G_1^t by adding the vertex v_t together with a single edge between v_t and v_i, where i is chosen randomly with

$$\mathbb{P}(i = s) = \begin{cases} d_{G_1^{t-1}}(v_s)/(2t-1) & 1 \leq s \leq t-1, \\ 1/(2t-1) & s = t. \end{cases}$$

In other words, we send an edge e from v_t to a random vertex v_i, where the probability that a vertex is chosen as v_i is proportional to its degree at the time, counting e as already contributing one to the degree of v_t. (We shall see why this is convenient later.) For $m > 1$ we add m edges from v_t one at a time, counting the previous edges as well as the 'outward half' of the edge being added as already contributing to the degrees. Equivalently, we define the process $(G_m^t)_{t\geq 0}$ by running the process (G_1^t) on a sequence $v_1', v_2', \ldots$; the graph G_m^t is formed from G_1^{mt} by identifying the vertices $v_1', v_2', \ldots, v_m'$ to form v_1, identifying $v_{m+1}', v_{m+2}', \ldots, v_{2m}'$ to form v_2, and so on.

From now on we shall take $v_i = i$, so G_m^t is a graph on $[t] = \{1, 2, \ldots, t\}$. We have defined G_m^t as an undirected graph, and it is as an undirected graph that we shall measure its diameter. We note, however, that G_m^t has a very natural orientation: direct each edge ij with $i > j$ from i to j, so each non-loop edge is directed from a later vertex to an earlier one. One can check that if any possible graph G_m^t is given *without* the vertex labels, the edge orientations can be uniquely reconstructed. As we shall always treat G_m^t as a graph with labelled vertices, this is of no relevance here.

Often we shall consider not the whole process, but just the graph obtained at one particular time: we shall write $\mathcal{G}_m^n$ for the probability space of graphs on $[n]$ where a random $G_m^n \in \mathcal{G}_m^n$ has the distribution derived from the process above.

As the process (G_m^t) is defined in terms of (G_1^t) we can deduce the properties of a random $G_m^n \in \mathcal{G}_m^n$ from those of a random $G_1^{mn} \in \mathcal{G}_1^{mn}$. For the moment we shall thus restrict our attention to the case $m = 1$. The reason for allowing loops is that this allows us to give a precise alternative description of the distribution of G_1^n, in terms of pairings.

An *n-pairing* is a partition of the set $\{1, 2, \ldots, 2n\}$ into pairs, so there are $(2n)!/(n!2^n)$ n-pairings. These objects are sometimes thought of as *linearized chord diagrams* (or *LCDs*) [26, 11], where an LCD with n chords consists of $2n$ distinct points on the x-axis paired off by semi-circular chords in the upper half plane. Two LCDs are considered to be the same when one can be turned into the other by moving the points on the x-axis without changing their order. Thinking of pairings as LCDs, we shall talk of chords and their left and right endpoints. We form a graph $\phi(L)$ from an LCD L as follows: starting from the left, identify all endpoints up to and including the first right endpoint reached to form vertex 1. Then identify all further endpoints up to the next right endpoint to form vertex 2, and so on. For the edges, replace each chord by an edge joining the vertex corresponding to its right endpoint to that corresponding to its left endpoint. We claim that if L is chosen uniformly at random from all $(2n)!/(n!2^n)$ LCDs with n chords (i.e., n-pairings), then $\phi(L)$ has the same distribution as a random $G_1^n \in \mathcal{G}_1^n$. To see this note that L can be obtained by taking a random LCD L' with $n-1$ chords and adding a new chord whose right endpoint is to the right of all $n-1$ chords, and whose left endpoint lies in one of the $2n-1$ possible places, each chosen with equal probability. This corresponds to adding a new vertex to $\phi(L')$ and joining it to another vertex with probabilities according to the degrees, exactly as in the description of (G_1^n).

We now study the diameter of G_m^n as an undirected graph, using this description from pairings.

3 Results

Our aim is to prove the following result, where $G_m^n \in \mathcal{G}_m^n$ is the random graph described above. Following standard terminology we shall say that *almost every* or *a.e.* $G_m^n \in \mathcal{G}_m^n$ has a certain property if the probability that a random $G_m^n \in \mathcal{G}_m^n$ has this property tends to 1 as $n \to \infty$, keeping m fixed.

Theorem 1. *Fix an integer $m \geq 2$ and a positive real number ϵ. Then a.e. $G_m^n \in \mathcal{G}_m^n$ is connected and has diameter $\mathrm{diam}(G_m^n)$ satisfying*

$$(1-\epsilon)\log n/\log\log n \leq \mathrm{diam}(G_m^n) \leq (1+\epsilon)\log n/\log\log n.$$

The lower bound is relatively easy to prove since we can proceed directly from the definition of the process (G_m^t). This bound is proved in section 4, in a tighter form than that given above. Our proof of the upper bound is much more complicated; for this we use an alternative method for generating a random G_m^n based on generating a random mn-pairing in a certain way, described in section 6.

In the rest of the paper we shall use the following standard notation. Given a graph G we shall write $\Delta(G)$ for its maximum degree, $E(G)$ for the set of its edges, and $e(G) = |E(G)|$ for the number of edges. We write $\mathrm{Bi}(n,p)$ for the binomial distribution with parameters n and p, and $\mathbb{I}_A$ for the indicator function of an event A. All logarithms are natural unless otherwise indicated.

4 The lower bound

To prove the lower bound in Theorem 1 we shall consider G_1^N with $N = mn$. We shall compare G_1^N with a random graph in which every edge ij is present with probability $C/\sqrt{ij}$ independently of the other possible edges, for some constant C. It turns out that 'small' graphs are not much more likely to occur in G_1^N than in this random graph, despite the dependence present in G_1^N.

Recall that G_1^N is defined as a graph on $[N] = \{1, 2, \ldots, N\}$. We shall write g_j for the vertex to which vertex j sends an edge, so $E(G_1^j) = E(G_1^{j-1}) \cup \{g_j j\}$. Now given G_1^{j-1}, the probability that $g_j = i$ is proportional to the degree of i in G_1^{j-1}. Thus, given that $g_j = i$, the vertex i is likely to have a relatively large degree in G_1^{j-1}, and hence in G_1^t for all t. This suggests that the events $g_j = i$ and $g_{j'} = i'$ are positively correlated if $i = i'$, and negatively correlated otherwise. However, as we shall now see, for $i = i'$ any positive correlation is not too strong. Later we shall prove the negative correlation in a more general context. All constants implicit in the $O()$ terms are absolute constants.

Lemma 2. *If $1 \le i < j$, then*

$$\mathbb{P}(g_j = i) = O\left((ij)^{-1/2}\right). \tag{1}$$

Also, if $1 \le i < j < k$, then

$$\mathbb{P}(g_j = i, g_k = i) = O\left(i^{-1}(jk)^{-1/2}\right). \tag{2}$$

Proof. We shall consider the process (G_1^t), applying repeatedly the fact that $\mathbb{E}(A) = \mathbb{E}(\mathbb{E}(A \mid B))$, with B the random variable G_1^{t-1}. Let $d_{t,i} = d_{G_1^t}(i)$ be the degree of the vertex i in the graph G_1^t. We start by evaluating the expectation of $d_{t,i}$ for all t. From the definition of the process (G_1^t) we have

$$\mathbb{P}\left(g_t = i \mid G_1^{t-1}\right) = \begin{cases} d_{t-1,i}/(2t-1) & t > i, \\ 1/(2i-1) & t = i. \end{cases} \tag{3}$$

For $t > i$ we have $d_{t,i} = d_{t-1,i} + \mathbb{I}_{\{g_t = i\}}$, so

$$\mathbb{E}\left(d_{t,i} \mid G_1^{t-1}\right) = d_{t-1,i} + \frac{d_{t-1,i}}{2t-1} = \left(1 + \frac{1}{2t-1}\right) d_{t-1,i}.$$

Taking expectations of both sides,

$$\mathbb{E}(d_{t,i}) = \left(1 + \frac{1}{2t-1}\right) \mathbb{E}(d_{t-1,i}). \tag{4}$$

Let us write $\mu_{t,i}$ for $\mathbb{E}(d_{t,i})$. Then, as $\mu_{i,i} = 2i/(2i-1)$, we have for all $t \geq i$ that

$$\mu_{t,i} = \prod_{s=i}^{t} \left(1 + \frac{1}{2s-1}\right) = O\left(\sqrt{t/i}\right),$$

where the estimate follows by taking logarithms and comparing the resulting sum with an integral. Taking the expectation of both sides of (3) for $t = j > i$ we obtain

$$\mathbb{P}(g_j = i) = \mu_{j-1,i}/(2j-1) = O\left((ij)^{-1/2}\right),$$

proving the first part of the lemma.

