EXPONENTIAL RANDOM GRAPHS AS MODELS OF OVERLAY NETWORKS

M. DRAIEF,* *Imperial College London*

A. GANESH,** *University of Bristol*

L. MASSOULIÉ,*** *Thomson Research*

Abstract

In this paper we give an analytic solution for graphs with n nodes and $E = cn \log n$ edges for which the probability of obtaining a given graph G is $\mu_n(G) = \exp(-\beta \sum_{i=1}^n d_i^2)$, where d_i is the degree of node i. We describe how this model appears in the context of load balancing in communication networks, namely peer-to-peer overlays. We then analyse the degree distribution of such graphs and show that the degrees are concentrated around their mean value. Finally, we derive asymptotic results for the number of edges crossing a graph cut and use these results (i) to compute the graph expansion and conductance, and (ii) to analyse the graph resilience to random failures.

Keywords: Exponential random graph; peer-to-peer network; overlay optimisation; load balancing; degree distribution; graph cut; expansion; conductance; failure resilience

2000 Mathematics Subject Classification: Primary 60K35; 68R10; 90B18
Secondary 05C07; 05C80; 05C85; 05C90

1. Introduction

Random graphs provide a way of modelling large and complex networks, and of studying stochastic processes on such networks [7, pp. 15–21]. Early work on this topic goes back to the famous random graph or Bernoulli graph introduced by Solomonoff and Rapoport [25] in the early 1950s and studied by Erdös–Rényi [8] a decade later. The Bernoulli random graph model is, however, rather simplistic and fails to capture important features of many real-world networks. This has stimulated work on a number of other random graph models. Exponential random graphs were first introduced in the early 1980s by Holland and Leinhardt [15] based on the work of Besag [2]. More recently, Frank and Strauss [9] studied a subclass of these graphs, namely Markov graphs. They correspond to log-linear statistical models of random graphs with general dependence structure and Markov dependence [4, Chapter 7] widely used by statisticians and social network analysts [24].

To motivate the study of such graphs, we consider the situation where we have measurements of a number of network properties, or observables, for a real-world network, and wish to come up with a network model that exhibits similar properties. Denote these observables by $(x_i)_{i=1,...,k}$ and denote by $(\bar{x}_i)_{i=1,...,k}$ their measured average values. Let $\mathcal{G}$ be a set of graphs, and let G be a graph in $\mathcal{G}$. To describe a family of graphs that reproduce the graph's observed properties,

we wish to choose a probability distribution μ on $\mathcal{G}$ such that

$$\sum_{G \in \mathcal{G}} \mu(G) x_i(G) = \bar{x}_i \quad \text{for all } i = 1, \ldots, k, \tag{1}$$

where $x_i(G)$ is the value taken by x_i in the graph G. Clearly, there are infinitely many such probability distributions; a popular choice is the one that maximises the Gibbs or Shannon entropy

$$S = -\sum_{G \in \mathcal{G}} \mu(G) \log \mu(G)$$

subject to (1) and the normalising condition $\sum_{G \in \mathcal{G}} \mu(G) = 1$. Introducing Lagrange multipliers we can easily show [23] that the maximum entropy is achieved for the distribution

$$\mu(G) = \frac{1}{Z} e^{-H(G)}, \qquad H(G) = \sum_{i=1}^{k} \theta_i x_i(G), \tag{2}$$

where $Z = \sum_{G \in \mathcal{G}} e^{-H(G)}$ is the normalising constant and the constants $(\theta_i)_{i=1,\ldots,k}$ can be expressed in terms of the measured average values $(\bar{x}_i)_{i=1,\ldots,k}$. Graphs drawn according to distributions defined by (2) are called *exponential random graphs*. They are random graphs with maximum entropy subject to the specified constraints.

Exponential random graphs can be generated using suitable random walks on the space of graphs, for which they arise as the stationary distribution. More precisely, given $H(G)$, a cost or energy function associated with the graph G, define the Markov chain on $\mathcal{G}$ with transition

$$p_{G,G'} = \min(1, e^{-(H(G') - H(G))}).$$

It can easily be shown that the transition matrix fulfills the detailed balance condition (the Markov chain is reversible) and the corresponding stationary distribution is given by the Boltzmann-type probability distribution $\mu(G) = Z^{-1} e^{-H(G)}$.

In this paper we study the particular case of graphs with n nodes and $E = cn \log n$ edges for which $H(G) = \sum_{i=1}^{n} d_i^2$, where d_i is the degree of node i. This model appears in the context of load balancing in certain communication networks, namely peer-to-peer overlays as described in Section 2. In Section 3 we present our main results. We analyse the degree distribution of such graphs in Section 4 and show that the degrees are concentrated around their mean value with high probability (w.h.p.). In Section 5 we derive asymptotic results on the number of edges crossing a graph cut and use these results (i) to compute the graph expansion and conductance in Subsection 5.1, and (ii) to analyse the graph resilience to random failures in Subsection 5.2. (This paper expands on an earlier short version which appeared in the proceedings of the 41st Allerton Conference on Communications, Control and Computing [11].)

2. Load balancing in overlay networks

Peer-to-peer (P2P) systems are decentralised networks enabling users to contribute resources for mutual benefit. File sharing using P2P networks has gained wide popularity and it has recently been suggested that it is the dominant consumer of bandwidth ahead of Web traffic [26]. In these systems users form a network or *overlay* of peers that can act as both clients and servers alleviating the bottlenecks that appear in centralised schemes relying on servers to provide the content. Early P2P systems relied on structured overlays wherein users are allocated identifiers using digital hash tables. These structured overlays are chosen to facilitate routeing between

nodes and searching for data. More recently, unstructured overlays have been advocated as a viable alternative to cope with some of the most important features of P2P networks [18].

In what follows we will model an overlay as a graph with n nodes representing the peers connected by edges describing whether two peers know each other or not. We assume that the 'who knows who' relationship is symmetric, i.e. the graph is undirected. We will call the nodes that a given node knows its neighbours, the degree of a node being the number of its acquaintances. In [12], an algorithm was described that ensures the construction of an Erdös–Rényi-like overlay, wherein any pair of peers is connected with a given probability independently from other pairs. It is known that such graphs are connected, w.h.p., if the order of the mean degree of nodes is higher than $\log n$ [3, Theorem 7.3], and the result is true for more general graphs [1]. These results highlight the fact that the overlay building protocol should ensure that each peer is provided with a set of neighbours of size $c \log n$, $c > 1$, so that the memory requirements on each member grows slowly in the overlay size whilst having the desired connectedness property.

Let $G = (V, \mathcal{E})$ be a graph where $V = \{1, \ldots, n\}$ is the set of nodes and $\mathcal{E}$ is the set of edges of size E. We will denote by d_i the degree of node i. In the rest of the paper we will focus on connected graphs with

$$\sum_{i=1}^{n} d_i = 2E = cn \log n, \qquad c > 1.$$

We now describe a distributed iterative scheme that will modify the initial overlay to provide the nodes with neighbourhoods of size *concentrated* around $\log n$. Define the energy of a graph G by

$$H(G) = \sum_{i=1}^{n} d_i^2.$$

Starting from a graph G, periodically at rate 1, each node i performs the following steps.

1. Choose uniformly at random two neighbours j and k of i.

2. Evaluate the cost of replacing link $\{i, j\}$ by link $\{j, k\}$, i.e.

$$\Delta H = 2(d_k - d_i + 1).$$

3. Replace link $\{i, j\}$ by link $\{j, k\}$ with probability

$$\min\left(\left(e^{-\beta \Delta H} \frac{d_i(d_i - 1)}{d_k(d_k + 1)}\right), 1\right).$$

The above algorithm is a particular instance of the *Metropolis algorithm* [21] (see [4, Chapter 7] for more details). It is parametrised by the constant β that corresponds to the inverse of a temperature and provides a trade-off between accuracy, in terms of how close we get to the optimal configuration, and speed of convergence. It defines a random walk on the set of connected graphs with a given number of edges $E = \frac{1}{2} \sum_{i=1}^{n} d_i = cn \log(n)/2$. In particular, the stationary distribution of this Markov chain is given by the following Gibbs distribution [4, Chapter 7]:

$$\mu_n(G) = \frac{1}{Z} \exp\left(-\beta \sum_{i=1}^{n} d_i^2\right) \mathbf{1}_{\{\sum_{i=1}^{n} d_i = cn \log n\}}, \tag{3}$$

where Z is a normalizing constant.

Starting from a connected graph, the above algorithm retains connectedness. In [10], the authors introduced a similar algorithm dubbed the *localiser* that takes into account locality awareness. More precisely, the authors included an additional term in the energy function that accounted for the cost of the communication between two nodes. Using simulations, they showed that in the graph obtained the set of neighbours of a node consists mainly of nearby nodes.

3. Main results

The main results of this paper state that, for graphs generated according to (3), we have the following assertions.

- (See Theorem 1, below.) Given $\overline{d} = \sum_{i=1}^{n} d_i / n$, the degrees of the nodes d_i are concentrated around $\overline{d} = \sum_{i=1}^{n} d_i / n$, i.e. there exists a constant $\alpha > 0$, such that

$$\max_{1 \leq i \leq n} |d_i - \overline{d}| \leq \sqrt{\alpha \log n} \quad \text{w.h.p.}$$

- (See Theorem 4, below.) Let U be a subset of the set of nodes V of size u, and let e_{U,U^c} be the number of links between U and its complement U^c. There exists a constant $\tilde{\delta}$ such that

$$e_{U,U^c} \geq (1 - \tilde{\delta}) cu \log n \quad \text{w.h.p.}$$

In the terms of the overlay networks this means that, if we assume that U is the set of nodes holding the data that all nodes are interested in, there is a large number of paths through which the data could be sent to the nodes in U^c that have not acquired it yet.

- (See Theorem 6, below.) The overlay network constructed above is resilient to node failures, i.e. it remains connected, w.h.p., despite nodes failing at random with probability p as long as $p < \exp(-1/c(1 - \tilde{\delta}))$.

In the rest of the paper we give detailed proofs of these results.

