Combinatorics, Probability and Computing (2017) 26, 697–719. © Cambridge University Press 2017
doi:10.1017/S0963548317000128

Bridge-Addability, Edge-Expansion and Connectivity

COLIN McDIARMID[1] and KERSTIN WELLER[1,2]

[1]Department of Statistics, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK
(e-mail: cmcd@stats.ox.ac.uk)
[2]Institut für Theoretische Informatik, Eigenössische Technische Hochschule Zürich,
8092 Zürich, Switzerland

Received 6 March 2014; revised 2 November 2016; first published online 11 May 2017

A class of graphs is called *bridge-addable* if, for each graph in the class and each pair u and v of vertices in different components, the graph obtained by adding an edge joining u and v must also be in the class. The concept was introduced in 2005 by McDiarmid, Steger and Welsh, who showed that, for a random graph sampled uniformly from such a class, the probability that it is connected is at least $1/e$.

We generalize this and related results to bridge-addable classes with edge-weights which have an edge-expansion property. Here, a graph is sampled with probability proportional to the product of its edge-weights. We obtain for example lower bounds for the probability of connectedness of a graph sampled uniformly from a relatively bridge-addable class of graphs, where some but not necessarily all of the possible bridges are allowed to be introduced. Furthermore, we investigate whether these bounds are tight, and in particular give detailed results about random forests in complete balanced multipartite graphs.

2010 *Mathematics subject classification*: Primary 05C80
Secondary 60C05

1. Introduction

Background. In recent years there has been a growing interest in random graphs sampled uniformly from a suitable structured class of (labelled) graphs, such as planar graphs. In particular, bridge-addable classes have received considerable attention. A set $\mathcal{A}$ of graphs is called *bridge-addable* if, for any graph G in $\mathcal{A}$ and any pair of vertices u, v in different components, the graph $G \cup uv$ obtained by adding the edge uv to G is also in $\mathcal{A}$. The concept of bridge-addability was introduced by McDiarmid, Steger and Welsh [13] in 2005 in the course of studying random planar graphs, and examples include forests, series-parallel graphs, planar graphs, and more generally graphs embeddable on any fixed surface, and any minor-closed class of graphs where the forbidden minors are all 2-connected.

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

If $\mathcal{S}$ is a finite non-empty set we write $R \in_u \mathcal{S}$ to mean that R is drawn uniformly at random from $\mathcal{S}$. When we use such notation we assume implicitly that $\mathcal{S}$ is non-empty. For a collection $\mathcal{A}$ of graphs we write $R_n \in_u \mathcal{A}$ to mean that R_n is drawn uniformly at random from $\mathcal{A}_n$, the set of all graphs in $\mathcal{A}$ on vertex set $[n] := \{1, \ldots, n\}$. By Theorem 2.2 in [13], for every finite bridge-addable set $\mathcal{A}$ of graphs and for $R \in_u \mathcal{A}$,

$$\mathbb{P}(R \text{ is connected}) \geqslant e^{-1}, \tag{1.1}$$

and indeed the random number $\kappa(R)$ of components is stochastically at most $1 + \mathrm{Po}(1)$, that is,

$$\kappa(R) \leqslant_s 1 + \mathrm{Po}(1). \tag{1.2}$$

Here Po(1) denotes a random variable which has the Poisson distribution with mean 1, and for random variables X and Y, we say that X is *stochastically at most* Y and write $X \leqslant_s Y$ if $\mathbb{P}(X \leqslant t) \geqslant \mathbb{P}(Y \leqslant t)$ for each t.

Let $\mathcal{F}$ be the class of forests (acyclic graphs). By a result of Rényi [18] in 1959, for $F_n \in_u \mathcal{F}$ we have

$$\mathbb{P}(F_n \text{ is connected}) \to e^{-1/2} \quad \text{as } n \to \infty,$$

and indeed

$$\kappa(F_n) \xrightarrow{d} 1 + \mathrm{Po}\left(\frac{1}{2}\right) \quad \text{as } n \to \infty.$$

Since forests are plausibly the 'least connected' bridge-addable class of graphs, it was natural to think that at least asymptotically (1.1) is not tight, and the following conjecture was made.

Conjecture 1.1 ([14]). *If the set $\mathcal{A}$ of graphs is bridge-addable and $R_n \in_u \mathcal{A}$, then*

$$\liminf_{n\to\infty} \mathbb{P}(R_n \textit{ is connected}) \geqslant e^{-1/2}.$$

It was proved independently in [1] and [9] that under a further assumption on the class $\mathcal{A}$ – the class $\mathcal{A}$ has to be *bridge-alterable* – Conjecture 1.1 holds. A class of graphs $\mathcal{A}$ is *bridge-alterable* if it is closed under adding and deleting bridges. Thus $\mathcal{A}$ is bridge-alterable exactly when it satisfies the condition that, for every graph H and every bridge e in H, the graph H belongs to $\mathcal{A}$ if and only if $H \setminus e$ (the graph obtained by deleting e) belongs to $\mathcal{A}$. For $\mathcal{A}$ bridge-alterable and $R_n \in \mathcal{A}$, we have [1, 9]

$$\liminf_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geqslant e^{-1/2}. \tag{1.3}$$

Clearly, the class of forests is bridge-alterable, and so the result in (1.3) is best possible. Also, all the examples of graph classes mentioned previously (that is, minor-closed classes of graphs defined by forbidding 2-connected minors and graphs embeddable on a fixed surface) are bridge-alterable. The first improvement on the bound in (1.1) was due to Balister, Bollobás and Gerke [4, 5], who proved that

$$\liminf_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geqslant e^{-0.7983}. \tag{1.4}$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

This was improved by Norine [17], who showed that

$$\liminf_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geqslant e^{-2/3}, \tag{1.5}$$

and indeed his arguments yield stronger results, for example

$$\liminf_{n\to\infty} \mathbb{P}(R_n \text{ is connected}) \geqslant e^{-9/14}. \tag{1.6}$$

Given a sequence of non-negative integer-valued random variables $X_1, X_2, \ldots$ and Y, we say that X_n is *stochastically at most Y asymptotically* and write $X_n \lesssim_s Y$ if, for each fixed $t \geqslant 0$,

$$\limsup_{n\to\infty} \mathbb{P}(X_n \geqslant t) \leqslant \mathbb{P}(Y \geqslant t).$$

In [11, Theorem 2.4] it was shown that the result (1.3) also corresponds to a better asymptotic bound for $\kappa(R_n)$: for any bridge-alterable class $\mathcal{A}$ of graphs and corresponding random graph $R_n \in_u \mathcal{A}$,

$$\kappa(R_n) \lesssim_s 1 + \mathrm{Po}\left(\frac{1}{2}\right).$$

Similarly, the proof of (1.4) together with [11, Lemma 3.3] gives that for any bridge-addable class $\kappa(R_n) \lesssim_s 1 + \mathrm{Po}(0.7983)$.

Let the *fragment* $\mathrm{Frag}(G)$ of a graph G be the subgraph induced on the vertices outside the largest component (with ties broken arbitrarily), and let the *fragment size* $\mathrm{frag}(G)$ be the number of vertices in $\mathrm{Frag}(G)$. In [11] (see inequality (7) in that paper) it was shown that for each bridge-addable class $\mathcal{A}$ and each n, for $R_n \in_u \mathcal{A}$,

$$\mathbb{E}[\mathrm{frag}(R_n)] < 2, \tag{1.7}$$

generalizing and improving Lemma 2.6 in [10]. For $F_n \in_u \mathcal{F}$, where $\mathcal{F}$ is the class of forests, we know that $\mathbb{E}[\mathrm{frag}(F_n)] \to 1$ (see for example [12]), which leads us to the next conjecture, extending Conjecture 1.1.

Conjecture 1.2 ([11]). *Let the set $\mathcal{A}$ of graphs be bridge-addable and let $R_n \in_u \mathcal{A}$. Then*
(a) $\kappa(R_n) \lesssim_s 1 + \mathrm{Po}(1/2)$ *and*
(b) $\limsup_{n\to\infty} \mathbb{E}[\mathrm{frag}(R_n)] \leqslant 1$.

Recently (after this paper was originally submitted), a proof of the full Conjecture 1.1 was presented by Chapuy and Perarnau [7].

Edge expansion. In the following we will generalize the definition of bridge-addable and show that the results (1.1), (1.2) and (1.7) are in fact special cases of a more general picture.

Given a set V, let $V^{(2)}$ denote the set of subsets of V of size 2. Let $\lambda = (\lambda_e : e \in [n]^{(2)}, 0 \leqslant \lambda_e \leqslant 1)$ be a family of edge-weights for the complete graph K_n. For a graph G and disjoint vertex sets B and C, we let $E(B, C)$ denote the set of edges between B and C. If $\sum_{e\in E(B,C)} \lambda_e \geqslant \alpha|B||C|$ for each partition $B \cup C$ of the vertex set, we say that the family λ is α-*edge-expanding*. We use the notation $\lambda(G) := \prod_{e\in E(G)} \lambda_e$ and, given a set $\mathcal{A}$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

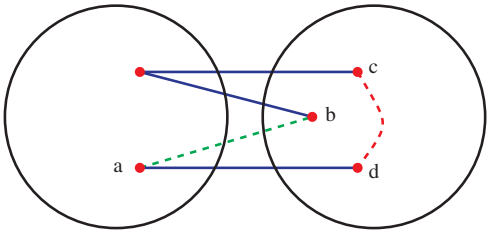


Figure 1. $K_{n/2,n/2}$ as a host graph H. Suppose the graph with the blue (solid) edges is in our bridge-addable class. Then the graph consisting of the blue edges and the green edge ab also has to be in the class, but not the graph with the blue edges and the red edge cd.

of graphs, $\lambda(\mathcal{A}_n) := \sum_{G\in\mathcal{A}_n}\lambda(G)$. We write $R_n \in_\lambda \mathcal{A}$ to denote the random graph in $\mathcal{A}_n$ sampled with respect to the edge-weights λ, that is,

$$\mathbb{P}(R_n = G) = \frac{\lambda(G)}{\lambda(\mathcal{A}_n)} \quad \text{for each graph } G \in \mathcal{A}_n.$$

We briefly recall the definition of 'relatively bridge-addable' introduced in [15]. Fix a graph H and call it the *host graph*. Let $\mathcal{A}'$ be a collection of spanning subgraphs of H (*i.e.* subgraphs with the same vertex set as H). We call $\mathcal{A}'$ *bridge-addable relative to* H if, for every $G \in \mathcal{A}'$ and every pair of vertices u, v which are adjacent in H and lie in different components of G, the graph $G \cup uv$ is also in $\mathcal{A}'$. An illustration can be found in Figure 1.

