Adv. Appl. Prob. **47,** 37–56 (2015)
Printed in Northern Ireland

THE MIXING TIME OF THE NEWMAN–WATTS SMALL-WORLD MODEL

LOUIGI ADDARIO-BERRY,* ** AND
TAO LEI,* *** *McGill University*

Abstract

'Small worlds' are large systems in which any given node has only a few connections to other points, but possessing the property that all pairs of points are connected by a short path, typically logarithmic in the number of nodes. The use of random walks for sampling a uniform element from a large state space is by now a classical technique; to prove that such a technique works for a given network, a bound on the mixing time is required. However, little detailed information is known about the behaviour of random walks on small-world networks, though many predictions can be found in the physics literature. The principal contribution of this paper is to show that for a famous small-world random graph model known as the Newman–Watts small-world model, the mixing time is of order $\log^2 n$. This confirms a prediction of Richard Durrett [5, page 22] , who proved a lower bound of order $\log^2 n$ and an upper bound of order $\log^3 n$.

Keywords: Random graph; small world; mixing time; conductance bound

2010 Mathematics Subject Classification: Primary 60C05
Secondary 05C81; 05C82

1. Introduction

The small-world phenomenon is a catchy name for an important physical phenomenon that shows up throughout the physical, biological, and social sciences. In brief, the term applies to large, locally sparse systems (usually possessing only a bounded number of connections from any given point) which nonetheless exhibit good long-range connectivity in the sense that there are short paths between all points in the system. The Erdős–Rényi random graph $G_{n,p}$ is perhaps the most mathematically famous model possessing small-world behaviour: when $p = c/n$ for $c > 1$ fixed, the average vertex degree is c, and the diameter of the largest connected component is

$$\frac{\log n}{\log c} + 2\frac{\log n}{\log(1/c^*)} + O_p(1)$$

(see [21]), where $c^* < 1$ satisfies $c\mathrm{e}^{-c} = c^*\mathrm{e}^{-c^*}$ and $O_p(1)$ denotes a random amount that remains bounded in probability as $n \to \infty$.

The Erdős–Rényi random graph is unsatisfactory as a real-world model in two ways: first, the network does not satisfy full connectivity (a constant proportion of vertices lie outside of the giant component); second, the graph is locally tree-like (for any fixed k, the probability that there is a cycle of length at most k through a randomly chosen node is $o(1)$). In real-world

Received 20 April 2012; revision received 6 January 2014.
* Postal address: Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montréal, Québec, H3A 0B9, Canada.
** Email address: louigi@math.mcgill.ca
*** Email address: tao.lei@mail.mcgill.ca

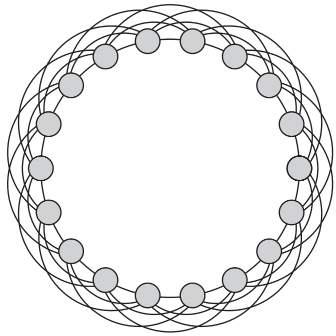

FIGURE 1.

networks showing small-world behaviour (e.g. social or business networks, gene regulatory networks, networks for modelling infectious disease spread, scientific collaboration networks, and many others — [19] contains many interesting examples), full or almost-full connectivity is common, and short cycles are plentiful. Several connected models have been proposed which in some respects capture the desired local structure as well as small-world behaviour, notably the Bollobás–Chung model [2], the Watts–Strogatz model [23], and the Newman–Watts model [14], [17], [18]. These models are closely related — all are based on adding sparse, long-range connections to a connected 'base network' which is essentially a cycle. Note that in [23], small-world models are connected random graph models having both short average distance and relatively high clustering coefficient.

Understanding the behaviour of random walks on small-world networks remains, in general, a challenging open problem. Numerical and nonrigorous results for return probabilities [11], relaxation times [20], spectral properties [6], hitting times [4], [10], and diffusivity [8] appear in the physics literature, but few rigorous results are known. In [5], Durrett considered the Newman–Watts small-world model, proving lower and upper bounds on the mixing time of order $\log^2 n$ and $\log^3 n$, respectively, and suggested that the lower bound should in fact be correct. The principle contribution of this paper is to confirm Durrett's prediction (see Theorem 1, below).

As stated in [16], the Newman–Watts small-world model was proposed independently by Monasson [14] and by Newman and Watts [17], [18]. In this paper, we follow the formulation as in [14]. To define the Newman–Watts small-world model, first fix integers $n \geq k \geq 1$. The *(n, k)-ring* $R_{n,k}$ is the graph with vertex set $[n] = \{1, \ldots, n\}$ and edge set $\{\{i, j\} \colon i + 1 \leq j \leq i + k\}$, where addition is interpreted modulo n. (A picture of $R_{18,3}$ appears in Figure 1.) In particular, an $(n, 1)$-ring is a cycle of length n, and whenever $n > 2k$ the (n, k)-ring is regular of degree $2k$. For $0 < p < 1$, the *(n, k, p)-Newman–Watts small-world model*, denoted by $H_{n,k,p}$, is the random graph obtained from the (n, k)-ring by independently replacing each non-edge of the (n, k)-ring by an edge with probability p. To be more precise, independently for each $1 \leq i < j \leq n$,

$$\mathbb{P}(\{i, j\} \text{ is an edge of } H_{n,k,p}) = \begin{cases} 1 & \text{if } \{i, j\} \in E(R_{n,k}), \\ p & \text{otherwise.} \end{cases}$$

We write H as shorthand for $H_{n,k,p}$ whenever the parameters are clear from the context.

Given a (finite, simple) graph $G = (V, E)$, by a *lazy simple random walk on G* we mean a random walk that at each step stays still with probability $\frac{1}{2}$, and otherwise moves to a uniformly

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

random neighbour. In other words, this is a Markov chain with state space V and transition probabilities

$$p_{x,y} = \begin{cases} \frac{1}{2} & \text{if } x = y, \\ 1/2d_G(x) & \text{if } y \in N_G(x), \\ 0 & \text{otherwise,} \end{cases}$$

where $d_G(x)$ denotes the number of neighbours of x in G and $N_G(x)$ denotes the collection of neighbours of x in G. (More generally, we shall say a chain is *lazy* if $p_{x,x} \geq \frac{1}{2}$ for all x in the state space.) If G is connected then this Markov chain has a unique stationary distribution π given by $\pi(x) = d_G(x)/2|E|$. (We note for later use the notation $\pi(S) = \sum_{v \in S} \pi(v)$ for S, a set of vertices in G.)

Given two probability distributions μ, ν on V, we define the *total variation distance* between μ and ν as

$$\|\mu - \nu\|_{\mathrm{TV}} = \sup_{S \subset V} |\mu(S) - \nu(S)| = \frac{1}{2} \sum_{v \in V} |\mu(v) - \nu(v)|,$$

where $\mu(S) = \sum_{v \in S} \mu(v)$. Now, let $(X_k)_{k \geq 0}$ be a lazy simple random walk on G, and write $\mu_{k,x}$ for the distribution of X_k when the walk is started from x; formally, for all $y \in G$, $\mu_{k,x}(y) = \mathbb{P}(X_k = y \mid X_0 = x)$. The *mixing time* of the lazy simple random walk on G is defined as

$$\tau_{\mathrm{mix}}(G) = \max_x \min\big\{k \colon \|\mu_{k,x} - \pi\|_{\mathrm{TV}} \leq \tfrac{1}{4}\big\}.$$

(There are many different notions of mixing time, many of which are known to be equivalent up to constant factors — [12] and [13] are both excellent references.) Note that the mixing time is a deterministic parameter of the underlying graph G, but $\tau_{\mathrm{mix}}(H)$ is a random variable since $H = H_{n,k,p}$ is random. We may now formally state our main result.

Theorem 1. *Fix $c > 0$, let $p = c/n$, and let H be an (n, k, p)-Newman–Watts small-world network. Then there exists $C_0 > 0$ depending only on c and k such that, with a probability of at least $1 - O(n^{-3})$,*

$$C_0^{-1} \log^2 n \leq \tau_{\mathrm{mix}}(H) \leq C_0 \log^2 n.$$

Furthermore, $\mathbb{E}[\tau_{\mathrm{mix}}(H)] \leq C_0(\log^2 n + 1)$.

The expectation bound in Theorem 1 follows easily from the probability bound. Indeed, given a finite reversible lazy chain $X = (X_t,\, t \geq 0)$ with state space Ω; for $x \in \Omega$ write $\tau_x = \min\{t \geq 0 \colon X_t = x\}$ for the *hitting time* of state x. Then $\tau_{\mathrm{mix}} \leq 2 \max_{x \in \Omega} \mathbb{E}_\pi(\tau_x) + 1$, where $\mathbb{E}_\pi$ denotes expectation starting from stationarity (see e.g. [12, Theorem 10.14 (ii)]). If X is a lazy simple random walk on a connected graph $G = (V, E)$ with $|V| = n$, then $\max_{x \in V} \mathbb{E}_\pi(\tau_x) \leq (\frac{4}{27} + o(1))n^3$ (see [3]). Assuming the probability bound of Theorem 1 and applying the two preceding facts, we obtain

$$\mathbb{E}[\tau_{\mathrm{mix}}(H)] \leq C_0 \log^2 n + \tfrac{8}{27} n^3 (1 + o(1)) \mathbb{P}(\tau_{\mathrm{mix}}(H) > C_0 \log^2 n) \leq C_0(\log^2 n + 1),$$

assuming that C_0 is chosen to be large enough.

