Combinatorics, Probability and Computing (2021), **30**, pp. 460–497
doi:10.1017/S0963548320000498

ARTICLE

Subgraph counts for dense random graphs with specified degrees†

Catherine Greenhill[1,*], Mikhail Isaev[2] and Brendan D. McKay[3]

[1]School of Mathematics and Statistics, UNSW Sydney, NSW 2052, Australia, [2]School of Mathematical Sciences, Monash University, VIC 3800, Australia, and [3]Research School of Computer Science, Australian National University, ACT 2601, Australia

*Corresponding author. Email: c.greenhill@unsw.edu.au

(Received 30 January 2018; revised 21 August 2020; accepted 13 August 2019; first published online 5 November 2020)

Abstract

We prove two estimates for the expectation of the exponential of a complex function of a random permutation or subset. Using this theory, we find asymptotic expressions for the expected number of copies and induced copies of a given graph in a uniformly random graph with degree sequence $(d_1, \ldots, d_n)$ as $n \to \infty$. We also determine the expected number of spanning trees in this model. The range of degrees covered includes $d_j = \lambda n + O(n^{1/2+\varepsilon})$ for some λ bounded away from 0 and 1.

2020 MSC Codes: Primary: 05A16; Secondary: 05C80

1. Introduction

For infinitely many natural numbers n, consider vectors

$$\boldsymbol{d}(n) = (d_1(n), \ldots, d_n(n)) \in \{0, \ldots, n-1\}^n.$$

Since we consider asymptotics with respect to $n \to \infty$, we will generally assume that n is sufficiently large and just write $\boldsymbol{d}$ in place of $\boldsymbol{d}(n)$, and similarly for other variables. Everywhere in the paper we assume that

$$\boldsymbol{d}(n) \text{ is a graphical degree sequence;}$$

that is, there exists a graph on the vertex set $\{1, \ldots, n\}$ such that $d_j(n)$ is the degree of vertex j, for $j = 1, \ldots, n$. Let $\mathcal{G}_{\boldsymbol{d}}$ denote the uniform random graph model of simple graphs on the vertex set $\{1, \ldots, n\}$ with degree sequence $\boldsymbol{d}$. By $G \sim \mathcal{G}_{\boldsymbol{d}}$ we mean that G is a random graph from $\mathcal{G}_{\boldsymbol{d}}$.

We study the occurrence of patterns in $G \sim \mathcal{G}_{\boldsymbol{d}}$ such as subgraphs or induced subgraphs isomorphic to a given graph. Using this theory, we find asymptotic expressions for the expected number of some more general structures, namely spanning trees and r-factors. Our aim is to provide formulae that cover sufficiently large and general structures so they could be subsequently used to estimate moments and derive tail bounds for the limiting distribution of the corresponding random variables.

†Research supported by Australian Research Council Discovery Project DP190100977.

For any vector $\boldsymbol{v} = (v_1, \ldots, v_t)$, let

$$\|\boldsymbol{v}\| = \max_{j=1\ldots t} |v_j|$$

denote the infinity norm of $\boldsymbol{v}$. We also use this norm for functions with finite domain. We will use the following parameters that depend only on $\boldsymbol{d}$:

$$\begin{aligned} d &= \frac{1}{n}\sum_{j=1}^{n} d_j, & \lambda &= \frac{d}{n-1}, \\ R &= \frac{1}{n}\sum_{j=1}^{n} (d_j - d)^2, & \delta &= \|(d_1 - d, \ldots, d_n - d)\|. \end{aligned} \tag{1.1}$$

We consider the range of $\boldsymbol{d}$ which satisfy the following assumptions for some constant $\eta \in (0, \frac{1}{2})$ and some constant $\varepsilon > 0$ which is sufficiently small depending on η:

$$\delta \leqslant n^{1/2+\varepsilon} \quad \text{and} \quad \min\{\lambda, 1-\lambda\} \geqslant \frac{1}{6\eta^2 \log n}. \tag{1.2}$$

The set of graphs with degrees $\boldsymbol{d}$ satisfying (1.2) is non-empty for sufficiently large n. This is implied by the enumeration results in [10] and also follows directly from the Erdős–Gallai characterization of graphical degree sequences [2]. The random graph model $\mathcal{G}_{\boldsymbol{d}}$ is thus well-defined.

Let $\mathcal{G}(n, p)$ denote the binomial model of random graph, in which each edge is present independently with probability p. Note that the degree sequence of a random graph from $\mathcal{G}(n, p)$ satisfies $\delta \leqslant n^{1/2+\varepsilon}$ with high probability for any $p = p(n)$. Our results show that counts of small subgraphs in $\mathcal{G}_{\boldsymbol{d}}$ closely match those in $\mathcal{G}(n, \lambda)$, but for larger subgraphs the two models diverge, and correction factors that we will determine are required.

Let G and H be graphs with the same vertex set $\{1, 2, \ldots, n\}$. The *number of copies of H in G* is the number of spanning subgraphs of G that are isomorphic to H. For given $\boldsymbol{d}$, H, the random variable $N_{\boldsymbol{d}}(H)$ is the number of copies of H in G when G is taken at random from $\mathcal{G}_{\boldsymbol{d}}$. The first problem we consider is the expectation $\mathbb{E}\, N_{\boldsymbol{d}}(H)$. If $\boldsymbol{h} = (h_1, \ldots, h_n)$ is the degree sequence of H, then we define

$$m = \frac{1}{2}\sum_{j=1}^{n} h_j, \quad \mu_t = \frac{1}{n}\sum_{j=1}^{n} h_j^t \quad \text{for } t \geqslant 1. \tag{1.3}$$

Note that $m = n\mu_1/2$ is the number of edges of H. Define Aut (H) to be the automorphism group of H, which is the set of permutations of the vertex set $\{1, \ldots, n\}$ that preserve the edge set of H.

Theorem 1.1. *For any constant $\eta \in (0, \frac{1}{2})$ there is some $\varepsilon_1(\eta) > 0$ such that the following holds for every fixed $\varepsilon \in (0, \varepsilon_1(\eta)]$. Let $\boldsymbol{d}$ be a degree sequence which satisfies (1.2). Suppose that H is a graph on vertex set $\{1, \ldots, n\}$ with m edges and degree sequence $\boldsymbol{h}$ such that*

$$m \leqslant n^{1+2\varepsilon}, \quad \frac{\delta^3 \mu_3}{\lambda^3 n^2} \leqslant n^{-1/2+\eta} \quad \textit{and} \quad \|\boldsymbol{h}\| \leqslant n^{1/2+\varepsilon}. \tag{1.4}$$

Then, as $n \to \infty$,

$$\begin{aligned} \mathbb{E}\, N_{\boldsymbol{d}}(H) = \frac{n!}{|\text{Aut}\,(H)|}\, \lambda^m \exp\biggl(& \frac{1-\lambda}{4\lambda}(\mu_1^2 + 2\mu_1 - 2\mu_2) - \frac{R}{2\lambda^2 n}(\mu_1^2 + \mu_1 - \mu_2) \\ & - \frac{1-\lambda^2}{6\lambda^2 n}\mu_3 - \frac{1-\lambda}{\lambda n^2}\sum_{jk\in E(H)} h_j h_k + O(n^{-1/2+\eta})\biggr). \end{aligned}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

The result of Theorem 1.1 simplifies for graphs H with moderate degrees, as shown in the following corollary.