For the second statement it turns out that we need to consider second moments. Suppose that $t > i$. Then from (3) we have

$$\begin{aligned}\mathbb{E}\left(d_{t,i}^2 \mid G_1^{t-1}\right) &= d_{t-1,i}^2\left(1 - \frac{d_{t-1,i}}{2t-1}\right) + (d_{t-1,i}+1)^2 \frac{d_{t-1,i}}{2t-1} \\ &= d_{t-1,i}^2\left(1 + \frac{2}{2t-1}\right) + \frac{d_{t-1,i}}{2t-1}.\end{aligned}$$

Taking the expectation of both sides we obtain

$$\mathbb{E}(d_{t,i}^2) = \left(1 + \frac{2}{2t-1}\right)\mathbb{E}(d_{t-1,i}^2) + \frac{\mathbb{E}(d_{t-1,i})}{2t-1}. \tag{5}$$

Adding (4) and (5), and writing $\mu_{t,i}^{(2)}$ for $\mathbb{E}(d_{t,i}^2) + \mathbb{E}(d_{t,i})$, we obtain

$$\mu_{t,i}^{(2)} = \left(1 + \frac{2}{2t-1}\right)\mu_{t-1,i}^{(2)},$$

while $\mu_{i,i}^{(2)} = O(1)$ as $d_{i,i}$ is either 1 or 2. Hence

$$\mu_{t,i}^{(2)} = \prod_{s=i+1}^{t}\left(1 + \frac{2}{2s-1}\right)\mu_{i,i}^{(2)} = \frac{2t+1}{2i+1}\mu_{i,i}^{(2)} = O(t/i).$$

Now from (3) we have

$$\mathbb{E}\left(d_{j,i}\mathbb{I}_{\{g_j=i\}} \mid G_1^{j-1}\right) = (d_{j-1,i}+1)\frac{d_{j-1,i}}{2j-1}.$$

Taking the expectation of both sides,

$$\mathbb{E}\left(d_{j,i}\mathbb{I}_{\{g_j=i\}}\right) = \frac{\mu_{j-1,i}^{(2)}}{2j-1} = O\left(i^{-1}\right).$$

Arguing as in the proof of (1) above, for all $t > j$ we have

$$\mathbb{E}\left(d_{t,i}\mathbb{I}_{\{g_j=i\}}\right) = O\left(\sqrt{t/j}\right)\mathbb{E}\left(d_{j,i}\mathbb{I}_{\{g_j=i\}}\right) = O\left(i^{-1}\sqrt{t/j}\right).$$

Now

$$\mathbb{P}\left(g_j = i, g_k = i \mid G_1^{k-1}\right) = \mathbb{I}_{\{g_j=i\}} \frac{d_{k-1,i}}{2k-1}.$$

Taking expectations gives

$$\mathbb{P}(g_j = i, g_k = i) = O\left((ik)^{-1}\sqrt{k/j}\right) = O\left(i^{-1}(jk)^{-1/2}\right),$$

completing the proof of the lemma. □

Note that a result similar to the above can be proved for the probability that $g_{j_1} = g_{j_2} = \cdots = g_{j_r} = i$, where r is fixed and $i < j_1 < \cdots < j_r$. In fact, our proof actually gives a formula this probability for $r = 1, 2$. Since this is rather cumbersome, it is more convenient to use the estimate stated in the lemma.

We now turn to the negative correlation case. For this we shall use the following alternative description of the process (G_1^t). We consider G_1^t as consisting of $2t$ half-edges numbered $1, 2, \ldots, 2t$, each attached to a vertex from $[t]$. Given G_1^{t-1}, we obtain G_1^t by adding a new half-edge (number $2t-1$) attached to vertex t, and then a half-edge (number $2t$) which must be attached to each vertex with probability proportional to its degree. As the degree of a vertex is the number of half-edges already attached to it, we can think of half-edge $2t$ as being associated to one of the $2t-1$ existing half-edges chosen uniformly at random, and attached to whichever vertex this half-edge is attached to. Let us write h_t for the number of the half-edge to which half-edge $2t$ is attached, so the h_t are independent, with h_t chosen uniformly from $\{1, 2, \ldots, 2t-1\}$.

Lemma 3. *Let E and E' be events of the form*

$$E = \bigcap_{s=1}^{r} \{g_{j_s} = i_s\}, \quad E' = \bigcap_{s=1}^{r'} \{g_{j'_s} = i'_s\},$$

where $i_s < j_s$, $i'_s < j'_s$ for all s. If the sets $\{i_1, \ldots, i_r\}$ and $\{i'_1, \ldots, i'_{r'}\}$ are disjoint then

$$\mathbb{P}(E \cap E') \le \mathbb{P}(E)\,\mathbb{P}(E').$$

Proof. In terms of the description above, $g_j = i$ if and only if either $h_j = 2i-1$, or for some j' we have $h_j = 2j'$, $h_{j'} = 2i-1$ or for some j', j'' .... More formally, the event $g_j = i$ is a disjoint union of events of the form

$$\exists s \ge 0, j_1, \ldots, j_s \text{ s.t. } (\forall a : 0 \le a < s) h_{j_a} = 2j_{a+1},\ h_{j_s} = 2i-1, \tag{6}$$

where $j_0 = j$. Thus we can write E as a disjoint union of events F_k, each an intersection of events of the form (6). Similarly we can write E' as a disjoint union of such events F'_l. Now for any k, l the events F_k and F'_l are either independent, if no h_t appears in both, or inconsistent; if some h_t appears in both, the corresponding half-edge must be attached to one of the vertices $i_1, \ldots, i_r$ if F_k holds, but to one of the vertices $i'_1 \ldots, i'_{r'}$ if F'_l holds. Thus $\mathbb{P}(F_k \cap F'_l) \le \mathbb{P}(F_k)\,\mathbb{P}(F'_l)$ in all cases. Summing over k and l completes the proof of the lemma. □

Combining the two lemmas above gives an upper bound on the probability that a simple enough graph is contained in G_1^N. In what follows, the subgraph S we are looking for is a graph on the same vertex set as G_1^N. To say that S is a subgraph of G_1^N, or occurs in G_1^N, is thus to say that the pairs of vertices joined in S are joined in G_1^N, not that G_1^N has a subgraph isomorphic to S. A vertex i is *earlier* than a vertex j if $i < j$, and *later* if $i > j$.

Lemma 4. *Let S be a loopless graph on $[N]$ in which each vertex is joined to at most one earlier vertex and at most two later vertices. Then*

$$\mathbb{P}\left(S \subset G_1^N\right) \leq C^{e(S)} \prod_{ij \in E(S)} \frac{1}{\sqrt{ij}},$$

for some absolute constant C.

Proof. Grouping the edges of S according to the earlier vertex of each, the event $S \subset G_1^N$ is the intersection of events

$$E_k = \{g_{j_{k,1}} = \cdots = g_{j_{k,n_k}} = i_k\},$$

where the i_k are distinct and $n_k = 1, 2$ for all k. By Lemma 2 there is an absolute constant C such that

$$\mathbb{P}(E_k) \leq C^{n_k} \prod_{l=1}^{n_k} (i_k j_{k,l})^{-1/2}$$

for every k. Also, by Lemma 3, the events E_k and $E_1 \cap \cdots \cap E_{k-1}$ are negatively correlated for every k, so

$$\mathbb{P}\left(S \subset G_1^N\right) = \mathbb{P}\left(\bigcap_k E_k\right) \leq \prod_k \mathbb{P}(E_k).$$

Combined with the bound on $\mathbb{P}(E_k)$ above this completes the proof. □

Note that we did not use the condition that S is loopless; this was taken for simplicity. A more general version of Lemma 2 would of course give a more general version of Lemma 4.

Recalling that G_m^n is obtained by identifying the vertices of G_1^{mn} in groups of m, we can deduce the lower bound in Theorem 1 in the following more precise form.

Theorem 5. *Let $m \geq 1$ be fixed. Then $\operatorname{diam}(G_m^n) > \log n / \log(3Cm^2 \log n)$ for a.e. $G_m^n \in \mathcal{G}_m^n$, where C is the constant in Lemma 4.*

Proof. We shall prove slightly more, namely that with probability tending to 1 the distance $\rho(n, n-1)$ between the vertices n and $n-1$ of G_m^n is greater than $L = \log n / \log(3Cm^2 \log n)$, with C as above.

Consider a particular path $P = v_0 v_1 \cdots v_l$ on $[n]$ of length $l \leq L$. A graph S on $[mn]$ is a *realization* of P if S consists of edges $x_t y_{t+1}$, $t = 0, 1, \ldots, l-1$, with

$\lceil x_t/m \rceil = \lceil y_t/m \rceil = v_t$. As P and hence S has maximum degree at most two, Lemma 4 tells us that one such realization S is present in G_1^{mn} with probability at most

$$C^l \prod_{t=0}^{l-1} \frac{1}{\sqrt{x_t y_{t+1}}} \leq C^l \prod_{t=0}^{l-1} \frac{1}{\sqrt{v_t v_{t+1}}} = \frac{C^l}{\sqrt{v_0 v_l}} \prod_{t=1}^{l-1} \frac{1}{v_t}.$$

As there are exactly m^{2l} realizations of P, and $P \subset G_m^n$ if and only if at least one is present in G_1^{mn}, we have

$$\mathbb{P}(P \subset G_m^n) \leq \frac{(Cm^2)^l}{\sqrt{v_0 v_l}} \prod_{t=1}^{l-1} \frac{1}{v_t}.$$

Thus, if n is large enough, the expected number of paths of length l between the vertices n and $n-1$ of G_m^n is bounded by

$$\begin{aligned} \frac{(Cm^2)^l}{\sqrt{n(n-1)}} \sum_{1 \leq v_1, \ldots, v_{l-1} \leq n} \prod_{t=1}^{l-1} \frac{1}{v_t} &= \frac{(Cm^2)^l}{\sqrt{n(n-1)}} \left(\sum_{v=1}^{n} \frac{1}{v} \right)^{l-1} \\ &\leq \frac{(2Cm^2)^l}{n} (\log n)^{l-1} \\ &\leq (2/3)^l (\log n)^{-1}, \end{aligned}$$

as $l \leq L$ implies that $(3Cm^2 \log n)^l \leq n$. Summing over $1 \leq l \leq L$, we find that the expected number of paths of length at most L joining the vertices n and $n-1$ tends to zero, so $\rho(n, n-1) > L$ in a.e. G_m^n. □

This proves the lower bound in Theorem 1. In the following sections we prove the upper bound for $m \geq 2$.