The lower bound of Theorem 1 is also straightforward, and we now provide its proof. For n sufficiently large, given $v \in [n]$ and $\ell \in \mathbb{N}$, the probability that all vertices w of $H_{n,k,p}$ with $w \in [v - \ell, v + \ell] \pmod n$ have degree exactly $2k$ is greater than

$$(1 - p)^{n(2\ell+1)} \geq \mathrm{e}^{-2c(2\ell+1)},$$

the inequality holding since $(1-c/n)^n > \mathrm{e}^{-2c}$ for large n. Taking $\alpha = 1/8c$, it follows easily that with a probability of at least $1 - O(n^{-3})$ there exists a $v \in [n]$ such that all vertices w with $w \in [v - \alpha \log n, v + \alpha \log n] \pmod n$ have degree exactly $2k$. Furthermore, the random walker starting from such a vertex v will, with high probability, take time of order $\log^2 n$ before first visiting a vertex in the complement of $[v - \alpha \log n, v + \alpha \log n] \pmod n$. Finally, under π, the set $[v - \alpha \log n, v + \alpha \log n] \pmod n$ has measure tending to zero with n; thus, it follows from the definition of τ_{mix} that the mixing time is of order at least $\log^2 n$ whenever such a vertex v exists.

Having taken care of the expectation upper bound and lower bound in probability from Theorem 1, the remainder of this paper is now devoted to proving that, with probability at least $1 - O(n^{-3})$, $\tau_{\text{mix}}(H) = O(\log^2 n)$. In Section 2 we explain the conductance-based mixing time bound of Fountoulakis and Reed [7] which will form the basis of our approach. The Fountoulakis–Reed bound requires control on sizes of the edge boundaries of connected subgraphs of $H_{n,k,p}$ (or *edge expansion* for short). To this end, in Section 3 we bound the expected number of connected subgraphs of $H_{n,k,p}$ of size j, for each $1 \le j \le n$; our probability bounds on the edge expansion of such subgraphs follow in Section 4. In Section 5 we finish the proof of Theorem 1. The proof is more straightforward when c is large; hence, we handle the large-c and small-c cases separately, as we will need to use the results for the large-c case in the proof of the small-c case. Finally, in Section 6 we provide some concluding remarks.

1.1. Notation

Given a graph G, write $V(G)$ for the set of vertices of G and $E(G)$ for the set of edges of G. Also, given $S \subset V(G)$, write $G[S]$ for the subgraph of G induced by S (i.e. $G[S]$ is the graph with vertex set S and edge set $\{e = (u, v) \colon u, v \in S, e \in E(G)\}$). We say that S is connected if $G[S]$ is connected. Finally, given a formal power series $F(z)$, we write $[z^j]F(z)$ to mean the coefficient of z^j in $F(z)$, so if $F(z) = \sum_{k \ge 0} a_k z^k$ then $[z^j]F(z) = a_j$.

Given sets $S, T \subset V$, write $E(S, T) = E_G(S, T)$ for the set of edges of G with one endpoint in S and the other in T, and write $e(S, T) = |E(S, T)|$. Also, given $S \subset V$ write $e(S) = \sum_{v \in S} d_G(v)$.

2. Mixing time via conductance bounds

A range of techniques are known for bounding mixing times ([15] is a recent survey of the available approaches), many of which are tailor-made to give sharp bounds for particular families of chains. One family of techniques is based on bounding the *conductance* of the underlying graph, a function which encodes the presence of bottlenecks at all scales. Informally, by a bottleneck at scale x we mean a set $S \subset V$ with $e(S, S) \approx x|E(G)|$ and with $e(S, S^c) \ll x|E(G)|$. For example, in the dumbbell graph formed by connecting two cliques of size $n/2$ with one edge, one of the cliques forms a bottleneck at scale $\frac{1}{2}$. In this case, there exists bottlenecks of scale $\frac{1}{2}$ but no bottlenecks of smaller scale. The precise bound we shall use is due to Fountoulakis and Reed [7]. The *conductance* of S, written $\Phi(S)$, is given by

$$\Phi(S) = \frac{e(S, S^c)}{e(S)}.$$

For $0 \le x \le \frac{1}{2}$, write

$$\Phi(x) = \min_{\substack{S \text{ connected} \\ x|E| \le e(S) \le 2x|E|}} \Phi(S).$$

The function $\Phi(x)$ is called the *conductance profile* of G. This definition is how we formalize the idea of bottleneck at scale x: $\Phi(x)$ is the smallest conductance of a connected set $S \subset V(G)$ with $\pi(S) \in [x/2, x]$. The quantity $\Phi(S)$ is sometimes called the *bottleneck ratio* (see e.g. [12, page 88]). Also, the conductance of S is sometimes defined to equal $Q(S, S^c)/\pi(S)\pi(S^c)$, where $Q(S, S^c) = \sum_{x \in S, y \notin S} \pi(x)P(x, y)$ and P is the transition matrix of the random walk (or, more generally, of some irreducible and aperiodic Markov chain with stationary distribution π). For a simple random walk on a graph G,

$$\frac{Q(S, S^c)}{\pi(S)\pi(S^c)} = \frac{e(S, S^c)}{e(S)} \frac{2|E|}{e(S^c)}. \tag{1}$$

Note that when $e(S) \leq |E|$, the quantity on the left-hand side of (1) differs from $\Phi(S)$ by at most a multiplicative factor of 2.

We will use the following theorem, a specialization of the main result from [7].

Theorem 2. ([7, Theorem 1].) *There exists a universal constant $C > 0$ so that, for any connected graph G,*

$$\tau_{\mathrm{mix}}(G) \leq C \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i}).$$

In fact, in [7] $\Phi(s)$ is defined as

$$\Phi(S) = \frac{Q(S, S^c)}{\pi(S)\pi(S^c)}$$

and $\Phi(x)$ is set as

$$\Phi(x) = \min_{\substack{S \text{ connected} \\ x/2 \leq \pi(S) \leq x}} \Phi(S).$$

However, by the comments preceding Theorem 2 this changes the precise upper bound in Theorem 2 by at most a multiplicative factor of 2.

With Theorem 2 at hand, proving mixing time bounds boils down to understanding what sorts of bottlenecks can exist in G. For the (n, k, p)-Newman–Watts small-world network with $p = c/n$, it is not hard to see that small sets can have low conductance. Indeed, in Section 1 we observed that, with high probability, the ring $R_{n,k}$ will contain connected sets S with $\Theta(\log n)$ nodes, to which no edges are added in $H_{n,k,p}$. Such a set S will have $e(S, S^c) < k^2$ and so will have conductance $\Phi(S) = O(1/\log n)$. It follows that, with high probability, $\Phi^{-2}(2^{-i}) = \Theta(\log^2 n)$ for some $i = \Theta(\log\log n)$. By Theorem 2, this suggests that to prove Theorem 1, we might try to show that, for i smaller than some threshold of order $\log\log n$ (corresponding to sets of size larger than S), the contribution to the sum is negligible compared with $\log^2 n$. We indeed take this approach.

To be more precise, we will prove the upper bound in the probability bound of Theorem 1 by showing that there are constants $\varepsilon > 0$, $C_0 > 0$ such that, with high probability, whenever $|S| \geq C_0 \log n$, we have $\Phi(S) \geq \varepsilon$. From a more precise version of this fact and Theorem 2, the upper bound in Theorem 1 will follow straightforwardly. To prove such bounds, we will need control on the likely number of connected subgraphs of $H_{n,k,p}$ of size s, for all $s \geq C \log n$ for some constant C. In Section 3, we bound the expected number of such subgraphs using Lagrange inversion and comparison with a branching process.

3. Counting connected subgraphs

Let $v \in [n]$, write $B_{j,v} = B_{j,v}(H)$ for the set of all $S \subset [n]$ containing v with $|S| = j$ such that $H[S]$ is connected, and let $B_j = \bigcup_{v \in [n]} B_{j,v}$. Our aim in this section is to establish the following proposition.

Proposition 1. *For any positive integer j and any $v \in [n]$, $\mathbb{E}|B_{j,v}| \le (4(c+2k))^j$, and $\mathbb{E}|B_j| \le n(4(c+2k))^j$.*

We will prove Proposition 1 by comparison with a Galton–Watson process. Recall that a Galton–Watson process can be described as follows. An initial individual — the progenitor — has a random number Z_1 of children, where Z_1 is some nonnegative, integer-valued random variable. The distribution of Z_1 is called the *offspring distribution*. Each child of the progenitor reproduces independently according to the offspring distribution, and this process continues recursively. The family tree of a Galton–Watson process is called a *Galton–Watson tree*, and is rooted at the progenitor.