Corollary 1.1. *Suppose the assumptions of Theorem 1.1 hold for some fixed* $\eta \in (0, \frac{1}{2})$ *and* $\varepsilon \in (0, \varepsilon_1(\eta)]$. *Suppose also that* $\mu_3 \leqslant \lambda^2 n^{1/2+\eta}$. *Then, as* $n \to \infty$,

$$\mathbb{E}\, N_{\boldsymbol{d}}(H) = \frac{n!}{|\mathrm{Aut}\,(H)|}\, \lambda^m \exp\biggl(\frac{1-\lambda}{4\lambda}(\mu_1^2 + 2\mu_1 - 2\mu_2) - \frac{R}{2\lambda^2 n}(\mu_1^2 + \mu_1 - \mu_2) + O(n^{-1/2+\eta})\biggr). \tag{1.5}$$

Furthermore, if $\mu_2 \leqslant n^{-3\varepsilon}$ *then*

$$\mathbb{E}\, N_{\boldsymbol{d}}(H) = \frac{n!}{|\mathrm{Aut}\,(H)|}\, \lambda^m (1 + O(n^{-1/2+\eta} + n^{-\varepsilon/2})),$$

which matches the binomial random graph model $\mathcal{G}(n, \lambda)$ *up to the error term.*

McKay [9, Theorem 2.8(a,b)] gave formulae for the number of perfect matchings and cycles of given size in $G \sim \mathcal{G}_{\boldsymbol{d}}$ when $\boldsymbol{d} = (d, \ldots, d)$ is regular (in other words, when $\delta = 0$). Applying (1.5) in these cases (H is a perfect matching, or a cycle of a given length) reproduces these expressions when $\boldsymbol{d}$ is regular, and generalizes them to irregular degree sequences. Kim, Sudakov and Vu [5] obtained a result overlapping the last part of Corollary 1.1 for the case that H has a constant number of edges and $\boldsymbol{d}$ is regular with $d = o(n)$.

For regular subgraphs H, we have the following result.

Corollary 1.2. *For any constant* $\eta \in (0, \frac{1}{2})$ *and every fixed* $\varepsilon \in (0, \varepsilon_1(\eta)]$, *the following holds as* $n \to \infty$, *where* $\varepsilon_1(\eta)$ *is provided by Theorem 1.1. Let* $\boldsymbol{d}$ *be a degree sequence which satisfies* (1.2). *Suppose also that* $h \leqslant n^{2\varepsilon}$ *is a positive integer and* nh *is even.*

(a) *Let* H *be an* h*-regular graph. Then*

$$\mathbb{E}\, N_{\boldsymbol{d}}(H) = \frac{n!}{|\mathrm{Aut}\,(H)|}\, \lambda^m \exp\biggl(-\frac{1-\lambda}{4\lambda}h(h-2) - \frac{Rh}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr).$$

(b) *The expected total number of* h*-regular spanning subgraphs of* $G \sim \mathcal{G}_{\boldsymbol{d}}$ *is*

$$\frac{\sqrt{2}}{(h!)^n}\biggl(\frac{2\lambda m}{e}\biggr)^m \exp\biggl(-\frac{h^2-1}{4} - \frac{1-\lambda}{4\lambda}h(h-2) - \frac{Rh}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr).$$

The proofs of Theorem 1.1 and Corollaries 1.1 and 1.2 are given in Section 4.

Our second main result concerns the expected number of (labelled) spanning trees in $\mathcal{G}_{\boldsymbol{d}}$. This extends, and corrects an error in, a result of McKay [9, Theorem 2.8(c)]. McKay considered the regular case only, and gave the first term as $7(1-\lambda)/(2\lambda)$. However, the correct value is $-(1-\lambda)/(2\lambda)$, as below.

Theorem 1.2. *For any constant* $\eta \in (0, \frac{1}{2})$ *there is some* $\varepsilon_2(\eta) > 0$ *such that the following holds for every fixed* $\varepsilon \in (0, \varepsilon_2(\eta)]$. *Let* $\boldsymbol{d}$ *be a degree sequence which satisfies* (1.2). *Then, as* $n \to \infty$, *the expected number of spanning trees in* $G \sim \mathcal{G}_{\boldsymbol{d}}$ *is*

$$n^{n-2}\lambda^{n-1} \exp\biggl(-\frac{1-\lambda}{2\lambda} - \frac{R}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr).$$

The proof of Theorem 1.2 is given in Section 5.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Let G and $H^{[r]}$ be graphs with vertex sets $\{1, \ldots, n\}$ and $\{1, \ldots, r\}$, respectively. The *number of induced copies of* $H^{[r]}$ *in* G is the number of induced subgraphs of G that are isomorphic to $H^{[r]}$. For given $\boldsymbol{d}$, $H^{[r]}$, the random variable $\widetilde{N}_{\boldsymbol{d}}(H^{[r]})$ is the number of induced copies of $H^{[r]}$ in G when G is taken at random from $\mathcal{G}_{\boldsymbol{d}}$. Our third main result estimates the expectation $\widetilde{N}_{\boldsymbol{d}}(H^{[r]})$ when r is not too large. If $\boldsymbol{h}^{[r]} = (h_1, \ldots, h_r)$ is the degree sequence of $H^{[r]}$, then we define

$$\omega_t = \sum_{j=1}^{r} (h_j - \lambda(r-1))^t \quad \text{for } t \geqslant 1. \tag{1.6}$$

Let $m = \frac{1}{2}\sum_{j=1}^{r} h_j$ be the number of edges of the graph $H^{[r]}$. Note that the automorphism group $\operatorname{Aut}(H^{[r]})$ is a subgroup of the group S_r of all permutations of $\{1, \ldots, r\}$.

Theorem 1.3. *For any constant* $\eta \in (0, \frac{1}{2})$ *there is some* $\varepsilon_3(\eta) > 0$ *such that the following holds for every fixed* $\varepsilon \in (0, \varepsilon_3(\eta)]$. *Let* $\boldsymbol{d}$ *be a degree sequence which satisfies* (1.2). *Suppose that* $H^{[r]}$ *is a graph on vertex set* $\{1, \ldots, r\}$ *with* m *edges and degree sequence* $\boldsymbol{h}^{[r]}$ *such that*

$$r \leqslant n^{1/2+\varepsilon} \quad \text{and} \quad \frac{\delta^3}{\lambda^3(1-\lambda)^3 n^3} \sum_{j=1}^{r} |h_j - \lambda(r-1)|^3 \leqslant n^{-1/2+\eta}. \tag{1.7}$$

Then, as $n \to \infty$,

$$\mathbb{E}\, \widetilde{N}_{\boldsymbol{d}}(H^{[r]}) = \frac{r!}{|\operatorname{Aut}(H^{[r]})|} \binom{n}{r} \lambda^m (1-\lambda)^{\binom{r}{2}-m} \exp\left(\Lambda_0 + \Lambda_1 + \Lambda_2 + O(n^{-1/2+\eta})\right),$$

where

$$\begin{aligned}
\Lambda_0 &= -\frac{\omega_2}{2\lambda(1-\lambda)n} + \frac{R\,\omega_2}{2\lambda^2(1-\lambda)^2 n^2}, \\
\Lambda_1 &= \frac{r^2}{2n} + \frac{(1-2\lambda)\omega_1}{2\lambda(1-\lambda)n} - \frac{\omega_1^2}{4\lambda(1-\lambda)n^2} \\
&\quad - \frac{r^2 R}{2\lambda(1-\lambda)n^2} - \frac{r\,\omega_2}{2\lambda(1-\lambda)n^2} - \frac{(1-2\lambda)\omega_3}{6\lambda^2(1-\lambda)^2 n^2} = O(n^{4\varepsilon}(\log n)^2), \\
\Lambda_2 &= -\frac{(1-2\lambda)R\,\omega_1}{2\lambda^2(1-\lambda)^2 n^2} - \frac{r\,\omega_1 \sum_{j=1}^{n}(d_j - d)^3}{2\lambda^2(1-\lambda)^2 n^4} = O(n^{-1/3+\eta/3+4\varepsilon}).
\end{aligned}$$

For induced subgraphs of more moderate order, the terms Λ_1 and Λ_2 fit into the $O(n^{-1/2+\eta})$ error term.