4. Degree distribution

We work with labelled graphs throughout. We consider random graphs on n nodes with E edges generated according to distribution (3). Our aim in this section is to show that the degree sequence of a graph generated according to (3) is concentrated around its mean.

The probability measure μ_n on graphs induces a probability measure on degree distributions, which we denote by π_n. For $\boldsymbol{d} = (d_1, \ldots, d_n)$,

$$\pi_n(\boldsymbol{d}) = \frac{1}{Z_n} G_n(\boldsymbol{d}) \exp\left(-\beta \sum_{i=1}^{n} d_i^2\right) \mathbf{1}_{\{\sum_{i=1}^{n} d_i = cn \log n\}},$$

where $G_n(\boldsymbol{d})$ is the number of graphs having the degree sequence $\boldsymbol{d}$ and Z_n is a normalizing constant. We can rewrite the above as

$$\pi_n(\boldsymbol{d}) = \frac{1}{Z_n(\gamma)} \left(\frac{E! \, 2^E}{(2E)!} G_n(\boldsymbol{d}) \prod_{i=1}^{n} d_i!\right) \prod_{i=1}^{n} \frac{1}{d_i!} \exp(-\beta d_i^2 + \gamma (\log n) d_i) \mathbf{1}_{\{\sum_{i=1}^{n} d_i = cn \log n\}}$$

$$= \frac{\tilde{G}_n(\boldsymbol{d})}{Z_n(\gamma)} \prod_{i=1}^{n} \frac{1}{d_i!} \exp(-\beta d_i^2 + \gamma (\log n) d_i) \mathbf{1}_{\{\sum_{i=1}^{n} d_i = cn \log n\}}, \tag{4}$$

where

$$\tilde{G}_n(\boldsymbol{d}) = \frac{E!\,2^E}{(2E)!} G_n(\boldsymbol{d}) \prod_{i=1}^n d_i!. \tag{5}$$

The introduction of the tilt parameter γ does not change the distribution as it multiplies $\pi_n(\boldsymbol{d})$ by $e^{2\gamma E \log n}$. This is a constant since the total number of edges is fixed. Thus, it can be absorbed into the normalization factor $Z_n(\gamma)$ along with the term $E!\,2^E/(2E)!$.

After sampling a degree sequence according to π_n, we use the *configuration model* to generate a graph $C(n, \boldsymbol{d})$ with this degree sequence [3, p. 52]. To each node i we associate d_i labelled half-edges, also called configuration points or stubs. All stubs need to be paired to construct the graph; this is done by randomly connecting them. When a stub of i is paired with a stub of j, we interpret this as an edge between i and j. The graph $C(n, \boldsymbol{d})$ obtained following this procedure may not be simple, i.e. it may contain self-loops due to the pairing of two stubs of i and multi-edges due to the existence of more than one pairing between two given nodes. The number of configurations with degree sequence $\boldsymbol{d}$ is given by

$$H_n(\boldsymbol{d}) = \frac{(2E)!}{E!\,2^E} \prod_{i=1}^n d_i!.$$

In fact, there are $(2E)!/E!\,2^E$ different ways of pairing the $2E$ configuration points, each corresponding to $\prod_{i=1}^n d_i!$ distinct configurations since the d_i edges incident on node i can be assigned to its d_i configuration points in $d_i!$ ways.

We denote the minimum and maximum degrees by $d_{\min}$ and $d_{\max}$, respectively. To restrict ourselves to the family of simple graphs, we define the *erased configuration model*. Starting from the multigraph obtained through the configuration model, we merge all multiple edges into a single edge and erase all self-loops. It was shown in [27] that, provided that the maximum degree of the graph $d_{\max}$ is such that $d_{\max} = o(\sqrt{n})$, the configuration model and the erased configuration model are asymptotically equivalent, in probability. We will show in Theorem 2, below, that the above condition is indeed satisfied.

Note that every simple graph obtained by the erased configuration model corresponds to exactly $\prod_{i=1}^n d_i!$ distinct configurations describing the number of ways stubs are assigned, and $\tilde{G}_n(\boldsymbol{d})$ introduced in (5) corresponds to the probability of obtaining a simple graph in the configuration model [16]. This implies the upper bound $\tilde{G}_n(\boldsymbol{d}) \leq 1$ for any degree sequence $\boldsymbol{d}$. When $d_{\max} = o(E^{1/4})$, McKay and Wormald [20] established the equivalence, for large n,

$$\tilde{G}_n(\boldsymbol{d}) \sim \exp(-\lambda - \lambda^2), \quad \text{where} \quad \lambda = \frac{1}{4E} \sum_{i=1}^n d_i(d_i - 1). \tag{6}$$

In [27, Proposition 7.5], it was shown that, for any degree sequence $\boldsymbol{d}$, conditional on the graph $C(n, \boldsymbol{d})$ being simple, $C(n, \boldsymbol{d})$ is a uniform simple random graph with degree sequence $\boldsymbol{d}$.

Given a degree sequence $\boldsymbol{d}$, we define the mean degree $\overline{d} = \sum_{i=1}^n d_i/n$ and the variance $\mathrm{var}(\boldsymbol{d}) = (1/n) \sum_{i=1}^n (d_i - \overline{d})^2$. For fixed positive constants α_1 and α_2, we define the following set of degree sequences:

$$A_1(\alpha_1, \alpha_2) = \{\boldsymbol{d} : -\sqrt{\alpha_1 \log n} \leq d_i - \overline{d} \leq \sqrt{\alpha_2 \log n} \text{ for all } i = 1, \ldots, n\}.$$

We now state the main result of this section.

Theorem 1. *For graphs generated according to (3), there exist two constants α_1 and α_2 such that*

$$\pi_n(A_1(\alpha_1, \alpha_2)) \to 1 \quad as \, n \to \infty.$$

The above theorem states that, for the random graph model defined by distribution (3), the node degrees concentrate about their mean value. Specifically, all node degrees are within order $\sqrt{\log n}$ of the mean, w.h.p. This is in contrast to the Erdös–Rényi model (with the same number of edges), where the maximum fluctuation of node degrees is typically of order $\log n$. The rest of the section is devoted to the proof of this theorem.

The main ingredient of the proof is to restrict our attention to sequences of degrees such that $d_i \le n^{1/4}$ for all $i = 1, \ldots, n$, for which estimate (6) holds.

Theorem 2. *Define the event $A_2 = \{\boldsymbol{d} : d_i \le n^{1/4} \text{ for all } i = 1, \ldots, n\}$. Then*

$$\pi_n(A_2^c) \to 0 \quad as \, n \to \infty.$$

To prove this result, we first state a series of lemmas which are proved in Appendix A.

For $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$, since $E = cn \log n / 2$, we have $d_{\max} = o(E^{1/4})$. Observe from (6) that

$$4E\lambda = n(\text{var}(\boldsymbol{d}) + \overline{d}^2 - \overline{d}).$$

Moreover, for $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$, we have $\text{var}(\boldsymbol{d}) \le \max\{\alpha_1, \alpha_2\} \log n$, so that

$$\lambda \le \frac{1}{2}\left(c \log n - 1 + \frac{1}{c} \max\{\alpha_1, \alpha_2\}\right). \tag{7}$$

Hence,

$$\boldsymbol{d} \in A_1(\alpha_1, \alpha_2) \quad \Longrightarrow \quad \frac{1}{\tilde{G}_n(\boldsymbol{d})} \sim \exp(\lambda + \lambda^2) \le \exp\left(\frac{c^2 \log^2 n}{2}\right) \tag{8}$$

for sufficiently large n. Recall that $\tilde{G}_n(\boldsymbol{d}) \le 1$ for all $\boldsymbol{d}$ and, in particular, for $\boldsymbol{d} \in A_2^c$, the complement of A_2.

It follows from (4) and (8) that, for sufficiently large n,

$$\frac{\pi_n(A_2^c)}{\pi_n(A_1(\alpha_1, \alpha_2))} \le \exp\left(\frac{c^2 \log^2 n}{2}\right) \frac{\sum_{\boldsymbol{d} \in A_2^c} \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}{\sum_{\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)} \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}. \tag{9}$$

Let $D_1, \ldots, D_n$ be independent and identically distributed (i.i.d.) random variables with

$$P(D_1 = k) = \frac{1}{F(\gamma)} \frac{1}{k!} \exp(-\beta k^2 + \gamma (\log n) k), \qquad k \in \mathbb{N}, \tag{10}$$

where $F(\gamma)$ is a normalization constant. The dependence of the D_i on n and γ has not been made explicit in the notation. We choose γ (depending on n, β, and c) so that $E[D_1] = c \log n$ for a specified constant c; this is possible by the following lemma.

Lemma 1. *Let*

$$d_\gamma = \frac{1}{2\beta}\left(\gamma \log n + \log \log n + \frac{\gamma}{2\beta}\right),$$

and let $k_\gamma - 1$ denote the integer part of d_γ. Then, $E[D_1] - k_\gamma$ and $\text{var}(D_1)$ remain bounded as n tends to ∞. In particular, γ can be chosen so that $E[D_1] = c \log n$.

Moreover, let $\rho = 2\beta(d_\gamma - k_\gamma + \frac{1}{2})$ and

$$\psi(\theta) = \frac{\sum_{j=-\infty}^{\infty} \exp(\theta j - \beta j^2)}{\sum_{j=-\infty}^{\infty} \exp(-\beta j^2)}.$$

Then, the moment generating function of D_1 satisfies

$$\mathrm{E}[\exp(\theta D_1)] \sim \exp(\theta k_\gamma) \frac{\psi(\theta + \rho)}{\psi(\rho)} \qquad as \ n \to \infty.$$

Proof. See Appendix A.