5 The upper bound

As the proof is rather lengthy, we give an outline before turning to the details. The basic idea is to use the random pairing model from section 2. Considering the pairing as an LCD pairing of independent uniformly chosen random points in $[0, 1]$, we shall choose the right endpoint of each chord first, according to the appropriate (non-uniform) distribution. Given these right endpoints, the left endpoints are independent of one another, each being constrained to come before its corresponding right endpoint. (We may of course ignore the probability 0 event that two endpoints coincide.) We shall first show that the distribution of right endpoints is likely to be 'nice', i.e., to have certain uniformity properties. For the rest of the proof we then fix a 'nice' distribution of right endpoints and work with this.

The main idea of the proof is to start with a given vertex v, and consider the set $N_k(v)$ of vertices within distance k of v, aiming to show that this increases in size at a certain rate. Rather than size, it turns out that one should consider 'weight', assigning weight $i^{-1/2}$ to the i^{th} vertex. Essentially, given the weight $w(N_{k-1})$ of N_{k-1}, the expected value of $w(N_k)$ is $(\log n)w(N_{k-1})$, as long as this is not too large. Thus in $\Theta(\log n/\log\log n)$ steps from v one can reach many vertices. We complete the proof by showing that from these many vertices one can almost certainly reach the first vertex in one step. Unfortunately there are several complications.

One essential complication is that when $w(N_{k-1})$ is fairly small, although $w(N_k)$ has expectation $(\log n)w(N_{k-1})$, it has a very skew distribution, and is almost always much smaller than its mean. This means that we have to work with a 'blow-up factor' which starts off as a constant and increases as we proceed.

Another complication is that the main argument we use does not work starting from a single vertex near the end. To deal we this we show separately than all such vertices are likely to be joined by short paths to sufficiently early vertices.

Finally, the 'nice' distributions of right endpoints are not really all that nice; globally they behave exactly as we would like, but it is impossible to avoid some local variation. This complicates the details of the proof.

The main probabilistic tool we shall use in the rest of the paper is the following lemma given by Janson [17], which may be deduced from the Chernoff bounds [15].

Lemma 6. *Let $X = X_1 + \cdots + X_k$ where the X_i are independent Bernoulli random variables with $\mathbb{P}(X_i = 1) = p_i$. Let $\mu = \mathbb{E}\,X = p_1 + \cdots + p_k$. Then for $t \geq 0$ we have*

$$\mathbb{P}(X \geq \mu + t) \leq \exp\left(-\frac{t^2}{2(\mu + t/3)}\right)$$

and

$$\mathbb{P}(X \leq \mu - t) \leq \exp\left(-\frac{t^2}{2\mu}\right).$$

□

6 Pairings on $[0,1]$

We start by describing the method we shall use to generate the random graph G_m^n. Using the results of section 2, this is equivalent to generating a random pairing of the integers $\{1, 2, \ldots, 2N\}$, where $N = mn$.

Let $x_1, \ldots, x_{2N}$ be $2N$ independent samples from the uniform distribution on $[0,1]$. Assuming that these $2N$ points are distinct, our LCD (linearized chord diagram) is given by pairing x_{2i-1} with x_{2i} for each i. Relabelling the x_i as $1, 2, \ldots$ in ascending order, this gives a random pairing with the correct distribution. To see this note that, loosely speaking, for any set $\{x_1, \ldots, x_{2N}\}$

of $2N$ distinct elements of $[0,1]$, all $(2N)!$ possibilities for the order in which $x_1, \dots, x_{2N}$ take these values are equally likely.

We now consider generating a pairing starting with the right endpoints. We call a random variable with density function $2x$, $0 < x < 1$, an $M_2(0,1)$ *random variable.* Let $\{l_i, r_i\} = \{x_{2i-1}, x_{2i}\}$ with $l_i < r_i$. As $\mathbb{P}(\max\{x_{2i-1}, x_{2i}\} \leq t) = t^2$, the r_i are independent $M_2(0,1)$ random variables. Also, given $r_1, \dots, r_N$, the random variables $l_1, \dots, l_N$ are independent with l_i uniformly distributed on $[0, r_i]$. To obtain the pairing we must sort all the endpoints together. We shall first sort the right endpoints, and then consider between which right endpoint each left endpoint lies. Our first lemma concerns properties of the sorted right endpoints.

For the rest of this section set $N = mn$. Let $r_1, \dots, r_N$ be independent $M_2(0,1)$ random variables, and let $R_1, \dots, R_N$ be obtained by sorting $r_1, \dots, r_N$ into ascending order. When it comes to constructing the actual graph G_m^n, we shall only be interested in every m^{th} endpoint and the spacings between them, so for $1 \leq i \leq n$ let $W_i = R_{mi}$, and let $w_i = W_i - W_{i-1}$, taking $W_0 = 0$. We shall show that, in a certain range, the W_i have the values one might expect. For the w_i we can only say that this holds on average over certain intervals.

For the rest of the paper we write $s = 2^a$ for the smallest power of 2 larger than $(\log n)^7$, and 2^b for the largest power of 2 smaller than $2n/3$. Note that $a < b$ if n is large enough. We shall consider the intervals $I_t = [2^t + 1, 2^{t+1}]$, for $a \leq t < b$.

Lemma 7. *Let $m \geq 2$ be fixed. Using the definitions above, each of the following five events has probability tending to 1 as $n \to \infty$:*

$$E_1 = \left\{ \left| W_i - \sqrt{\frac{i}{n}} \right| \leq \frac{1}{10}\sqrt{\frac{i}{n}} \text{ for } s \leq i \leq n \right\},$$

$$E_2 = \left\{ I_t \text{ contains at least } 2^{t-1} \text{ vertices } i \text{ with } w_i \geq \frac{1}{10\sqrt{in}},\ a \leq t < b \right\},$$

$$E_3 = \left\{ w_1 \geq \frac{4}{\log n \sqrt{n}} \right\},$$

$$E_4 = \left\{ w_i \geq \frac{(\log n)^2}{n} \text{ for } i < n^{1/5} \right\},$$

$$E_5 = \{ w_i \leq n^{-4/5} \text{ for } i > n/(\log n)^5 \}.$$

The events E_1 and E_2 describe the main properties of the W_i and w_i we shall use. The remaining properties are simple technicalities we shall need in section 7. In the following proof, and in the rest of the paper, we shall always assume that n is larger than some sufficiently large constant, even when this is not explicitly stated.

Proof of Lemma 7. Let us write $R(x)$ for $|\{i : r_i \le x\}| = |\{i : R_i \le x\}|$. For $0 \le x \le 1$ we have $R(x) \sim \mathrm{Bi}(mn, x^2)$. As $W_i = R_{mi}$, we have that $W_i \le x$ if and only if $R(x) \ge mi$.

Let $x = \frac{9}{10}\sqrt{i/n}$. Then as $\mathbb{E}\, R(x) = mnx^2 = 81mi/100$, Lemma 6 implies that

$$\mathbb{P}(W_i \le x) = \mathbb{P}(R(x) \ge mi) \le e^{-mi/60}.$$

Similarly, with $x = \frac{11}{10}\sqrt{i/n}$ either $x \ge 1$ or $\mathbb{E}\, R(x) = 121mi/100$, and in either case

$$\mathbb{P}(W_i > x) = \mathbb{P}(R(x) < mi) \le e^{-mi/60}.$$

Thus

$$\mathbb{P}(E_1^c) \le 2\sum_{i=s}^{\infty} e^{-i/60} = o(1),$$

since $s \to \infty$ as $n \to \infty$.

It now suffices to show that $\mathbb{P}(E_r^c \cap E_1) = o(1)$ for $r = 2, \dots, 5$.

Given R_{mi} and the set $\mathcal{J}$ of indices j for which $r_j > R_{mi}$, each of the $m(n-i)$ variables r_j, $j \in \mathcal{J}$, is independently distributed with density $2x/(1 - R_{mi}^2)$, $R_{mi} < x < 1$. For $R_{m(i+1)} = R_{mi+m}$ to be less than $R_{mi} + y$ it must happen that at least m of these r_j fall in the interval $[R_{mi}, R_{mi} + y]$. Conditional on R_{mi}, the expected number of r_j in this interval is exactly

$$m(n-i)\frac{(R_{mi} + y)^2 - R_{mi}^2}{1 - R_{mi}^2}.$$

Thus

$$\mathbb{P}(R_{m(i+1)} \le R_{mi} + y \mid R_{mi}) \le (n-i)\frac{(R_{mi} + y)^2 - R_{mi}^2}{1 - R_{mi}^2}.$$

Translating this to a statement about the W_i, and replacing i by $i - 1$ for convenience later,

$$\mathbb{P}(w_i \le y \mid W_{i-1}) \le (n - i + 1)\frac{(W_{i-1} + y)^2 - W_{i-1}^2}{1 - W_{i-1}^2}. \tag{7}$$

For $s + 1 \le i \le 2n/3$, provided $W_{i-1} \le \frac{11}{10}\sqrt{\frac{i-1}{n}}$ this probability is at most $2n(2.2y\sqrt{i/n} + y^2)$. Thus, for i in this range,

$$\mathbb{P}(w_i < \frac{1}{10\sqrt{in}} \mid W_{i-1}) \le \frac{9}{20}$$

when $W_{i-1} \le \frac{11}{10}\sqrt{\frac{i-1}{n}}$. As W_{i-1} satisfies this condition when E_1 holds, writing B_i for the 'bad' event that $w_i < \frac{1}{10\sqrt{in}}$ we have

$$\mathbb{P}(B_i \cap E_1 \mid W_{i-1}) \le \frac{9}{20}$$

on the whole probability space.