Corollary 1.3. *Suppose the assumptions of Theorem 1.3 hold for some fixed* $\eta \in (0, \frac{1}{2})$ *and* $\varepsilon \in (0, \varepsilon_3(\eta)]$. *Suppose also that*

$$r^2(1 + \delta^2/n) \leqslant \lambda^2(1-\lambda)^2 n^{1/2+\eta}.$$

Then, as $n \to \infty$,

$$\mathbb{E}\, \widetilde{N}_{\boldsymbol{d}}(H^{[r]}) = \frac{r!}{|\operatorname{Aut}(H^{[r]})|} \binom{n}{r} \lambda^m (1-\lambda)^{\binom{r}{2}-m} \exp\left(\Lambda_0 + O(n^{-1/2+\eta})\right). \tag{1.8}$$

Furthermore, if $r \leqslant n^{1/3-\varepsilon}$ *then*

$$\mathbb{E}\, \widetilde{N}_{\boldsymbol{d}}(H^{[r]}) = \frac{r!}{|\operatorname{Aut}(H^{[r]})|} \binom{n}{r} \lambda^m (1-\lambda)^{\binom{r}{2}-m} \left(1 + O(n^{-1/2+\eta} + n^{-\varepsilon/2})\right),$$

which matches the binomial random graph model $\mathcal{G}(n, \lambda)$ *up to the error term.*

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Note that assumption (1.7) is always satisfied if $r \leqslant n^{1/3-\varepsilon}$ and $\eta \geqslant \frac{1}{3}$. The proofs of Theorem 1.3 and Corollary 1.3 are given in Section 6.

Xiao, Yan, Wu and Ren [12] obtained a result overlapping the last part of Corollary 1.3 for the case that H has constant size and regular $\boldsymbol{d} = (d, \ldots, d)$ with $d = o(n)$. The relationship between the two random graph models $\mathcal{G}_{\boldsymbol{d}}$ and $\mathcal{G}(n, \lambda)$ was also studied by Krivelevich, Sudakov, Vu and Wormald, who established concentration near the mean when $r = O(1)$ and $\boldsymbol{d} = (n/2, \ldots, n/2)$; see [6, Corollary 2.11]. The following includes their result as a special case.

Corollary 1.4. *For any constant $\eta \in (0, \frac{1}{2})$ and every fixed $\varepsilon \in (0, \varepsilon_3(\eta)]$, the following holds, where $\varepsilon_3(\eta)$ is provided by Theorem 1.3. Define $\lambda_{\min} = \min\{\lambda, 1-\lambda\}$. Suppose also that*

$$r \leqslant (2-\varepsilon)\frac{\log n}{\log \lambda_{\min}^{-1}}.$$

Then $\mathbb{E}\,\widetilde{N}_{\boldsymbol{d}}(H^{[r]}) \to \infty$ as $n \to \infty$, and

$$\mathbb{P}\left(\left| \frac{\widetilde{N}_{\boldsymbol{d}}(H^{[r]})}{\mathbb{E}\,\widetilde{N}_{\boldsymbol{d}}(H^{[r]})} - 1 \right| \geqslant n^{-\varepsilon/6} + n^{-1/6+\eta/3} \right) = O(n^{-\varepsilon/6} + n^{-1/6+\eta/3}).$$

Since a clique is a subgraph if and only if it is an induced subgraph, we can use either Theorem 1.1 or Theorem 1.3 to estimate the expected number of r-cliques. Taking H to be K_r plus $n-r$ isolated vertices in Theorem 1.1, or $H^{[r]} = K_r$ in Theorem 1.3, we obtain the following corollary.

Theorem 1.4. *For any constant $\eta \in (0, \frac{1}{2})$ there is some $\varepsilon_4(\eta) > 0$ such that the following holds for every fixed $\varepsilon \in (0, \varepsilon_4(\eta)]$. Let $\boldsymbol{d}$ be a degree sequence which satisfies (1.2). Then, as $n \to \infty$, for any positive integer r such that $r \leqslant n^{1/2+\varepsilon}$ and $\delta^3 r^4/(\lambda^3 n^3) \leqslant n^{-1/2+\eta}$, the expected number of r-cliques in $G \in \mathcal{G}_{\boldsymbol{d}}$ is*

$$\binom{n}{r}\lambda^{\binom{r}{2}} \exp\left(-\frac{(1-\lambda)r^2(r-3)}{2\lambda n} + \frac{Rr^3}{2\lambda^2 n^2} - \frac{(1-\lambda)(2+5\lambda)r^4}{12\lambda^2 n^2} + O(n^{-1/2+\eta}) \right).$$

The formula for the number of independent subsets of size r can be obtained from the formula given in Corollary 1.4 by simply swapping the roles of λ and $1-\lambda$.

1.1 Outline of our approach

Our proofs are based on the asymptotic enumeration results of McKay [9]. To illustrate the nature of our task, the proof of Theorem 1.1 relies on a theorem from [9], here quoted as Theorem 4.1, that the probability of a subgraph H appearing in a fixed location in $G \sim \mathcal{G}_{\boldsymbol{d}}$ has the form

$$\mathbb{P}(\boldsymbol{d}, H) = \lambda^m e^{F(\boldsymbol{d},H) + o(1)} \tag{1.9}$$

for a certain function F. In order to find the expectation of the number of all appearances of isomorphs of H as subgraphs, we need to sum $\mathbb{P}(\boldsymbol{d}, H')$ over all $H' \cong H$. Clearly, this is equivalent to finding the expectation of $e^{F(\boldsymbol{d}^\sigma,H)+o(1)}$, where $\boldsymbol{d}^\sigma$ is a uniformly random permutation σ of the entries of $\boldsymbol{d}$.

Since the function $F(\boldsymbol{d}, H)$ in (1.9) is too large to allow useful expansion of the exponential, we must estimate $\mathbb{E}\, e^{F(\boldsymbol{d}^\sigma,H)}$ directly. We do this by applying the theory of exponentials of martingales developed in [4], which we summarize in Section 2. In order to facilitate similar applications in the future, in Section 2.2 we prove some general theorems about the expectations of the exponentials of functions of random permutations. Theorem 1.1 and its corollaries are proved in Section 4.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

The proof of Theorem 1.3 is given in Section 6. It follows by a similar argument starting from [9, Theorem 2.4], which is quoted here as Theorem 6.1.

The first k entries in a random permutation form a random k-subset, so the same theorems can be used to estimate the expectations of the exponentials of functions of random subsets, and thereby functions of hypergeometric and multinomial distributions. We use this theory to prove Theorem 1.2 in Section 5, as multinomial distributions appear naturally for counts of trees with given degrees.

2. Expectations of exponentials

First, in Section 2.1 we review some notation and results from [4]. Then, in Sections 2.2 and 2.3, we prove some auxiliary results which will help us to apply the machinery from [4] in the discrete setting.

In this paper we will only apply the machinery of this section to real-valued martingales. However, the complex-valued discrete setting is also covered in this section, in order to provide bounds which may be useful for future applications. In particular, such bounds can be useful for determining asymptotic distributions by analysis of the corresponding characteristic functions (Fourier inversion).