Let $\boldsymbol{D}$ denote the random vector $(D_1, \ldots, D_n)$, and consider the events $A = \{\boldsymbol{d} : \bar{d} = c \log n\}$ and $\hat{A}_1(\alpha_1, \alpha_2) = A \cap A_1(\alpha_1, \alpha_2)$. Since we are interested in graphs generated according to (3), $\pi_n(A) = 1$ and $\pi_n(\hat{A}_1(\alpha_1, \alpha_2)) = \pi_n(A_1(\alpha_1, \alpha_2))$. We can now rewrite (9) as

$$\frac{\pi_n(A_2^c)}{\pi_n(\hat{A}_1(\alpha_1, \alpha_2))} \leq \exp\left(\frac{c^2 \log^2 n}{2}\right) \frac{\mathrm{P}(\boldsymbol{D} \in A_2^c)}{\mathrm{P}(\boldsymbol{D} \in \hat{A}_1(\alpha_1, \alpha_2))}. \tag{11}$$

Lemma 2. *There exists a constant $K > 0$, independent of n, such that*

$$\mathrm{P}(\boldsymbol{D} \in A_2^c) \leq K n \mathrm{e}^{-\beta\sqrt{n}/4}. \tag{12}$$

Proof. See Appendix A.

Let $(\tilde{D}_1, \ldots, \tilde{D}_n)$ have the joint distribution of $(D_1, \ldots, D_n)$ conditional on $\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)$. Equivalently, $\tilde{D}_1, \ldots, \tilde{D}_n$ are i.i.d., with $\tilde{D}_j$ having the distribution of D_j conditional on

$$-\sqrt{\alpha_1 \log n} \leq D_j - \mathrm{E}[D_j] \leq \sqrt{\alpha_2 \log n}.$$

Now

$$\mathrm{P}(\boldsymbol{D} \in \hat{A}_1(\alpha_1, \alpha_2)) = \mathrm{P}(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)) \mathrm{P}\left(\sum_{j=1}^{n} D_j = cn \log n \mid \boldsymbol{D} \in A_1(\alpha_1, \alpha_2)\right)$$

$$= \mathrm{P}(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)) \mathrm{P}\left(\sum_{j=1}^{n} \tilde{D}_j = cn \log n\right). \tag{13}$$

Suppose that $\alpha_1, \alpha_2 > 0$ are chosen large enough so that, for large n, $\mathrm{E}[\tilde{D}_1] = \mathrm{E}[D_1] = c \log n$.

We wish to estimate the probability that $\tilde{D}_1 + \tilde{D}_2 + \cdots + \tilde{D}_n = cn \log n$. We will do this using a result from [19]. For $j = 1, \ldots, n$, define the centred random variables, $X_{nj} = \tilde{D}_j - \mathrm{E}[\tilde{D}_j]$; we have made the dependence of the distribution of $\tilde{D}_j$ on n explicit in the notation. Thus, $X_{n1}, X_{n2}, \ldots, X_{nn}$ is an array of integer-valued zero-mean random variables such that, for each n, $X_{n1}, \ldots, X_{nn}$ are i.i.d. To apply [19, Theorem 1], we need the following result.

Lemma 3. *The random variables $\{X_{nj}, \ j = 1, \ldots, n, \ n \in \mathbb{N}\}$ satisfy the following conditions:*

(i) $\limsup_{n\to\infty} \mathrm{E}[\exp(\theta|X_{n1}|)] < \infty$ *for some $\theta > 0$,*

(ii) $\liminf_{n\to\infty} \mathrm{var}(X_{n1}) > 0$,

(iii) $\liminf_{n\to\infty} \sum_{j=-\infty}^{\infty} \min\{\mathrm{P}(X_{n1} = j), \mathrm{P}(X_{n1} = j + 1)\} > 0$.

Proof. See Appendix A.

An immediate corollary of [19, Theorem 1] is as follows.

Theorem 3. *If a sequence of independent random variables,* $\{X_{nj}, \ j = 1, \ldots, n, \ n \in \mathbb{N}\}$, *satisfies conditions (i), (ii), and (iii) of Lemma 3, then*

$$P\left(\sum_{j=1}^{n} X_{nj} = \sum_{j=1}^{n} E[X_{nj}]\right) = \frac{1}{\sqrt{2\pi \sum_{j=1}^{n} \operatorname{var}(X_{nj})}} \left(1 + O\left(\frac{1}{n}\right)\right).$$

A direct application of the above result yields

$$P\left(\sum_{j=1}^{n} \tilde{D}_j = cn \log n\right) = \frac{1}{\sqrt{2\pi n}\tilde{\sigma}} \left(1 + O\left(\frac{1}{n}\right)\right), \tag{14}$$

where $\tilde{\sigma} = \operatorname{var}(\tilde{D}_1)$ remains bounded as $n \to \infty$. Combining this with (11), (12), and (13), we obtain

$$\pi_n(A_2^c) \leq \frac{\pi_n(A_2^c)}{\pi_n(\hat{A}_1(\alpha_1, \alpha_2))} \leq \exp\left(\frac{c^2 \log^2 n}{2}\right) \frac{\sqrt{2\pi}\tilde{\sigma} K n^{3/2} e^{-\beta\sqrt{n}/4}}{P(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2))} \left(1 + O\left(\frac{1}{n}\right)\right). \tag{15}$$

Lemma 4. *Let* $\boldsymbol{D}$ *denote the random vector* $(D_1, \ldots, D_n)$. *Given any* $K > 0$, *we can choose* α_1 *and* α_2 *such that* $P(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)^c) < e^{-K \log n}$ *for all sufficiently large n.*

Proof. See Appendix A.

Combining the above lemma with the bound in (15), it is immediate that $\pi_n(A_2^c) \to 0$ as $n \to \infty$, which establishes the claim of Theorem 2. Thus, to prove Theorem 1, we can restrict our attention to graphs with degree sequences in A_2, for which we can use the estimate in (6).

Proof of Theorem 1. Observe that

$$\begin{aligned} \pi_n(\hat{A}_1(\alpha_1, \alpha_2)) &= \pi_n(A) - \pi_n(A \setminus A_1(\alpha_1, \alpha_2)) \\ &\geq \pi_n(A) - \pi_n((A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2) - \pi_n(A_2^c). \end{aligned}$$

But, $\pi_n(A) = 1$ by definition, and we have shown above that $\pi_n(A_2^c) \to 0$ as $n \to \infty$. Hence, it suffices to show that

$$\pi_n((A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2) \to 0 \quad \text{as } n \to \infty. \tag{16}$$

Recall from (6) that if $\boldsymbol{d} \in A_2$ then $\tilde{G}_n(\boldsymbol{d}) \sim \exp(-\lambda(\boldsymbol{d}) - \lambda(\boldsymbol{d})^2)$. Now,

$$\lambda(\boldsymbol{d}) = \frac{\operatorname{var}(\boldsymbol{d}) + \overline{d}^2 - \overline{d}}{2\overline{d}} \geq \frac{c \log n - 1}{2} \quad \text{for all } \boldsymbol{d} \in A,$$

since the mean degree $\overline{d} = c \log n$. In particular, the above lower bound on $\lambda(\boldsymbol{d})$ holds for all degree sequences $\boldsymbol{d}$ in $(A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2$, since this is a subset of A.

In addition, we saw earlier in (7) that if $\boldsymbol{d} \in \hat{A}_1(\alpha_1, \alpha_2)$ then

$$\lambda(\boldsymbol{d}) \leq \frac{1}{2}\left(c \log n - 1 + \frac{1}{c}\alpha\right),$$

where $\alpha = \max\{\alpha_1, \alpha_2\}$, and the estimate in (6) holds.

Now, by (4),

$$\frac{\pi_n((A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2)}{\pi_n(\hat{A}_1(\alpha_1, \alpha_2))}$$

$$= \frac{\sum_{\boldsymbol{d} \in (A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2} \exp(-\lambda(\boldsymbol{d}) - \lambda(\boldsymbol{d})^2) \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}{\sum_{\boldsymbol{d} \in \hat{A}_1(\alpha_1, \alpha_2)} \exp(-\lambda(\boldsymbol{d}) - \lambda(\boldsymbol{d})^2) \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}$$

$$\leq \exp\left(\frac{\alpha}{2c}\left(c \log n + \frac{\alpha}{2c}\right)\right) \frac{\sum_{\boldsymbol{d} \in (A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2} \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}{\sum_{\boldsymbol{d} \in \hat{A}_1(\alpha_1, \alpha_2)} \prod_{i=1}^n (1/d_i!) \exp(-\beta d_i^2 + \gamma (\log n) d_i)}.$$

In other words, there are constants κ_1 and κ_2 such that

$$\frac{\pi_n((A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2)}{\pi_n(\hat{A}_1(\alpha_1, \alpha_2))} \leq \kappa_1 \exp(\kappa_2 \log n) \frac{\mathrm{P}(\boldsymbol{D} \in (A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2)}{\mathrm{P}(\boldsymbol{D} \in \hat{A}_1(\alpha_1, \alpha_2))}$$

$$\leq \kappa_1 \exp(\kappa_2 \log n) \frac{\mathrm{P}(\boldsymbol{D} \in A \setminus A_1(\alpha_1, \alpha_2))}{\mathrm{P}(\boldsymbol{D} \in \hat{A}_1(\alpha_1, \alpha_2))}. \tag{17}$$

Now, by Lemma 4, for any given $K > 0$, we can choose α_1 and α_2 such that $\mathrm{P}(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)^c) \leq \mathrm{e}^{-K \log n}$. Thus,

$$\mathrm{P}(\boldsymbol{D} \in A \setminus A_1(\alpha_1, \alpha_2)) \leq \mathrm{P}(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)^c) \leq \mathrm{e}^{-K \log n}. \tag{18}$$

Moreover, analogous to (14), we have

$$\mathrm{P}(\boldsymbol{D} \in A) = \mathrm{P}\left(\sum_{j=1}^n D_j = cn \log n\right) = \frac{1}{\sqrt{2\pi n}\sigma}\left(1 + O\left(\frac{1}{n}\right)\right),$$

where $\sigma = \mathrm{var}(D_1)$ remains bounded as $n \to \infty$. Therefore, for large n,

$$\mathrm{P}(\boldsymbol{D} \in \hat{A}_1(\alpha_1, \alpha_2)) = \mathrm{P}(\boldsymbol{D} \in A) - \mathrm{P}(\boldsymbol{D} \in A \cap A_1(\alpha_1, \alpha_2)^c)$$

$$\geq \mathrm{P}(\boldsymbol{D} \in A) - \mathrm{P}(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)^c)$$

$$= \frac{1}{\sqrt{2\pi n}\sigma}\left(1 + O\left(\frac{1}{n}\right)\right). \tag{19}$$

Substituting (18) and (19) into (17), we have

$$\pi_n((A \setminus A_1(\alpha_1, \alpha_2)) \cap A_2) \leq \frac{\pi_n(A \setminus A_1(\alpha_1, \alpha_2))}{\pi_n(\hat{A}_1(\alpha_1, \alpha_2))}$$

$$\leq \kappa_1 \sigma \sqrt{2\pi n} \exp((\kappa_2 - K) \log n)\left(1 + O\left(\frac{1}{n}\right)\right).$$

Since K can be chosen arbitrarily large, the above quantity goes to 0 as $n \to \infty$, which establishes (16) and the claim of the theorem.