Observe that 'bridge-addable relative to K_n' is equivalent to 'bridge-addable'. Observe also that, for any graph H, the set of acyclic subgraphs of H is bridge-addable relative to H.

If $0 < \alpha \leqslant 1$ and $|E(B,C)| \geqslant \alpha|B||C|$ for each partition $B \cup C$ of $V(H)$, we say that H is an *α-edge-expander*. Thus the complete graph K_n is a 1-edge-expander. Every α-edge-expander corresponds to an α-edge-expanding family λ of edge-weights, where for each e present in the host graph $\lambda_e = 1$ and $\lambda_e = 0$ otherwise. A natural example of an expander is a d-regular graph on n vertices. In this case we may take $\alpha = (d-\mu)/n$, where μ is the second largest eigenvalue of the adjacency matrix of H; see [2, Theorem 9.1.2]. Hence (or otherwise) $K_{d,d}$ is a $\frac{1}{2}$-edge-expander. Another example is the classical random graph $G_{n,p}$: with high probability $G_{n,p}$ is a $(1-\epsilon)p$-edge-expander for $p \geqslant c\log n/n$, where c is a large constant. To see this, observe that for any fixed partition with $|B| = i$ and $|C| = n - i$, the probability that $|E(B,C)| < (1-\epsilon)p\, i(n-i)$ is at most $e^{-\Theta(\epsilon^2 i(n-i)p)}$ by standard Chernoff bounds; see for example [6]. Hence, for $p \geqslant c\log n/n$, with $c = c(\epsilon)$ large enough, we see that the probability that $G_{n,p}$ is not a $(1-\epsilon)p$-edge-expander is at most $\sum_{1\leqslant i\leqslant n/2} n^i n^{-2i} = o(1)$.

Conversely, let $\mathcal{A}$ be a bridge-addable class of graphs, and let λ be an α-edge-expanding family of edge-weights such that $\lambda_e \in \{0,1\}$ for each edge e in K_n. Then the graph H obtained from K_n by deleting edges of weight 0 is an α-edge-expander, and the set $\mathcal{A}'$ of subgraphs of H in $\mathcal{A}_n$ is bridge-addable relative to H.

Plan of the paper. The paper is organized as follows. Section 2 contains the generalization of (1.1), (1.2) and (1.7) to bridge-addable classes with edge-weights, where the family λ

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

of edge-weights is α-edge-expanding. The results on relatively bridge-addable classes of graphs given in [15, Theorems 2.1, 2.2] will follow immediately. In Section 3 we look at a particular family of natural examples: we consider random forests in a balanced complete k-partite host graph. In particular, this section will include full proofs for the results on forests in complete bipartite graphs stated in [15]. We also obtain similar results for random k-coloured forests. Finally, there are brief concluding remarks in Section 4.

2. Bridge-addable graph classes and edge-expanding weights

Recall that K_n is an α-edge-expander for $\alpha = 1$, and hence the family λ of edge-weights with $\lambda_e = 1$ for every edge e is α-edge-expanding for $\alpha = 1$. The following result thus generalizes inequalities (1.1) and (1.2).

Theorem 2.1. *Let $0 < \alpha \leqslant 1$, let $n \in \mathbb{N}$, and let the edge-weights $\lambda = (\lambda_e : e \in [n]^{(2)}, 0 \leqslant \lambda_e \leqslant 1)$ be α-edge-expanding. Let $\mathcal{A}$ be a bridge-addable set of graphs on $[n]$, and let $R_n \in_\lambda \mathcal{A}$. Then*

$$\kappa(R_n) \leqslant_s 1 + \mathrm{Po}\left(\frac{1}{\alpha}\right),$$

and in particular

$$\mathbb{P}(R_n \text{ is connected}) \geqslant e^{-1/\alpha}.$$

To prove this result we will use one preliminary lemma.

Lemma 2.2 (Lemma 3.4. in [11]). *Let the random variable X take non-negative integer values. Let $\alpha > 0$ and suppose*

$$\mathbb{P}(X = k+1) \leqslant \frac{\alpha}{k+1}\mathbb{P}(X = k) \quad \text{for each } k = 0, 1, 2, \ldots.$$

Then $X \leqslant_s Y$, where $Y \sim \mathrm{Po}(\alpha)$. □

Proof of Theorem 2.1. For a graph H let $\mathrm{Bridge}(H)$ denote the set of bridges of H, with $|\mathrm{Bridge}(H)| = \mathrm{bridge}(H)$; and let $\mathrm{Cross}(H)$ denote the set of edges of G between components of H. Let $\mathcal{A}^k$ denote the set of graphs in $\mathcal{A}$ with k components. For $k = 1, \ldots, n$, define

$$a_k = \min_{H \in \mathcal{A}^k} \sum_{e \in \mathrm{Cross}(H)} \lambda_e$$

and

$$b_k = \max_{H \in \mathcal{A}^k} \mathrm{bridge}(H).$$

Then, for each $k = 1, \ldots, n-1$,

$$a_{k+1}\lambda(\mathcal{A}^{k+1}) \leqslant \sum_{H \in \mathcal{A}^{k+1}} \lambda(H) \sum_{e \in \mathrm{Cross}(H)} \lambda_e \leqslant \sum_{H' \in \mathcal{A}^k} \lambda(H')\,\mathrm{bridge}(H') \leqslant b_k\,\lambda(\mathcal{A}^k).$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

To see the middle inequality here, note that for each $H \in \mathcal{A}^{k+1}$ and $e \in \mathrm{Cross}(H)$, the graph $H' = H \cup e$ is in $\mathcal{A}^k$, $\lambda(H') = \lambda(H)\lambda_e$, and H' is counted at most bridge(H') times.

Observe that $b_k \leqslant n-k$ since the number of edges in a spanning forest of a graph on n vertices with k components is $n-k$. We shall show that

$$a_{k+1} \geqslant \alpha k(n-k). \tag{2.1}$$

Then it will follow that

$$\lambda(\mathcal{A}^{k+1}) \leqslant \frac{b_k}{a_{k+1}}\lambda(\mathcal{A}^k) \leqslant \frac{1}{\alpha k}\lambda(\mathcal{A}^k),$$

and so

$$\mathbb{P}(\kappa(R_n) = k+1) = \frac{\lambda(\mathcal{A}^{k+1})}{\lambda(\mathcal{A})} \leqslant \frac{1}{\alpha k}\mathbb{P}(\kappa(R_n) = k). \tag{2.2}$$

Applying Lemma 2.2 together with (2.2) then will give the theorem.

It remains to show that (2.1) holds. We want to find a graph H with $k+1$ components which minimizes $\sum\{\lambda_e : e \in \mathrm{Cross}(H)\}$. Take an arbitrary graph $H \in \mathcal{A}$ with $k+1$ components, where the components have $n_1, \ldots, n_{k+1}$ vertices. As λ is α-edge-expanding, we get

$$\sum_{e \in \mathrm{Cross}(H)} \lambda_e \geqslant \frac{\alpha}{2}\sum_{i=1}^{k+1} n_i(n-n_i) = \frac{\alpha}{2}\left(n^2 - \sum_{i=1}^{k+1} n_i^2\right).$$

Thus we want to maximize $\sum_{i=1}^{k+1} n_i^2$ subject to $\sum_{i=1}^{k+1} n_i = n$ and $n_i \geqslant 1$ for all i. The maximum is attained at $n_1 = n-k$ and $n_2 = \cdots = n_{k+1} = 1$ with value $(n-k)^2 + k = n^2 - 2kn + k^2 + k$. Hence

$$\sum_{e \in \mathrm{Cross}(H)} \lambda_e \geqslant \frac{\alpha}{2}(2nk - k^2 - k) = \frac{\alpha}{2}(2k(n-k) + k^2 - k) \geqslant \alpha k(n-k),$$

and so $a_{k+1} \geqslant \alpha k(n-k)$, as required. □

As remarked in the Introduction, every α-edge-expanding host graph corresponds to $\{0,1\}$-valued α-edge-expanding edge-weights. Hence by applying Theorem 2.1 to these particular cases we obtain the following result, which is Theorem 2.1 in [15].

Corollary 2.3. *Let $0 < \alpha \leqslant 1$, let G be an α-edge-expander, let the set $\mathcal{A}$ of spanning subgraphs of G be bridge-addable relative to G, and let $R_n \in_u \mathcal{A}$. Then*

$$\kappa(R_n) \leqslant_s 1 + \mathrm{Po}\left(\frac{1}{\alpha}\right),$$

and in particular

$$\mathbb{P}(R_n \text{ is connected}) \geqslant e^{-1/\alpha}.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Recall that $\mathrm{frag}(G)$ is the number of vertices which are not in the largest component. The following result on the expected number of vertices not in the largest component generalizes the result (1.7) above (from [11]).