The number of neighbours of a vertex in $H_{n,k,p}$ is distributed as $\mathrm{bin}(n-2k-1, p) + 2k$. From this it is easily seen that $|B_{j,v}|$ is stochastically dominated by the number of subtrees of size j containing the root in $\mathcal{T}$, where $\mathcal{T}$ is a Galton–Watson tree with offspring distribution $\mathrm{bin}(n-2k-1, p)+2k$. To bound the expectations of the latter random variables, we will first encode these expectations as the coefficients of a generating function, then use the Lagrange inversion formula ([22, Theorem 5.4.2]), which we now recall.

Theorem 3. (Lagrange inversion formula, in [22, Theorem 5.4.2].) *If $G(x)$ is a formal power series and $f(x) = xG(f(x))$, then*

$$n[x^n]f(x)^k = k[x^{n-k}]G(x)^n.$$

Fix a nonnegative, integer-valued random variable B, and for $m \ge 0$ write $p_m = \mathbb{P}(B = m)$. Given a Galton–Watson tree $\mathcal{T}$ with offspring distribution B, let $\mu_j = \mu_j(B)$ denote the expected number of subtrees of $\mathcal{T}$ containing the root of $\mathcal{T}$ and having exactly j vertices (so $\mu_0 = 0$). Also, write

$$q_j = \sum_{m \ge j} p_m (m)_j,$$

where $(m)_j = m!/(m-j)!$ is the falling factorial. Note that q_j is the expected number of ways to choose and order j children of the root in $\mathcal{T}$. Let $F(z) = \sum_{j=0}^{\infty} \mu_j z^j$ and $Q(z) = \sum_{j=0}^{\infty} q_j z^j$ be the generating functions of μ_j and q_j respectively, viewed as formal power series.

Lemma 1. *Let $F(z) = zQ(F(z))$.*

Proof. We have

$$\begin{aligned} Q(F(z)) &= \sum_{j\ge 0} q_j \left(\sum_{r \ge 1} \mu_r z^r \right)^j \\ &= \sum_{j \ge 0} q_j \left(\sum_{r \ge j} \sum_{\substack{r_1+\cdots+r_j=r \\ r_1,\cdots,r_j \in \mathbb{N}^+}} z^r \mu_{r_1} \cdots \mu_{r_j} \right) \\ &= \frac{1}{z} \sum_{r \ge 0} z^{r+1} \sum_{j \le r} q_j \left(\sum_{\substack{r_1+\cdots+r_j=r \\ r_1,\cdots,r_j \in \mathbb{N}^+}} \mu_{r_1} \cdots \mu_{r_j} \right). \end{aligned} \tag{2}$$

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

The rth term in the outer sum on the right-hand side of (2) encodes subtrees of $\mathcal{T}$ with $r+1$ vertices that contain the root, as follows. First specify the degree j of the root of the tree T to be embedded. Then choose which j children of the root of $\mathcal{T}$ will form part of the embedding, and the order in which the children of the root of T will be mapped to these nodes (there are q_j ways to do this on average). Next, choose the sizes $r_1, \dots, r_j$ of the subtrees of the children of the root in the embedded tree; finally, embed each such subtree in the respective subtree of $\mathcal{T}$; on average, there are μ_{r_i} ways to do this. It follows that

$$\frac{1}{z}\sum_{r\geq 0} z^{r+1}\sum_{j\leq r} q_j\Bigg(\sum_{\substack{r_1+\cdots+r_j=r\\ r_1,\cdots,r_j\in\mathbb{N}^+}} \mu_{r_1}\cdots\mu_{r_j}\Bigg) = \frac{1}{z}\sum_{r\geq 0} z^{r+1}\mu_{r+1} = \frac{1}{z}F(z),$$

which proves the lemma. (We remark that verifying that the identity in the first line can be done purely formally; however, we find the preceding explanation more instructive.)

Lemma 2. *Fix* $C>0$. *If* $q_j \leq C^j$ *for all* $j\geq 0$ *then*

$$\mu_j \leq \frac{1}{j}\binom{2j-2}{j-1}C^{j-1} < (4C)^{j-1},$$

for all $j\geq 1$.

Proof. By Lemma 1 and Theorem 3, we have $j[z^j]F(z)^k = k[z^{j-k}]Q(z)^j$. In particular, taking $k=1$, we have $\mu_j = [z^j]F(z) = (1/j)[z^{j-1}]Q(z)^j$. Now,

$$Q(z)^j = \bigg(\sum_{l\geq 0} q_l z^l\bigg)^j = \sum_{r\geq 0}\Bigg(\sum_{\substack{l_1+\cdots+l_j=r\\ l_1,\cdots,l_j\in\mathbb{N}}} q_{l_1}q_{l_2}\cdots q_{l_j}\Bigg)z^r.$$

Therefore,

$$[z^{j-1}]Q(z)^j = \sum_{\substack{l_1+\cdots+l_j=j-1\\ l_1,\cdots,l_j\in\mathbb{N}}} q_{l_1}q_{l_2}\cdots q_{l_j}.$$

Each summand $q_{l_1}\cdots q_{l_j}$ is at most C^{j-1} by assumption. There are $\binom{2j-2}{j-1}$ nonnegative integer solutions to the equation $l_1+\cdots+l_j=j-1$, so we obtain that $[z^{j-1}]Q(z)^j \leq \binom{2j-2}{j-1}C^{j-1}$. The result follows.

The next lemma controls the growth of q_j for some important special offspring distributions, which allows us to use Lemma 2 to prove Proposition 1.

Lemma 3. *If the offspring distribution* B *is* $\mathrm{Poi}(c)$ *distributed then* $q_j = c^j$ *for all* j. *Also, if* B *is* $\mathrm{bin}(n, c/n)$ *distributed then, for all* $j\geq 0$, $q_j\leq c^j$. *Finally, if* $B-\ell$ *is* $\mathrm{bin}(n,c/n)$ *distributed for some fixed* $\ell\geq 0$ *then, for all* $j\geq 0$, $q_j\leq (c+\ell)^j$.

Proof. If $B \overset{\mathrm{D}}{=} \mathrm{Poi}(c)$, then

$$q_j = \sum_{m\geq j} p_m\frac{m!}{(m-j)!} = \sum_{m\geq j}\frac{c^m}{m!}\mathrm{e}^{-c}\frac{m!}{(m-j)!} = \mathrm{e}^{-c}c^j\sum_{l\geq 0}\frac{c^l}{l!} = \mathrm{e}^{-c}c^j\mathrm{e}^c = c^j.$$

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

If $B \overset{\mathrm{D}}{=} \mathrm{bin}(n, c/n)$ then, using the binomial theorem, we have

$$\begin{aligned} q_j &= \sum_{j \le m \le n} \binom{n}{m} \left(\frac{c}{n}\right)^m \left(1 - \frac{c}{n}\right)^{n-m} \frac{m!}{(m-j)!} \\ &= \left(\frac{c}{n}\right)^j n(n-1)\cdots(n-j+1) \sum_{j \le m \le n} \binom{n-j}{n-m} \left(\frac{c}{n}\right)^{m-j} \left(1 - \frac{c}{n}\right)^{n-m} \\ &\le c^j, \end{aligned}$$

Finally, if $B \overset{\mathrm{D}}{=} \mathrm{bin}(n, c/n) + \ell$ then we consider three Galton–Watson trees, $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$, with offspring distributions $B_1 \equiv l$, $B_2 \overset{\mathrm{D}}{=} \mathrm{bin}(n, c/n)$, and $B_3 \overset{\mathrm{D}}{=} \mathrm{bin}(n, c/n) + \ell$ respectively. For $i = 1, 2, 3$ write $q_j^{(i)} = q_j(B_i)$. Since $q_j^{(1)}$ is the expected number of ways to choose and order precisely j children of the root in $\mathcal{T}_1$, we have $q_j^{(1)} = (l)_j \le l^j$ and by the previous argument we know that $q_j^{(2)} \le c^j$. Finally, since $q_j^{(3)}$ is the expected number of ways to choose and order precisely j children of the root in $\mathcal{T}_3$, by independence we have

$$q_j^{(3)} = \sum_{s=0}^{j} \binom{j}{s} q_s^{(1)} q_{j-s}^{(2)} \le \sum_{s=0}^{j} \binom{j}{s} l^s c^{j-s} = (l+c)^j.$$

The factor $\binom{j}{s}$ in the first equality arises because as long as we choose s positions for children coming from the deterministic component of offspring distribution, the order of all j children are fixed since the order among s children and the order among the other $j - s$ are both fixed.