Given a complex random variable Z, two types of squared variation are commonly defined. The variance is

$$\mathrm{Var}\, Z = \mathbb{E}\, |Z - \mathbb{E}\, Z|^2 = \mathbb{E}\, |Z|^2 - |\mathbb{E}\, Z|^2 = \mathrm{Var}\, \Re Z + \mathrm{Var}\, \Im Z,$$

while the pseudovariance is

$$\mathbb{V} Z = \mathbb{E}\ (Z - \mathbb{E}\, Z)^2 = \mathbb{E}\, Z^2 - (\mathbb{E}\, Z)^2 = \mathrm{Var}\, \Re Z - \mathrm{Var}\, \Im Z + 2i\, \mathrm{Cov}\, (\Re Z, \Im Z).$$

We will need both. Of course, they are equal for real random variables.

2.1 Complex martingales

Let $P = (\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A sequence $\boldsymbol{\mathcal{F}} = \mathcal{F}_0, \ldots, \mathcal{F}_n$ of σ-subfields of $\mathcal{F}$ is a *filter* if $\mathcal{F}_0 \subseteq \cdots \subseteq \mathcal{F}_n$. A sequence $Z_0, \ldots, Z_n$ of random variables on $P = (\Omega, \mathcal{F}, \mathbb{P})$ is a *martingale with respect to* $\boldsymbol{\mathcal{F}}$ if

(i) Z_j is $\mathcal{F}_j$-measurable and has finite expectation, for $j = 0, \ldots, n$,
(ii) $\mathbb{E}\, [Z_j \mid \mathcal{F}_{j-1}] = Z_{j-1}$ for $j = 1, \ldots, n$.

Observe that $Z_j = \mathbb{E}\, [Z_n \mid \mathcal{F}_j]$ a.s. for each $j = 0, \ldots, n$.

Let Z be a random variable on P. We use the following notation for statistics conditional on $\mathcal{F}_j$, for $j = 0, \ldots, n$:

$$\begin{aligned} \mathbb{E}_j\, Z &= \mathbb{E}\, [Z \mid \mathcal{F}_j], \\ \mathbb{V}_j\, Z &= \mathbb{E}\, [(Z - \mathbb{E}_j\, Z)^2 \mid \mathcal{F}_j] = \mathbb{E}_j\, Z^2 - (\mathbb{E}_j\, Z)^2, \\ \mathrm{diam}_j\, Z &= \mathrm{diam}\, [Z \mid \mathcal{F}_j]. \end{aligned}$$

Here the *conditional diameter* of Z with respect to σ-subfield $\mathcal{F}'$ of $\mathcal{F}$ is defined as

$$\mathrm{diam}\, [Z \mid \mathcal{F}'] = \sup_{\theta \in (-\pi, \pi]} \left[\mathrm{ess\,sup}\, [\Re(e^{-i\theta} Z) \mid \mathcal{F}'] + \mathrm{ess\,sup}\, [-\Re(e^{-i\theta} Z) \mid \mathcal{F}'] \right], \tag{2.1}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

where the conditional essential supremum of a real random variable X with $|X| \leqslant c$ a.s. can be defined (see [1]) by

$$\text{ess sup}\,[X \mid \mathcal{F}'] = -c + \lim_{r\to\infty} (\mathbb{E}\,[(X+c)^r \mid \mathcal{F}'])^{1/r}.$$

When Z is real, we can restrict (2.1) to $\theta = 0$ and then diam $[Z \mid \mathcal{F}']$ is the same as the conditional range defined by McDiarmid [7]. In the trivial case $\mathcal{F}' = \{\emptyset, \Omega\}$, the (unconditional) diameter can be alternatively defined by

$$\text{diam}\, Z = \text{diam}\,[Z \mid \mathcal{F}'] = \text{ess sup}\,|Z - Z'|, \quad \text{where } Z' \text{ is an independent copy of } Z. \tag{2.2}$$

For more information about conditional essential supremum and conditional diameter, see *e.g.* [1] and [4, Section 2.1]. We will use the fact that the diameter and conditional diameter are seminorms and so, in particular, they are subadditive.

The following first-order and second-order estimates were proved in [4, Theorem 2.7 and Theorem 2.9], and are stated below for convenience.

Theorem 2.1. *Let $Z = Z_0, Z_1, \ldots, Z_n$ be an a.s. bounded complex-valued martingale with respect to a filter $\mathcal{F}_0, \ldots, \mathcal{F}_n$. For $j = 1, \ldots, n$, define*

$$R_j = \text{diam}_{j-1}\, Z_j, \qquad Q_j = \max\{\text{diam}_{j-1}\, \mathbb{E}_j\,(Z_n - Z_j)^2, \text{diam}_{j-1}\, \mathbb{E}_j\,(\Re Z_n - \Re Z_j)^2\}.$$

Then the following estimates hold.

(a) $\mathbb{E}_0\, e^{Z_n} = e^{Z_0}(1 + K(Z))$, *where $K(Z)$ is an $\mathcal{F}_0$-measurable random variable with*

$$|K(Z)| \leqslant \text{ess sup}\left[e^{\frac{1}{8}\sum_{j=1}^n R_j^2} \mid \mathcal{F}_0\right] - 1 \ \ a.s.$$

(b) $\mathbb{E}_0\, e^{Z_n} = e^{Z_0 + \frac{1}{2}\mathbb{V}_0\, Z_n}(1 + L(Z)e^{\frac{1}{2}\mathbb{V}_0\,[\Im Z_n]})$, *where $L(Z)$ is an $\mathcal{F}_0$-measurable random variable with*

$$|L(Z)| \leqslant \text{ess sup}\left[\exp\left(\sum_{j=1}^n \left(\frac{1}{6}R_j^3 + \frac{1}{6}R_jQ_j + \frac{5}{8}R_j^4 + \frac{5}{32}Q_j^2\right)\right) \,\middle|\, \mathcal{F}_0\right] - 1 \ \ a.s.$$

The following lemma, proved in [4, Lemma 2.8], is useful for bounding the quantities Q_j when applying Theorem 2.1(b).

Lemma 2.1. *Under the conditions of Theorem 2.1, we have*

$$\mathbb{E}_j\,(Z_n - Z_j)^2 = \sum_{k=j+1}^n \mathbb{E}_j\,(Z_k - Z_{k-1})^2$$

for $0 \leqslant j \leqslant n$.

An important example of a martingale is made by the Doob martingale process. Suppose $X = (X_1, \ldots, X_n)$ is a random vector on P and $f(X)$ is a complex random variable of bounded expectation. Consider the filter $\mathcal{F}_0, \ldots, \mathcal{F}_n$ defined by $\mathcal{F}_j = \sigma(X_1, \ldots, X_j)$, where $\sigma(X_1, \ldots, X_j)$ denotes the σ-field generated by the random variables $X_1, \ldots, X_j$. In particular, $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathbb{E}_0$ is the ordinary expectation. Then we have the martingale

$$Z_j = \mathbb{E}\,[f(X_1, \ldots, X_n) \mid \mathcal{F}_j], \quad j = 0, \ldots, n.$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

It was shown in [4, Lemma 3.1] that for this case the conditional diameter satisfies the following property:

$$\begin{aligned}&\mathrm{diam}_j f(X) \text{ has the same distribution as } \delta_j(X_1, \ldots, X_j), \text{ where}\\ &\delta_j(x_1, \ldots, x_j) = \mathrm{diam}\,[f(x_1, x_2, \ldots, x_j, X_{j+1}, \ldots, X_n)].\end{aligned} \tag{2.3}$$

Here the variables $X_j, \ldots, X_n$ are random and $x_1, \ldots, x_j$ are fixed.

2.2 Random permutations

Let S_n denote the set of permutations of $\{1, \ldots, n\}$. We will write a permutation as a vector: if $\omega \in S_n$ maps j to ω_j for $j = 1, \ldots, n$, then we write $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$. For any $\omega, \sigma \in S_n$, define

$$\omega \circ \sigma = (\omega_{\sigma_1}, \ldots, \omega_{\sigma_n}).$$

That is, σ acts on ω on the right by permuting the *positions* of ω, not the values.