Theorem 2.4. *Let $0 < \alpha \leqslant 1$, let $n \in \mathbb{N}$, and let the edge-weights $\lambda = (\lambda_e : e \in [n]^{(2)}, 0 \leqslant \lambda_e \leqslant 1)$ be α-edge-expanding. Let $\mathcal{A}$ be a bridge-addable set of graphs on $[n]$, and let $R_n \in_\lambda \mathcal{A}$. Then*

$$\mathbb{E}[\mathrm{frag}(R_n)] < \frac{2}{\alpha}.$$

Proof. Let us prove the following claim first:

$$\sum_{e \in \mathrm{Cross}(H)} \lambda_e \geqslant \alpha \frac{n}{2} \mathrm{frag}(H) \tag{2.3}$$

for each graph H in $\mathcal{A}$.

Observe that it is trivially true for $\mathrm{frag}(H) \leqslant n/2$ as λ is α-edge-expanding. For the general case, let $H \in \mathcal{A}$ have k components, with $n_1 \geqslant \cdots \geqslant n_k$ vertices. As λ is α-edge-expanding, as earlier,

$$\sum_{e \in \mathrm{Cross}(H)} \lambda_e \geqslant \frac{\alpha}{2} \sum_{i=1}^{k} n_i(n - n_i) = \frac{\alpha}{2} \left(n^2 - \sum_{i=1}^{k} n_i^2 \right).$$

But

$$\sum_{i=1}^{k} n_i^2 \leqslant \sum_{i=1}^{k} n_1 n_i = n_1 n.$$

Thus, since $\mathrm{frag}(H) = n - n_1$, we get

$$\sum_{e \in \mathrm{Cross}(H)} \lambda_e \geqslant \frac{\alpha}{2} n(n - n_1) = \frac{\alpha n}{2} \mathrm{frag}(H),$$

which establishes the claim (2.3).

By (2.3), and arguing as earlier, we have

$$\frac{\alpha n}{2} \sum_{H \in \mathcal{A}} \lambda(H) \mathrm{frag}(H) \leqslant \sum_{H \in \mathcal{A}} \lambda(H) \sum_{e \in \mathrm{Cross}(H)} \lambda_e \leqslant \sum_{H \in \mathcal{A}} \lambda(H)\, \mathrm{bridge}(H).$$

Hence

$$\begin{aligned} \mathbb{E}[\mathrm{frag}(R_n)] &= \sum_{H \in \mathcal{A}} \frac{\lambda(H)\, \mathrm{frag}(H)}{\lambda(\mathcal{A})} \\ &\leqslant \frac{1}{\lambda(\mathcal{A})} \cdot \frac{2}{\alpha n} \sum_{H \in \mathcal{A}} \lambda(H)\, \mathrm{bridge}(H) < \frac{2}{\alpha} \end{aligned}$$

since $\mathrm{bridge}(H) < n$, completing the proof of the theorem. □

As before, we may recover the corresponding result for relatively bridge-addable classes of graphs: the following result is Theorem 2.2 in [15].

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Corollary 2.5. *Let H be an α-edge-expander, let the set $\mathcal{A}$ of spanning subgraphs of H be bridge-addable relative to H, and let $R_n \in_u \mathcal{A}$. Then*

$$\mathbb{E}[\mathrm{frag}(R_n)] < \frac{2}{\alpha}.$$

3. Balanced complete multipartite graphs as host

What happens for particular α-edge-expanding edge-weights and for α-edge-expanders? How tight are the bounds in Theorem 2.1 and Theorem 2.4? In the (fully) bridge-addable case, Conjectures 1.1 (now proved) and 1.2 say that the best bounds correspond to forests. In this section we consider the connectivity of the forests in complete multipartite graphs. First results on forests in the bipartite graph $K_{n/2,n/2}$ were stated without proof in [15]. In this section, we will extend these results to forests in multipartite graphs and provide complete proofs, using the results of the last section.

Fix an integer $k \geqslant 2$. Given a k-tuple $\mathbf{n} = (n_1, \ldots, n_k)$ of positive integers, let $K_{\mathbf{n}}$ denote the complete k-partite graph with parts of sizes $n_1, \ldots, n_k$. For definiteness, we assume that the vertex set is $[n]$ where $n = \sum_i n_i$, that the first part is $\{1, \ldots, n_1\}$, the second is $\{n_1 + 1, \ldots, n_1 + n_2\}$ and so on. We want to consider the balanced or nearly balanced case. Let H_n denote $K_{\mathbf{n}}$ when $\lceil n/k \rceil = n_1 \geqslant \cdots \geqslant n_k = \lfloor n/k \rfloor$; this is the balanced case. When $k = 2$ and n is even this graph H_n is $K_{n/2,n/2}$. When $k|n$ the graph H_n is regular of degree $((k-1)/k)n$, and is a $(k-1)/k$-edge-expander. In general, H_n is an α-edge-expander where $\alpha = (k-1)/k + o(1)$. The result on α-edge-expanders (Theorem 2.1) gives that $\mathbb{P}(R_n \text{ connected}) \geqslant e^{-1/\alpha}$ and more generally that $\kappa(R_n) \leqslant_s 1 + \mathrm{Po}(1/\alpha)$.

Consider $\mathbf{n} = (n_1, \ldots, n_k)$ where each $n_i = n_i(n) \sim n/k$ and $\sum_i n_i = n$. We call $\mathbf{n}$ a *near-balanced* sequence of partitions of n, and call $K_{\mathbf{n}}$ a *near-balanced* sequence of complete k-partite graphs. It seems natural to focus mainly on the balanced case. For most of the following theorem it suffices to assume that $\mathbf{n}$ is near-balanced. For the last part, concerning $\mathbb{E}[\mathrm{frag}(R_n)]$, we assume that $\mathbf{n}$ is closer to being balanced, so that we can give a proof of modest length.

Theorem 3.1. *Fix an integer $k \geqslant 2$. Let $K_{\mathbf{n}}$ be a near-balanced sequence of complete k-partite graphs, and let the random graph R_n be sampled uniformly at random from the forests in $K_{\mathbf{n}}$. Let $\lambda = k/(2(k-1))$. Then, as $n \to \infty$:*

(a) *$\kappa(R_n)$ tends to $1 + \mathrm{Po}(\lambda)$ in distribution, and in particular*

$$\mathbb{P}(R_n \text{ is connected}) \to e^{-\lambda};$$

(b) *the unlabelled trees T appear in $\mathrm{Frag}(R_n)$ asymptotically independently, with distribution $\mathrm{Po}(\lambda(T))$ where $\lambda(T) = k/(k-1) \cdot (e^{v(T)}\,\mathrm{aut}(T))^{-1}$; and*

(c) *if each $n_i = n/k + O(\sqrt{n})$ then*

$$\mathbb{E}[\mathrm{frag}(R_n)] \to \frac{k}{k-1}.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

For comparison, let F_n be sampled uniformly from the n-vertex forests in K_n. Then

$$\mathbb{P}(F_n \text{ is connected}) \to e^{-1/2} \quad \text{as } n \to \infty,$$

and indeed $\kappa(F_n)$ converges in distribution to $1 + \mathrm{Po}(1/2)$, as mentioned earlier. Furthermore, as $n \to \infty$, $\mathbb{E}[\mathrm{frag}(F_n)] \to 1$, and the unlabelled trees T appear in $\mathrm{Frag}(F_n)$ asymptotically independently, with distribution $\mathrm{Po}(\mu(T))$ where $\mu(T) = 1/(e^{v(T)}\,\mathrm{aut}(T))$ (see [13]). Thus by Theorem 3.1, not only is it true that

$$\mathbb{E}[\kappa(\mathrm{Frag}(R_n))] \sim \frac{k}{k-1}\,\mathbb{E}[\kappa(\mathrm{Frag}(F_n))]$$

and

$$\mathbb{E}[\mathrm{frag}(R_n)] \sim \frac{k}{k-1}\,\mathbb{E}[\mathrm{frag}(F_n)],$$

but it is true at the detailed level that for each unlabelled tree T the expected number of appearances of T in $\mathrm{Frag}(R_n)$ is asymptotically $k/(k-1)$ times that in $\mathrm{Frag}(F_n)$.

For simplicity, consider the case $k|n$, where $K_{\mathbf{n}}$ is a $(k-1)/k$-edge-expander. By Theorem 2.1 we have $\kappa(R_n) \leqslant_s 1 + \mathrm{Po}(2\lambda)$. Hence it will follow from part (a) of Theorem 3.1 that, for each fixed positive integer k, $\mathbb{E}[\kappa(R_n)^k] \to \mathbb{E}[(1 + \mathrm{Po}(\lambda))^k]$ as $n \to \infty$. In particular, $\mathbb{E}[\kappa(R_n)] \to 1 + \lambda$ and the variance of $\kappa(R_n)$ tends to λ as $n \to \infty$.

In order to prove Theorem 3.1, we introduce a variant S_n of the random forest R_n, and estimate the probability that $\mathrm{Frag}(S_n)$ is a given forest F, in terms of the probability that R_n is connected; see (3.1) below. By summing over F we estimate the probability that $\mathrm{Frag}(R_n)$ consists of j trees each of size at most ℓ, again in terms of the probability that R_n is connected; see (3.2) below. From that result we determine the limiting probability that R_n is connected, and complete the proof of part (a). We may then prove part (b) by noting that we can ignore large components, and again summing results like (3.1) approximating the probability that $\mathrm{Frag}(S_n)$ is a given forest, except with the probability that R_n is connected replaced by its limit $e^{-\lambda}$. Finally, part (c) follows from part (b), once we show that the expected contribution from large components is negligible.

For the proof of parts (a) and (b) of Theorem 3.1 we need one preliminary lemma, the following exact counting result. Before we move on to prove part (c) of the theorem we will introduce several more lemmas.

Lemma 3.2 ([3, 19]). *Given an integer $k \geqslant 2$ and a k-tuple $\mathbf{n} = (n_1, \ldots, n_k)$ of positive integers with sum n, the number $t(\mathbf{n})$ of spanning trees in the complete k-partite graph $K_{\mathbf{n}}$ (with a fixed partition) is*

$$t(\mathbf{n}) = n^{k-2} \prod_{i=1}^{k} (n - n_i)^{n_i - 1}.$$

More references and a proof of this result can be found in [16, pp. 30–32].