We remark that if $\mathcal{T}$ has deterministic d-ary branching (every node has exactly d children with probability one), then for all j, the number of subtrees containing the root and having precisely j nodes is exactly $\binom{dj}{j-1}/j$ (see [22, Theorem 5.3.10]) which is bounded above by $(edj/(j-1))^{j-1}/j \le (ed)^j$. Thus, Lemma 2 shows that when factorial moments grow only exponentially quickly, the values μ_j behave roughly as in the case of deterministic branching. We also note that when $\mathcal{T}$ has $\mathrm{Poi}(c)$ branching distribution, Lemma 2 and the argument of Lemma 3 together yield the exact formula $\mu_j = (c^{j-1}/j)\binom{2j-2}{j-1}$.

Proof of Proposition 1. Let $\mathcal{T}$ be a Galton–Watson tree with offspring distribution $B \overset{\mathrm{D}}{=} \mathrm{bin}(n, c/n) + 2k$. Then, for any $v \in V(H_{n,k,p})$, the random variable $|B_{j,v}|$ is stochastically dominated by the number of subtrees of $\mathcal{T}$ containing the root of $\mathcal{T}$ and having exactly j vertices. As above, we write $\mu_j(B)$ for the expected number of such subtrees. By Lemma 3, we have that $q_j(B) \le (c + 2k)^j$ for all j, and it then follows from Lemma 2 that $\mu_j(B) \le (4(c + 2k))^{j-1}$ for all j, proving the proposition.

4. Bounding the expansion of connected subgraphs of $H_{n,k,p}$

Recall from Section 3 that B_j is the collection of connected subsets S of $V(H_{n,k,p})$ with $|S| = j$. We will show that, with high probability, for all $j \ge \log n$, all elements S of B_j have conductance uniformly bounded away from zero. Many of our proofs are easier when c is large, and we treat this case first.

In the course of the proofs we will make regular use of the standard Chernoff bounds (see, e.g. [9] Theorem 2.1) which we summarize here.

Theorem 4. *If* $X \stackrel{\mathrm{D}}{=} \mathrm{bin}(m, q)$ *then*

$$\mathbb{P}(X \le (1-x)mq) \le \exp(-mq\phi(-x)) \le \exp\left(\frac{-mqx^2}{2}\right) \quad \textit{for all } 0 < x < 1$$

and

$$\mathbb{P}(X \ge (1+x)mq) \le \exp(-mq\phi(x)) \le \exp\left(\frac{-mqx^2}{2(1+x)}\right) \quad \textit{for all } x > 0,$$

where $\phi(x) = (1+x)\log(1+x) - x$.

We will use the coarser bounds most of the time. The finer bound will only be used twice, once in the proof of Lemma 7 and once in the proof of Theorem 1.

We will also need the following lemma, which is an easy consequence of the Fortuin-–Kasteleyn—Ginibre (FKG) inequality (see, e.g. [9, Theorem 2.12]). Let $\Gamma = [n] = \{1, 2, \dots, n\}$. Given $0 \le p_1, \dots, p_n \le 1$, $\Gamma_{p_1,\dots,p_n} \subset [n]$ is obtained by including element i with probability p_i independently for all i. We say that a function $f : 2^\Gamma \to \mathbb{R}$ is *increasing* if $f(A) \le f(B)$ for $A \subset B$, and f is *decreasing* if $f(A) \ge f(B)$ for $A \subset B$.

Lemma 4. *If C is an increasing event and A is a decreasing event, then*

$$\mathbb{P}(C \mid A) \le \mathbb{P}(C).$$

We begin by bounding the edge expansion of all but the very large connected sets, in the case that c is large.

Lemma 5. *Fix c large enough such that $c/720 - \log(4(c+2k)) > 5$. Then, for all n,*

$$\mathbb{P}\left(\textit{there exists } S \in \bigcup_{\log n \le j \le 9n/10} B_j,\ e(S, S^c) \le \frac{c|S|}{12}\right) \le \frac{1}{n^3}.$$

Proof. Fix $j \in [\log n, 9n/10]$, and $S \subset [n]$ with $|S| = j$. Note that $E(S, S^c)$ is independent of $H[S]$, so given that $S \in B_j$, $e(S, S^c)$ stochastically dominates a $\mathrm{bin}(j(n-j), p)$ random variable. Since $j(n-j) \ge jn/10$, it follows that under this conditioning $e(S, S^c)$ also stochastically dominates X, a $\mathrm{bin}(nj/10, p)$ random variable. By a union bound it follows that

$$\begin{aligned}
&\mathbb{P}\left(\text{there exists } S, |S| = j, H[S] \text{ connected}, e(S, S^c) \le \frac{cj}{12}\right) \\
&\quad \le \sum_{S, |S|=j} \mathbb{P}\left(e(S, S^c) \le \frac{cj}{12} \,\middle|\, H[S] \text{ connected}\right)\mathbb{P}(H[S] \text{ connected}) \\
&\quad \le \mathbb{P}\left(X \le \frac{cj}{12}\right)\mathbb{E}|B_j| \\
&\quad \le \mathrm{e}^{-cj/720} n(4(c+2k))^j,
\end{aligned}$$

where the last line follows by a Chernoff bound and by Proposition 1. (This is a typical example of our use of Proposition 1 in the remainder of the paper.)

By our assumption that $c/720 - \log(4(c+2k)) > 5$ and since $j \ge \log n$, we obtain that this probability is at most

$$\exp\left(\log n + j\left(\log(4(c+2k)) - \frac{c}{720}\right)\right) \le \frac{1}{n^4}.$$

The result follows by a union bound over $j \in [\log n, 9n/10]$.

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

The next lemma provides a lower bound on the edge expansion of very large sets, again in the case that c is sufficiently large.

Lemma 6. *If $c > 40$ then, for all n sufficiently large,*

$$\mathbb{P}\bigg(\textit{there exists } S \subset [n]\colon |S| > \frac{9n}{10}, e(S) \leq |E(H)|\bigg) \leq \left(\frac{2}{e}\right)^n.$$

Proof. In this proof write $E = E(H)$. Since, for any set $S \subset [n]$, $e(S) + e(S^c) = 2|E|$, the claim of the lemma is equivalent to

$$\mathbb{P}\bigg(\text{there exists } S \subset [n]\colon |S| < \frac{n}{10}, e(S) \geq |E|\bigg) \leq \left(\frac{2}{e}\right)^n.$$

Fix any set S with $|S| < n/10$. Write $e^*(S) = \sum_{v \in S} |\{e \ni v\colon e \notin E(R_{n,k})\}|$ for the total degree incident to S not including edges of the ring $R_{n,k}$, and similarly let $E^* = E \setminus E(R_{n,k})$. Since $|S| < n/10 < n/2$, in order to have $e(S) \geq |E|$ we must in fact have $e^*(S) \geq |E^*|$. Also, $e^*(S)$ is stochastically dominated by $\mathrm{bin}(n^2/10, p)$, and $|E^*| \stackrel{\mathrm{D}}{=} \mathrm{bin}(n(n-1-2k)/2, p)$. When n is large enough such that $n - 1 - 2k > 4n/5$, we have

$$\begin{aligned}
\mathbb{P}(e(S) \geq |E(H)|) &\leq \mathbb{P}\bigg(|E^*| \leq \frac{cn}{5}\bigg) + \mathbb{P}\bigg(e^*(S) > \frac{cn}{5}\bigg) \\
&\leq \mathbb{P}\bigg(|E^*| \leq \frac{c(n-1-2k)}{4}\bigg) + \mathbb{P}\bigg(e^*(S) > \frac{cn}{5}\bigg) \\
&< \exp\bigg(\frac{-c(n-1-2k)}{16}\bigg) + \exp\bigg(\frac{-cn}{40}\bigg),
\end{aligned}$$

by a Chernoff bound. For n sufficiently large, the last line is at most $2\mathrm{e}^{-cn/40} < 2\mathrm{e}^{-n}$, and the result follows by a union bound over all S with $|S| \leq n/10$ (there are less than 2^{n-1} such sets).

A similar but slightly more involved argument yields the following result, which will be useful for dealing with smaller values of c.

Lemma 7. *For any $c > 0$ there exists $\beta = \beta(c) > 0$ such that, for all n sufficiently large,*

$$\mathbb{P}(\textit{there exists } S \subset [n]\colon |S| > (1-\beta)n,\ e(S) \leq |E(H)|) \leq (1-\beta)^n.$$

Proof. As in the proof of Lemma 6, it suffices to prove that, for some $\beta > 0$,

$$\mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e(S) > |E|) \leq (1-\beta)^n.$$

Furthermore, since $\mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e(S) > |E|)$ decreases as β decreases, it suffices to find $\beta > 0$ and $\varepsilon > 0$ such that, for n sufficiently large,

$$\mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e(S) > |E|) = O(\mathrm{e}^{-\varepsilon n}).$$

We fix $0 < \beta < 1/(3\mathrm{e})$ small enough such that $1/2\beta - 8k/c > 1 + 1/3\beta$. Additionally, recalling the function $\phi(x) = (1+x)\log(1+x) - x$ from Theorem 4, we choose β small enough such that $\phi(1/3\beta) > \log(1/3\beta)/6\beta$. Finally, we assume that $\beta < c/36$.