5. Graph cuts

Given a graph G and a subset U of its vertex set, let $e_U(G)$ denote the number of edges incident within U (i.e. having both their vertices within U), let $e_{U,U^c}(G)$ denote the number of

edges having one vertex in U and the other in its complement U^c (i.e. crossing the cut (U, U^c)), and denote by u or $|U|$ the number of vertices or the size of U.

Let $\boldsymbol{d}(G) = (d_1, d_2, \ldots, d_n)$ denote the degree sequence of G, and define the volume of a subset of vertices U by

$$\text{vol}(U) = \sum_{i \in U} d_i.$$

Note that

$$2e_U(G) + e_{U,U^c}(G) = \text{vol}(U). \tag{20}$$

In the remainder of this section we derive lower bounds for graph cuts in a graph generated according to (3). To this end, we will show that there exists a constant $\tilde{\delta}$ such that $e_{U,U^c}(G) > (1 - \tilde{\delta})uc \log n$, w.h.p., using different techniques depending on the size of U, when $|U| \leq n/2$.

Proposition 1. *For any $\varepsilon > 0$, there exists $\delta_1 \in (0, 1)$, independent of n, such that if the subset of vertices U is such that $u \leq 2\varepsilon c \log n$ then $e_{U,U^c}(G) \geq (1 - \delta_1)uc \log n$, w.h.p.*

Proof. Denote $|U|$ by u. Suppose first that $u \leq 2\varepsilon c \log n$ for a given $\varepsilon > 0$. The number of edges incident within U can be at most $\binom{u}{2}$, so $e_U(G) \leq \varepsilon uc \log n$ for all U. Now, for any degree sequence $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$, $\text{vol}(U) \geq cu \log n - u\sqrt{\alpha_1 \log n}$. By Theorem 1, it is not restrictive to consider only graphs with degree sequences belonging to the set $A_1(\alpha_1, \alpha_2)$. Hence, using (20) for graphs G with such degree sequences,

$$e_{U,U^c}(G) \geq u[(1 - 2\varepsilon)c \log n - \sqrt{\alpha_1 \log n}].$$

Let $\delta_1 = 3\varepsilon$. Then, for sufficiently large n, $e_{U,U^c}(G) \geq (1 - \delta_1)uc \log n$, w.h.p., whenever $u \leq 2\varepsilon c \log n$, and the claim of the proposition is established.

To prove a similar result for subsets U such that $u \leq n/2$, first recall the definition of the configuration model from Section 4. For constants $\delta \in (0, 1)$, $\varepsilon > 0$, and $\tau > 0$, for $n \in \mathbb{N}$ and a degree sequence $\boldsymbol{d}$, we define the following subsets of graphs on a vertex set V of cardinality n:

$$\mathcal{E}_1(n, \delta, \tau, \boldsymbol{d}) = \{G : \boldsymbol{d}(G) = \boldsymbol{d} \text{ and } e_{U,U^c}(G) < (1 - \delta)uc \log n$$
$$\text{for some } U \subseteq V \text{ with } 2\varepsilon c \log n < u \leq \tau n\}, \tag{21}$$
$$\mathcal{E}_2(n, \delta, \tau, \boldsymbol{d}) = \{G : \boldsymbol{d}(G) = \boldsymbol{d} \text{ and } e_{U,U^c}(G) < (1 - \delta)uc \log n$$
$$\text{for some } U \subseteq V \text{ with } \tau n < u \leq n/2\}. \tag{22}$$

We also define

$$\mathcal{E}_1(n, \delta, \tau) = \bigcup_{\boldsymbol{d}} \mathcal{E}_1(n, \delta, \tau, \boldsymbol{d}), \qquad \mathcal{E}_2(n, \delta, \tau) = \bigcup_{\boldsymbol{d}} \mathcal{E}_2(n, \delta, \tau, \boldsymbol{d}).$$

We will derive bounds on the probabilities of these sets using the definition of the configuration model.

Given a degree sequence $\boldsymbol{d} = (d_1, d_2, \ldots, d_n)$, and for H, a configuration on V, we define the analogous sets of configurations $\hat{\mathcal{E}}_1(n, \delta, \tau, \boldsymbol{d})$, $\hat{\mathcal{E}}_2(n, \delta, \tau, \boldsymbol{d})$, $\hat{\mathcal{E}}_1(n, \delta, \tau)$, and $\hat{\mathcal{E}}_2(n, \delta, \tau)$, that is, instead of considering sets of graphs G, we consider sets of configurations H fulfilling the conditions in (21) and (22).

Note that, given the degree sequence $\boldsymbol{d}$ such that $\sum_{i=1}^n d_i = cn \log n$, $\mu_n(\cdot \mid \boldsymbol{d})$ corresponds to the probability with respect to the uniform distribution on graphs with degree sequence $\boldsymbol{d}$. Let $\Pr(\cdot \mid \boldsymbol{d})$ denote the probability with respect to the uniform distribution on configurations with

degree sequence $\boldsymbol{d}$. Under μ_n, $\boldsymbol{d} \in A(\alpha_1, \alpha_2)$ w.h.p. and estimate (6) holds, i.e. the probability that a configuration yields a (simple) graph is asymptotically equivalent to $\exp(-\lambda - \lambda^2)$. Hence,

$$\mu_n(\mathcal{E}_i(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d}) \leq \exp(\lambda + \lambda^2) \Pr(H \in \hat{\mathcal{E}}_i(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d}). \tag{23}$$

Recall that λ was defined in (6) to be $\sum_{i=1}^n d_i(d_i - 1)/4E$, where E is the number of edges, i.e. $2E = \sum_{i=1}^n d_i$. The dependence of λ on $\boldsymbol{d}$ has been suppressed for notational convenience.

Proposition 2. *If* $\tau \in (0, 1/(1 + 4e))$ *then there exists* $\delta_2 \in (0, 1)$*, independent of n, such that*

$$\lim_{n \to \infty} \mu_n(\mathcal{E}_1(n, \delta_2, \tau)) = 0,$$

where the distribution μ_n was defined in (3).

Proof. First note that

$$\mu_n(\mathcal{E}_1(n, \delta, \tau)) \leq \mu_n(\mathcal{E}_1(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d} \in A_1(\alpha_1, \alpha_2)) + \mu_n(\boldsymbol{d} \notin A_1(\alpha_1, \alpha_2)).$$

By Theorem 1, $\mu_n(\boldsymbol{d} \notin A_1(\alpha_1, \alpha_2))$ goes to 0 as well. In what follows we will use (23) to prove that

$$\lim_{n \to \infty} \mu_n(\mathcal{E}_1(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d} \in A_1(\alpha_1, \alpha_2)) = 0.$$

For degree sequences $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$ and any subset U of the vertex set, $\mathrm{vol}(U) \sim uc \log n$ for large n. Hence, by (20), $e_{U, U^c}(H) < u(1 - \delta)c \log n$ for a subset U implies that $e_U(H) > (\delta/2)\mathrm{vol}(U)$ for sufficiently large n. To prove the proposition, it therefore suffices to show that there exists $\delta_2 \in (0, 1)$ such that $\Pr(e_U(H) > (\delta_2/2)\mathrm{vol}(U))$ tends to 0 when n tends to ∞.

Recall that, for subset U of V, the volume of U is given by $\mathrm{vol}(U) = \sum_{i \in U} d_i$. As the half-edges in the configuration model are matched uniformly, $e_U(H)$, the number of edges incident within U in a random configuration, is bounded above by a binomial random variable X with parameters $\mathrm{vol}(U)$ and $\mathrm{vol}(U)/(2E - \mathrm{vol}(U))$. The dependence of X on U has been suppressed for notational convenience. For $\delta \in (0, 1)$, by Chernoff's bound we have

$$\log \Pr\left(X > \frac{\delta}{2}\mathrm{vol}(U)\right)$$
$$\leq -\mathrm{vol}(U)\left(\frac{\delta}{2} \log \frac{(\delta/2)(2E - \mathrm{vol}(U))}{\mathrm{vol}(U)} + \left(1 - \frac{\delta}{2}\right) \log \frac{(1 - \delta/2)(2E - \mathrm{vol}(U))}{2E - 2\mathrm{vol}(U)}\right)$$
$$\leq -\mathrm{vol}(U)\left(\frac{\delta}{2} \log \frac{(\delta/2)(2E - \mathrm{vol}(U))}{\mathrm{vol}(U)} + \left(1 - \frac{\delta}{2}\right) \log\left(1 - \frac{\delta}{2}\right)\right).$$

Applying the inequality $\log x \leq x - 1$ for $x \geq 1$ to $x = 1/(1 - \delta/2)$, we have $\log(1 - \delta/2) \geq -(\delta/2)/(1 - \delta/2)$. Using the fact that

$$\left|\frac{\mathrm{vol}(U)}{uc \log n} - 1\right| < \frac{\sqrt{\alpha}}{c} \frac{1}{\sqrt{\log n}},$$

we have

$$\log \Pr\left(X > \frac{\delta}{2}\mathrm{vol}(U)\right) \leq -uc \log n \left(\frac{\delta}{2} \log\left(\frac{\delta(n - u)}{2u}\right) - \frac{\delta}{2}\right)\left(1 + O\left(\frac{1}{\sqrt{\log n}}\right)\right). \tag{24}$$

Suppose first that $2\varepsilon c \log n < u \leq \sqrt{n}$.