Proof of Theorem 3.1, part (a). Call a partition $\mathbf{U} = (U_1, \ldots, U_k)$ of $[n]$ an $\mathbf{n}$-partition if $|U_i| = n_i$ for each i. Consider the random graph S_n generated by choosing an $\mathbf{n}$-partition

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

$\mathbf{U} = (U_1, \ldots, U_k)$ of $[n]$ uniformly at random, letting $\tilde{H}_n$ be the complete multipartite graph with parts U_i, and picking a forest in $\tilde{H}_n$ uniformly at random. Clearly $\kappa(S_n)$ and $\kappa(R_n)$ have the same distribution, and indeed the unlabelled graphs corresponding to S_n and R_n have the same distribution.

Fix a forest F with vertex set a subset of $[n]$, which consists of j trees with $t_1, \ldots, t_j$ vertices, and let $t = \sum_i t_i$. Then we claim that

$$\mathbb{P}(\mathrm{Frag}(S_n) = F) \sim \mathbb{P}(R_n \text{ is connected}) \cdot \left(\frac{k}{k-1}\right)^j (en)^{-t}. \tag{3.1}$$

It will take us some time to prove (3.1).

Note first that the probability that $\mathbf{U} = (U_1, \ldots, U_k)$ gives a proper k-colouring of F is asymptotically equal to $((k-1)/k)^{t-j}$. To see this, consider the trees in F one after another, each time picking a root vertex and then exploring the tree using, say, breadth-first search. At each step when we arrive at a non-root vertex v, whatever has happened so far, the probability that the edge to v is properly coloured is $\sim (k-1)/k$. The result now follows, since there are $t - j$ edges.

Fix an $\mathbf{n}$-partition $\mathbf{U}^0 = (U_1^0, \ldots, U_k^0)$ of $[n]$ that gives a proper k-colouring of F. Let x_i be the number of vertices of F in U_i^0, so that $\sum_i x_i = t$. Then by Lemma 3.2 and the fact that $(1 + y/m)^m \to e^y$ as $m \to \infty$,

$$\begin{aligned}
&\mathbb{P}(\mathrm{Frag}(S_n) = F \mid \mathbf{U} = \mathbf{U}^0)/\mathbb{P}(S_n \text{ is connected}) \\
&\quad = \frac{(n-t)^{k-2} \prod_{i=1}^k (n - t - (n_i - x_i))^{n_i - x_i - 1}}{n^{k-2} \prod_{i=1}^k (n - n_i)^{n_i - 1}} \\
&\quad = \left(1 - \frac{t}{n}\right)^{k-2} \frac{\prod_{i=1}^k (n - n_i)^{n_i - x_i - 1} \left(1 - \frac{t - x_i}{n - n_i}\right)^{n_i - x_i - 1}}{\prod_{i=1}^k (n - n_i)^{n_i - 1}} \\
&\quad \sim \prod_{i=1}^k (n - n_i)^{-x_i} \left(1 - \frac{t - x_i}{n - n_i}\right)^{n_i - x_i - 1} \\
&\quad \sim \left(\frac{k-1}{k} n\right)^{-\sum_i x_i} \cdot \prod_{i=1}^k \left(1 - \frac{k}{k-1} \frac{t - x_i}{n}\right)^{n/k - x_i - 1} \\
&\quad \sim \left(\frac{k-1}{k} n\right)^{-t} \cdot \prod_{i=1}^k e^{-\frac{(t - x_i)}{(k-1)}} \\
&\quad = \left(\frac{k-1}{k} en\right)^{-t}
\end{aligned}$$

since

$$\sum_i \frac{t - x_i}{k-1} = t.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

This holds uniformly for each relevant partition $\mathbf{U}^0$, and hence we can relax the conditioning, to obtain

$$\frac{\mathbb{P}(\mathrm{Frag}(S_n) = F \mid \mathbf{U} \text{ properly colours } F)}{\mathbb{P}(S_n \text{ is connected})} \sim \left(\frac{k-1}{k}\, en\right)^{-t}.$$

Hence

$$\begin{aligned}
&\frac{\mathbb{P}(\mathrm{Frag}(S_n) = F)}{\mathbb{P}(S_n \text{ is connected})} \\
&\quad = \frac{\mathbb{P}(\mathrm{Frag}(S_n) = F) \mid \mathbf{U} \text{ properly colours } F)}{\mathbb{P}(S_n \text{ is connected})} \cdot \mathbb{P}(\mathbf{U} \text{ properly colours } F) \\
&\quad \sim \left(\frac{k-1}{k}\, en\right)^{-t} \cdot \left(\frac{k-1}{k}\right)^{t-j} = \left(\frac{k-1}{k}\right)^{-j} \cdot (en)^{-t}.
\end{aligned}$$

Finally we obtain (3.1), on recalling that

$$\mathbb{P}(S_n \text{ is connected}) = \mathbb{P}(R_n \text{ is connected}).$$

Our next task is to use (3.1) to prove (3.2) below. We need some notation. Let

$$\beta_\ell = \sum_{j=1}^{\ell} \frac{j^{j-2}}{j!\, e^j},$$

which is the truncated generating function of (unrooted labelled) trees, evaluated at its radius of convergence $1/e$. Let $\tilde{\beta}_\ell = (k/(k-1))\beta_\ell$. It is well known (see for example [8, p. 404 *ff.*]) that β_ℓ increases to $1/2$ as $\ell \to \infty$, and so $\tilde{\beta}_\ell$ increases to λ. Thus $\sum_{j=0}^{\ell}((\tilde{\beta}_\ell)^j/j!)$ increases to e^λ as $\ell \to \infty$.

Let $p(n, j, \leqslant \ell)$ be the probability that $\mathrm{Frag}(R_n)$ consists of j component trees, each of order at most ℓ. We claim that, for each fixed j and ℓ, as $n \to \infty$,

$$p(n, j, \leqslant \ell) \sim \mathbb{P}(R_n \text{ is connected}) \cdot \frac{(\tilde{\beta}_\ell)^j}{j!}. \tag{3.2}$$

Let us prove this. Let $\mathcal{F}_n(j, \leqslant \ell)$ denote the set of forests on $[n]$ consisting of j trees each with at most ℓ vertices. Then by (3.1) we have

$$\begin{aligned}
&p(n, j, \leqslant \ell)/\mathbb{P}(R_n \text{ is connected}) \\
&\quad = \sum_{F \in \mathcal{F}_n(j, \leqslant \ell)} \mathbb{P}(\mathrm{Frag}(S_n) = F)/\mathbb{P}(R_n \text{ is connected}) \\
&\quad \sim \sum_{F \in \mathcal{F}_n(j, \leqslant \ell)} \left(\frac{k-1}{k}\right)^j \cdot (en)^{-v(F)} \\
&\quad = \frac{1}{j!} \sum_{1 \leqslant t_1 \leqslant \ell} \cdots \sum_{1 \leqslant t_j \leqslant \ell} \binom{n}{t_1} \binom{n-t_1}{t_2} \cdots \binom{n-(t_1+\cdots+t_{j-1})}{t_j} \\
&\qquad\qquad \times \prod_{i=1}^{j} t_i^{t_i-2} \cdot \left(\frac{k}{k-1}\right)^j (en)^{-\sum_i t_i}
\end{aligned}$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

$$\sim \frac{1}{j!} \sum_{1 \leqslant t_1 \leqslant \ell} \cdots \sum_{1 \leqslant t_j \leqslant \ell} \prod_{i=1}^{j} \frac{n^{t_i}}{t_i!} t_i^{t_i-2} \cdot \left(\frac{k}{k-1}\right)^j (en)^{-\sum_i t_i}$$

$$= \frac{1}{j!} \sum_{1 \leqslant t_1 \leqslant \ell} \cdots \sum_{1 \leqslant t_j \leqslant \ell} \prod_{i=1}^{j} \left(\frac{t_i^{t_i-2}}{t_i! e^{t_i}} \cdot \frac{k}{k-1}\right)$$

$$= \frac{1}{j!} \left(\sum_{i=1}^{\ell} \frac{i^{i-2}}{i! e^i} \cdot \frac{k}{k-1}\right)^j = \frac{1}{j!} (\tilde{\beta}_\ell)^j,$$

which completes the proof of (3.2).