Now suppose $X = (X_1, \ldots, X_n)$ is a uniformly random element of S_n. Although the random variables $X_1, \ldots, X_n$ are dependent, the Doob martingale process is still applicable: for a given permutation $\omega = (\omega_1, \ldots \omega_n) \in S_n$ and the function $f: S_n \to \mathbb{C}$, define

$$Z_k(\omega) = \mathbb{E}\,[f(X) \mid X_j = \omega_j,\ 1 \leqslant j \leqslant k]. \tag{2.4}$$

The sequence $Z_0(X), Z_1(X), \ldots, Z_n(X)$ is a martingale with respect to the filter $\mathcal{F}_0, \ldots, \mathcal{F}_n$, where for each k, the σ-field $\mathcal{F}_k$ is generated by the sets

$$\Omega_{k,\sigma} = \{\omega \in S_n \mid \omega_j = \sigma_j,\ 1 \leqslant j \leqslant k\}$$

for all k-tuples $(\sigma_1, \ldots, \sigma_k)$ with distinct components. From now on we simply write Z_k instead of $Z_k(X)$, for $k = 0, \ldots, n$.

Since $Z_n = Z_{n-1}$ and $\mathcal{F}_n = \mathcal{F}_{n-1}$, we will find it convenient to stop the martingale at Z_{n-1}. In the following we will use the notations of Section 2.1 for statistics conditional on $\mathcal{F}_k$.

Given a function $f: S_n \to \mathbb{C}$, we use the infinity norm

$$\|f\| = \max_{\omega \in S_n} |f(\omega)|.$$

For any $j, a \in \{1, \ldots, n\}$, and any $\omega \in S_n$, define

$$D^{(j\,a)} f(\omega) = f(\omega) - f(\omega \circ (j\,a)).$$

Here $(j\,a) \in S_n$ is the transposition which exchanges j and a. Now, let

$$\alpha_j[f, S_n] = \frac{1}{n-j} \sum_{a=j+1}^{n} \|D^{(j\,a)} f\|, \qquad 1 \leqslant j \leqslant n-1$$

$$\Delta_{jk}[f, S_n] = \frac{1}{(n-j)(n-k)} \sum_{a=j+1}^{n} \sum_{b=k+1}^{n} \|D^{(k\,b)} D^{(j\,a)} f\|, \quad 1 \leqslant j \neq k \leqslant n-1.$$

Note that the parameters α_j and Δ_{jk} satisfy the triangle inequality:

$$\begin{aligned}\alpha_j[f + f', S_n] &\leqslant \alpha_j[f, S_n] + \alpha_j[f', S_n],\\ \Delta_{jk}[f + f', S_n] &\leqslant \Delta_{jk}[f, S_n] + \Delta_{jk}[f', S_n].\end{aligned} \tag{2.5}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

The following lemma provides bounds on the quantities that arise in Theorem 2.1.

Lemma 2.2. *Let* $X = (X_1, \ldots, X_n)$ *be a uniformly random element of* S_n. *Let* $f\colon S_n \to \mathbb{C}$ *and let* $(Z_0, Z_1, \ldots, Z_{n-1})$ *be the Doob martingale sequence given by* (2.4). *Write* $\alpha_k = \alpha_k[f, S_n]$ *and* $\Delta_{jk} = \Delta_{jk}[f, S_n]$. *Then*

$$\operatorname{diam}_{j-1} Z_j \leqslant \alpha_j, \qquad 1 \leqslant j \leqslant n-1, \tag{2.6}$$

$$\operatorname{diam}_{j-1} \mathbb{E}_j\, (Z_k - Z_{k-1})^2 \leqslant 2\alpha_k \Delta_{jk}, \quad 1 \leqslant j < k \leqslant n-1. \tag{2.7}$$

Proof. First, observe that Z_j can be represented by a function of j arguments:

$$Z_j(\omega) = f_j(\omega_1, \ldots, \omega_j), \quad \omega \in S_n.$$

Recalling (2.2) and (2.3), we have

$$\operatorname{diam}_{j-1} Z_j = \max |f_j(\sigma_1, \ldots, \sigma_j) - f_j(\sigma_1, \ldots, \sigma_{j-1}, \sigma_j')|,$$

where the maximum is taken over all j-tuples $(\sigma_1, \ldots, \sigma_j)$ with distinct components, and $\sigma_j' \neq \sigma_1, \ldots, \sigma_{j-1}$. By definition of f_j and Z_j, we have

$$\begin{aligned}
&|f_j(\sigma_1, \ldots, \sigma_j) - f_j(\sigma_1, \ldots, \sigma_{j-1}, \sigma_j')| \\
&\quad = |\mathbb{E}\,[f(X) \mid X_1 = \sigma_1, \ldots, X_j = \sigma_j] - \mathbb{E}\,[f(X) \mid X_1 = \sigma_1, \ldots, X_{j-1} = \sigma_{j-1},\ X_j = \sigma_j']| \\
&\quad = \left| \frac{1}{n-j} \sum_{a=j+1}^{n} \mathbb{E}\,[D^{(j\,a)} f(X) \mid X_1 = \sigma_1, \ldots, X_j = \sigma_j,\ X_a = \sigma_j'] \right|,
\end{aligned}$$

since σ_j' must occupy some position $a \in \{j+1, \ldots, n\}$ in σ, and by symmetry each possibility is equally likely. Therefore

$$|f_j(\sigma_1, \ldots, \sigma_j) - f_j(\sigma_1, \ldots, \sigma_{j-1}, \sigma_j')| \leqslant \frac{1}{n-j} \sum_{a=j+1}^{n} \|D^{(j\,a)} f\| = \alpha_j, \tag{2.8}$$

which implies the bound (2.6) for $\operatorname{diam}_{j-1} Z_j$.

Now we proceed to the bound for $\operatorname{diam}_{j-1} \mathbb{E}_j\, (Z_k - Z_{k-1})^2$. Define $\widetilde{f}\colon S_n \to \mathbb{C}$ by

$$\widetilde{f}(\omega) = (Z_k(\omega) - Z_{k-1}(\omega))^2 = (f_k(\omega_1, \ldots, \omega_k) - f_{k-1}(\omega_1, \ldots, \omega_{k-1}))^2.$$

Since $D^{(j\,a)} \widetilde{f}(\omega)$ is the difference of two squares, we have

$$\begin{aligned}
D^{(j\,a)} \widetilde{f}(\omega) &= \widetilde{f}(\omega) - \widetilde{f}(\omega \circ (j\,a)) \\
&= (Z_k(\omega) - Z_{k-1}(\omega) + Z_k(\omega \circ (j\,a)) - Z_{k-1}(\omega \circ (j\,a))) \\
&\quad \times (Z_k(\omega) - Z_{k-1}(\omega) - Z_k(\omega \circ (j\,a)) + Z_{k-1}(\omega \circ (j\,a))).
\end{aligned}$$

Using (2.6) applied to f, we have $|Z_k(\omega) - Z_{k-1}(\omega)| \leqslant \operatorname{diam}_{k-1} Z_k$ and hence

$$|Z_k(\omega) - Z_{k-1}(\omega) + Z_k(\omega \circ (j\,a)) - Z_{k-1}(\omega \circ (j\,a))| \leqslant 2\ \operatorname{diam}_{k-1} Z_k \leqslant 2\alpha_k.$$

Therefore, applying (2.6) to $\widetilde{f}$ gives

$$\begin{aligned}
&\operatorname{diam}_{j-1} \mathbb{E}_k\, (Z_k - Z_{k-1})^2 \\
&\leqslant \alpha_j[\widetilde{f}, S_n] = \frac{1}{n-j} \sum_{a=j+1}^{n} \|D^{(j\,a)} \widetilde{f}\,\| \\
&\leqslant \frac{2\alpha_k}{n-j} \sum_{a=j+1}^{n} \max_{\omega \in S_n} |Z_k(\omega) - Z_{k-1}(\omega) - Z_k(\omega \circ (j\,a)) + Z_{k-1}(\omega \circ (j\,a))|.
\end{aligned} \tag{2.9}$$

In the remainder of the proof we work towards an upper bound on the summand.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

For any $c \in \{1, \ldots, n\}$ and any permutation $(k\, b)$, with $1 \leqslant k \leqslant b \leqslant n$ (either a transposition or the identity permutation), write $c^{(k\, b)}$ for the image of c under the action of $(k\, b)$. Given $k \in \{1, \ldots, n\}$, define the set

$$I_k = \{(b, c) \mid k \leqslant b \neq c \leqslant n\}$$

of distinct ordered pairs with both entries at least k.