For any $S \subset [n]$ with $|S| < \beta n$ and with $e(S) > |E|$, defining $e^*(S)$ and E^* as in the proof of Lemma 6, we then have $e^*(S) \geq |E| - 2k|S| \geq |E| - 2k\beta n$ and

$$\begin{aligned}
&\mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e(S) > |E|) \\
&\quad\leq \mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e^*(S) \geq |E| - 2k\beta n) \\
&\quad\leq \mathbb{P}\bigg(|E| \leq \binom{n}{2}p - 2k\beta n\bigg) \\
&\qquad + \mathbb{P}\bigg(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e^*(S) > \binom{n}{2}p - 4k\beta n\bigg). \qquad (3)
\end{aligned}$$

Since $|E|$ stochastically dominates $\mathrm{bin}(\binom{n}{2}, p)$, by a Chernoff bound we have

$$\begin{aligned}
\mathbb{P}\bigg(|E| \leq \binom{n}{2}p - 2k\beta n\bigg) &\leq \exp\bigg(-\frac{1}{2}\binom{n}{2}p\bigg(\frac{2k\beta n}{\binom{n}{2}p}\bigg)^2\bigg) \\
&< \exp\bigg(-\frac{4k^2\beta^2}{c}n\bigg), \qquad (4)
\end{aligned}$$

which handles the first summand in (3). For the second summand, let X be $\mathrm{bin}(\beta n(n-1), p)$ distributed, and note that, for all $S \subset [n]$ with $|S| \leq \beta n$, $e^*(S)$ is stochastically dominated by X. Also, for $\beta < \frac{1}{3}$ there are less than $2\binom{n}{\lfloor \beta n \rfloor}$ subsets of $[n]$ of size less than βn, and it follows by a union bound that

$$\begin{aligned}
&\mathbb{P}\bigg(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e^*(S) > \binom{n}{2}p - 4k\beta n\bigg) \\
&\quad\leq 2\binom{n}{\lfloor \beta n \rfloor}\mathbb{P}\bigg(X > \binom{n}{2}p - 4k\beta n\bigg). \qquad (5)
\end{aligned}$$

Since $1/(p(n-1)) = n/(c(n-1)) \leq 2/c$ for all $n \geq 2$, and by our assumption that $1/(2\beta) - 8k/c > 1 + 1/3\beta$, we have

$$\binom{n}{2}p - 4k\beta n = \beta n(n-1)p\bigg(\frac{1}{2\beta} - \frac{4k}{p(n-1)}\bigg) > \beta n(n-1)p\bigg(1 + \frac{1}{3\beta}\bigg).$$

By the sharper of the Chernoff upper bounds in Theorem 4 and by our assumption that $\phi(1/3\beta) > \log(1/3\beta)/6\beta$, it follows that

$$\begin{aligned}
\mathbb{P}\bigg(X > \binom{n}{2}p - 4k\beta n\bigg) &\leq \exp\bigg(-\beta n(n-1)p\phi\bigg(\frac{1}{3\beta}\bigg)\bigg) \\
&\leq \exp\bigg(-\beta n(n-1)p\frac{\log(1/3\beta)}{6\beta}\bigg) \\
&< \exp\bigg(-\frac{c\log(1/3\beta)}{12}n\bigg),
\end{aligned}$$

for n sufficiently large.

Combined with (5) this yields

$$\begin{aligned}
&\mathbb{P}\bigg(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e^*(S) > \binom{n}{2}p - 4k\beta n\bigg) \\
&\quad< 2\binom{n}{\lfloor \beta n \rfloor}\exp\bigg(-\frac{c\log(1/3\beta)}{12}n\bigg)
\end{aligned}$$

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

$$\leq 2\left(\frac{e}{\beta}\right)^{\beta n}\exp\left(-\frac{c\log(1/3\beta)}{12}n\right)$$
$$= 2\exp\left(n\left(\beta+\beta\log\left(\frac{1}{\beta}\right)-\frac{c}{12}\log\left(\frac{1}{3\beta}\right)\right)\right)$$
$$< 2\exp\left(-\frac{c}{36}n\right),$$

where in the last inequality we used that $\beta+\beta\log(1/\beta) < 2\beta\log(1/\beta) < (c/18)\log(1/\beta)$ and that $\log(1/3\beta) > 1$. Together with (3) and (4), we obtain

$$\mathbb{P}(\text{there exists } S \subset [n]\colon |S| < \beta n,\ e(S) > |E|) \leq \exp\left(-\frac{4k^2\beta^2}{c}n\right)+2\exp\left(-\frac{c}{36}n\right),$$

which completes the proof.

In order to use Lemma 5 to bound the conductance of connected subsets S of $V(H_{n,k,p})$ of size at most $9n/10$, we need to know that for such subsets we have $e(S) = O(|S|)$ with high probability. Such a bound is provided by Lemma 8, below.

For given k, let $x = x_k$ be the positive solution of the equation $x/720 - \log(4(x+2k)) = 5$, and let $M = M(c,k) = k+1+10\max(x_k, c)$. We remark that x_k is much larger than 40 for all $k \geq 1$.

Lemma 8. *For all $c > 0$ and for all n,*

$$\mathbb{P}\left(\text{there exists } S \in \bigcup_{1\leq j\leq n} B_j,\ e(S,S) > M\max(|S|,\log n)\right) \leq \frac{1}{n^3}.$$

Proof. First note that the event whose probability we aim to bound is increasing in p (and in c), so increasing c only increases its probability of occurrence. Since $M(c,k)$ is constant for $c \leq x_k$, it suffices to prove the bound for $c = x_k$, and the case $c < x_k$ follows. Hence, we now assume that $c \geq x_k$. Note that in this case $c/720 - \log(4(c+2k)) \geq 5$. For all n, and any $j \in [n]$, we have

$$\begin{aligned}&\mathbb{P}(\text{there exists } S \in B_j,\ e(S,S) > (k+1+10c)\max(j,\log n))\\ &\quad\leq \sum_{S\subset[n],\,|S|=j}\mathbb{P}(e(S,S) > (k+1+10c)\max(j,\log n),\ S \in B_j).\end{aligned}\tag{6}$$

Write $\mathbf{T}_S$ for the set of all possible trees on vertex set S (so $|\mathbf{T}_S| = |S|^{|S|-2}$) and list the elements of $\mathbf{T}_S$ as $t_1,\ldots,t_r$. For $i \in [r]$, let F_i be the event that t_i is a subgraph of H, and let $E_i = F_i \setminus \bigcup_{j<i} F_j$ be the event that t_i is a subgraph of H but none of $t_1,\ldots,t_{i-1}$ are subgraphs of H. The events E_i partition the event that $S \in B_j$, so

$$\begin{aligned}&\mathbb{P}(e(S,S) > (k+1+10c)\max(j,\log n) \mid S \in B_j)\\ &\quad\leq \max_{i\in[r]}\mathbb{P}(e(S,S) > (k+1+10c)\max(j,\log n) \mid E_i).\end{aligned}\tag{7}$$

For $i \in [r]$, write $\mathbb{P}_i(\cdot)$ for the conditional probability measure $\mathbb{P}(\cdot \mid F_i)$.

Write C for the event that $|E(S,S)\setminus E(t_i)| > (k+1+10c)\max(j,\log n)-(j-1)$. Then we have

$$\mathbb{P}(e(S,S) > (k+1+10c)\max(j,\log n) \mid E_i) = \mathbb{P}_i\left(C \,\middle|\, \bigcap_{j<i} F_j^c\right).$$

Since C is increasing and $\bigcap_{j<i} F_j^c$ is decreasing, it follows from Lemma 4 that

$$\mathbb{P}_i\Bigg(C\,\Bigg|\,\bigcap_{j<i} F_j^c\Bigg) \le \mathbb{P}_i(C)$$
$$= \mathbb{P}_i(|E(S,S)\setminus E(t_i)| > (k+1+10c)\max(j,\log n) - (j-1)).$$

Under $\mathbb{P}_i$, the set $E(S,S)\setminus(E(t_i)\cup E(R_{n,k}))$ is distributed as a $\text{bin}(p)$ random subset of $\{uv\colon u,v\in S\}\setminus(E(t_i)\cup E(R_{n,k}))$ since, after conditioning that t_i is a subgraph of H, those edges not in t_i or $R_{n,k}$ still appear independently. Furthermore, $|E(S,S)\cap E(R_{n,k})|\le kj$. It follows that, under $\mathbb{P}_i$, $|E(S,S)\setminus E(t_i)|$ is stochastically dominated by $kj+\text{bin}(\max(j,\log n)n/2,p)$. Letting X have distribution $\text{bin}(n\max(j,\log n)/2,p)$. For all $i\in[r]$ we thus have

$$\begin{aligned}
&\mathbb{P}(e(S,S) > (k+1+10c)\max(j,\log n) \mid E_i)\\
&\quad\le \mathbb{P}(kj+(j-1)+X > (k+1+10c)\max(j,\log n))\\
&\quad\le \mathbb{P}(X > 10c\max(j,\log n))\\
&\quad= \mathbb{P}(X > 20\mathbb{E}X)\\
&\quad\le \mathrm{e}^{-9c\max(j,\log n)/2}.
\end{aligned}$$