Now, given W_s and $w_{s+1}, \ldots, w_{i-1}$, the distribution of w_i depends only on $W_{i-1} = W_s + w_{s+1} + \cdots + w_{i-1}$. Thus, if for $i = s+1, \ldots, 2^b$ we examine the w_i in turn, at each stage the conditional probability of $B_i \cap E_1$ is at most $9/20$. Restricting our attention to $i \in I_t$ for some $a \le t < b$, it follows that the quantity

$$Z_t = \left| \left\{ i \in I_t : w_i < \frac{1}{10\sqrt{in}} \right\} \right| \times \mathbb{I}_{E_1}$$

is stochastically dominated by a $\mathrm{Bi}(2^t, 9/20)$ distribution, and so from Lemma 6

$$\mathbb{P}(Z_t > 2^{t-1}) \le \exp(-2^t/400).$$

As $E_2^c \cap E_1$ is exactly the event that for some $a \le t < b$ we have $Z_t > 2^{t-1}$,

$$\mathbb{P}(E_2^c \cap E_1) \le \sum_{t=a}^{b-1} \exp(-2^t/400) = o(1),$$

as $a \to \infty$.

Turning to E_3, note that $w_1 = W_1$. Taking W_0 to be identically zero, the argument giving (7) applies with $i = 1$ to give $\mathbb{P}(w_1 \le x) \le nx^2$ and hence $\mathbb{P}(E_3) = 1 - o(1)$.

Suppose that $E_4^c \cap E_1$ holds, and let $\delta = (\log n)^2/n$. Then as E_1 holds we have $W_{\lfloor n^{1/5} \rfloor} \le 2n^{-2/5}$. As E_4 does not hold it follows that some interval $[x, x+\delta]$ with $0 \le x < 2n^{-2/5}$ contains two of the W_i, and hence two of the r_i. Each such interval is contained in one of the intervals $J_t = [t\delta, (t+2)\delta]$, $0 \le t \le n^{3/5}$. The probability that a particular r_i lies in one such J_t is exactly $(4t+4)\delta^2$, so the probability that at least two r_i lie in J_t is at most $mn(mn-1)(4t+4)^2\delta^4/2 < (\log n)^9 (t+1)^2/n^2$. Thus

$$\mathbb{P}(E_4^c \cap E_1) \le \sum_{t=0}^{n^{3/5}} (\log n)^9 (t+1)^2/n^2 = o(1),$$

so $\mathbb{P}(E_4) = 1 - o(1)$.

Finally, suppose that $E_5^c \cap E_1$ holds. This time let $\delta = n^{-4/5}$. Then for some x, $(\log n)^{-3} < x < 1 - \delta$, the interval $[x, x+\delta]$ contains no W_i, and hence contains at most $m-1$ of the r_i. Setting $\delta' = \delta/(m+1)$, each such interval contains m disjoint intervals of the form $[t\delta', (t+1)\delta')$ with t an integer and $(\log n)^{-3} < t\delta' < 1 - \delta'$, one of which must contain no r_i. For a given t, the number of r_i in $[t\delta', (t+1)\delta')$ has a $\mathrm{Bi}(mn, p_t)$ distribution with

$$p_t = (2t+1)\delta'^2 > (\log n)^{-3}\delta' > n^{-5/6}.$$

The probability that no r_i lies in this interval is thus

$$(1 - p_t)^{mn} \le e^{-mnp_t} < e^{-n^{1/6}} = o(n^{-1}).$$

Summing over the $O(n^{1/5})$ values of t shows that $\mathbb{P}(E_5^c \cap E_1) = o(1)$, and hence that $\mathbb{P}(E_5) = 1 - o(1)$, completing the proof of the lemma. $\square$

From now on we shall assume that the right endpoints $R_1, \ldots, R_{mn}$ are given, writing W_i for R_{mi}. The remaining randomness is given by taking independent random variables L_i, $1 \leq i \leq mn$, uniformly distributed on $[0, R_i]$, to obtain an LCD given by chords $\{L_i, R_i\}$ already sorted by their right endpoints. Recall that the graph G_1^{mn} is constructed from this LCD by taking an edge from each vertex i to a vertex k, where $R_{k-1} < L_i < R_k$ (taking $R_0 = 0$). Thus the graph G_m^n we shall study is given by taking m edges from each vertex i, $1 \leq i \leq n$, joining i to $l_{i,j}$, $1 \leq j \leq m$, where $l_{i,j} = k$ if $W_{k-1} < L_{m(i-1)+j} < W_k$.

Speaking somewhat imprecisely, we would like to forget the R_i and consider only the W_i. The only problem is that the value of $l_{i,j}$ is determined by a variable $L_{m(i-1)+j}$ distributed uniformly on $[0, R_{m(i-1)+j}]$. However, as $W_{i-1} < R_{m(i-1)+j} \leq W_i$, if we take $L_{m(i-1)+j}$ uniformly distributed on $[0, W_i]$ we increase the probability that the diameter is large, by increasing the probability of loops at i. After making this change we can essentially restrict our attention to the case $m = 2$. More precisely, we shall only consider two edges from each vertex i, to the vertices $l_{i,1}$ and $l_{i,2}$. This change just involves deleting edges from the graph, which again can only increase the diameter.

Fix $m \geq 2$. A precise description of the model we shall use from now on is as follows. Suppose that n distinct real numbers $0 < W_1 < \cdots < W_n < 1$ are given, and let $w_i = W_i - W_{i-1}$, taking $W_0 = 0$. We construct a random graph $G(W_1, \cdots, W_n)$ by taking two edges from each vertex i, joining i to $l_{i,1}$ and $l_{i,2}$, where the $l_{i,j}$ are independent with $\mathbb{P}(l_{i,j} = k) = w_k/W_i$ for $k \leq i$. Suppose now that $R_1, \ldots, R_{mn}$ are obtained by sorting mn independent $M_2(0,1)$ random variables, and the random graph $G = G(W_1, \cdots, W_n)$, $W_i = R_{mi}$, is then constructed. For $m = 2$ the random graphs G and G_2^n have almost the same distribution. More precisely and more generally, for any fixed $m \geq 2$ the random graph G can be coupled with G_m^n so that G may be obtained from G_m^n by deleting some edges and adding some loops.

Let $m \geq 2$ and $\epsilon > 0$ be fixed. For the rest of the paper we concentrate on the W_i rather than the R_i. We shall assume the W_i are given and have the properties $E_1, \ldots, E_5$ above, and form a random graph $G = G(W_1, \ldots, W_n)$ as above, writing $\mathbb{P}_{\mathrm{L}}$ for the corresponding probability measure. To simplify the notation in the following results and proofs, the implicit bound in each occurrence of $o(.)$ should be taken as a function of n only, not depending on the W_i or any other parameters, tending to 0 as $n \to \infty$. (Of course these implicit functions depend on our fixed parameters m and ϵ; they are functions of n, m and ϵ which tend to 0 as $n \to \infty$ with m and ϵ fixed.) In this light, the results are most simply seen as proving an (implicit) bound on some probability that holds for *every sufficiently large* n, and *every sequence* $0 < W_1 < \cdots < W_n < 1$ satisfying the stated conditions. With these conventions we shall show that

$$\mathbb{P}_{\mathrm{L}}(\mathrm{diam}(G) \leq (1+\epsilon)\log n/\log\log n) = 1 - o(1)$$

whenever the W_i satisfy $E_1, \ldots, E_5$. Together with Lemma 7 this is sufficient to prove the upper bound in Theorem 1, namely that for $m \geq 2$ a.e. $G_m^n \in \mathcal{G}_m^n$ has $\mathrm{diam}(G_m^n) \leq (1+\epsilon)\log n/\log\log n$.

7 Getting started

We shall say that a vertex i is *useful* if $i \le n/(\log n)^5$ and $w_i \ge (\log n)^2/n$. As we are assuming $E_1, \ldots, E_5$, if n is large enough then from E_4 every i with $i < n^{1/5}$ is useful; we shall use this fact several times. Our aim in this section is to prove that the following lemma holds for all sufficiently large n and all $0 < W_1 < \cdots < W_n < 1$ satisfying the stated conditions.

Lemma 8. *Assuming $E_1, \ldots, E_5$, with $\mathbb{P}_{\mathrm{L}}$ probability $1 - o(1)$ every vertex v of G is joined by a (descending) path of length at most $8 \log\log n$ to a useful vertex.*

Lemma 8 will be proved in two stages, each of which will be given as a separate lemma. For this section it will be convenient to think of G as a directed graph, directing each edge ij with $i > j$ from i to j.

Let v be fixed, and consider the vertices reached from v after k steps. More precisely, let $D_0 = \{v\}$, and for $k \ge 1$ let D_k be the set of those vertices outside $D_0 \cup \cdots \cup D_{k-1}$ incident with an edge directed from some vertex of D_{k-1}. Our first lemma shows that $|D_k|$ is likely to be not much smaller than $2|D_{k-1}|$, as long as neither D_{k-1} nor D_k contains any useful vertices.