We may now use (3.2) to complete the proof of part (a). Let

$$p(n, \leqslant \ell, \leqslant \ell) = \sum_{j=0}^{\ell} p(n, j, \leqslant \ell),$$

the probability that $\mathrm{Frag}(R_n)$ consists of at most ℓ components each of size at most ℓ. By Theorem 2.4, $\mathbb{E}[\mathrm{frag}(R_n)] \leqslant c$ for a suitable constant c. Hence by Markov's inequality

$$1 - p(n, \leqslant \ell, \leqslant \ell) \leqslant \mathbb{P}(\mathrm{frag}(R_n) > \ell) < \mathbb{E}[\mathrm{frag}(R_n)]/\ell \leqslant c/\ell.$$

So by taking ℓ large enough we may ensure that $p(n, \leqslant \ell, \leqslant \ell)$ is arbitrarily close to 1. But by (3.2), for each (fixed) ℓ,

$$p(n, \leqslant \ell, \leqslant \ell) \sim \mathbb{P}(R_n \text{ is connected}) \cdot \sum_{j=0}^{\ell} \frac{1}{j!} (\tilde{\beta}_\ell)^j.$$

Also, as we noted earlier, the sum on the right side tends to e^λ as $\ell \to \infty$. Thus

$$\mathbb{P}(R_n \text{ is connected}) \to e^{-\lambda} \quad \text{as } n \to \infty. \tag{3.3}$$

Furthermore, we may argue as above to see that

$$p(n, j, \leqslant \ell) \leqslant \mathbb{P}(\kappa(R_n) = 1 + j) \leqslant p(n, j, \leqslant \ell) + c/\ell,$$

and so for each fixed j, by taking ℓ large enough we can ensure that $p(n, j, \leqslant \ell)$ is arbitrarily close to $\mathbb{P}(\kappa(R_n) = 1 + j)$, and $(\tilde{\beta}_\ell)^j / j!$ is arbitrarily close to $\lambda^j / j!$. But by (3.2) and (3.3), as $n \to \infty$,

$$p(n, j, \leqslant \ell) \sim e^{-\lambda} \cdot \frac{(\tilde{\beta}_\ell)^j}{j!},$$

and so

$$\mathbb{P}(\kappa(R_n) = 1 + j) \to e^{-\lambda} \cdot \frac{\lambda^j}{j!}.$$

Thus $\kappa(R_n)$ converges in distribution to $1 + \mathrm{Po}(\lambda)$. This completes the proof of part (a). □

Proof of Theorem 3.1, part (b). Let $\mathcal{U}$ denote the class of all unlabelled trees. For $T \in \mathcal{U}$ let $\kappa(G, T)$ count the number of components of the graph G isomorphic to T. Let $A \subseteq \mathcal{U}$ be finite and non-empty. Let n_T be a non-negative integer for each $T \in A$, and let E_n^1 be

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

the event that $\kappa(R_n, T) = n_T$ for each $T \in A$. We wish to show that

$$\mathbb{P}(E_n^1) \to p := \prod_{T \in A} e^{-\lambda(T)} \frac{\lambda(T)^{n_T}}{n_T!} \quad \text{as } n \to \infty. \tag{3.4}$$

Let $0 < \epsilon < 1$. We shall show that $|\mathbb{P}(E_n^1) - p| < 4\epsilon$ for n sufficiently large, to complete the proof. Enumerate the trees in $\mathcal{U}$ as $T_1, T_2, \ldots$. We shall choose (large) constants m and b. For each positive integer j let $B_j = \{T_1, \ldots, T_j\}$, and let $B_j' = B_j \setminus A$ and $\overline{B_j} = \mathcal{U} \setminus B_j$.

Choose m large enough that $A \subseteq B_m$, and each tree in $\overline{B_m}$ has size at least c/ϵ. Then, since $\mathbb{E}[\text{frag}(R_n)] \leqslant c$,

$$\mathbb{P}(\kappa(\text{Frag}(R_n), T) > 0 \text{ for some } T \in \overline{B_m}) \leqslant \epsilon.$$

Let $\mathcal{T}$ denote the class of labelled trees. Recall that for the corresponding exponential generating function

$$T(z) = \sum_{T \in \mathcal{T}} \frac{z^{v(T)}}{v(T)!} = \sum_{n \geqslant 1} \frac{n^{n-2} z^n}{n!}$$

we have $T(1/e) = 1/2$.

Let $T \in \mathcal{U}$ (and think of T as an equivalence class of trees in $\mathcal{T}_n$ where $n = v(T)$): then

$$\frac{z^{v(T)}}{\text{aut}(T)} = \sum_{T' \in T} \frac{z^{v(T)}}{v(T)!}.$$

Hence

$$\sum_{T \in \mathcal{U}} \lambda(T) = \frac{k}{k-1} \sum_{T \in \mathcal{U}} \frac{e^{-v(T)}}{\text{aut}(T)} = \frac{k}{k-1} \cdot T\left(\frac{1}{e}\right) = \lambda.$$

Now increase m if necessary so that it is large enough that also

$$\prod_{T \in \overline{B_m}} e^{-\lambda(T)} \geqslant (1+\epsilon)^{-1}.$$

Choose b large enough that both $\mathbb{P}(\kappa(R_n) > b) \leqslant \epsilon$ for all n (which we can do by Corollary 2.3 or 2.5), and

$$\prod_{T \in B_m} \mathbb{P}(\text{Po}(\lambda(T)) \leqslant b) \geqslant (1+\epsilon)^{-1}.$$

Let S be a finite non-empty subset of $\mathcal{U}$, and let the unlabelled forest F consists of k_T copies of tree T for each T in S. Then the number of labelled copies of F on $[n]$ is

$$\sim \prod_{T \in S} \left(\left(\frac{n^{v(T)}}{\text{aut}(T)} \right)^{k_T} \frac{1}{k_T!} \right).$$

Also note that by (3.3) and the above result that $\sum_{T \in \mathcal{U}} \lambda(T) = \lambda$, we have

$$\mathbb{P}(R_n \text{ is connected}) \to e^{-\lambda} = e^{-\sum_{T \in \mathcal{U}} \lambda(T)} \quad \text{as } n \to \infty.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Hence by (3.1) (for n sufficiently large)

$$\begin{aligned}
&\mathbb{P}(\kappa(R_n,T)=k_T\ \forall T\in S,\ \kappa(\mathrm{Frag}(R_n),T)=0\ \forall T\in\mathcal{U}\setminus S)\\
&\quad=\mathbb{P}(\mathrm{Frag}(R_n)\cong F)\\
&\quad\sim\prod_{T\in S}\left(\left(\frac{n^{v(T)}}{\mathrm{aut}(T)}\right)^{k_T}\frac{1}{k_T!}(en)^{-v(T)k_T}\left(\frac{k}{k-1}\right)^{k_T}\right)\cdot e^{-\sum_{T'\in\mathcal{U}}\lambda(T')}\\
&\quad=\prod_{T\in S}\left(\left(\frac{n^{v(T)}}{\mathrm{aut}(T)}(en)^{-v(T)}\frac{k}{k-1}\right)^{k_T}\frac{1}{k_T!}e^{-\lambda(T)}\right)\cdot\prod_{T'\in\mathcal{U}\setminus S}e^{-\lambda(T')}\\
&\quad=\prod_{T\in S}e^{-\lambda(T)}\frac{\lambda(T)^{k_T}}{k_T!}\ \cdot\prod_{T'\in\mathcal{U}\setminus S}e^{-\lambda(T')}.
\end{aligned}$$

Next, we are going to use this result to deduce (3.4). Let m_T be a non-negative integer for each $T\in B'_m$ $(=B_m\setminus A)$, and let E_n^2 be the event that $\kappa(R_n,T)=m_T$ for each $T\in B'_m$. Let E_n^3 be the event that $\kappa(\mathrm{Frag}(R_n),T)=0$ for each $T\in\overline{B_m}$.

From the previous result with $S=A\cup B'_m$, we know that

$$\mathbb{P}(E_n^1\wedge E_n^2\wedge E_n^3)\to p\cdot\prod_{T\in B'_m}e^{-\lambda(T)}\frac{\lambda(T)^{m_T}}{m_T!}\prod_{T\in\overline{B_m}}e^{-\lambda(T)}\quad\text{as } n\to\infty.$$

Summing over the finite set of values $m_T=0,1,\ldots,b$ for $T\in B'_m$, we see that, as $n\to\infty$,

$$\mathbb{P}(E_n^1\wedge(\kappa(R_n,T)\leqslant b\ \forall T\in B'_m)\wedge E_n^3)\to p\cdot\prod_{T\in B'_m}\mathbb{P}(\mathrm{Po}(\lambda(T)\leqslant b)\prod_{T\in\overline{B_m}}e^{-\lambda(T)}.$$

Call the left side x_n and the right side p'. Then by our choice of m and b we have $x_n\leqslant\mathbb{P}(E_n^1)\leqslant x_n+2\epsilon$, and

$$p'\leqslant p\leqslant p'(1+\epsilon)^2\leqslant p'+3\epsilon.$$

Hence

$$|\mathbb{P}(E_n^1)-p|\leqslant|x_n-p'|+3\epsilon<4\epsilon$$

for n sufficiently large, as required to complete the proof of part (b). □

We now introduce several more lemmas, to prepare for the proof of part (c) of Theorem 3.1.

Lemma 3.3. *In Lemma 3.2, when the n_i differ by at most 1 we have*

$$t(\mathbf{n})\sim n^{n-2}\left(\frac{k-1}{k}\right)^{n-k}\quad\text{as } n\to\infty,$$

and this gives the maximum value, that is,

$$n^{k-2}\prod_{i=1}^{k}(n-n_i)^{n_i-1}\leqslant n^{n-2}\left(\frac{k-1}{k}\right)^{n-k}.\tag{3.5}$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Proof. The first part follows easily from Lemma 3.2. To show (3.5), let $x_i = n_i/n$, so that $0 < x_i < 1$ for each i and $\sum_i x_i = 1$. Then

$$n^{k-2} \prod_i (n - n_i)^{n_i - 1} = n^{n-2} \prod_i (1 - x_i)^{nx_i - 1}.$$

Now

$$\ln \prod_i (1 - x_i)^{nx_i - 1} = \sum_i f(x_i),$$

where $f(x) = (nx - 1)\ln(1 - x)$ for $0 < x < 1$. But f is concave on $(0, 1)$, so by Jensen's inequality

$$\sum_i f(x_i) \leqslant k f\left(\frac{1}{k}\right) = (n - k)\log\left(1 - \frac{1}{k}\right),$$

which gives (3.5). □

Lemma 3.4. *Assume that each* $n_i = n/k + O(\sqrt{n})$*, as in part* (c) *of Theorem 3.1. Then*

$$t(\mathbf{n}) = \Theta(1) \cdot n^{n-2} \left(\frac{k-1}{k}\right)^{n-k}.$$

Proof. Write $n_i = (1 + \epsilon_i)n/k$, and $\sum_i$ for $\sum_{i=1}^k$. Then, by Lemma 3.2,

$$\begin{aligned}
\ln t(\mathbf{n}) &= (k-2)\ln n + \sum_i (n_i - 1)\left(\ln n + \ln\left(1 - \frac{1 + \epsilon_i}{k}\right)\right) \\
&= (n-2)\ln n + \sum_i (n_i - 1)\left(\ln\left(1 - \frac{1}{k}\right) + \ln\left(1 - \frac{\epsilon_i}{k-1}\right)\right) \\
&= (n-2)\ln n + (n-k)\ln\left(1 - \frac{1}{k}\right) + \sum_i (n_i - 1)\ln\left(1 - \frac{\epsilon_i}{k-1}\right).
\end{aligned}$$