For all sufficiently large n, (24) becomes

$$\log \Pr\left(X > \frac{\delta}{2}\mathrm{vol}(U)\right) \le -\frac{u\delta c}{6}\log^2 n.$$

Since X stochastically dominates $e_U(H)$ (conditional on $\boldsymbol{d}$), we have, by the union bound, for sufficiently large n,

$$\Pr\left(\text{there exists } U,\ 2\varepsilon c \log n < u \le \sqrt{n},\ e_U(H) > \frac{\delta}{2}\mathrm{vol}(U)\right)$$

$$\le \sum_{u=2\varepsilon c \log n}^{\sqrt{n}} \binom{n}{u} \exp\left(-\frac{u\delta c}{6}\log^2 n\right)$$

$$\le \sum_{u=2\varepsilon c \log n}^{\sqrt{n}} \frac{1}{u!} \exp\left(u \log n - \frac{u\delta c}{6}\log^2 n\right)$$

$$\le \kappa_3 \exp(-\kappa_4 \varepsilon \delta c^2 \log^3 n) \tag{25}$$

for two constants $\kappa_3, \kappa_4 > 0$. We have used the inequality $\binom{n}{u} \le n^u/u!$ to obtain the second inequality above.

Next, consider $\sqrt{n} < u \le \tau n$.

In this case, (24) becomes

$$\log \Pr\left(X > \frac{\delta}{2}\mathrm{vol}(U)\right) \le -\frac{1}{2}uc \log n\left(\delta \log\left(\frac{\delta(1-\tau)}{2\tau}\right) - \delta\right)\left(1 + O\left(\frac{1}{\sqrt{\log n}}\right)\right).$$

If $\tau < 1/(1 + 4e)$ then there exists $\delta_2 \in (0, 1)$ such that

$$\delta_2 \log\left(\frac{\delta_2(1-\tau)}{2\tau}\right) - \delta_2 > \frac{2}{c},$$

and, subsequently, for all sufficiently large n and $u \le \tau n$, we have

$$\log \Pr\left(X > \frac{\delta_2}{2}\mathrm{vol}(U)\right) \le -2u \log n.$$

Hence, by the union bound,

$$\Pr\left(\text{there exists } U: \sqrt{n} < u < \tau n,\ e_U(H) > \frac{\delta_2}{2}\mathrm{vol}(U)\right) \le \sum_{u=\sqrt{n}}^{\tau n} \binom{n}{u} e^{-2u \log n}$$

$$\le \sum_{u=\sqrt{n}}^{\tau n} \frac{1}{u!} e^{-u \log n}$$

$$\le \kappa_5 e^{-\sqrt{n}\log n}. \tag{26}$$

By (23), (25), and (26), for large n, we can find two constants $\kappa_6, \kappa_7 > 0$ such that

$$\mu_n(\mathcal{E}_1(n, \delta_2, \tau, \boldsymbol{d}) \mid \boldsymbol{d}) \le \exp(\lambda + \lambda^2)\kappa_6 \exp(-\kappa_7 \log^3 n).$$

Since $\lambda = O(\log n)$, it is readily checked that $\mu_n(\mathcal{E}_1(n, \delta_2, \tau, \boldsymbol{d}) \mid \boldsymbol{d} \in A_1(\alpha_1, \alpha_2))$ goes to 0 as $n \to \infty$.

The claim of the proposition is established.

Next, we find a similar lower bound for $e_{U,U^c}(G)$ that holds, w.h.p., for subsets U with $\tau n < u \leq n/2$.

Proposition 3. *For $\tau > 0$, there exists $\delta_3 \in (0, 1)$, independent of n, such that*

$$\lim_{n \to \infty} \mu_n(\mathcal{E}_2(n, \delta_3, \tau)) = 0.$$

Proof. As in the proof of Proposition 2, we fix a degree sequence $\boldsymbol{d}$ and a subset U, and bound the probability that $e_{U,U^c}(G) < u(1 - \delta)c \log n$ in terms of the probability that $e_{U,U^c}(H) < u(1 - \delta)c \log n$, where H is drawn uniformly at random from configurations with degree sequence $\boldsymbol{d}$, i.e.

$$\mu_n(\mathcal{E}_2(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d}) \leq \exp(\lambda + \lambda^2) \Pr(H \in \hat{\mathcal{E}}_2(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d}). \tag{27}$$

Fix constants $\tau > 0$ and $\delta \in (0, 1)$, and a degree sequence $\boldsymbol{d}$. Let U be a subset of the vertex set with $\tau n < u \leq n/2$, and let $j < (1 - \delta)uc \log n \leq \frac{1}{2}(1 - \delta)cn \log n$. Recall that the number of configurations with degree sequence $\boldsymbol{d}$ is

$$H_n(\boldsymbol{d}) = \frac{(2E)!}{E! \, 2^E} \prod_{i=1}^{n} d_i!, \tag{28}$$

where $E = \sum_{i=1}^{n} d_i/2$ is the total number of edges.

Let H be a configuration such that we choose j configuration points each from U and U^c to pair, and pair the remaining $2E - 2j$ configuration points within the sets U and U^c: $\mathrm{vol}(U) - j$ configuration points are paired in U and $2E - \mathrm{vol}(U) - j$ configuration points are paired in U^c. For such a configuration to exist, $\mathrm{vol}(U) - j$ must be even, as should $2E - \mathrm{vol}(U) - j$. The number of these configurations with exactly j edges crossing the cut between U and U^c is

$$\mathcal{H}_{U,U^c}(j) \leq \binom{\mathrm{vol}(U)}{j}\binom{2E - \mathrm{vol}(U)}{j} j! \, \frac{(\mathrm{vol}(U) - j)!}{((\mathrm{vol}(U) - j)/2)! \, 2^{(\mathrm{vol}(U) - j)/2}}$$
$$\times \frac{(2E - \mathrm{vol}(U) - j)!}{(E - (\mathrm{vol}(U) - j)/2)! \, 2^{(2E - \mathrm{vol}(U) - j)/2}} \prod_{i=1}^{n} d_i!. \tag{29}$$

The dependence of $\mathcal{H}$ on $\boldsymbol{d}$ has been suppressed for notational convenience. The first two terms on the right-hand side above count the number of ways we can choose j configuration points each from U and U^c to match up. The term $j!$ counts the number of ways of matching them. The remaining configuration points have to be matched within the sets U and U^c as there are only j edges crossing the cut. The number of ways of doing this is equal to the number of configurations on U with $\mathrm{vol}(U) - j$ points times the number of configurations on U^c with $2E - \mathrm{vol}(U) - j$ points, and with a degree sequence strictly bounded by $\boldsymbol{d}$ (since j points each in U and U^c have been used up). This yields the remaining terms in the bound above. We obtain from (28) and (29), after some simplification,

$$\Pr(e_{U,U^c}(H) = j) = \frac{\mathcal{H}_{U,U^c}(j)}{H_n(\boldsymbol{d})} \leq \frac{\binom{E}{\mathrm{vol}(U)/2}\binom{\mathrm{vol}(U)/2}{j/2}\binom{E - \mathrm{vol}(U)/2}{j/2}}{\binom{2E}{\mathrm{vol}(U)}\binom{j}{j/2}} 2^j.$$

Taking logarithms and using Stirling's formula, we obtain

$$\log \Pr(e_{U,U^c}(H) = j) \le E h\left(\frac{\text{vol}(U)}{2E}\right) + \frac{\text{vol}(U)}{2} h\left(\frac{j}{\text{vol}(U)}\right)$$
$$+ \frac{2E - \text{vol}(U)}{2} h\left(\frac{j}{2E - \text{vol}(U)}\right)$$
$$- 2E h\left(\frac{\text{vol}(U)}{2E}\right) + O(\log n), \qquad (30)$$

where, for $x \in [0, 1]$, $h(x) = -x \log x - (1-x) \log(1-x)$ is the binary entropy of x. Now, $2E = cn \log n$ and, since it was assumed that $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$, $|\text{vol}(U) - cu \log n| \le u\sqrt{\alpha \log n}$. Moreover, $\tau n < u \le n/2$, while $j < \frac{1}{2}(1 - \delta)cn \log n$. Hence, for some $\hat{\delta}_1$ and large enough n, we have, for all $\delta \ge \hat{\delta}_1$

$$h\left(\frac{j}{\text{vol}(U)}\right) < h\left(\frac{(1-\delta)n \log n}{2\tau n \log n}\right) = h\left(\frac{(1-\delta)}{2\tau}\right),$$

and it can likewise be shown that, for some $\hat{\delta}_2$ and large enough n, we have, for all $\delta \ge \hat{\delta}_2$

$$h\left(\frac{j}{2E - \text{vol}(U)}\right) < h(1 - \delta).$$

On the other hand, as $|U| < n/2$, for large n,

$$h\left(\frac{\text{vol}(U)}{2E}\right) \ge h(\tau).$$

Using the fact that $\text{vol}(U) \le 2E$ for all U, it follows from (30) that, for sufficiently large n,

$$\log \Pr(e_{U,U^c}(H) = j) \le -E\left(h(\tau) - h\left(\frac{1-\delta}{2\tau}\right) - h(1 - \delta)\right) \le -\kappa n \log n,$$

where δ is chosen big enough so that $h(\tau) - h((1-\delta)/2\tau) - h(1-\delta) > 0$, i.e. $\kappa > 0$.