Now we consider two cases.

Case 1. First, suppose that $a \in \{k+1, \ldots, n\}$. To begin, observe that

$$Z_{k-1}(\omega) = \frac{1}{(n-k)(n-k+1)} \times \sum_{(b,c)\in I_k} \mathbb{E}\left[f(X) \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1}, X_b = \omega_k,\ X_c = \omega_a\right] \tag{2.10}$$

using arguments similar to those which led to (2.8). Next, let $\widetilde{X} = X \circ (k\, b)$, which is also a uniformly random element of S_n, and write

$$\mathbb{E}\left[f(X \circ (k\, b)) \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1},\ X_b = \omega_k,\ X_c = \omega_a\right] = \mathbb{E}\left[f(\widetilde{X}) \mid \widetilde{X}_1 = \omega_1, \ldots, \widetilde{X}_k = \omega_k,\ \widetilde{X}_{c^{(k\,b)}} = \omega_a\right].$$

Note that $c' = c^{(k\, b)}$ ranges over $\{k+1, \ldots, n\}$ as c ranges over $\{k, \ldots, n\} \setminus \{b\}$. Therefore

$$\begin{aligned} &\frac{1}{(n-k)(n-k+1)} \sum_{(b,c)\in I_k} \mathbb{E}\left[f(X \circ (k\, b)) \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1},\ X_b = \omega_k,\ X_c = \omega_a\right] \\ &= \frac{1}{(n-k)(n-k+1)} \sum_{b=k}^{n} \sum_{c'=k+1}^{n} \mathbb{E}\left[f(\widetilde{X}) \mid \widetilde{X}_1 = \omega_1, \ldots, \widetilde{X}_k = \omega_k,\ \widetilde{X}_{c'} = \omega_a\right] \qquad (2.11) \\ &= \frac{1}{n-k} \sum_{c'=k+1}^{n} \mathbb{E}\left[f(\widetilde{X}) \mid \widetilde{X}_1 = \omega_1, \ldots, \widetilde{X}_k = \omega_k,\ \widetilde{X}_{c'} = \omega_a\right] \\ &= \mathbb{E}\left[f(\widetilde{X}) \mid \widetilde{X}_1 = \omega_1, \ldots, \widetilde{X}_k = \omega_k\right] = Z_k(\omega), \qquad (2.12) \end{aligned}$$

similarly to (2.8), since the summand in (2.11) is independent of b.

Arguing as above with $\widetilde{X} = X \circ (j\, c)$ gives

$$\begin{aligned} &Z_{k-1}(\omega \circ (j\, a)) \\ &= f_{k-1}(\omega_1, \ldots, \omega_{j-1}, \omega_a, \omega_{j+1}, \ldots, \omega_{k-1}) \\ &= \frac{1}{(n-k)(n-k+1)} \\ &\quad \times \sum_{(b,c)\in I_k} \mathbb{E}\left[f(X \circ (j\, c)) \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1},\ X_b = \omega_k,\ X_c = \omega_a\right]. \qquad (2.13) \end{aligned}$$

Finally, let $\widetilde{X} = X \circ (j\, c) \circ (k\, b)$, which is a uniformly random element of S_n, and recall that $c^{(k\, b)}$ ranges over $\{k+1, \ldots, n\}$ as c runs over $\{k, \ldots, n\} \setminus \{b\}$. Arguing as above gives

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

$$\frac{1}{(n-k)(n-k+1)}$$
$$\times \sum_{(b,c)\in I_k} \mathbb{E}\,[f(X\circ(j\,c)\circ(k\,b)) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k,\ X_c=\omega_a]$$
$$= \frac{1}{(n-k)(n-k+1)}$$
$$\times \sum_{(b,c)\in I_k} \mathbb{E}\,[f(\widetilde{X}) \mid \widetilde{X}_1=\omega_1,\ldots,\widetilde{X}_{j-1}=\omega_{j-1},\ \widetilde{X}_j=\omega_a,\ \widetilde{X}_{j+1}=\omega_{j+1},\ldots,$$
$$\widetilde{X}_{k-1}=\omega_{k-1},\widetilde{X}_k=\omega_k,\ \widetilde{X}_{c^{(k\,b)}}=\omega_j]$$
$$= \frac{1}{n-k+1}\sum_{b=k}^{n} Z_k(\omega\circ(j\,a)) = Z_k(\omega\circ(j\,a)). \tag{2.14}$$

Combining (2.10)–(2.14), we find that when $a\in\{k+1,\ldots,n\}$,

$$|Z_k(\omega)-Z_{k-1}(\omega)-Z_k(\omega\circ(j\,a))+Z_{k-1}(\omega\circ(j\,a))|$$
$$= \frac{1}{(n-k)(n-k+1)}$$
$$\times \sum_{(b,c)\in I_k} \mathbb{E}\,[D^{(k\,b)}D^{(j\,a)}f(X) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k,\ X_c=\omega_a]$$
$$\leqslant \frac{1}{(n-k)^2}\sum_{(b,c)\in I_k} \|D^{(k\,b)}D^{(j\,c)}f\|. \tag{2.15}$$

Case 2. Now suppose that $a\in\{j+1,\ldots,k\}$. Define $z=b$ if $a=k$, and $z=a$ if $a\in\{j+1,\ldots,k-1\}$. Arguing as above, we have

$$Z_k(\omega) = \frac{1}{n-k+1}\sum_{b=k}^{n} \mathbb{E}\,[f(X\circ(k\,b)) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k],$$
$$Z_{k-1}(\omega) = \frac{1}{n-k+1}\sum_{b=k}^{n} \mathbb{E}\,[f(X) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k],$$
$$Z_{k-1}(\omega\circ(j\,a)) = \frac{1}{n-k+1}\sum_{b=k}^{n} \mathbb{E}\,[f(X\circ(j\,z)) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k],$$
$$Z_k(\omega\circ(j\,a)) = \frac{1}{n-k+1}$$
$$\times \sum_{b=k}^{n} \mathbb{E}\,[f(X\circ(j\,z)\circ(k\,b)) \mid X_1=\omega_1,\ldots,X_{k-1}=\omega_{k-1},\ X_b=\omega_k].$$