But, since $\sum_i \epsilon_i = 0$ and each $\sum_i \epsilon_i^2 = O(1/n)$,

$$\begin{aligned}
\sum_i n_i \ln\left(1 - \frac{\epsilon_i}{k-1}\right) &= \frac{n}{k} \sum_i \left((1 + \epsilon_i)\left(-\frac{\epsilon_i}{k-1} + O(\epsilon_i^2)\right)\right) + O(1) \\
&= \frac{n}{k} \cdot O\left(\sum_i \epsilon_i^2\right) + O(1) = O(1),
\end{aligned}$$

and the lemma follows. □

Lemma 3.5. *Let* $k \geqslant 2$ *be an integer and let* $0 < \epsilon < 1$*. Then there exists* $\alpha > 0$ *such that the following holds. Let* $\mathbf{n} = (n_1, \ldots, n_k)$ *be a* k*-tuple of non-negative integers with sum* n *such that* $\max_i n_i \leqslant (1 - \epsilon)n$*. Then the complete* k*-partite graph* $K_{\mathbf{n}}$ *is an* α*-edge-expander.*

Proof. Assume without loss of generality that $n_1 \geqslant \cdots \geqslant n_k$. Let $\eta = \epsilon/(k-1)$ and note that $n_2 \geqslant \eta n$. Let S be a set of $s \leqslant n/2$ vertices, with $s_1, \ldots, s_k$ vertices in the k parts. We break it into three cases.

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

(a) Suppose that $s_2 \geqslant \frac{1}{2}n_2$. Then $s - s_2 \leqslant \frac{1}{2}(n - n_2)$, and so

$$|V \setminus (V_2 \cup S)| = (n - n_2) - (s - s_2) \geqslant \frac{1}{2}(n - n_2) \geqslant \frac{1}{4}n.$$

Hence

$$e(S, \bar{S}) \geqslant e(V_2 \cap S, V \setminus (V_2 \cup S)) \geqslant \frac{1}{2}n_2 \cdot \frac{1}{4}n \geqslant \frac{1}{8}\eta n^2 \geqslant \frac{1}{4}\eta sn.$$

(b) Suppose that $s_2 \geqslant \frac{2}{3}s$. Then $n - n_2 \geqslant n_2 \geqslant s_2 \geqslant \frac{2}{3}s$, so, much as before,

$$|V \setminus (V_2 \cup S)| \geqslant (n - n_2) - \frac{1}{3}s \geqslant \frac{1}{2}(n - n_2) \geqslant \frac{1}{4}n.$$

Hence

$$e(S, \bar{S}) \geqslant e(V_2 \cap S, V \setminus (V_2 \cup S)) \geqslant \frac{2}{3}s \cdot \frac{1}{4}n \geqslant \frac{1}{6}sn.$$

(c) Now suppose that neither of the above cases apply, so $s_2 < \frac{1}{2}n_2$ and $s_2 < \frac{2}{3}s$. Then

$$e(S, \bar{S}) \geqslant e(S \setminus V_2, V_2 \setminus S_2) > \frac{1}{3}s \cdot \frac{1}{2}n_2 \geqslant \frac{1}{6}\eta sn.$$

Putting the above together, we see that in each case $e(S, \bar{S}) \geqslant \frac{1}{6}\eta sn$, and so we may take

$$\alpha = \frac{1}{6}\eta = \frac{\epsilon}{6(k-1)}.$$ □

Now let us denote the number of (spanning) forests in the complete k-partite graph $K_{\mathbf{n}}$ (with a fixed partition) by $f(\mathbf{n})$.

Lemma 3.6. *Let $k \geqslant 2$ be an integer and let $0 < \epsilon < 1$, and let the constant $\alpha > 0$ be as in the last lemma. Let $\mathbf{n} = (n_1, \ldots, n_k)$ be a k-tuple of non-negative integers with sum n such that $\max_i n_i \leqslant (1 - \epsilon)n$. Then*

$$t(\mathbf{n}) \leqslant f(\mathbf{n}) \leqslant e^{1/\alpha}\, t(\mathbf{n}).$$

Proof. Since $K_{\mathbf{n}}$ is a α-edge-expanding by our choice of α, it follows from Theorem 2.1 that $t(\mathbf{n})/f(\mathbf{n}) \geqslant e^{-1/\alpha}$, and the lemma follows. □

Lemma 3.7. *Let $n = n_1 + n_2$, where n_1, n_2 are positive integers, and let H be the complete bipartite graph K_{n_1,n_2}. Let $0 < s \leqslant n/2$ and let S be a set of s vertices in H. Then in H*

$$e(S, \bar{S}) \geqslant \frac{1}{2}s(n - s) - \frac{1}{8}(n_1 - n_2)^2.$$

Proof. Let S have x elements in the first part of H. Then $e(S, \bar{S}) \geqslant f(x)$, where

$$f(x) = x(n_2 - s + x) + (n_1 - x)(s - x).$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Treating x as a continuous variable,

$$f'(x) = n_2 - s + 2x - n_1 - s + 2x = 4x - 2s - (n_1 - n_2).$$

Thus f takes its minimum value $f_{\min}$ at

$$x = \frac{1}{2}s + \frac{1}{4}(n_1 - n_2).$$

Then, temporarily setting $y = \frac{1}{4}(n_1 - n_2)$,

$$\begin{aligned} f_{\min} &= \left(\frac{s}{2} + y\right)\left(n_2 - \frac{s}{2} + y\right) + \left(n_1 - \frac{s}{2} - y\right)\left(\frac{s}{2} - y\right) \\ &= \frac{s}{2}\left(n_2 - \frac{s}{2}\right) + n_2 y + y^2 + \frac{s}{2}\left(n_1 - \frac{s}{2}\right) - n_1 y + y^2 \\ &= \frac{1}{2}s(n-s) - (n_1 - n_2)y + 2y^2 \\ &= \frac{1}{2}s(n-s) - \frac{1}{8}(n_1 - n_2)^2, \end{aligned}$$

and the lemma follows. □

Lemma 3.8. *Let $n/3 < s \leqslant n/2$, and for $i = 0, 1$ let $\mathcal{G}_n^i$ be the set of graphs on $[n]$ such that* Frag(G) *has exactly i components with at least s vertices. (Note that* Frag(G) *must have 0 or 1 such components.) Let $G \in \mathcal{G}_n^0$, and let F be the set of bridges e in G such that $G - e \in \mathcal{G}_n^1$. Then $|F| \leqslant n - 2s + 1$.*

Proof. For each bridge $e = uv$ in G, let $C(e,u)$ be the vertex set of the component of $G - e$ containing u, and similarly let $C(e,v)$ be the vertex set of the component of $G - e$ containing v. Thus F is the set of bridges $e = uv$ such that $\min\{|C(e,u)|, |C(e,v)|\}| \geqslant s$.

Assume that $F \neq \emptyset$, and let $e_1 = u_1v_1$ minimize $\min\{|C(e,u)|, |C(e,v)|\}|$ over all $e = uv$ in F. Without loss of generality, assume that $|C(e_1,u_1)| \leqslant |C(e_1,v_1)|$, and denote $C(e_1,u_1)$ by U. Then for each $e = uv$ in F, e is not in the induced subgraph $G[U]$, so either $C(e,u) \supseteq U$ and $C(e,v)$ is disjoint from U, or the other way around. For each $e = uv \in F$ consider the 'end-set' $C(e,u)$ or $C(e,v)$ disjoint from U and let $e_2 = u_2v_2$ minimize the size of this set. Without loss of generality, we may assume the set is $C(e_2,v_2)$, and denote the set by W. Now each $e = uv$ in F has either $C(e,u) \supseteq U$ and $C(e,v) \supseteq W$, or $C(e,u) \supseteq W$ and $C(e,v) \supseteq U$. Thus e separates U and W, so there is a path between U and W with at most $n - (|U| - 1) - (|W| - 1) \leqslant n - 2s + 2$ vertices, which contains each edge in F. Hence $|F| \leqslant n - 2s + 1$ as required. □

Proof of Theorem 3.1, part (c). Let $N_\ell = N_\ell(G)$ be the number of components of size ℓ in a graph G. We shall show that the contribution to $\mathbb{E}[\text{frag}(R_n)]$ from large components is negligibly small. In particular we shall prove the following claim.

Claim. For any $\eta > 0$ there is an ℓ_1 such that, for all n,

$$\sum_{\ell_1 \leqslant \ell \leqslant n/2} \ell\, \mathbb{E}[N_\ell(R_n)] < \eta. \tag{3.6}$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Clearly we may assume that $\eta \leqslant 1$, and we do so from now on. Suppose temporarily that the claim holds. By part (b), for each unlabelled tree T,

$$\mathbb{E}[v(T)\kappa(R_n, T)] \to \frac{k}{k-1}\frac{v(T)}{e^{v(T)}\operatorname{aut}(T)} \quad \text{as } n \to \infty.$$

Summing up contributions from the finite set $\mathcal{U}^1$ of unlabelled trees T with size less than ℓ_1, as $n \to \infty$,

$$\mathbb{E}\left[\sum_{T\in\mathcal{U}^1} v(T)\kappa(R_n, T)\right] \to \frac{k}{k-1}\sum_{T\in\mathcal{U}^1}\frac{v(T)}{e^{v(T)}\operatorname{aut}(T)}.$$

But since the exponential generating function for rooted labelled trees takes value 1 at $1/e$, we have

$$\sum_{T\in\mathcal{U}}\frac{v(T)}{e^{v(T)}\operatorname{aut}(T)} = \sum_{T\in\mathcal{T}}\frac{v(T)}{e^{v(T)}v(T)!} = \sum_{n\geqslant 1}\frac{n^{n-1}}{n!e^n} = 1.$$

Now, using (3.6), it follows easily that

$$\mathbb{E}[\operatorname{frag}(R_n)] \to \frac{k}{k-1} \quad \text{as } n \to \infty.$$

Thus it remains to establish the claim (3.6). To do this we shall upper-bound $\mathbb{E}[N_\ell(R_n)]$ by comparing it to $\mathbb{E}[N_\ell(F_n)]$, where $F_n \in_u \mathcal{F}$ (*i.e.* F_n is sampled uniformly from the forests on $[n]$).