The above bound applies for all subsets U of V, of size u where $n < u < n/2$. The number of subsets U with cardinality between τn and $n/2$ is smaller than the total number of subsets, which is 2^n. Hence, by the union bound,

$$\Pr(H \colon \text{ there exists } U \text{ with } \tau n < u < n/2 \text{ and } e_{U,U^c}(H) = j) \le 2^n e^{-\kappa n \log n}.$$

The above holds for each $j < \frac{1}{2}(1 - \delta)cn \log n$. Applying the union bound once more,

$$\Pr(H \in \hat{\mathcal{E}}_2(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d}) \le (1 - \delta)cn \log(n) 2^{n-1} e^{-\kappa n \log n}$$

for all $\boldsymbol{d} \in A_1(\alpha_1, \alpha_2)$. Substituting this into (27) and noting that $\lambda = O(\log n)$, we see that, for large enough δ,

$$\mu_n(\mathcal{E}_2(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d} \in A_1(\alpha_1, \alpha_2)) \to 0 \quad \text{as } n \to \infty.$$

We also know from Theorem 1 that $\mu_n(\boldsymbol{d} \notin A_1(\alpha_1, \alpha_2))$ goes to 0. Since

$$\mu_n(\mathcal{E}_2(n, \delta, \tau)) \le \mu_n(\mathcal{E}_2(n, \delta, \tau, \boldsymbol{d}) \mid \boldsymbol{d} \in A_1(\alpha_1, \alpha_2)) + \mu_n(\boldsymbol{d} \notin A_1(\alpha_1, \alpha_2)),$$

there exists $\delta_3 > 0$ such that $\mu_n(\mathcal{E}_2(n, \delta_3, \tau)) \to 0$ as $n \to \infty$, as claimed.

Fix $\varepsilon > 0$ and $\tau < 1/(1+4e)$. Then, by Propositions 1, 2, and 3, there exists $\tilde{\delta}$, independent of n, which is the maximum of δ_1, δ_2, and δ_3 for which the three propositions hold. Hence, we have the following lower bound for the graph cut.

Theorem 4. *For graphs G drawn according to (3), there exists $\tilde{\delta} \in (0, 1)$ such that, for a subset U of V with $u = |U| \leq n/2$, the number of edges crossing the cut (U, U^c) is such that*

$$e_{U,U^c} \geq (1 - \tilde{\delta})cu \log n \quad w.h.p.$$

5.1. Conductance and expansion

Using Theorem 4, we can easily recover asymptotic results for the conductance and the expansion of a graph drawn according to (3), which are relevant for phenomena such as routeing congestion analysis [14], the behaviour of random walks in terms of the mixing and cover times [17], and epidemic thresholds [6], [13].

Let $A = (a_{ij})_{i, j=1,...,n}$ be the adjacency matrix of a graph G, and let $D = \text{diag}(d_1, \ldots, d_n)$ be the diagonal matrix of the degree distribution of G. First, we define the *isoperimetric constant* or *expansion* of a graph G by

$$\phi = \inf_{\substack{U \subset V \\ u \leq n/2}} \frac{e_{U,U^c}}{u}.$$

It is related to $\lambda_2(L)$, the second (smallest) eigenvalue of the Laplacian $L = D - A$ of the graph, through the following inequality [5], [22]:

$$\frac{\phi^2}{2d_{\max}} \leq \lambda_2(L) \leq 2\phi.$$

The lower bound in the above inequality is known as Cheeger's inequality.

The *conductance* of a graph G is defined by

$$\Phi = \inf_{\substack{U \subset V \\ \text{vol}(U) \leq E}} \frac{e_{U,U^c}}{\text{vol}(U)}.$$

Let $\lambda_2(P)$ be the second (largest) eigenvalue of P, the transition matrix of the simple random walk on a graph $p_{ij} = a_{ij}/d_i$. By Cheeger's inequality [17, Theorem 5.3],

$$\frac{\Phi^2}{8} \leq 1 - \lambda_2(P) \leq \Phi.$$

Theorem 5. *For graphs G drawn according to (3), and for the constant $\tilde{\delta}$ of Theorem 4, the expansion ϕ and the conductance Φ satisfy*

$$(1 - \tilde{\delta})c \log n \leq \phi \leq c \log n, \qquad (1 - \tilde{\delta}) \leq \Phi \leq 1, \quad w.h.p.$$

Proof. First note that if $d_{\min}$ is the minimum degree of G then, by Theorem 1, $d_{\min} = c \log n - \sqrt{\alpha_1 \log n}$, w.h.p. Hence,

$$\phi \leq (1 + o(1))c \log n, \qquad \Phi \leq (1 + o(1)), \quad w.h.p.$$

The lower bounds follow from Theorem 4.

5.2. Failure resilience

In the following we work with graphs whose degree sequence belongs to the set $A_1(\alpha_1, \alpha_2)$ for some specified α_1 and α_2. We are interested in the probability that the graph remains connected when links fail independently with probability p. It is straightforward to compute the probability that a given node i becomes isolated due to link failures; it is simply p^{d_i}. Thus, by the union bound, the probability that some node becomes isolated is at most

$$\sum_{i=1}^{n} p^{d_i} \leq np^{c \log n - \sqrt{\alpha_1 \log n}} = \exp((1 + c \log p) \log n - \sqrt{\alpha_1 \log n} \log p).$$

Hence, if $c \log p < -1$ or, equivalently, $p < \exp(-1/c)$, then the probability that some node becomes isolated goes to 0 as n increases to ∞.

By way of comparison, consider the classical random graph model of Erdös and Rényi [8] with the same mean degree. Here, an edge is present between each pair of nodes with probability $c \log n/n$, independent of all other edges. Here we should assume that $c > 1$ to ensure that the Erdös–Rényi graph is connected, w.h.p. After taking failures into account, the edge probability becomes $(1 - p)c \log n/n$, and the presence of edges continues to be mutually independent. It is well known for this model that if $(1 - p)c < 1$ then the graph is disconnected, w.h.p. Moreover, in a sense that can be made precise, the main reason for disconnection when $(1 - p)c$ is 'close to' 1 is the isolation of individual nodes. Intuitively, these arguments suggest that graphs generated according to (3) can tolerate link failure rates up to $e^{-1/c}$ while retaining connectivity, whereas classical random graphs can tolerate failure rates only up to $(c - 1)/c$. We now rigorously establish a weaker result.

We will use Theorem 4 to show that random graphs drawn from the distribution μ_n can tolerate link failure rates up to $\exp(-1/c(1 - \tilde{\delta}))$, where $\tilde{\delta}$ is defined in Theorem 4, without losing connectivity.

Theorem 6. *For any $p < \exp(-1/c(1 - \tilde{\delta}))$, a graph G chosen at random from the distribution μ_n and subjected to independent link failures with probability p remains connected, w.h.p.*

Proof. Fix $p < \exp(-1/c(1 - \tilde{\delta}))$. For a subset U of the vertex set, let $\hat{e}_{U,U^c}$ denote the number of edges between U and U^c that have not failed. We will show that, w.h.p., $\hat{e}_{U,U^c} > 0$ for all subsets U, i.e. the graph is connected. Now,

$$\mu_n(\hat{e}_{U,U^c}(G) = 0 \mid e_{U,U^c}(G)) = p^{e_{U,U^c}(G)}.$$

Assume that $e_{U,U^c}(G) \geq (1 - \tilde{\delta})uc \log n$ for all $U \subseteq V$ with $u \leq \tau n$. Then

$$\mu_n(\text{there exists } U : u \leq \tau n, \ \hat{e}_{U,U^c}(G) = 0) \leq \sum_{u=1}^{\tau n} \binom{n}{u} p^{(1-\tilde{\delta})uc \log n}.$$

Since $p < \exp(-1/c(1 - \tilde{\delta}))$ is given, then, for some $\varepsilon > 0$ and large n, $p^{(1-\tilde{\delta})c \log n} < e^{-(1+\varepsilon) \log n}$. Using the inequality $\binom{n}{u} \leq n^u/u!$, we obtain

$$\mu_n(\text{there exists } U : u \leq \tau n, \ e_{U,U^c}(G) = 0) \leq \sum_{u=1}^{\tau n} \frac{1}{u!} (np^{(1-\tilde{\delta})c \log n})^u$$

$$\leq \exp(np^{(1-\tilde{\delta})c \log n}) - 1$$

$$\leq \exp(ne^{-(1+\varepsilon) \log n}) - 1, \qquad (31)$$

which goes to 0 as $n \to \infty$.

Suppose that $e_{U,U^c}(G) \geq (1 - \tilde{\delta})cu \log n$ for all $U \subseteq V$ with $\tau n < u \leq n/2$. Then

$$\mu_n(\text{there exists } U: \tau n < u \leq n/2, \hat{e}_{U,U^c}(G) = 0)$$
$$\leq \sum_{U:\, \tau n < u \leq n/2} p^{(1-\tilde{\delta})cu \log n}$$
$$\leq 2^n p^{(1-\tilde{\delta})\tau cn \log n}. \tag{32}$$

We see from (31) and (32) that

$$\mu_n(\text{there exists } U: \hat{e}_{U,U^c}(G) = 0 \mid e_{U,U^c}(G) \geq (1 - \tilde{\delta})cu \log n) \to 0 \quad \text{as } n \to \infty.$$

Also, by Theorem 4,

$$\mu_n\left(e_{U,U^c}(G) < (1 - \tilde{\delta})cu \log n \text{ for all } U \subseteq V,\ 0 < u \leq \frac{n}{2}\right) \to 0 \quad \text{as } n \to \infty,$$

when G is chosen according to the distribution μ_n, which establishes the claim of the theorem.