Let n_k be $|D_k|$ if $D_0 \cup \cdots \cup D_k$ contains no useful vertices, and ∞ otherwise.

Lemma 9. *For $1 \le v \le n$, $1 \le k \le 8 \log\log n$ and $1 \le c \le |D_{k-1}|$ we have*

$$\mathbb{P}_{\mathrm{L}}(n_k \le 2n_{k-1} - c \mid D_0, D_1, \ldots, D_{k-1}) \le n^{-3c/5+o(1)}.$$

Proof. Let $t = |D_{k-1}|$, and let $i_1, \ldots, i_t$ be the vertices of D_{k-1} in some order. We may assume that no vertex of $D_0 \cup \cdots \cup D_{k-1}$ is useful as otherwise $n_{k-1} = \infty$ and there is nothing to prove. Thus $n_{k-1} = t$, and for each s we have $i_s \ge n^{1/5}$.

Suppose that i_s sends its first and second edges to j_{2s-1} and j_{2s}. Then $j_1, \ldots, j_{2t}$ are independent, and for $1 \le s \le 2t$ and $1 \le j' \le i_{\lceil s/2 \rceil}$ we have

$$\mathbb{P}_{\mathrm{L}}(j_s = j') = w_{j'}/W_{i_{\lceil s/2 \rceil}}. \tag{8}$$

Now $|D_k|$ is the number of distinct j_s which lie outside $D_0 \cup \cdots \cup D_{k-1}$, a set of at most $2^k - 1$ vertices none of which is useful. Considering the j_s in order, we say that a possible value of j_s is *repetitive* if it coincides with a previous $j_{s'}$ which is not useful, or lies in $D_0 \cup \cdots \cup D_{k-1}$. Recalling that $n_k = \infty$ if any vertex in D_k is useful, we see that n_k is at least $2n_{k-1}$ minus the number of s for which a repetitive choice is made. For a fixed s the number of possible repetitive choices is at most 2^{k+1}. Consider one such choice j'. As j' is not useful, either $w_{j'} < (\log n)^2/n$ or $j' > n/(\log n)^5$. In the first case, using E_1,

$$W_{i_{\lceil s/2 \rceil}} \ge W_{\lceil n^{1/5} \rceil} \ge \sqrt{n^{1/5}/n}/2 = n^{-2/5}/2,$$

so

$$\mathbb{P}_{\mathrm{L}}(j_s = j') = w_{j'}/W_{i_{\lceil s/2 \rceil}} \le n^{-3/5+o(1)}.$$

In the second case we have $w_{j'} \le n^{-4/5}$ as E_5 holds. Also, as $i_{\lceil s/2\rceil} \ge j' \ge n^{4/5}$ we have $W_{i_{\lceil s/2\rceil}} \ge n^{-1/10}/2$, and again $\mathbb{P}_{\mathrm{L}}(j_s = j') \le n^{-3/5+o(1)}$.

Thus at each stage the probability of making a repetitive choice is at most $2^{k+1}n^{-3/5+o(1)}$. Since there are only $2|D_{k-1}| < 2^k = n^{o(1)}$ stages, the lemma follows. $\square$

We shall also need the following simple lemma showing that from any non-useful vertex we have a reasonable chance of hitting a useful vertex. As usual we are assuming that n is sufficiently large.

Lemma 10. *Assume $E_1, \ldots, E_5$ hold and let i be a fixed non-useful vertex. Then the probability that $l_{i,1}$ is useful is at least $(\log n)^{-3}$.*

Proof. Let i be a fixed non-useful vertex. Note that we must have $i \ge n^{1/5}$ as E_4 holds.

Let $2^{b'}$ be the largest power of 2 smaller than $\min\{i, n/(\log n)^5\}$. Consider the intervals $I_t = [2^t+1, 2^{t+1}]$ for $a \le t < b'$, recalling that $s = 2^a$ is the smallest power of 2 larger than $(\log n)^7$. The probability that $l_{i,1}$ lies in the union of these I_t is exactly

$$\rho = (W_{2^{b'}} - W_{2^a})/W_i.$$

As E_1 holds,

$$W_{2^a} \le 2\sqrt{s/n} = n^{-1/2+o(1)}, \tag{9}$$

while

$$W_i \ge W_{\lceil n^{1/5}\rceil} \ge n^{-2/5}/2. \tag{10}$$

Also, either $i \le n/(\log n)^5$, or $i > n/(\log n)^5$. In the first case $2^{b'} \ge i/2$, so, using E_1 twice, $W_{2^{b'}}/W_i \ge 1/8$. In the second case $2^{b'} \ge n/(\log n)^5/2$, so $W_{2^{b'}} \ge (\log n)^{-5/2}/2 \ge 100(\log n)^{-3}$. As we always have $W_i \le 1$, in either case $W_{2^{b'}}/W_i \ge 100(\log n)^{-3}$. Combined with (9) and (10) this shows that $\rho \ge 99(\log n)^{-3}$.

Let us say that a vertex i is *good* if $w_i \ge 1/(10\sqrt{in})$. Then E_2 says exactly that for $a \le t < b$ each interval I_t contains at least 2^{t-1} good vertices. As each good vertex i in I_t has weight $w_i \ge 1/(10\sqrt{2^{t+1}n})$, the total weight of good vertices within I_t is at least $2^{t-1}/(10\sqrt{2^{t+1}n}) \ge \sqrt{2^t/n}/30$. On the other hand, the total weight of I_t is $W_{2^{t+1}} - W_{2^t} \le W_{2^{t+1}} \le 3\sqrt{2^t/n}$. For $a \le t \le b'$, the whole of I_t lies to the left of i, and the probability that $l_{i,1}$ is good given that $l_{i,1} \in I_t$ is thus at least $1/90$. As $2^{b'} \le n/(\log n)^5$, any good vertex of I_t is also useful. This shows that the probability that $l_{i,1}$ is useful is at least $\rho/90$, from which the lemma follows. $\square$

Using the two lemmas above, it is now easy to prove Lemma 8.

Proof of Lemma 8. Fix a vertex v, and let $K = \lfloor 8 \log\log n \rfloor - 1$. From Lemma 9, with probability at least $1 - n^{-6/5+o(1)} = 1 - o(n^{-1})$ either v is within distance K of a useful vertex, or we have $|D_k| = 2|D_{k-1}|$ for all but at most one value of k, $1 \le k \le K$, and $|D_k| \ge 2|D_{k-1}| - 1$ for all k, $1 \le k \le K$. Suppose that the second case holds. Then $|D_K| \ge 2^{K-1} \ge 2(\log n)^4$, provided n is sufficiently large. Given the sequence $D_0, \ldots, D_K$, the vertices $l_{i,1}$, $i \in D_K$, are independent with their original distribution. If D_K does not contain a useful vertex then, from Lemma 10, the probability that no i in D_K has a useful neighbour is at most

$$(1 - (\log n)^{-3})^{|D_K|} \le \exp(-|D_K|/(\log n)^3) \le n^{-2}.$$

Thus with probability at least $1 - o(n^{-1})$ the given vertex v is within distance $8 \log\log n$ of a useful vertex. As there are only n vertices v to consider, Lemma 8 follows. $\square$

8 Expanding neighbourhoods

The lemma below gives the heart of the argument. As before, $o(.)$ notation refers to an implicit function depending on n and ϵ only. The statement of the lemma then applies for all sufficiently large n and all $0 < W_1 < \cdots < W_n < 1$ satisfying the stated conditions.

Lemma 11. *Assume that $E_1, \ldots, E_5$ hold, let $\epsilon > 0$ be fixed, and let v, $1 \le v \le n$, be a useful vertex. With $\mathbb{P}_{\mathrm{L}}$ probability $1 - o(n^{-1})$ there is a path in G between v and 1 of length at most $(1/2 + \epsilon) \log n / \log\log n$.*

The basic strategy of the proof is to divide $[s+1, n]$ into intervals of the form $I_t = [2^t + 1, 2^{t+1}]$. Given the set of vertices Γ_k at distance k from v, we consider the expected size of $\Gamma_{k+1} \cap I_t$. Whenever this is bigger than $\log n$ we will have enough independence to show that the actual size of $\Gamma_{k+1} \cap I_t$ will be close to its expected value. This will show that the weight $f_k = \sum_{i \in \Gamma_k} 1/\sqrt{in}$ almost certainly grows at a certain rate, depending on its current value. In fact we only consider the intervals I_t for $a \le t < b$ where, as before, $s = 2^a$ is the smallest power of 2 larger than $(\log n)^7$, and 2^b is the largest power of 2 smaller than $2n/3$. Thus $a \sim 7 \log_2 \log n$ and $b \sim \log_2 n$.

As the proof of Lemma 11 is a little complicated, we state part of the estimation as separate lemma. What this lemma claims is that if the f_k satisfy a certain condition then they grow fast enough.

Lemma 12. *Suppose that $\epsilon > 0$, and set $K = (1/2 + \epsilon) \log n / \log\log n - 1$. Let $f_0, f_1, \ldots$ be a sequence of real numbers with $f_0 \ge (\log n)^2/n$ and*

$$f_{k+1} \ge \min\{2 \log_2(f_k n / \log n) - 31, b - a - 1\} f_k / 1000, \tag{11}$$

for $k \ge 0$. Then, provided n is sufficiently large, $\ell = \min\{k : f_k \ge (\log n)^2/\sqrt{n}\}$ exists and is at most K.