Now

$$\mathbb{E}[N_\ell(F_n)] = \binom{n}{\ell}\frac{|\mathcal{T}_\ell||\mathcal{F}_{n-\ell}|}{|\mathcal{F}_n|}.$$

But $|\mathcal{F}_{n-\ell}| \geqslant |\mathcal{T}_{n-\ell}|$, and $|\mathcal{F}_n| \leqslant e|\mathcal{T}_n| = en^{n-2}$ by (1.1), so

$$\mathbb{E}[N_\ell(F_n)] \geqslant \binom{n}{\ell}\ell^{\ell-2}\frac{(n-\ell)^{n-\ell-2}}{n^{n-2}}\cdot\frac{1}{e}. \tag{3.7}$$

(Similarly, we also have $|\mathcal{F}_{n-\ell}| \leqslant e|\mathcal{T}_{n-\ell}|$ and $|\mathcal{F}_n| \geqslant |\mathcal{T}_n|$, so $\mathbb{E}[N_\ell(F_n)]$ is at most e^2 times the bound in (3.7).)

Let $0 < \eta \leqslant 1$ be fixed. Later we shall choose ϵ (depending on η) with $0 < \epsilon \leqslant 1/4$. We may assume that n is sufficiently large that

$$(1-\epsilon)\frac{n}{k} \leqslant n_i \leqslant (1+\epsilon)\frac{n}{k} \quad \text{for each } i = 1,\ldots,k.$$

Let $s_1,\ldots,s_k$ be integers with $0 \leqslant s_i \leqslant n_i$ for each i, and let $n_i' = n_i - s_i$ for each i. If $k \geqslant 3$ then

$$\max_i n_i' \leqslant (1+\epsilon)\frac{n}{3} \leqslant \frac{5}{12}n \leqslant \frac{5}{6}(n-\ell) \quad \text{for } \ell \leqslant \frac{n}{2}.$$

If $k = 2$ and

$$\ell \leqslant (1-2\epsilon)\frac{n}{2},$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

then

$$n - \ell \geqslant \left(\frac{1}{2} + \epsilon\right) n,$$

and so

$$\max_i n_i' \leqslant \frac{(1+\epsilon)}{2} \frac{n-\ell}{1/2+\epsilon} \leqslant \left(1 - \frac{\epsilon}{2}\right)(n-\ell).$$

We are led to define $\ell_0 = \ell_0(n) = \lfloor n/2 \rfloor$ if $k \geqslant 3$, and $\ell_0 = \lfloor (1-2\epsilon)n/2 \rfloor$ if $k = 2$, and to set $L_n = \{1, \ldots, \ell_0\}$. Then by Lemma 3.5, for a suitable constant $\alpha > 0$, for each n, each $\ell \in L_n$ and each ℓ-set $S \subseteq [n]$, the complete k-partite graph $K_{\mathbf{n}'}$ is an α-edge-expander. Hence by Theorem 2.1, for $R_{\mathbf{n}'}$ sampled uniformly at random from the set of forests in $K_{\mathbf{n}'}$, we have

$$\mathbb{P}(R_{\mathbf{n}'} \text{ is connected}) \geqslant e^{-1/\alpha},$$

that is, $f(\mathbf{n}') \leqslant e^{1/\alpha} t(\mathbf{n}')$.

Now

$$\mathbb{E}[N_\ell(R_n)] = \frac{1}{f(\mathbf{n})} \sum_{s_1 + \cdots + s_k = \ell} \prod_i \binom{n_i}{s_i} t(\mathbf{s})\, f(\mathbf{n} - \mathbf{s})$$

where the sum is over non-negative integers $s_1, \ldots, s_k$ with sum ℓ, and $\mathbf{s} = (s_1, \ldots, s_k)$. But $f(\mathbf{n}) \geqslant t(\mathbf{n})$, and for $\ell \in L_n$ we have just seen that $f(\mathbf{n} - \mathbf{s}) \leqslant e^{1/\alpha} t(\mathbf{n} - \mathbf{s})$. Hence by Lemma 3.3 (the second part, then the first),

$$\begin{aligned}
\mathbb{E}[N_\ell(R_n)] &\leqslant \frac{e^{1/\alpha}}{t(\mathbf{n})} \sum_{s_1 + \cdots + s_k = \ell} \prod_i \binom{n_i}{s_i} t(\mathbf{s})\, t(\mathbf{n} - \mathbf{s}) \\
&\leqslant \frac{e^{1/\alpha}}{t(\mathbf{n})} \sum_{s_1 + \cdots + s_k = \ell} \prod_i \binom{n_i}{s_i} \ell^{\ell-2} (n-\ell)^{n-\ell-2} \left(\frac{k-1}{k}\right)^{n-2k} \\
&\leqslant O(1) \cdot \sum_{s_1 + \cdots + s_k = \ell} \prod_i \binom{n_i}{s_i} \ell^{\ell-2} (n-\ell)^{n-\ell-2} / n^{n-2} \\
&= O(1) \cdot \binom{n}{\ell} \ell^{\ell-2} (n-\ell)^{n-\ell-2} / n^{n-2}.
\end{aligned}$$

Thus by (3.7) there is a constant c_0 such that

$$\mathbb{E}[N_\ell(R_n)] \leqslant c_0 \cdot \mathbb{E}[N_\ell(F_n)]$$

for all n and all $\ell \in L_n$. But there exists a constant ℓ_1 such that

$$\mathbb{E}\left[\sum_{\ell_1 \leqslant \ell \leqslant n/2} \ell N_\ell(F_n)\right] \leqslant \frac{\eta}{2} c_0$$

(see for example [12, p. 9]), and so

$$\mathbb{E}\left[\sum_{\ell \geqslant \ell_1, \ell \in L_n} \ell N_\ell(R_n)\right] \leqslant \frac{\eta}{2}.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

To complete the proof of the claim (3.6), it remains to consider the case $k = 2$, and to show that, with a suitable choice of ϵ,

$$\mathbb{E}\Big[\sum_{\ell_0(n)<\ell\leqslant n/2} \ell N_\ell(R_n)\Big] \leqslant \frac{\eta}{2} \tag{3.8}$$

(recall that $\ell_0 = \lfloor (1-2\epsilon)n/2 \rfloor$ when $k = 2$). Let $\mathcal{G}_n$ be the set of forests in $K_{\mathbf{n}}$; much as in Lemma 3.8, for $i = 0, 1$ let $\mathcal{G}_n^i$ be the set of graphs in $\mathcal{G}_n$ such that Frag(G) has i components with at least $\ell_0 + 1$ vertices. Form a bipartite graph B with parts $\mathcal{G}_n^0$ and $\mathcal{G}_n^1$, where node $G \in \mathcal{G}_n^0$ is adjacent to node $G' \in \mathcal{G}_n^1$ if G' can be obtained from G by deleting an edge. Then each node in $\mathcal{G}_n^0$ has degree in B at most $n - 2\ell_0 + 1 \leqslant 2\epsilon n + 1$ by Lemma 3.8. Suppose that $\epsilon \leqslant 1/5$. We claim that, once n is sufficiently large that $|n_1 - n_2| \leqslant \epsilon n$, each node G in $\mathcal{G}_n^1$ has degree in B at least

$$\frac{1}{8}n^2(1-5\epsilon). \tag{3.9}$$

To see this, let H be the component of Frag(G) of size at least $\ell_0 + 1$, and let H^+ be the other component of G of size at least that of H (the 'giant' component). Let X be the set of vertices of G in Frag(G) but not in H, and let $x = |X|$, with x_1 in the first part and $x_2 = x - x_1$ in the second part. Then $0 \leqslant x \leqslant n - 2(\ell_0 + 1) < 2\epsilon n$. Let $n_1' = n_1 - x_1$, $n_2' = n_2 - x_2$ and $n' = n - x = n_1' + n_2'$. Note that

$$\begin{aligned} s(n'-s) &\geqslant (1-2\epsilon)\frac{1}{2}n\Big(n' - (1-2\epsilon)\frac{1}{2}n\Big) \\ &\geqslant (1-2\epsilon)\frac{1}{2}n\Big((1-2\epsilon)n - (1-2\epsilon)\frac{1}{2}n\Big) \\ &= \frac{1}{8}(1-2\epsilon)^2n^2 \end{aligned}$$

since $(1-2\epsilon)\frac{1}{2}n < s \leqslant \frac{1}{2}n'$ and $x \leqslant 2\epsilon n$. Now observe that the degree d in B of node G is the number of edges in $K_{\mathbf{n}}$ between H and H^+. Hence, by Lemma 3.7 with the vertices in X deleted and with $S = V(H)$,

$$\begin{aligned} d &\geqslant \frac{1}{2}s(n'-s) - \frac{1}{8}(n_1' - n_2')^2 \\ &\geqslant \frac{1}{8}(1-2\epsilon)^2n^2 - \frac{1}{8}(|n_1 - n_2| + 2\epsilon n)^2, \end{aligned}$$

again using $x \leqslant 2\epsilon n$. Hence, once $|n_1 - n_2| \leqslant \epsilon n$,

$$d \geqslant \frac{1}{8}n^2((1-2\epsilon)^2 - (3\epsilon)^2),$$

and (3.9) follows.