Appendix A

Let $D_1, \ldots, D_n$ be i.i.d. random variables with distribution given by (10). Define

$$f(j, \gamma) = \frac{1}{j!} \exp(-\beta j^2 + \gamma j \log n) \quad \text{and} \quad F(\gamma) = \sum_{j=0}^{\infty} f(j, \gamma), \tag{33}$$

so that $P(D_1 = j) = f(j, \gamma)/F(\gamma)$. Now, the ratio

$$\frac{f(j+1, \gamma)}{f(j, \gamma)} = \frac{1}{j+1} e^{-(2j+1)\beta + \gamma \log n}$$

is a decreasing function of j. Define k_γ to be the smallest value of j for which $f(j + 1, \gamma)/f(j, \gamma) \leq 1$, and note that the maximum of $f(j, \gamma)$ over j is attained at k_γ. Now, $k_\gamma - 1$ is the integer part of the (unique) solution of the equation

$$h(x, \gamma) := -\log(x + 1) - (2x + 1)\beta + \gamma \log n = 0.$$

It is readily verified that the solution is

$$d_\gamma = \frac{1}{2\beta}\left(\gamma \log n + \log \log n + \frac{\gamma}{2\beta}\right) + o(1).$$

Let $k_\gamma = \lfloor d_\gamma \rfloor + 1$, where $\lfloor x \rfloor$ is the integer value of x. Then, for any $j > 0$,

$$\frac{f(k_\gamma + j + 1, \gamma)}{f(k_\gamma + j, \gamma)} = \frac{1}{k_\gamma + j + 1} \exp(-\beta(2k_\gamma + 2j + 1) + \gamma \log n)$$
$$= \frac{f(k_\gamma + 1, \gamma)}{f(k_\gamma, \gamma)} \frac{k_\gamma + 1}{k_\gamma + j + 1} e^{-2\beta j}$$
$$\leq e^{-2\beta j},$$

where we have used the fact that $f(k_\gamma + 1, \gamma)/f(k_\gamma, \gamma) \leq 1$ to obtain the last inequality. Iterating this inequality yields $f(k_\gamma + j, \gamma)/f(k_\gamma, \gamma) \leq e^{-\beta j(j-1)}$. Similarly, we obtain

$$\frac{f(k_\gamma - j - 1, \gamma)}{f(k_\gamma - j, \gamma)} = \frac{f(k_\gamma - 1, \gamma)}{f(k_\gamma, \gamma)}\left(1 - \frac{j}{k_\gamma}\right)e^{-2\beta j} \leq e^{-2\beta j},$$

since $f(k_\gamma, \gamma)/f(k_\gamma - 1, \gamma) > 1$ by the definition of k_γ. Iterating this inequality yields $f(k_\gamma - j, \gamma)/f(k_\gamma, \gamma) \leq e^{-\beta j(j-1)}$. Thus, for all integers $j \geq -k_\gamma$, we have the inequality

$$\frac{f(k_\gamma + j, \gamma)}{f(k_\gamma, \gamma)} \leq e^{-\beta |j|(|j|-1)} \leq \exp(-\beta(|j| - 1)^2). \tag{34}$$

Next, we derive an equivalent for the above ratio. Observe that, for any fixed j,

$$\frac{f(k_\gamma + j, \gamma)}{f(k_\gamma, \gamma)} = \frac{k_\gamma!}{(k_\gamma + j)!} \exp(-\beta j(2k_\gamma + j) + \gamma j \log n)$$

$$= \frac{1}{k_\gamma^j} \exp(-\beta j(2k_\gamma + j) + \gamma j \log n)\left(1 + O\left(\frac{j^2}{k_\gamma}\right)\right).$$

Taking logarithms,

$$\log \frac{f(k_\gamma + j, \gamma)}{f(k_\gamma, \gamma)} = -j \log k_\gamma - \beta j(2k_\gamma + j) + \gamma j \log n + O\left(\frac{j^2}{\log n}\right)$$

$$= jh(x_\gamma, \gamma) + \rho j - \beta j^2 + O\left(\frac{j^2}{\log n}\right),$$

where $\rho = 2\beta(x_\gamma - k_\gamma + \frac{1}{2})$. Note that $\rho \in [-\beta, \beta]$ for all n because $k_\gamma \in [d_\gamma, d_\gamma + 1]$. Since $h(d_\gamma, \gamma) = 0$ by the definition of d_γ, we can now write

$$g(j, \gamma) := \frac{f(k_\gamma + j, \gamma)}{f(k_\gamma, \gamma)} = (1 + \lambda_j) \exp(\rho j - \beta j^2), \quad \text{where} \quad \lambda_j = O\left(\frac{j^2}{\log n}\right). \tag{35}$$

Thus, by (33),

$$F(\gamma) = f(k_\gamma, \gamma) \sum_{j=-k_\gamma}^{\infty} g(j, \gamma) = K_0(\rho, \beta) f(k_\gamma, \gamma), \tag{36}$$

where $K_0(\rho, \beta) \sim \sum_{j=-\infty}^{\infty} \exp(\rho j - \beta j^2)$ is bounded uniformly in γ and n.

Proof of Lemma 1. We obtain from (10) and (35) that

$$\mathrm{E}[D_1] = \frac{\sum_{j=0}^{\infty} jf(j, \gamma)}{\sum_{j=0}^{\infty} f(j, \gamma)}$$

$$= k_\gamma \frac{\sum_{j=-k_\gamma}^{\infty} (1 + j/k_\gamma)g(j, \gamma)}{\sum_{j=-k_\gamma}^{\infty} g(j, \gamma)}$$

$$= k_\gamma \left(1 + \frac{1}{k_\gamma} \frac{\sum_{j=-k_\gamma}^{\infty} j(1 + \lambda_j)\exp(\rho j - \beta j^2)}{\sum_{j=-k_\gamma}^{\infty} (1 + \lambda_j)\exp(\rho j - \beta j^2)}\right)$$

$$= k_\gamma + K_1(\rho, \beta),$$

where

$$K_1(\rho, \beta) \sim \frac{\sum_{j=-\infty}^{\infty} j \exp(\rho j - \beta j^2)}{\sum_{j=-\infty}^{\infty} \exp(\rho j - \beta j^2)}.$$

Note that $K_1(\rho, \beta)$ is bounded uniformly in γ and n. It is also easy to see that $\mathrm{E}[D_1]$ is a continuous and increasing function of γ. This yields the first claim of the lemma.

A similar calculation yields

$$\begin{aligned}
\mathrm{E}[(D_1)^2] &= \frac{\sum_{j=0}^{\infty} j^2 f(j, \gamma)}{\sum_{j=0}^{\infty} f(j, \gamma)} \\
&= k_\gamma^2 \frac{\sum_{j=-k_\gamma}^{\infty} (1 + j/k_\gamma)^2 g(j, \gamma)}{\sum_{j=-k_\gamma}^{\infty} g(j, \gamma)} \\
&= k_\gamma^2 + 2k_\gamma K_1(\rho, \beta) + K_2(\rho, \beta),
\end{aligned}$$

where

$$K_2(\rho, \beta) \sim \frac{\sum_{j=-\infty}^{\infty} j^2 \exp(\rho j - \beta j^2)}{\sum_{j=-\infty}^{\infty} \exp(\rho j - \beta j^2)}$$

remains bounded, uniformly in γ and n. Hence,

$$\mathrm{var}(D_1) = K_2(\rho, \beta) - K_1(\rho, \beta)^2$$

remains bounded. In fact, we see that $\mathrm{var}(D_1)$ is asymptotic to the variance of a discrete Gaussian distribution; this distribution is nondegenerate for any finite β. Hence, $\mathrm{var}(D_1)$ remains bounded below by some strictly positive constant as n goes to ∞.

Next, we evaluate the moment generating function of D_1. Proceeding as in the calculations of the mean and variance, we have

$$\begin{aligned}
\mathrm{E}[e^{\theta D_1}] &= \frac{\sum_{j=0}^{\infty} e^{\theta j} f(j, \gamma)}{\sum_{j=0}^{\infty} f(j, \gamma)} \\
&= \exp(\theta k_\gamma) \frac{\sum_{j=-k_\gamma}^{\infty} e^{\theta j} g(j, \gamma)}{\sum_{j=-k_\gamma}^{\infty} g(j, \gamma)} \\
&= \exp(\theta k_\gamma) \frac{\sum_{j=-k_\gamma}^{\infty} (1 + \lambda_j) \exp((\theta + \rho)j - \beta j^2)}{\sum_{j=-k_\gamma}^{\infty} (1 + \lambda_j) \exp(\rho j - \beta j^2)} \\
&\sim \exp(\theta k_\gamma) \frac{\psi(\theta + \rho)}{\psi(\rho)},
\end{aligned} \tag{37}$$

where

$$\psi(\theta) = \frac{\sum_{j=-\infty}^{\infty} \exp(\theta j - \beta j^2)}{\sum_{j=-\infty}^{\infty} \exp(-\beta j^2)} \tag{38}$$

is the moment generating function of the discrete Gaussian distribution which adds mass proportional to $e^{-\beta j^2}$ at each $j \in \mathbb{Z}$.

Proof of Lemma 2. We obtain, using (34) and (36), for large n,

$$P(D_1 > n^{1/4}) = \frac{\sum_{j=n^{1/4}+1}^{\infty} f(j, \gamma)}{F(\gamma)}$$

$$\leq \frac{1}{K_0(\rho, \beta)} \sum_{j=0}^{\infty} \exp(-\beta(j + n^{1/4} - k_\gamma)^2)$$

$$\leq \frac{1}{K_0(\rho, \beta)} \sum_{j=0}^{\infty} \exp\left(-\beta\left(j + \frac{1}{2}n^{1/4}\right)^2\right)$$

$$\leq \frac{\sum_{j=0}^{\infty} \exp(-\beta j^2)}{K_0(\rho, \beta)} e^{-\beta\sqrt{n}/4}.$$

By the union bound,

$$P(\boldsymbol{D} \in A_2^c) \leq \sum_{i=1}^{n} P(D_i > n^{1/4}) \leq Kn e^{-\beta\sqrt{n}/4},$$

which establishes the claim of the lemma.

Proof of Lemma 3. In what follows we prove the result for the sequence D_i. Following the same lines, we can prove the lemma for $\tilde{D}_i$.

Since $E[D_1] = k_\gamma + K_1(\rho, \beta)$, it follows from (37) that

$$E[\exp(\theta X_{n1})] = \exp(-\theta E[D_1]) E[\exp(\theta D_1)] \sim \exp(-\theta K_1(\rho, \beta)) \frac{\psi(\theta + \rho)}{\psi(\rho)}. \qquad (39)$$

For fixed θ, this is bounded uniformly in n since $K_1(\rho, \beta)$ is so bounded, and ψ does not depend on n. The first claim of the lemma now follows from the inequality $E[\exp(\theta|X_{n1}|)] \leq E[\exp(\theta X_{n1})] + E[\exp(-\theta X_{n1})]$.

Since $X_{n1} = D_1 - E[D_1]$, $\text{var}(X_{n1}) = \text{var}(D_1)$, and the second claim of the lemma is immediate from Lemma 1.