Proof. Provided n is sufficiently large we have $\log_2(f_0 n/\log n) \geq \log_2(\log n) \geq 1016$. Thus (11) implies that $f_1 \geq 2f_0$. It follows inductively that $f_{k+1} \geq 2f_k$ and $f_k \geq 2^k f_0$ hold for all $k \geq 0$. This shows that ℓ exists.

As $b-a-1 \geq \log_2 n - 8\log_2 \log n$, the minimum in (11) is different from the first term only when $f_k \geq (\log n)^{-3}/\sqrt{n}$. If this first happens at $k = k_0$, say, then for $k \geq k_0$ we have

$$f_{k+1} \geq (b-a-1)f_k/1000 = (\log_2 n)^{1+o(1)} f_k,$$

which implies that $\ell \leq k_0 + 6$. Thus

$$f_{k+1} \geq \frac{\log_2(f_k n/\log n) - 16}{500} f_k \tag{12}$$

for $0 \leq k < \ell - 6$. As $f_k \geq 2^k f_0$ and $\log_2(f_0 n/\log n) \geq 17$ (if n is large enough), we have $\log_2(f_k n/\log n) \geq k + 17$. Combined with (12) this implies that $f_{k+1} \geq f_k(k+1)/500$ for $0 \leq k < \ell - 6$, and hence that

$$f_{\ell-6} \geq \frac{(\ell-6)!}{500^{\ell-6}} f_0 \geq \left(\frac{\ell-6}{500e}\right)^{\ell-6} f_0,$$

using Stirling's formula. As $f_{\ell-6} < (\log n)^2/\sqrt{n} \leq \sqrt{n} f_0$ this implies that $\ell - 6 \leq (1/2 + \epsilon/2)\log n/\log\log n < K - 6$, and the lemma follows. □

With this straightforward but slightly messy calculation behind us, we turn to the proof of Lemma 11.

Proof of Lemma 11. Roughly speaking, we consider the set of vertices within distance k of v, showing that the weight of this set increases rapidly with k. In fact, to keep independence, when working outwards from v we only allow ourselves to use an edge ij from a vertex i at distance k to a new vertex j if $i > j$ and ij is the first edge from i or if $i < j$ and ij is the second edge from j. As before we say that a vertex i is *good* if

$$w_i \geq \frac{1}{10\sqrt{in}}.$$

Let $\Gamma_0 = \{v\}$, and for $k \geq 1$ let Γ_k consist of those j in $[s+1, 2^b] \setminus (\Gamma_0 \cup \cdots \cup \Gamma_{k-1})$ such that j is good and either $l_{j,2} \in \Gamma_{k-1}$ or there is an $i \in \Gamma_{k-1}$ with $l_{i,1} = j$. We shall write N_k for $\Gamma_0 \cup \cdots \cup \Gamma_k$.

Rather than the actual weight (sum of the w_i) we consider the 'ideal weight' of Γ_k, given by

$$f_k = \sum_{i \in \Gamma_k} \frac{1}{\sqrt{in}}.$$

Note that for $k \geq 1$ the set Γ_k contains only good vertices, so its actual weight is at least $f_k/10$. For the sake of simplicity we shall assume that v itself is good, as well as useful. The case that v is useful but not good can be dealt with by redefining f_0 to be $(\log n)^2/n$. It is easy to check that the first step in the argument that follows requires only that $1/\sqrt{vn} \geq f_0$ and that $w_v \geq f_0/10$. At the end of the argument we only use $f_0 \geq (\log n)^2/n$, and nowhere else is the goodness of v relevant.

Given $\Gamma_0, \ldots, \Gamma_k$, for each $i \in \Gamma_k$ the random variable $l_{i,1}$ has its original unconditioned distribution, since we have 'looked at' $l_{i,2}$ but not $l_{i,1}$. Similarly, for $i \notin N_k$ all we know about $l_{i,2}$ is that $l_{i,2} \notin N_{k-1}$, and the variables $\{l_{i,1} : i \in \Gamma_k\} \cup \{l_{i,2} : i \notin N_k\}$ are independent.

Let us fix a t with $a \leq t < b$, and consider the interval $I_t = [2^t + 1, 2^{t+1}]$. Let X be the number of vertices of $N_k \cap I_t$ (these vertices are 'excluded' for our present purpose), and let S be the set of good vertices of $I_t \setminus N_k$. Now E_2 states exactly that I_t contains at least 2^{t-1} good vertices, so $|S| \geq 2^{t-1} - X$. Let $i_1, \ldots, i_d$ be the vertices of $\Gamma_k \cap [2^{t+1} + 1, n]$, listed in any order. We consider examining each i in turn to see whether $l_{i,1}$ lies in S. More precisely, let C_r be the number of distinct elements of S equal to $l_{i,1}$ for $i \in \{i_1, \ldots, i_r\}$, so $|\Gamma_{k+1} \cap I_t| \geq C_d$. Conditioning on everything so far, i.e., on $\Gamma_0, \ldots, \Gamma_k$ and $l_{i_1,1}, \ldots, l_{i_{r-1},1}$, provided $C_{r-1} \leq 2^{t-2} - X$ the probability P_r that $C_r = C_{r-1} + 1$ satisfies

$$P_r \geq 2^{t-2} \frac{1}{10\sqrt{2^{t+1}n}} W_{i_r}^{-1},$$

as there are at least $|S| - C_{r-1} \geq 2^{t-2}$ vertices j of S unused, each of weight $w_j \geq \frac{1}{10\sqrt{jn}} \geq \frac{1}{10\sqrt{2^{t+1}n}}$. Using $W_i \leq \sqrt{2i/n}$ for $i \geq s$,

$$P_r \geq \frac{2^{t-2}}{20\sqrt{2^t n}} \sqrt{\frac{n}{i_r}} = \frac{\sqrt{2^t}}{80\sqrt{i_r}} = p_r, \tag{13}$$

say. Let Y be a random variable given by the sum of d independent Bernoulli random variables $\mathrm{Bi}(1, p_r)$ with means $p_1, \ldots, p_d$. Then C_d stochastically dominates $\min\{Y, 2^{t-2} - X\}$. (Roughly speaking, (13) implies by induction on r that $C_r \geq \sum_{r'=1}^{r} \mathrm{Bi}(1, p_{r'})$ until $C_r = 2^{t-2} - X$.)

Let

$$\mu_1 = \sum_{i \in \Gamma_k,\, i > 2^{t+1}} \frac{\sqrt{2^t}}{80\sqrt{i}} = \sum_{r=1}^{d} p_r.$$

Provided $\mu_1 \geq 10 \log n$ we have from Lemma 6 that $\mathbb{P}(Y \leq \mu_1/2) \leq n^{-5/4}$, and hence that

$$\mathbb{P}_{\mathrm{L}}(C_d \leq \min\{\mu_1/2, 2^{t-2} - X\} \mid \Gamma_0, \ldots, \Gamma_k) \leq n^{-5/4}. \tag{14}$$

In other words, if μ_1 is large enough, then $\mu_1/2$ vertices of I_t are likely to be hit by edges coming out of $\Gamma_k \cap [2^{t+1} + 1, n]$, unless this would mean that more than 2^{t-2} vertices of I_t would be contained in N_{k+1}.

We now consider edges from I_t into $\Gamma_k \cap [s + 1, 2^t]$.

Let $X' = X + C_d$ and let S' be the set of good vertices of $I_t \setminus N_k$ not counted in C_d, so $|S'| \geq 2^{t-1} - X'$. We consider the number C' of $j \in S'$ for which $l_{j,2}$ lies in Γ_k, noting that $|\Gamma_{k+1} \cap I_t| \geq C'$. Given $\Gamma_0, \ldots, \Gamma_k$, for each $j \in S'$ the random variable $l_{j,2}$ has its original distribution conditioned on $l_{j,2} \notin N_{k-1}$. Thus, as Γ_k is disjoint from N_{k-1},

$$\begin{aligned}
\mathbb{P}(l_{j,2} \in \Gamma_k) &\geq \sum_{i \in \Gamma_k,\, i \leq j} w_i / W_j \\
&\geq \sum_{i \in \Gamma_k,\, i \leq 2^t} w_i / W_{2^{t+1}} \\
&\geq \sum_{i \in \Gamma_k,\, i \leq 2^t} \frac{1}{10\sqrt{in}} \sqrt{\frac{n}{2^{t+2}}} \\
&= \sum_{i \in \Gamma_k,\, i \leq 2^t} \frac{1}{20\sqrt{2^t i}} = p,
\end{aligned}$$

say. As the $l_{j,2}$ are independent, the number C' of $j \in S'$ for which $l_{j,2}$ is in Γ_k stochastically dominates a $\mathrm{Bi}(2^{t-1} - X', p)$ distribution. Let $\mu_2 = 2^{t-2} p$. Then provided $\mu_2 \geq 10 \log n$ we have from Lemma 6 that

$$\mathbb{P}_{\mathrm{L}}(C' \leq \mu_2/2 \mid \Gamma_0, \ldots, \Gamma_k, X' < 2^{t-2}) \leq n^{-5/4}. \tag{15}$$

In other words, if $\mu_2 \geq 10 \log n$ and we have not yet (in N_k together with the part of Γ_{k+1} we found above) reached 2^{t-2} vertices of I_t, then it is likely that $\mu_2/2$ new vertices of I_t can be reached by following back edges from $\Gamma_k \cap [s+1, 2^t]$.