The last inequality uses the upper Chernoff bound with $x=19$ (using $19^2/40>9$). It then follows from (6) and (7) that

$$\begin{aligned}
&\mathbb{P}(\text{there exists } S\in B_j,\ e(S,S) > (k+1+10c)\max(j,\log n))\\
&\quad\le \sum_{S\subset[n],\,|S|=j} \mathrm{e}^{-9c\max(j,\log n)/2}\mathbb{P}(S\in B_j)\\
&\quad= \mathrm{e}^{-9c\max(j,\log n)/2}\mathbb{E}|B_j|\\
&\quad\le \mathrm{e}^{-9c\max(j,\log n)/2} n(4(c+2k))^j,
\end{aligned}$$

the last inequality follows by Proposition 1. By assumption, c is large enough such that $c/720-\log(4(c+2k))>5$, and it follows that

$$\begin{aligned}
&\mathbb{P}(\text{there exists } S\in B_j,\ e(S,S) > (k+1+10c)\max(j,\log n))\\
&\quad\le \exp\bigg(\log n + \log(4(c+2k))\max(j,\log n) - \frac{9c\max(j,\log n)}{2}\bigg)\\
&\quad< n^{-4}.
\end{aligned}$$

A union bound over $j\in[n]$ completes the proof.

5. Proof of Theorem 1

As noted in Section 1, the case when c is large is more straightforward, and we handle it first.

Proof of Theorem 1 assuming $c > x_k$. We begin by summarizing the structure of the proof. Recall that Theorem 2 yields an upper bound on the mixing time of $H=H_{n,k,p}$ in terms of the conductance profile function $\Phi(x)$, i.e.

$$\tau_{\text{mix}}(H) \le C \sum_{i=1}^{\lceil \log_2 |E|\rceil} \Phi^{-2}(2^{-i}), \tag{8}$$

where

$$\Phi(x) = \min_{\substack{S \text{ connected} \\ x|E| \le e(S) \le 2x|E|}} \Phi(S),$$

for $0 \le x \le \frac{1}{2}$.

Let A be the event that for all $S \subset [n]$ with $|S| > 9n/10$ we have $e(S) > |E(H)|$. Let A' be the event that $n(c/2+k)/2 \le |E(H)| \le 2n(c/2+k)$.

Note that $|E(H)| \stackrel{\mathrm{D}}{=} nk + \mathrm{bin}(n(n-2k-1)/2, p)$, so by a Chernoff bound, for all n large enough, $\mathbb{P}(A') \ge 1 - n^{-3}$. Also, by Lemma 6, $\mathbb{P}(A) \ge 1 - n^{-3}$ for all n sufficiently large. Thus, $\mathbb{P}(A \cap A') \ge 1 - 2n^{-3}$. In analyzing the upper bound in (8), we would like to work conditionally upon $A \cap A'$. However, it turns out to be technically more straightforward to instead replace Φ by a function Φ_0 with the property that, for all x, $\Phi_0(x) \le \Phi(x)$ whenever $A \cap A'$ occurs. More precisely, for $x \ge 0$ write

$$\Phi_0(x) = \min\left\{ \frac{e(S, S^c)}{e(S)} : S \text{ connected}, |S| \le \frac{9n}{10}, \frac{xn(c/2+k)}{2} \le e(S) \le 4xn\left(\frac{c}{2}+k\right)\right\}. \tag{9}$$

If A occurs then, for all $0 \le x \le \frac{1}{2}$ and all $\hat{S} \subset [n]$ with $e(\hat{S}) \le 2x|E(H)|$, we have $e(\hat{S}) \le |E(H)|$, so $|\hat{S}| \le 9n/10$. If A' occurs then, for all $x \ge 0$ and all $\hat{S}$ with $x|E(H)| \le e(\hat{S}) \le 2x|E(H)|$, we have $xn(c/2+k)/2 \le e(\hat{S}) \le 4xn(c/2+k)$. It follows that, on $A \cap A'$, for all $0 \le x \le \frac{1}{2}$ we have $\Phi_0(x) \le \Phi(x)$ since $\Phi_0(x)$ minimizes over a larger set than $\Phi(x)$. From the preceding sentence and (8) (on the event $A \cap A'$), it follows that, with a probability of at least $1 - 2n^{-3}$,

$$\tau_{\mathrm{mix}}(H) \le \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i}) \le \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi_0^{-2}(2^{-i}).$$

We now focus on bounding the latter quantity. Note that for $i < i'$, if S and S', respectively, appear in the minimum in (9), then $e(S) > e(S')$. Also, recall the definitions of x_k and $M = M(c,k)$ from just before the statement of Lemma 8. Let

$$k_n := \left\lfloor \log_2\left(\frac{n(c/2+k)}{8M \log n}\right)\right\rfloor = \left\lfloor \log_2 n - \log_2 \log n + \log_2\left(\frac{c+2k}{16M}\right)\right\rfloor.$$

We split the sum into two cases: $i \ge k_n$ and $i \le k_n$.

First we consider $i \ge k_n$, and let $j = i - k_n$. For all S considered when bounding $\Phi_0(2^{-i})$ we have $e(S) \le 2^{-j+6} M \log n$ and $e(S, S^c) \ge 2$, so

$$\begin{aligned}
\sum_{i=k_n}^{\lceil \log_2 |E| \rceil} \Phi_0^{-2}(2^{-i}) &= \sum_{j=0}^{\lceil \log_2 |E| \rceil - k_n} \Phi_0^{-2}(2^{-k_n-j}) \\
&\le 4^5 M^2 \log^2 n \sum_{j=0}^{\infty} 4^{-j} \\
&= \frac{4^6 M^2}{3} \log^2 n.
\end{aligned}$$

Next suppose that $i \le k_n$. In this case we have $e(S) \ge 4M \log n$. By Lemma 8, with a probability of at least $1 - n^{-3}$, for all connected sets S with $|S| \le \log n$ we have

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

$e(S, S) \le M \log n$. Since $e(S) \ge 4M \log n$, this implies that

$$e(S, S^c) = e(S) - 2e(S, S) \ge e(S) - 2M \log n \ge \frac{e(S)}{2},$$

so $\Phi(S) \ge \frac{1}{2}$. Also, by Lemmas 5 and 8, with a probability of at least $1 - 2n^{-3}$, for all connected sets S with $|S| \ge \log n$, we have

$$\frac{e(S, S^c)}{e(S)} = \frac{e(S, S^c)}{e(S, S^c) + 2e(S, S)} \ge \frac{c|S|}{12} \frac{1}{c|S|/12 + 2M|S|} \ge \frac{c}{36M},$$

so $\Phi(S) \ge c/(36M)$. Since $\frac{1}{2} > c/36M$ it follows that, with a probability of at least $1 - 3n^{-3}$, for all $i \le k_n$ we have $\Phi_0(2^{-i}) \ge c/36M$, and in this case

$$\sum_{i=1}^{k_n} \Phi_0^{-2}(2^{-i}) \le \frac{6^4 M^2}{c^2} \log_2 n.$$

Combining these bounds, we see that, with a probability of at least $1 - 3n^{-3}$,

$$\sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi_0^{-2}(2^{-i}) \le \frac{4^6 M^2}{3} \log^2 n + \frac{6^4 M^2}{c^2} \log_2 n,$$

so with a probability of at least $1 - 5n^{-3}$, $\sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i})$ is at most the same quantity. By Theorem 2 it follows that, with a probability of at least $1 - 5n^{-3}$,

$$\tau_{\text{mix}}(G) \le C\left(\frac{4^6 M^2 \log^2 n}{3} + \frac{6^4 M^2}{c^2} \log_2 n\right).$$

This completes the proof in the case $c > x_k$.

For the remainder of this paper, we fix $0 < c < x_k$ and let $R = \lceil \max(k, 2x_1/c) \rceil$. Also, recall the constant $\beta = \beta(c)$ from Lemma 7. The remaining case of Theorem 1 follows straightforwardly from the following lemma.

Lemma 9. *There exists a $\alpha = \alpha(c) > 0$ such that, for all n sufficiently large,*

$$\mathbb{P}\left(\text{there exists } S \in \bigcup_{R \log n \le j \le (1-\beta)n} B_j(H),\ e_H(S, S^c) \le \alpha |S|\right) \le \frac{3R^3}{n^3}.$$

We provide the proof of Lemma 9 at the end of this section.