Combining these, we find that when $a\in\{j+1,\ldots,k\}$,

$$|Z_k(\omega)-Z_{k-1}(\omega)-Z_k(\omega\circ(j\,a))+Z_{k-1}(\omega\circ(j\,a))| \leqslant \frac{1}{n-k}\sum_{b=k}^{n} \|D^{(k\,b)}D^{(j\,z)}f\|. \tag{2.16}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Consolidation. Now we perform the sum over a. From (2.15) and (2.16) we have

$$\begin{aligned}\sum_{a=j+1}^{n} &|Z_k(\omega)-Z_{k-1}(\omega)-Z_k(\omega\circ(j\,a))+Z_{k-1}(\omega\circ(j\,a))|\\ &\leqslant \frac{1}{n-k}\sum_{(b,c)\in I_k}\|D^{(k\,b)}\,D^{(j\,c)}f\|+\frac{1}{n-k}\sum_{b=k}^{n}\|D^{(k\,b)}D^{(j\,b)}f\|\\ &\quad+\frac{1}{n-k}\sum_{a=j+1}^{k-1}\sum_{b=k}^{n}\|D^{(k\,b)}D^{(j\,a)}f\|,\end{aligned} \tag{2.17}$$

using the fact that (2.15) is independent of a in Case 1. Replacing the dummy variable c in the first sum by a, and observing that any term with $k=b$ equals zero, we can rewrite the right-hand side of (2.17) as

$$\frac{1}{n-k}\sum_{a=j+1}^{n}\sum_{b=k}^{n}\|D^{(k\,b)}D^{(j\,a)}f\|=\frac{1}{n-k}\sum_{a=j+1}^{n}\sum_{b=k+1}^{n}\|D^{(k\,b)}D^{(j\,a)}f\|=(n-j)\,\Delta_{jk}.$$

Substituting this into (2.9), we conclude that

$$\operatorname{diam}_{j-1}\mathbb{E}_j\,(Z_k-Z_{k-1})^2\leqslant 2\alpha_k\,\Delta_{jk}$$

as required. □

Combining the bounds proved above with Theorem 2.1 gives the following.

Theorem 2.2. *Let X be a uniformly random element of S_n and let $f\colon S_n\to\mathbb{C}$. Write $\alpha_k=\alpha_k(f,S_n)$ and $\Delta_{jk}=\Delta_{jk}(f,S_n)$. Then we have the following.*

(a) $\mathbb{E}\,e^{f(X)}=e^{\mathbb{E}f(X)}(1+K(f))$, *where* $K(f)\in\mathbb{C}$ *satisfies*

$$|K(f)|\leqslant e^{\frac{1}{8}\sum_{j=1}^{n-1}\alpha_j^2}-1.$$

(b) $\mathbb{E}\,e^{f(X)}=e^{\mathbb{E}f(X)+\frac{1}{2}\mathbb{V}f(X)}(1+L(f)e^{\frac{1}{2}\operatorname{Var}\Im f(X)})$, *where* $\beta_j=\sum_{k=j+1}^{n-1}\alpha_k\Delta_{jk}$ *and* $L(f)\in\mathbb{C}$ *satisfies*

$$|L(f)|\leqslant\exp\Bigg(\sum_{j=1}^{n-1}\bigg(\frac{1}{6}\alpha_j^3+\frac{1}{3}\alpha_j\beta_j+\frac{5}{8}\alpha_j^4+\frac{5}{8}\beta_j^2\bigg)\Bigg)-1.$$

Proof. Let $Z(X)=(Z_0,Z_1,\dots,Z_{n-1})$ be the Doob martingale sequence given by (2.4). By applying Theorem 2.1 to $Z(X)$, it remains to show that

$$R_j\leqslant\alpha_j,\qquad Q_j\leqslant 2\beta_j.$$

The first bound is given by (2.6) and the definition of R_j.

Observe that $D^{(ja)}(\Re f(\omega))=\Re D^{(ja)}f(\omega)$ for any $j,a\in\{1,\dots,n\}$. Therefore

$$\begin{aligned}\alpha_j[\Re f,S_n]&\leqslant\alpha_j[f,S_n], && 1\leqslant j\leqslant n-1,\\ \Delta_{jk}[\Re f,S_n]&\leqslant\Delta_{jk}[f,S_n], && 1\leqslant j\neq k\leqslant n-1.\end{aligned}$$

Using (2.7) twice (for f and $\Re f$), we find that both quantities $\operatorname{diam}_{j-1}\mathbb{E}_j\,(Z_k-Z_{k-1})^2$ and $\operatorname{diam}_{j-1}\mathbb{E}_j\,(\Re Z_k-\Re Z_{k-1})^2$ are bounded above by $2\alpha_k\,\Delta_{jk}$. Since the conditional diameter is subadditive, we can apply Lemma 2.1 to obtain the remaining bound on Q_j. □

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

2.3 Random subsets and other discrete distributions

Using our estimates for random permutations, we can also apply Theorem 2.1 for functions of random subsets of given size, as well as functions of random vectors with standard multidimensional discrete distributions, such as the hypergeometric distribution or the multinomial distribution. We now define analogues of the operator $D^{(j\,a)}$ for these cases.

Subsets. Let $2^{[n]}$ denote the set of all subsets of $\{1, 2, \ldots, n\}$. For a given $f\colon 2^{[n]} \to \mathbb{C}$, and for every $A \in 2^{[n]}$, let

$$D_B^{(j\,a)} f(A) = f(A) - f(A \oplus \{j,\ a\}),$$

where $\oplus$ denotes the symmetric difference. Note that if $|A \cap \{j,\ a\}| = 1$ then $A \oplus \{j, a\}$ has the same size as A. Let $B_{n,m}$ denote the set of m-subsets of $\{1, \ldots, n\}$, and define

$$\alpha_{\max}[f, B_{n,m}] = \max |D_B^{(j\,a)} f(A)|,$$

where the maximum is taken over all $A \in B_{n,m}$ and all $j, a \in \{1, \ldots, n\}$ such that $j \in A$ and $a \notin A$. Similarly, define

$$\Delta_{\max}[f, B_{n,m}] = \max |D_B^{(k\,b)}\, D_B^{(j\,a)} f(A)|,$$

where the maximum is over all distinct $j, k, a, b \in \{1, \ldots, n\}$ and all $A \in B_{n,m}$ such that $j, k \in A$ and $a, b \notin A$. Note that $\alpha_{\max}[f, B_{n,m}]$ and $\Delta_{\max}[f, B_{n,m}]$ depend only on the values of f on the set $B_{n,m}$.

Sequences. For a given function $f\colon \mathbb{Z}^\ell \to \mathbb{C}$, and for every $\boldsymbol{x} = (x_1, \ldots, x_\ell) \in \mathbb{Z}^\ell$, define

$$D_N^{(j\,a)} f(\boldsymbol{x}) = f(\boldsymbol{x}) - f(\boldsymbol{x}'),$$

where $\boldsymbol{x}'$ has all entries equal to those of $\boldsymbol{x}$, except that the jth entry is increased by 1 and the ath entry is decreased by 1. For positive integers ℓ, m, define

$$N_{\ell,m} = \{(x_1, \ldots, x_\ell) \in \{0, 1, 2, \ldots\}^\ell : x_1 + \cdots + x_\ell = m\}.$$

Note that if $\boldsymbol{x} \in N_{\ell,m}$ with $x_a > 0$ then $\boldsymbol{x}'$, defined above, also belongs to $N_{\ell,m}$. (If $\boldsymbol{x}$ has any positive entry then no other entry can equal m.) Define

$$\alpha_{\max}[f, N_{\ell,m}] = \max |D_N^{(j\,a)} f(\boldsymbol{x})|, \tag{2.18}$$

where the maximum is over all $\boldsymbol{x} \in N_{\ell,m}$ and all distinct j, a such that $x_a > 0$. Also, define

$$\Delta_{\max}[f, N_{\ell,m}] = \max |D_N^{(k\,b)}\, D_N^{(j\,a)} f(\boldsymbol{x})|, \tag{2.19}$$

where the maximum is taken over all distinct j, k, a, b, such that $\min\{x_a, x_b\} > 0$ and all $\boldsymbol{x} \in N_{\ell,m}$. Again, observe that $\alpha_{\max}[f, N_{\ell,m}]$ and $\Delta_{\max}[f, N_{\ell,m}]$ depend only on the values of f on $N_{\ell,m}$.