We claim that

$$\mathbb{P}_{\mathrm{L}}\left(|\Gamma_{k+1} \cap I_t| \leq \min\left\{ \frac{\mu_1 + \mu_2}{4}, 2^{t-2} - |N_k \cap I_t| \right\} \mid \Gamma_0, \ldots, \Gamma_k \right) \leq n^{-5/4} \tag{16}$$

holds whenever $\mu_1 + \mu_2 \geq 20 \log n$. To see this, suppose first that $\mu_1 \geq \mu_2$. Then $\mu_1 \geq (\mu_1 + \mu_2)/2$ and $\mu_1 \geq 10 \log n$. As $X = |N_k \cap I_t|$ by definition, while $|\Gamma_{k+1} \cap I_t| \geq C_d$, in this case (16) follows from (14).

Suppose instead that $\mu_2 > \mu_1$. Then $\mu_2 \geq (\mu_1 + \mu_2)/2$ and $\mu_2 \geq 10 \log n$. Thus (15) implies that

$$\mathbb{P}_{\mathrm{L}}(C' \leq (\mu_1 + \mu_2)/4 \mid \Gamma_0, \ldots, \Gamma_k, X' < 2^{t-2}) \leq n^{-5/4}. \tag{17}$$

Now $|\Gamma_{k+1} \cap I_t| \geq C'$. Also, $X' \leq |\Gamma_{k+1} \cap I_t| + |N_k \cap I_t|$. Thus if $X' \geq 2^{t-2}$ then $|\Gamma_{k+1} \cap I_t| \geq 2^{t-2} - |N_k \cap I_t|$. Thus (17) implies (16) in this case also, showing that (16) holds whenever $\mu_1 + \mu_2 \geq 20 \log n$.

We shall now vary t in the range $a \leq t < b$ to obtain a lower bound on the weight of Γ_{k+1}. Note that

$$\mu_1 + \mu_2 = \mu_1(t) + \mu_2(t) = \frac{\sqrt{2^t}}{80} \sum_{i \in \Gamma_k \setminus I_t} \frac{1}{\sqrt{i}}.$$

Recalling that

$$f_k = \frac{1}{\sqrt{n}} \sum_{i\in\Gamma_k} \frac{1}{\sqrt{i}},$$

it follows that for all but at most one value of t we have

$$\mu_1 + \mu_2 \geq \frac{\sqrt{2^t n}}{160} f_k,$$

and the proof is nearly complete.

Given $\Gamma_0, \ldots, \Gamma_k$, let $\mathcal{T}$ be the set of indices t, $a \leq t < b$, for which $\sqrt{2^t n} f_k/160 \geq 20 \log n$. As $2^b \geq n/3$, considering the smallest element of $\mathcal{T}$ shows that

$$|\mathcal{T}| \geq \min\{2\log_2(f_k n/\log n) - 30, b-a\}.$$

Let us write $\mathcal{T}'$ for the set of $t \in \mathcal{T}$ for which the quantity $\mu_1 + \mu_2$ defined above (which depends on t) is at least $\sqrt{2^t n} f_k/160$, so $|\mathcal{T}'| \geq |\mathcal{T}| - 1$. We shall say that an interval I_t is *full* at stage $k+1$ if $|N_{k+1} \cap I_t| \geq 2^{t-2}$. From (16), for each $t \in \mathcal{T}'$ with probability $1 - O(n^{-5/4})$ either I_t is full in N_{k+1} or $|\Gamma_{k+1} \cap I_t| \geq \sqrt{2^t n} f_k/640$. In the latter case we have

$$\sum_{i\in\Gamma_{k+1}\cap I_t} \frac{1}{\sqrt{in}} \geq \frac{\sqrt{2^t n} f_k}{640\sqrt{2^{t+1} n}} \geq \frac{f_k}{1000}.$$

Thus, with probability $1 - O((\log n) n^{-5/4}) = 1 - O(n^{-6/5})$ either some I_t is full at stage $k+1$, or

$$f_{k+1} \geq \frac{|\mathcal{T}'| f_k}{1000} \geq \frac{\min\{2\log_2(f_k n/\log n) - 31, b-a-1\}}{1000} f_k. \tag{18}$$

Starting with $\Gamma_0 = \{v\}$ let us construct $\Gamma_1, \Gamma_2, \ldots$, and let K be the minimum k for which either $f_k \geq (\log n)^2/\sqrt{n}$ or (18) does not hold. As v is useful we have $f_0 = w_v \geq (\log n)^2/n$. As (18) holds for $0 \leq k < K$, Lemma 12 implies that $K \leq (1/2 + \epsilon)\log n/\log\log n - 1$.

We claim that

$$f_K \geq (\log n)^2/\sqrt{n} \tag{19}$$

holds with probability $1 - o(n^{-1})$. The only possible problem is that some I_t may become full at some stage.

Suppose first that some I_t, $a \leq t < b$, is full at some stage k, $1 \leq k \leq K$. Then (leaving out Γ_0 as v is not necessarily good), in $N_K \setminus \{v\}$ we have at least $2^{t-2} - 1$ good vertices in I_t. This gives

$$f_1 + \cdots + f_K \geq \frac{2^{t-2} - 1}{10\sqrt{2^{t+1} n}} \geq \frac{\sqrt{2^t}}{160\sqrt{n}} \geq \frac{\sqrt{2^a}}{160\sqrt{n}} = \frac{\sqrt{s}}{160\sqrt{n}} > \frac{(\log n)^3}{\sqrt{n}}.$$

As $K < \log n$ and $f_k < (\log n)^2/\sqrt{n}$ for $k < K$, this implies (19). On the other hand, if no I_t becomes full at any stage, then (18) holds with probability at least $1 - O(n^{-6/5})$ at each of $K \leq \log n$ stages, so (19) holds with probability $1 - o(n^{-1})$, as claimed.

If (19) holds we stop the construction of the sequence $\Gamma_1, \Gamma_2, \dots$ at Γ_K. Given the sequence so far, the $l_{i,1}$, $i \in \Gamma_K$, are independent. As $K > 0$ we have $\Gamma_K \subset [s+1, n]$ by construction. Thus, since E_1 and E_3 hold, for each $i \in \Gamma_K$ we have

$$\mathbb{P}_{\mathrm{L}}(l_{i,1} = 1 \mid \Gamma_0, \dots, \Gamma_K) = \frac{w_1}{W_i} \geq \frac{4}{\log n\sqrt{n}}\sqrt{\frac{n}{2i}} > \frac{2}{\log n\sqrt{i}}.$$

Conditioning on the sequence $\Gamma_0, \dots, \Gamma_K$, if (19) holds the probability that Γ_K does not send an edge to vertex 1 is at most

$$\prod_{i \in \Gamma_K}\left(1 - \frac{2}{\log n\sqrt{i}}\right) \leq \exp\left(-\sum_{i \in \Gamma_K}\frac{2}{\log n\sqrt{i}}\right) = \exp\left(-\frac{2f_K\sqrt{n}}{\log n}\right) \leq n^{-2}.$$

Since (19) holds with probability $1 - o(n^{-1})$, it follows that with probability $1 - o(n^{-1})$ the vertex v is connected to 1 by a path of length at most $K + 1 \leq (1/2 + \epsilon)\log n/\log\log n$, completing the proof of Lemma 11. □

Combining the results above proves Theorem 1.

Proof of Theorem 1. We have already proved the lower bound as Theorem 5. Fix $m \geq 2$ and $\epsilon > 0$. Together Lemmas 8 and 11, applied with $\epsilon/4$ in place of ϵ, imply that given $W_1, \dots, W_n$

$$\mathbb{P}_{\mathrm{L}}(\mathrm{diam}(G(W_1, \dots, W_n)) > (1 + \epsilon/2)\log n/\log\log n + 16\log\log n) = o(1),$$

whenever $E_1, \dots, E_5$ hold. Let $R_1, \dots, R_{mn}$ be obtained by sorting mn independent $M_2(0,1)$ random variables, set $W_i = R_{mi}$, and consider the graph $G = G(W_1, \dots, W_n)$. As each E_i holds with probability $1 - o(1)$ by Lemma 7, it follows that

$$\mathbb{P}(\mathrm{diam}(G) > (1 + \epsilon/2)\log n/\log\log n + 16\log\log n) = o(1).$$

Finally, we have shown at the end of section 6 that the random graph G defined in this way may be coupled with G_m^n so that G can be obtained from G_m^n by (perhaps) deleting some edges and adding some loops. Under this coupling $\mathrm{diam}(G) \geq \mathrm{diam}(G_m^n)$, completing the proof. □

We next turn briefly to the case $m = 1$, which has been studied earlier and behaves rather differently.

9 The case $m = 1$

One particular step in the proof of the upper bound on $\mathrm{diam}(G_m^n)$ given above appears at first sight to be rather wasteful, namely the step where we split the edges into two separate classes, considering at every point only edges from one class. Something like this is necessary to achieve independence when considering how the $(k+1)$-neighbourhood N_{k+1} of a given vertex relates to N_k. However, as the neighbourhoods grow quite quickly, one would imagine that we only rarely try to use edges already used. This suggests that a version of the proof above should apply with $m = 1$, losing about a factor of two in the growth rate, which would have negligible effect on the diameter. This turns out to be not the case – for $m = 1$ the diameter is in fact $\Theta(\log n)$, as shown by Pittel [24] in a slightly different context described below.