Proof of Theorem 1 assuming $c \le x_k$. For $x \ge 0$ write

$$\Phi_0(x) = \min\left\{\frac{e(S, S^c)}{e(S)} : S \text{ connected}, |S| \le (1-\beta)n, \frac{xn(c/2 + k)}{2} \le e(S) \le 4xn\left(\frac{c}{2} + k\right)\right\}.$$

As in the case $c > x_k$, by a Chernoff bound and by Lemma 7, for all n sufficiently large, with a probability of at least $1 - 2n^{-3}$ we have

$$\sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi^{-2}(2^{-i}) \le \sum_{i=1}^{\lceil \log_2 |E| \rceil} \Phi_0^{-2}(2^{-i}),$$

and the remainder of the proof is just as in the case $c > x_k$, but using Lemma 9 in place of Lemma 5.

It now remains to prove Lemma 9; before doing so, we briefly describe our approach. We shall divide vertices of H into groups of size R, each containing R consecutive vertices. We view each group as a new single vertex; two new vertices are connected if there is an edge connecting their constituent sets. This yields an auxiliary graph H', whose distribution is that of an $(n', 1, p')$ Newman–Watts small-world model, for suitable n' and p'. We will shortly see that $p' = c'/n'$ for some $c' > x_1$, so all 'large-c' results can be applied to H'.

To translate edge expansion results from H' into corresponding results for H, we proceed as follows. Given a set S of vertices of H, we consider the *blow-up* S^+ of S, which is the collection of all vertices of H belonging to the same group as some element of S. The idea is that in most cases, the event that $e(S, S^c)$ is small relative to $|S|$ should be nearly identical to the event that $e(S^+, (S^+)^c)$ is small relative to $|S^+|$. If this were always true, Lemma 5 would then yield bounds for the number of edges leaving S^+, which would in turn yield strong bounds on the probability that $e(S^+, (S^+)^c) < \varepsilon\,|S^+|$, where $\varepsilon > 0$ will be a function of R.

The above line of argument relies upon the intuition that the size of the blow-up S^+ should be essentially a constant factor greater than that of S. Since the ratio $|S^+|/|S|$ is in fact a random quantity, to make the above argument work, we end up needing to additionally show that $e(S^+, (S^+)^c)$ is very unlikely to be large if $e(S, S^c)/|S|$ is extremely small. In order to quantify the notion of 'extremely small', we are forced to introduce a third parameter $\delta > 0$, with δ much smaller than ε. We now turn to the details.

Proof of Lemma 9. We assume for simplicity that R divides n (the general case is practically identical) and write $n' = n/R$. For $i \in [n']$ let $w_i = \{(i-1)R + j,\ 1 \le j \le R\}$. We form an auxiliary graph $H' = (V', E')$ with $V' = \{w_i,\ i \in [n']\}$ by adding an edge between w_i and w_j if there is some edge from an element of w_i to an element of w_j in H. It is easily verified (using the fact that $R > k$) that H' is an $(n', 1, p')$-Newman–Watts small-world model, with $p' = \mathbb{P}(\mathrm{bin}(R^2, p) > 0) > Rc/2n' = c'/n'$ (where $c' = Rc/2$) for all n sufficiently large. Note that since $c' = Rc/2 > x_1$, it follows that we may apply Lemma 5 to H' (this is the only way we will use this bound on R).

Given $S \subset [n]$, write $I' = \{i \in n' \colon w_i \cap S \ne \varnothing\}$, let $S' = \{w_i,\ i \in I'\}$, and write $S^+ = \bigcup_{i \in S'} w_i$. Now, write $j = |S|$. As in the proof of Lemma 8, we partition the event that $S \in B_j(H)$ into $E_1(S), \dots, E_r(S)$ according to the first spanning tree appearing in S. Fix $\varepsilon = c/12R(2Rc+1)$ and let $\delta > 0$ be small enough so that $\varepsilon c \ge 2k\delta$ and that $\varepsilon Rc(\log(\varepsilon/\delta) - 1) \ge 5 + \log(4(c+2k))$. For all $1 \le i \le r$ we then have

$$
\begin{aligned}
&\mathbb{P}(e_H(S^+ \setminus S, (S^+)^c) > 2\varepsilon Rcj \mid e_H(S, S^c) \le \delta j, E_i(S)) \\
&\quad \le \mathbb{P}(e_H(S^+ \setminus S, (S^+)^c) > (\varepsilon c + 2k\delta)Rj \mid e_H(S, S^c) \le \delta j, E_i(S)) \\
&\quad \le \mathbb{P}(e_H(S^+ \setminus S, (S^+)^c) > (\varepsilon c + 2k\delta)Rj \mid e_H(S, S^c) \le \delta j, t_i \subset H).
\end{aligned}
$$

The first inequality is true since we pick δ such that $\varepsilon c \ge 2k\delta$ and the second inequality holds by Lemma 4. Now write $g(S) = |\{i \in S' \colon |S \cap w_i| < |w_i|\}|$, so $g(S)$ is the number of sets w_i that intersect S but are not covered by S. It is easily checked that $e_H(S, S^c) \ge g(S)$; the extremal case is that for each $i \in S'$, $e_H(S \cap w_i, w_i \setminus S) = 1$, while for $i, j \in S'$ with $i \ne j$, $e_H(w_i, w_j) = 0$.

Next, suppose that $S \subset [n]$ satisfies $e_H(S, S^c) \le \delta j$. Then we must have $g(S) \le \delta j$, and it follows that $|S^+ \setminus S| \le R\delta j$. Under this conditioning, $|E_H(S^+ \setminus S, (S^+)^c) \setminus E(R_{n,k})|$ is independent of the event that $t_i \subset H$ since they are determined by disjoint sets of edges and it is stochastically dominated by $\mathrm{bin}(R\delta jn, p)$. It follows that $e_H(S^+ \setminus S, (S^+)^c)$ is stochastically dominated by $2kR\delta j + \text{bin}(R\delta jn, p)$, so by the finer of the Chernoff upper bounds,

$$\mathbb{P}(e_H(S^+ \setminus S, (S^+)^c) > (\varepsilon c + 2k\delta)Rj \mid e_H(S, S^c) \le \delta j, t_i \subset H) \\ \le \exp\left(-\varepsilon Rcj\left(\log\left(\frac{\varepsilon}{\delta}\right) - 1\right)\right).$$

It follows that, for $j \ge R \log n$, writing

$$\mathcal{S}_1 = \{S \subset [n], S \in B_j(H), e_H(S, S^c) \le \delta j, e_H(S^+ \setminus S, (S^+)^c) > 2\varepsilon Rcj\},$$

we have

$$\begin{aligned}\mathbb{E}|\mathcal{S}_1| &\le \sum_{S\subset[n],\, |S|=j} \sum_{1\le i\le r} \mathbb{P}(E_i(S), e_H(S, S^c) \le \delta|S|, e_H(S^+ \setminus S, (S^+)^c) > 2\varepsilon Rc|S|) \\ &\le \exp\left(-\varepsilon Rcj\left(\log\left(\frac{\varepsilon}{\delta}\right) - 1\right)\right)\mathbb{E}|\{S \subset [n], |S| = j, H[S] \text{ connected}\}| \\ &\le \exp\left(-\varepsilon Rcj\left(\log\left(\frac{\varepsilon}{\delta}\right) - 1\right)\right)n(4(c + 2k))^j \\ &\le n^{-4},\end{aligned}$$

the second-to-last inequality is by Proposition 1, and the last inequality follows since we chose δ such that

$$\varepsilon Rc\left(\log\left(\frac{\varepsilon}{\delta}\right) - 1\right) \ge 5 + \log(4(c + 2k))$$

and since $j \ge R \log n \ge \log n$. It follows by a union bound over $R \log n \le j \le 9n/10$ and Markov's inequality that

$$\mathbb{P}\left(\text{there exists } S \in \bigcup_{R\log n\le j\le 9n/10} B_j(H), e_H(S, S^c) \le \delta|S|, e_H(S^+ \setminus S, (S^+)^c) > 2\varepsilon Rc|S|\right) \\ \le \frac{1}{n^3}. \tag{10}$$

Next, write

$$\mathcal{S}_2 = \left\{S \in \bigcup_{R\log n\le j\le 9n/10R} B_j(H), e_H(S, S^c) \le \varepsilon|S|, e_H(S^+ \setminus S, (S^+)^c) \le 2\varepsilon Rc|S|\right\}.$$