By considering the degrees in B, we have

$$|\mathcal{G}_n^1|\,\frac{1}{8}n^2(1-5\epsilon) \leqslant |\mathcal{G}_n^0|\,(2\epsilon n + 1).$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Hence

$$\frac{|\mathcal{G}_n^1|}{|\mathcal{G}_n|} \leqslant \frac{8(2\epsilon n+1)}{(1-5\epsilon)n^2} \leqslant \frac{\eta}{n}$$

if we set $\epsilon = \eta/25$, for all $n \geqslant 50/\eta$; for then

$$\frac{8(2\epsilon n+1)}{(1-5\epsilon)n^2} \leqslant \frac{8(2\eta n/25)(1+25/(2\eta n))}{(4/5)n^2} \leqslant \frac{4\eta}{5n}\left(1+\frac{25}{2\eta n}\right) \leqslant \frac{\eta}{n}.$$

But now

$$\mathbb{E}\left[\sum_{\ell_0(n)<\ell\leqslant n/2} \ell N_\ell(R_n)\right] \leqslant \frac{(n/2)|\mathcal{G}_n^1|}{|\mathcal{G}_n|} \leqslant \eta/2,$$

which establishes (3.8), and thus the claim (3.6). This completes the proof of part (c) and thus of the whole of Theorem 3.1. □

Random *k*-coloured forests. Following the investigations above, it is natural to consider also the connectivity of random k-coloured forests (with k fixed). We have been picking near-balanced partitions of $[n]$ into k parts deterministically; but it is easy to see that, if we pick a partition of $[n]$ into k parts uniformly at random, then with high probability the partition is near-balanced. We may think of the random forest S_n as coming from picking a k-colouring of $[n]$, and then picking a properly coloured forest on $[n]$ uniformly at random. What if we pick uniformly at random a pair consisting of a forest on $[n]$ and a proper k-colouring of the forest, and then ignore the colours? We shall see quickly that the conclusions in parts (a) and (b) of Theorem 3.1 apply also to such random forests.

Proposition 3.9. *Fix an integer $k \geqslant 2$, and consider a uniformly random pair consisting of a forest $\tilde{R}_n$ on $[n]$ and a proper k-colouring of $\tilde{R}_n$. Let $\lambda = k/(2(k-1))$. Then, as $n \to \infty$:*

(a) *$\kappa(\tilde{R}_n)$ tends to $1 + \mathrm{Po}(\lambda)$ in distribution;*

(b) *the unlabelled trees T appear in $\mathrm{Frag}(\tilde{R}_n)$ asymptotically independently, with distribution $\mathrm{Po}(\lambda(T))$ where*

$$\lambda(T) = \frac{k}{k-1} \cdot (e^{v(T)}\,\mathrm{aut}(T))^{-1}.$$

Proof. We apply results on weighted random graphs from [12]. Let $\tau = (\mu, \nu)$, where $\mu > 0$ is a constant edge-weighting and $\nu > 0$ is a component-weighting. The class of forests (with any weighting) is the most basic example of a (very) well-behaved class; see [12]. Let F_n^τ be a random forest on $[n]$ where

$$\mathbb{P}(F_n^\tau = F) \propto \mu^{e(F)}\nu^{\kappa(F)} \quad \text{for each } F \in \mathcal{F}_n.$$

Take $\mu = 1$ and $\nu = k/(k-1)$ (or $\mu = k-1$ and $\nu = k$). Then $\tilde{R}_n$ and F_n^τ have exactly the same distribution. For, let $F \in \mathcal{F}_n$, and note that the number of k-colourings of F is

$$k^{\kappa(F)}(k-1)^{n-\kappa(F)} = \nu^{\kappa(F)}(k-1)^n.$$

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

Thus

$$\mathbb{P}(\tilde{R}_n = F) = \frac{v^{\kappa(F)}(k-1)^n}{\sum_{F' \in \mathcal{F}_n} v^{\kappa(F')}(k-1)^n} = \mathbb{P}(F_n^\tau = F).$$

Now we see by Corollary 2.3(c) of [12] that $\kappa(\tilde{R}_n)$ converges in distribution to $1 +$ Po$(v/(2\mu))$, and $v/(2\mu) = \lambda$, so part (a) of the proposition holds. Similarly, by Theorems 2.1 and 2.2 of [12], and noting that for each tree T

$$\rho^{v(T)}\mu^{e(T)}v^{\kappa(T)}/\operatorname{aut}(T) = e^{-v(T)}v/\operatorname{aut}(T) = \lambda(T),$$

we see that part (b) of the proposition holds. □

4. Concluding remarks

Let n be even, and recall that $K_{n/2,n/2}$ is an α-edge-expander with $\alpha = 1/2$. By Theorem 2.1, if $\mathcal{A}$ is any relatively bridge-addable class of spanning subgraphs of $K_{n/2,n/2}$, then for $R_n \in_u \mathcal{A}$

$$\mathbb{P}(R_n \text{ is connected}) \geqslant e^{-1/\alpha} = e^{-2}.$$

However, Theorem 3.1 (in the case $k = 2$) suggests that this bound may not be tight, and we could replace e^{-2} by $e^{-1+o(1)}$. Also, by Theorem 2.4,

$$\mathbb{E}(\text{frag}(R_n)) < 2/\alpha = 4,$$

but Theorem 3.1 suggests that we could replace the 4 by $2 + o(1)$. Similar comments hold for general $k \geqslant 2$. These observations suggest an extension of Conjectures 1.1 and 1.2. Perhaps, in the conclusions of Theorems 2.1 and 2.4, we may replace α by $2\alpha + o(1)$?

It has been conjectured that forests are the bridge-addable class of graphs that are the least likely to be connected; see for example Conjecture 1.2 of [5]. Perhaps forests are the relatively bridge-addable class of graphs that are the least likely to be connected? A strong but quite natural conjecture extending that of [5] is as follows.

Conjecture 4.1. *Let G be a finite connected graph, let the set $\mathcal{A}$ of spanning subgraphs of G be bridge-addable relative to G, and let $\mathcal{F}$ be the set of forests in G. For $R \in_u \mathcal{A}$ and $F \in_u \mathcal{F}$,*

$$\mathbb{P}(R \textit{ is connected}) \geqslant \mathbb{P}(F \textit{ is connected})$$

and indeed

$$\kappa(R) \leqslant_s \kappa(F).$$

Acknowledgement

We would like to thank careful referees, whose comments have led to a much improved paper.

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press

References

[1] Addario-Berry, L., McDiarmid, C. and Reed, B. (2012) Connectivity for bridge-addable monotone graph classes. *Combin. Probab. Comput.* **21** 803–815.

[2] Alon, N. and Spencer, J. H. (2008) *The Probabilistic Method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley.

[3] Austin, T. L. (1960) The enumeration of point labelled chromatic graphs and trees. *Canad. J. Math.* **12** 535–545.

[4] Balister, P., Bollobás, B. and Gerke, S. (2008) Connectivity of addable graph classes. *J. Combin. Theory Ser. B* **98** 577–584.

[5] Balister, P., Bollobás, B. and Gerke, S. (2010) Connectivity of random addable graphs. In *Proc. ICDM 2008*, number 13, pp. 127–134.

[6] Bollobás, B. (2001) *Random Graphs*, Cambridge Studies in Advanced Mathematics, Cambridge University Press.

[7] Chapuy, G. and Perarnau, G. (2015) Connectivity in bridge-addable graph classes: The McDiarmid–Steger–Welsh conjecture. arXiv:1504.06344

[8] Flajolet, P. and Sedgewick, R. (2009) *Analytic Combinatorics*, Cambridge University Press.

[9] Kang, M. and Panagiotou, K. (2013) On the connectivity of random graphs from addable classes. *J. Combin. Theory Ser. B* **103** 306–312.

[10] McDiarmid, C. (2009) Random graphs from a minor-closed class. *Combin. Probab. Comput.* **18** 583–599.

[11] McDiarmid, C. (2012) Connectivity for random graphs from a weighted bridge-addable class. *Electron. J. Combin.* **19** P53.

[12] McDiarmid, C. (2013) Random graphs from a weighted minor-closed class. *Electron. J. Combin.* **20** P52.

[13] McDiarmid, C., Steger, A. and Welsh, D. J. A. (2005) Random planar graphs. *J. Combin. Theory Ser. B* **93** 187–205.

[14] McDiarmid, C., Steger, A. and Welsh, D. J. A. (2006) Random graphs from planar and other addable classes. In *Topics in Discrete Mathematics*, Vol. 26 of *Algorithms and Combinatorics*, Springer, pp. 231–246.

[15] McDiarmid, C. and Weller, K. (2014) Relatively bridge-addable classes of graphs (extended abstract). In *LATIN 2014, Theoretical Informatics: Proc. 11th Latin American Symposium* (A. Pardo and A. Viola, eds), Vol. 8392 of *Lecture Notes in Computer Science*, Springer, pp. 391–398.

[16] Moon, J. W. (1970) *Counting Labelled Trees*. From lectures delivered to the Twelfth Biennial Seminar of the Canadian Mathematical Congress (Vancouver) 1969: Canadian Mathematical Congress, Montreal, Quebec.

[17] Norine, S. (2013) Connectivity of addable classes of forests (draft). http://www.math.mcgill.ca/snorin/papers/Forests.pdf

[18] Rényi, A. (1959) Some remarks on the theory of trees. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* **4** 73–85.

[19] Scoins, H. I. (1962) The number of trees with nodes of alternate parity. *Proc. Cambridge Philos. Soc.* **58** 12–16.

https://doi.org/10.1017/S0963548317000128 Published online by Cambridge University Press