The last claim of the lemma follows from the fact that

$$\sum_{j=-\infty}^{\infty} \min\{P(X_{n1} = j), P(X_{n1} = j + 1)\}$$

$$= \sum_{j=0}^{\infty} \min\{P(D_1 = j), P(D_1 = j + 1)\}$$

$$\geq \sum_{j=0}^{\infty} P(D_1 = j) P(D_1 = j + 1)$$

$$\sim \frac{\sum_{j=-\infty}^{\infty} (\exp(\rho j - \beta j^2))(\exp(\rho(j + 1) - \beta(j + 1)^2))}{\sum_{j=-\infty}^{\infty} \exp(\rho j - \beta j^2)}$$

$$> 0.$$

This completes the proof of the lemma.

Proof of Lemma 4. We will bound $P(\boldsymbol{D} \in A_1(\alpha_1, \alpha_2)^c)$ using the moment generating function of $X_{n1} := D_1 - E[D_1]$ and Chernoff's bound. Observe from (39) that

$$E[\exp((\sqrt{\theta \log n})X_{n1})] = \exp(-\sqrt{\theta \log n}\,K_1(\rho, \beta))\frac{\psi(\sqrt{\theta \log n} + \rho)}{\psi(\rho)}, \qquad (40)$$

where ψ is defined in (38). Here, ρ and β are constants, and $K_1(\rho, \beta)$ remains bounded as $n \to \infty$. Let

$$y^* = \frac{\sqrt{\theta \log n} + \rho}{2\beta}, \qquad j^* = \lfloor y^* \rfloor.$$

We have

$$\left(\sum_{j=-\infty}^{\infty} \exp(-\beta j^2)\right)\psi(\sqrt{\theta \log n} + \rho)$$

$$= \exp((\sqrt{\theta \log n} + \rho)j^* - \beta(j^*)^2) \sum_{j=-\infty}^{\infty} \exp((\sqrt{\theta \log n} + \rho)(j - j^*)$$
$$- \beta(j^2 - (j^*)^2))$$

$$= e^{\beta j^*(2y^* - j^*)} \sum_{k=-\infty}^{\infty} \exp(2\beta(y^* - j^*)k - \beta k^2)$$

$$= \exp(\beta(y^*)^2)\exp(-\beta(y^* - j^*)^2) \sum_{k=-\infty}^{\infty} \exp(2\beta(y^* - j^*)k - \beta k^2),$$

and so

$$\psi(\sqrt{\theta \log n} + \rho) = \kappa(\rho, \beta, \theta)\exp\left(\frac{(\sqrt{\theta \log n} + \rho)^2}{4\beta}\right),$$

where $\kappa(\rho, \beta, \theta)$ is bounded, uniformly in n and θ. Substituting this into (40) yields

$$E[\exp((\sqrt{\theta \log n})X_{n1})] = \kappa_1 \exp\left(\frac{\theta \log n}{4\beta} + \kappa_2 \sqrt{\theta \log n}\right),$$

where κ_1 and κ_2 may depend on ρ, β, θ, and n, but are bounded. Thus, we obtain, using Chernoff's bound,

$$P(X_{n1} > \sqrt{\alpha_2 \log n}) \le \kappa_1 \exp\left(-\sqrt{\theta \alpha_2}\log n + \frac{\theta \log n}{4\beta} + \kappa_2\sqrt{\theta \log n}\right) \quad \text{for all } \theta > 0.$$

Take $\theta = 4\alpha_2\beta^2$. Now, by the union bound,

$$P\left(\bigcup_{j=1}^{n}\{X_{nj} > \sqrt{\alpha_2 \log n}\}\right) \le \kappa_1 \exp(-(\alpha_2\beta - 1)\log n + 2\kappa_2\beta\sqrt{\alpha_2 \log n}).$$

The constant α_2 can be chosen large enough so that $\alpha_2\beta - 1 > K$. Hence, the right-hand side above decreases to 0 faster than $e^{-K\log n}$ as $n \to \infty$. A similar bound can be obtained on the probability that $X_{nj} < -\sqrt{\alpha_1 \log n}$ for some $j \in \{1, \ldots, n\}$. Thus, we have shown that, given $K > 0$, we can choose $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$ so that

$$P\left(\bigcup_{j=1}^{n}\{X_{nj} > \sqrt{\tilde{\alpha}_2 \log n}\} \cup \bigcup_{j=1}^{n}\{X_{nj} < -\sqrt{\tilde{\alpha}_1 \log n}\}\right) < \frac{e^{-K\log n}}{2} \qquad (41)$$

for all sufficiently large n. Here, $X_{nj} = D_j - \mathrm{E}[D_j]$, and the D_j are i.i.d. with mean $c \log n$. Let $\overline{D}$ denote the empirical mean of $D_1, \ldots, D_n$. The event, $|\overline{D} - \mathrm{E}[D_1]| > \sqrt{\eta \log n}$ is the same as the event $|X_{n1} + \cdots + X_{nn}| > n\sqrt{\eta \log n}$. Using the same Chernoff bound techniques as above, we can show that η can be chosen so that, for sufficiently large n, this event has probability at most $\mathrm{e}^{-K \log n}/2$. Combining this with (41) yields the claim of the lemma: simply take $\sqrt{\alpha_1} = \sqrt{\tilde{\alpha}_1} + \sqrt{\eta}$ and $\sqrt{\alpha_2} = \sqrt{\tilde{\alpha}_2} + \sqrt{\eta}$.

References

[1] BALL, F. AND BARBOUR, A. (1990). Poisson approximation for some epidemic models. *J. Appl. Prob.* **27,** 479–490.

[2] BESAG, J. E. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. R. Statist. Soc. Ser. B* **36,** 192–236.

[3] BOLLOBÀS, B. (2001). *Random Graphs* (Camb. Stud. Adv. Math. **73**). Cambridge Univeristy Press.

[4] BRÉMAUD, P. (1999). *Markov chains, Gibbs fields, Monte Carlo Simulation, and Queues*. Springer, New York.

[5] CHUNG, F. (1996). Laplacians of graphs and Cheeger's inequalities. In *Combinatorics, Paul Erdös is Eighty*, Vol. 2 (Keszthely, 1993), János Bolyai Mathematical Society, Budapest, pp. 157–172.

[6] DRAIEF, M., GANESH, A. AND MASSOULIÉ, L. (2008). Thresholds for virus spread on networks. *Ann. Appl. Prob.* **18,** 359–378.

[7] DURRETT, R. (2007). *Random Graph Dynamics*. Cambridge University Press.

[8] ERDÖS, P. AND RÉNYI, A. (1960). On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **5,** 17–61.

[9] FRANK, O. AND STRAUSS, D. (1986). Markov graphs. *J. Amer. Statist. Assoc.* **81,** 832–842.

[10] GANESH, A. J., KERMARREC, A.-M. AND MASSOULIÉ, L. (2003). Network awareness and failure resilience in self-organizing overlay networks. In *Proc. IEEE Symp. Reliab. Distributed Systems*, IEEE, pp. 47–55.

[11] GANESH, A. J. AND MASSOULIÉ, L. (2003). Failure resilience in balanced overlay networks. In *Proc. 41st Allerton Conf. Commun., Control Comput.*, IEEE.

[12] GANESH, A. J., KERMARREC, A.-M. AND MASSOULIÉ, L. (2003). Probabilistic reliable dissemination in large-scale systems. *IEEE Trans. Parallel Distributed Systems* **14,** 248–258.

[13] GANESH, A. J., MASSOULIÉ, L. AND TOWSLEY, D. (2005). The effect of network topology on the spread of epidemics. In *Proc. IEEE INFOCOM* 2005, IEEE, pp. 1455–1466

[14] GKANTSIDIS, C., MIHAIL, M. AND SABERI, A. (2003). Conductance and congestion in power law graphs. In *Proc. ACM SIGMETRICS* 2003, ACM, pp. 148–159.

[15] HOLLAND, P. W. AND LEINHARDT, S. (1981). An exponential family of probability densities for directed graphs. *J. Amer. Statist. Assoc.* **76,** 33–65.

[16] JANSON, S. (2009). The probability that a random multigraph is simple. To appear in *Combinatorics Prob. Comput.*

[17] LOVÁSZ, L. (1996). Random walks on graphs: a survey. In *Combinatorics, Paul Erdös is Eighty*, Vol. 2, (Keszthely, 1993), János Bolyai Mathematical Society, Budapest, pp. 353–397.

[18] LUA, E. K. *et al.* (2005). A survey and comparison of peer-to-peer overlay network schemes. *IEEE Commun. Surveys Tutorials* **7,** 72–93.

[19] MCDONALD, D. (1979). A local limit theorem for large deviations of sums of independent, nonidentically distributed random variables. *Ann. Prob.* **7,** 526–531.

[20] MCKAY, B. D. AND WORMALD, N. C. (1990). Asymptotic enumeration by degree sequence of graphs of high degree. *Europ. J. Combinatorics* **11,** 565–580.

[21] METROPOLIS, N. *et al.* (1953). Equations of state calculations by fast computing machines. *J. Chemical Physics* **21,** 1087–1092.

[22] MOHAR, B. (1997). Some applications of Laplace eigenvalues of graphs. In *Graph Symmetry*, eds G. Hahn and G. Sabidussi, Kluwer, Dordrecht, pp. 225–275.

[23] PARK, J. AND NEWMAN, M. E. J. (2004). Statistical mechanics of networks. *Phys. Rev. E* **70,** 066117.

[24] SNIJDERS, T. A. B. (2002). Markov chain Monte Carlo estimation of exponential random graph models. *J. Social Structure* **3,** 1–40.

[25] SOLOMONOFF, R. AND RAPOPORT, A. (1951). Connectivity of random nets. *Bull. Math. Biophysics* **13,** 107–117.

[26] STEINMETZ, R. AND WEHRLE, K. (2005). What is this peer-to-peer about? In *Peer-to-Peer Systems and Applications*, Springer, Berlin, pp. 9–16.

[27] VAN DER HOFSTAD, R. (2009). *Random Graphs and Complex Networks*. Available at http://www.win.tue.nl/ rhofstad/NotesRGCN2009.pdf.