For any $S \in \mathcal{S}_2$, we have

$$\begin{aligned}e_H(S^+, (S^+)^c) &= e_H(S, (S^+)^c) + e_H(S^+\setminus S, (S^+)^c) \\ &\le e_H(S, S^c) + e_H(S^+\setminus S, (S^+)^c) \\ &\le \varepsilon(2Rc + 1)|S| \\ &\le \varepsilon(2Rc + 1)R|S'| \\ &= \frac{c'|S'|}{12},\end{aligned}$$

https://doi.org/10.1239/aap/1427814580 Published online by Cambridge University Press

where the last equality follows by our choice of ε. Then we obtain that $e_{H'}(S', (S')^c) \le c'|S'|/12$. Furthermore, since $|S| \ge R\log n$ we have $|S'| \ge \log n \ge \log n'$, and since $|S| \le 9n/10R$ we have $|S'| \le 9n/10R = 9n'/10$. It follows by Lemma 5 that

$$\begin{aligned}
&\mathbb{P}\Bigg(\text{there exists } S \in \bigcup_{R\log n \le j \le 9n/10R} B_j(H), e_H(S, S^c) \le \varepsilon|S|, \\
&\quad e_H(S^+ \setminus S, (S^+)^c) \le 2\varepsilon Rc|S|\Bigg) \\
&\quad \le \mathbb{P}\Bigg(\text{there exists } S' \in \bigcup_{\log n' \le j \le 9n'/10} B_j(H'), e_{H'}(S', (S')^c) \le \frac{c'|S'|}{12}\Bigg) \\
&\quad \le \frac{1}{(n')^3}.
\end{aligned} \tag{11}$$

Next, for any $m \ge 1$, if $e_H(S, S^c) \le m$ then, viewed as a subset of a cycle of length n, S must have at most m connected components. The number of subsets of an n-cycle with at most m connected components is $2n\binom{n+2m-1}{2m-1}$. (This is a straightforward combinatorial exercise but may be seen as follows: the factor n chooses a starting point on the cycle, the factor $\binom{n+2m-1}{2m-1}$ chooses the points on the cycle at which membership in S alternates, and the factor 2 accounts for whether or not the starting point belongs to S.)

It follows that, for any $\gamma > 0$, the number of subsets of an n-cycle with at most γn connected components is at most

$$\begin{aligned}
2n\binom{n+\lfloor 2\gamma n\rfloor}{\lfloor 2\gamma n\rfloor} &\le 2n\left(\frac{e(1+2\gamma)n}{2\gamma n}\right)^{2\gamma n} \\
&= \exp\left(\log(2n) + n2\gamma\left(1 + \log\left(1 + \frac{1}{2\gamma}\right)\right)\right).
\end{aligned} \tag{12}$$

Since $x(1+\log(1/2x)) \to 0$ as $x \downarrow 0$, we may choose $0 < \gamma < 9\beta c/20R$ small enough such that (12) is at most $\exp(n9\beta c/160R)$ for all n sufficiently large.

Finally, for a fixed set S with $9n/10R \le |S| \le (1-\beta)n$, the probability that $e_H(S, S^c) \le \gamma n$ is bounded above by $\mathbb{P}(\mathrm{bin}((9n/10R)\beta n, p) \le \gamma n)$, since $e_H(S, S^c)$ stochastically dominates $\mathrm{bin}(|S|(n-|S|), p)$ and $|S| \ge 9n/10R$, $n - |S| \ge \beta n$. Since $10R\gamma/9\beta c \le \frac{1}{2}$, by a Chernoff bound we have

$$\mathbb{P}(e_H(S, S^c) \le \gamma n) \le \mathbb{P}\left(\mathrm{bin}\left(\frac{9n}{10R}\beta n, p\right) \le \gamma n\right) \le \exp\left(-\frac{9\beta c}{80R}n\right). \tag{13}$$

Since *all* sets $S \subset [n]$ with at least γn components (still viewed as subsets of the n-cycle) have $e(S, S^c) \ge \gamma n$, it follows by (12), (13), and a union bound over sets with at most γn components that, for all n sufficiently large,

$$\begin{aligned}
&\mathbb{P}\Bigg(\text{there exists } S \in \bigcup_{9n/10R \le j \le (1-\beta)n} B_j(H), e_H(S, S^c) \le \gamma n\Bigg) \\
&\quad \le 2n\binom{n+\lfloor 2\gamma n\rfloor}{\lfloor 2\gamma n\rfloor}\exp\left(-\frac{9\beta c}{80R}n\right)
\end{aligned}$$

$$\leq \exp\left(-\frac{9\beta c}{160R}n\right)$$
$$\leq \frac{1}{n^3}. \tag{14}$$

Writing $\alpha = \min(\gamma, \varepsilon, \delta)$, it follows from (10), (11), and (14) that

$$\mathbb{P}\left(\text{there exists } S \in \bigcup_{R\log n \leq j \leq (1-\beta)n} B_j(H), e_H(S, S^c) \leq \alpha|S|\right) \leq \frac{3}{(n')^3}.$$

Since $n' = n/R$ this completes the proof.

6. Conclusion

We remark that a slightly different model was proposed in [17] in which random edges can be added between any pair of sites of the ring. This results in a multi-graph model. The current work is very easily adapted to the latter model. For both models, the only difference appears in Lemma 6 of Section 4 and bounding $E(H)$ in the proof of Theorem 1 in Section 5. In both cases we would have to use $\text{bin}(n(n-1)/2, p)$ instead of $\text{bin}(n(n-2k-1)/2, p)$ for the 'multi-graph' version. But when we actually work on the computation with a Chernoff bound, this makes no difference for our estimation (e.g. we can always use the stochastic dominating random variable $\text{bin}(n^2, p)$ for both estimations).

Acknowledgements

The authors are grateful to Omer Angel for suggesting the approach to use Lagrange inversion formula to count the number of connected subgraphs of the Newman–Watts small-world model [1]. The authors also would like to thank an anonymous referee for a detailed reading of the paper and for many helpful remarks.

References

[1] Angel, O. (2011). Answer to: Counting subtrees of a random tree ("random Catalan numbers"). Available at http://mathoverflow.net/questions/66595.
[2] Bollobás, B. and Chung, F. R. K. (1988). The diameter of a cycle plus a random matching. *SIAM J. Discrete Math.* **1,** 328–333.
[3] Brightwell, G. and Winkler, P. (1990). Maximum hitting time for random walks on graphs. *Random Structures Algorithms* **1,** 263–276.
[4] Condamin, S. *et al.* (2007). First-passage times in complex scale-invariant media. *Nature* **450,** 77–80.
[5] Durrett, R. (2007). *Random Graph Dynamics.* Cambridge University Press.
[6] Farkas, I. J., Derényi, I., Barabási, A. L. and Vicsek, T. (2001). Spectra of 'real-world' graphs: beyond the semicircle law. *Phys. Rev. E* **64,** 026704.
[7] Fountoulakis, N. and Reed, B. A. (2007). Faster mixing and small bottlenecks. *Prob. Theory Relat. Fields* **137,** 475–486.
[8] Gallos, L. K., Song, C., Havlin, S. and Makse, H. A. (2007). Scaling theory of transport in complex biological networks. *Proc. Nat. Acad. Sci. USA* **104,** 7746–7751.
[9] Janson, S., Łuczak, T. and Ruciński, A. (2000). *Random Graphs*. John Wiley, New York.
[10] Jasch, F. and Blumen, A. (2001). Target problem on small-world networks. *Phys. Rev. E* **63,** 041108.
[11] Jespersen, S., Sokolov, I. M. and Blumen, A. (2000). Relaxation properties of small-world networks. *Phys. Rev. E* **62,** 4405.
[12] Levin, D. A., Peres, Y. and Wilmer, E. L. (2009). *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI.
[13] Lovász, L. and Winkler, P. (1998). Mixing times. In *Microsurveys in Discrete Probability* (Princeton, NJ, 1997; DIMACS Ser. Discrete Math. Theoret. Comput. Sci. **41**), American Mathematical Society, Providence, RI, pp. 85–133.

[14] Monasson, R. (1999). Diffusion, localization and dispersion relations on 'small-world' lattices. *Europ. Phys. J. B* **12,** 555–567.
[15] Montenegro, R. and Tetali, P. (2006). Mathematical aspects of mixing times in Markov chains. *Found. Trends Theoret. Comput. Sci.* **1,** 121pp.
[16] Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Rev.* **45,** 167–256.
[17] Newman, M. E. J. and Watts, D. J. (1999). Renormalization group analysis of the small-world network model. *Phys. Lett. A* **263,** 341–346.
[18] Newman, M. E. J. and Watts, D. J. (1999). Scaling and percolation in the small-world network model. *Phys. Rev. E* **60,** 7332.
[19] Newman, M. E. J., Barabási, A.-L. and Watts, D. J. (eds) (2006). *The Structure and Dynamics of Networks*. Princeton University Press.
[20] Noh, J. D. and Rieger, H. (2004). Random walks on complex networks. *Phys. Rev. Lett.* **92,** 118701.
[21] Riordan, O. and Wormald, N. (2010). The diameter of sparse random graphs. *Combin. Prob. Comput.* **19,** 835–926.
[22] Stanley, R. P. (1999). *Enumerative Combinatorics*, Vol 2. Cambridge University Press.
[23] Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* **393,** 440–442.