For $m = 1$ the graph G_m^n has a rather simple structure. Each vertex sends an edge either to itself or to an earlier vertex, so the graph consists of components which are trees, each with a loop attached. It is much more natural to consider trees rather than forests. Trees on $\{1, 2, \ldots, n\}$ where each vertex other than the first sends an edge to an earlier vertex are known as *recursive* trees; random recursive trees have been studied for some time, see, for example, [22]. When a random recursive tree is constructed one vertex at a time, each new vertex must be joined to an earlier vertex chosen according to some rule. The most studied case is that of a uniform choice, but probabilities essentially proportional to degrees have also been considered. Such objects (where for the first vertex degree plus one is used) are known as *random plane-oriented recursive trees*, see [27, 21], for example. Pittel [24] showed that the height (maximum distance of a vertex from the root) of such an object is $(c + o(1)) \log n$ with probability $1 - o(1)$, where $c = (2\gamma)^{-1}$ for γ the solution of $\gamma e^{1+\gamma} = 1$. It is easy to see that his method also shows that the diameter is $(2c + o(1)) \log n$. Now G_1^n has on average $\Theta(\log n)$ loops. Thus with high probability G_1^n has $O(\log n)$ loops and hence $O(\log n)$ components. It is easy to see that Pittel's method can be applied to prove the following result.

Theorem 13. *Let γ be the solution of $\gamma e^{1+\gamma} = 1$, and let $\epsilon > 0$ be fixed. Then for almost every $G_1^n \in \mathcal{G}_1^n$ the largest distance between two vertices in the same component is between $(\gamma^{-1} - \epsilon) \log n$ and $(\gamma^{-1} + \epsilon) \log n$.* $\square$

Note that this result is given only for completeness – for $m = 1$ the random plane-oriented recursive tree model is actually more natural than G_1^n.

10 Conclusions

In several papers random graphs modeling complex real-world networks are studied, in particular with respect to their diameter or *average diameter*, the average distance between two pairs of vertices. One model studied is the process G_m^n considered here. Another model is given by taking a function $P(d)$, $d \geq 1$, to represent the probability that a vertex has degree d, and choosing a graph on n vertices at random from all graphs with this distribution of degrees.

In [1, 5, 23] results from computer experiments have been published suggesting that for the second model the diameter of the random graph will vary as $A + B\log n$ when n is large, for a wide variety of functions $P(d)$. In [23] a heuristic argument is given, based on the standard neighbourhood expansion method; roughly speaking, if one follows a random edge to one of its endvertices v, the probability $P'(d)$ that v will have degree d should be proportional to $dP(d)$. Thus, writing $N_k(v)$ for the set of vertices within distance k of a given vertex v, one would expect $|N_{k+1}(v)|$ to be about

$$f = \sum dP'(d) = \left(\sum d^2P(d)\right) / \left(\sum dP(d)\right)$$

times larger than $|N_k(v)|$. One then expects the diameter of the graph to be about $\log n/\log f$.

In the context of G_m^n, which has $P(d)$ proportional to d^{-3}, this argument actually predicts an expansion factor of $\log n$, and hence a diameter of around $\log n/\log\log n$. The same expansion factor is predicted by approximating G_m^n by a graph in which each edge ij is present with probability proportional to $1/\sqrt{ij}$, as in the proof of Theorem 5. As shown by Theorem 1, this prediction is correct if $m \geq 2$. The explanation for the prediction of $\Theta(\log n)$ given by the more sophisticated heuristic in [23] is that a cutoff is imposed on the degree distribution so that the formalism used there applies. In the light of our results, we believe that the heuristic given above will apply in many situations, in particular when the graph is chosen at random subject to a power law distribution being imposed on the degrees. On the other hand, in any given case it can be highly non-trivial to take the actual structure into account. Furthermore, this structure can cause the heuristic argument to fail, as shown by the case of G_1^n, for which exactly the same heuristics apply as for G_2^n.

We finish with two more remarks concerning possible extensions of Theorem 1. For the first, note that the upper bound on the diameter of G_m^n given by Theorem 1 is a pure existence result. An interesting question is whether a short path between two given vertices can be constructed quickly using only 'local' information (see [20], for example). In G_m^n we believe that it will take $\Theta(\log n)$ steps to reach vertex 1 from vertex n using local information. In the other direction, given any set S of $o(n)$ vertices of a typical G_m^{n-1} it is easy to check that the probability of vertex n sending an edge to S is $o(1)$. It follows that at least $\Theta(n)$ steps are needed to locally construct a path from 1 to n in G_m^n.

The second remark concerns the 'robustness' of the graph G_m^n. In [2] the following question is raised, and answered experimentally: how many vertices can be deleted at random from a 'scale-free' random graph without the diameter increasing too much, ignoring small components if necessary? It is easy to see that the expanding neighbourhoods argument of section 8 is unaffected if the vertices of G_m^n are deleted independently with probability p, for any constant $p < 1$. In the terminology of section 7, from any 'useful' vertex there will be a short path to the first surviving vertex of G_m^n. In general, after deleting vertices in this manner, what remains of G_m^n will consist of many small components, together with a giant component whose diameter is asymptotically $\log n/\log\log n$, for *any* $p < 1$.

Finally, we finish by mentioning an interesting problem to which one of the referees has drawn our attention. Suppose we are given a rooted tree of height n in which each non-leaf node has $k \geq 2$ children, and suppose that the edges are weighted with independent standard exponential random variables. What is the asymptotic distribution of the minimum weight of a path from root to leaf as $n \to \infty$ with k fixed? For results on this question, related to the question of which vertices of G_m^n can be reached from an initial vertex by descending paths of a certain length, see [16, 25].

References

[1] R. Albert, H. Jeong and A.-L. Barabási, Diameter of the world-wide web, *Nature* **401** (1999), 130–131.

[2] R. Albert, H. Jeong and A.-L. Barabási, Error and attack tolerance of complex networks, *Nature* **406** (2000), 378–382.

[3] A.-L. Barabási and R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.

[4] A.-L. Barabási, R. Albert and H. Jeong, Mean-field theory for scale-free random networks, *Physica A* **272** (1999), 173-187.

[5] A.-L. Barabási, R. Albert and H. Jeong, Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A* **281** (2000), 69–77.

[6] B. Bollobás, The diameter of random graphs, *Trans. Amer. Math. Soc.* **267** (1981), 41–52.

[7] B. Bollobás, *Random Graphs, Second Edition,* Cambridge studies in advanced mathematics, vol. 73, Cambridge University Press, Cambridge, 2001, xviii + 498 pp.

[8] B. Bollobás and F.R.K. Chung, The diameter of a cycle plus a random matching, *SIAM J. Discrete Math.* **1** (1988), 328–333.

[9] B. Bollobás and W. Fernandez de la Vega, The diameter of random regular graphs, *Combinatorica* **2** (1982), 125–134.

[10] B. Bollobás and V. Klee, Diameters of random bipartite graphs, *Combinatorica* **4** (1984), 7–19.

[11] B. Bollobás and O.M. Riordan, Linearized chord diagrams and an upper bound for Vassiliev invariants, *J. Knot Theory Ramifications* **9** (2000), 847–853.

[12] B. Bollobás, O.M. Riordan, J. Spencer, and G. Tusnády, The degree sequence of a scale-free random graph process, *Random Structures and Algorithms* **18** (2001), 279–290.

[13] Ju.D. Burtin, Asymptotic estimates of the diameter and the independence and domination numbers of a random graph, *Dokl. Akad. Nauk SSSR* **209** (1973), 765–768, translated in *Soviet Math. Dokl.* **14** (1973), 497–501.

[14] Ju.D. Burtin, Extremal metric characteristics of a random graph. I, *Teor. Verojatnost. i Primenen.* **19** (1974), 740–754.

[15] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, *Ann. Math. Statistics* **23** (1952) 493–507.

[16] L. Devroye, Branching processes and their applications in the analysis of tree structures and tree algorithms, in Probabilistic methods for algorithmic discrete mathematics, *Algorithms Combin.* **16** (1998), 249–314.

[17] S. Janson, On concentration of probability, *Combinatorics, Probability and Computing* **11** (2002).

[18] V. Klee and D.G. Larman, Diameters of random graphs, *Canad. J. Math.* **33** (1981), 618–640.

[19] V. Klee, D.G. Larman and E.M. Wright, The diameter of almost all bipartite graphs, *Studia Sci. Math. Hungar.* **15** (1980), 39–43.

[20] J. Kleinberg, The small-world phenomenon: an algorithmic perspective, Cornell Computer Science Technical Report 99-1776 (October 1999).

[21] H.M. Mahmoud, R.T. Smythe and J. Szymański, On the structure of random plane-oriented recursive trees and their branches, *Random Structures and Algorithms* **4** (1993), 151–176.

[22] H.M. Mahmoud and R.T. Smythe, A survey of recursive trees, *Th. of Probability and Math. Statistics* **51** (1995), 1–27.

[23] M.E.J. Newman, S.H. Strogatz and D.J. Watts, Random graphs with arbitrary degree distribution and their applications, *Physical Review E* **64** (2001), 026118.

[24] B. Pittel, Note on the heights of random recursive trees and random m-ary search trees, *Random Structures and Algorithms* **5** (1994), 337–347.

[25] P. Révész, Critical branching Wiener process in $\mathbb{R}^d$, in Random walks (Budapest, 1998), *Bolyai Soc. Math. Stud.* **9** (1999), 299–348.

[26] A. Stoimenow, Enumeration of chord diagrams and an upper bound for Vassiliev invariants, *J. Knot Theory Ramifications* **7** (1998), 93–114.

[27] J. Szymański, On a nonuniform random recursive tree, *Annals of Discrete Math.* **33** (1987), 297–306.

[28] D. J. Watts and S. H. Strogatz, Collective dynamics of 'small-world' networks, *Nature* **393** (1998), 440–442.