Theorem 2.3. *Consider any one of the following three possibilities.*

(i) *X is a uniformly random element of $B_{n,m}$, where $m \leqslant n/2$.*
(ii) *$X = (X_1, \ldots, X_\ell)$ is an $N_{\ell,m}$-valued random variable with the hypergeometric distribution with parameters $n_1, \ldots, n_\ell \geqslant 0$ such that $n_1 + \cdots + n_\ell = n \geqslant 2m$, that is,*

$$\mathbb{P}(X = (x_1, \ldots, x_\ell)) = \binom{n}{m}^{-1} \prod_{j=1}^{\ell} \binom{n_j}{x_j}, \qquad (x_1, \ldots, x_\ell) \in N_{\ell,m}.$$

(iii) $X = (X_1, \ldots, X_\ell)$ *is an* $N_{\ell,m}$*-valued random variable with the multinomial distribution with parameters* $p_1, \ldots, p_\ell > 0$ *such that* $p_1 + \cdots + p_\ell = 1$*, that is,*

$$\mathbb{P}(X = (x_1, \ldots, x_\ell)) = m! \prod_{j=1}^{\ell} \frac{p_j^{x_j}}{x_j!}, \quad (x_1, \ldots, x_\ell) \in N_{\ell,m}. \tag{2.20}$$

With $\Lambda = B_{n,m}$ *or* $\Lambda = N_{\ell,m}$*, and given a function* $f \colon \Lambda \to \mathbb{C}$*, let* $\alpha_{\max} = \alpha_{\max}[f, \Lambda]$ *and* $\Delta_{\max} = \Delta_{\max}[f, \Lambda]$*. Then we have the following.*

(a) $\mathbb{E}\, e^{f(X)} = e^{\mathbb{E} f(X)}(1 + K(f))$*, where* $K(f) \in \mathbb{C}$ *satisfies* $|K(f)| \leqslant e^{\frac{1}{8} m\, \alpha_{\max}^2} - 1$.

(b) $\mathbb{E}\, e^{f(X)} = e^{\mathbb{E} f(X) + \frac{1}{2} \mathbb{V} f(X)}(1 + L(f)\, e^{\frac{1}{2} \operatorname{Var} \Im f(X)})$*, where* $L(f) \in \mathbb{C}$ *satisfies*

$$|L(f)| \leqslant \exp\left(\frac{1}{2} m\alpha_{\max}^3 + \frac{1}{6} m^2 \alpha_{\max}^2 \Delta_{\max} + 2m\alpha_{\max}^4 + \frac{5}{8} m^3 \alpha_{\max}^2 \Delta_{\max}^2\right) - 1.$$

Proof. First suppose that X has the distribution described in (i), and define $\widetilde{f} \colon S_n \to \mathbb{C}$ by

$$\widetilde{f}(\omega_1, \ldots, \omega_n) = f(\{\omega_1, \ldots, \omega_m\}), \quad \omega \in S_n.$$

Let Y be a uniformly random element of S_n. Observe that $f(X)$ and $\widetilde{f}(Y)$ have the same distribution, and hence

$$\mathbb{E}\, e^{f(X)} = \mathbb{E}\, e^{\widetilde{f}(Y)}, \quad \mathbb{E} f(X) = \mathbb{E} \widetilde{f}(Y),$$

and similarly for $\mathbb{V} f(X)$ and $\operatorname{Var} \Im f(X)$.

Let $\alpha_j = \alpha_j[\widetilde{f}, S_n]$ and $\Delta_{jk} = \Delta_{jk}[\widetilde{f}, S_n]$ denote the parameters used in Lemma 2.2, defined with respect to the function $\widetilde{f}$ and set S_n. We will apply Theorem 2.2 to the function $\widetilde{f}$. Then the bound (a) follows immediately from Theorem 2.2(a), since

$$\alpha_j \leqslant \begin{cases} \alpha_{\max} & \text{for } j = 1, \ldots, m, \\ 0 & \text{for } j = m+1, \ldots, n. \end{cases}$$

Next, note that $\beta_j = 0$ for $j = m+1, \ldots, n$, and $\Delta_{jk} = 0$ if $k > m$. We now estimate Δ_{jk} when $j < k \leqslant m$.

If $b \leqslant m$ or $a \leqslant m$ then $D^{(k\,b)}\, D^{(j\,a)} \widetilde{f} = 0$, since $\widetilde{f}$ depends only on the set of the first m components of the input permutation. Next, observe that if $a = b > m$ then

$$\|D^{(k\,a)}\, D^{(j\,a)} \widetilde{f}\| \leqslant 2\alpha_{\max},$$

while if $a \neq b$ and $a, b > m$ then

$$\|D^{(k\,b)}\, D^{(j\,a)} \widetilde{f}\| \leqslant \Delta_{\max}.$$

Therefore

$$\Delta_{jk} \leqslant \frac{2(n-m)\alpha_{\max} + (n-m)(n-m-1)\Delta_{\max}}{(n-j)(n-k)} \leqslant \frac{2}{n-m}\alpha_{\max} + \Delta_{\max}.$$

Hence using Lemma 2.2 it follows that

$$\beta_j \leqslant (m-j)\alpha_{\max}\left(\frac{2}{n-m}\alpha_{\max} + \Delta_{\max}\right)$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

for $j = 1, \dots, m$. Using these bounds and the fact that $2m \leqslant n$, we find that

$$\begin{aligned}
\sum_{j=1}^{n-1} &\left(\frac{1}{6}\alpha_j^3 + \frac{1}{3}\alpha_j\beta_j + \frac{5}{8}\alpha_j^4 + \frac{5}{8}\beta_j^2\right) \\
&\leqslant \frac{1}{6}m\alpha_{\max}^3 + \frac{1}{3}\,\alpha_{\max}^2\left(\frac{2}{n-m}\alpha_{\max} + \varDelta_{\max}\right)\sum_{j=1}^{m}(m-j) + \frac{5}{8}m\alpha_{\max}^4 \\
&\quad + \frac{5}{8}\alpha_{\max}^2\left(\frac{2}{n-m}\alpha_{\max} + \varDelta_{\max}\right)^2\sum_{j=1}^{m}(m-j)^2 \\
&\leqslant \frac{1}{2}m\alpha_{\max}^3 + \frac{1}{6}m^2\alpha_{\max}^2\varDelta_{\max} + +\frac{5}{8}m\alpha_{\max}^4 + \frac{5}{24}m^3\,\alpha_{\max}^2\left(\frac{2}{n-m}\alpha_{\max} + \varDelta_{\max}\right)^2 \\
&\leqslant \frac{1}{2}m\alpha_{\max}^3 + \frac{1}{6}m^2\alpha_{\max}^2\varDelta_{\max} + 2m\alpha_{\max}^4 + \frac{5}{8}m^3\alpha_{\max}^2\varDelta_{\max}^2.
\end{aligned}$$

We used the inequality $(a+b)^2 \leqslant \frac{3}{2}a^2 + 3b^2$ in the final line. Applying Theorem 2.2(b) completes the proof for when X has the distribution described in (i).

Next, suppose that X has the hypergeometric distribution described in (ii). Take disjoint sets $A_1, \dots, A_\ell$ with $|A_j| = n_j$ for each j. If we choose a random subset $B \subseteq A_1 \cup \dots \cup A_\ell$ with size m, then $X = (|B \cap A_1|, \dots, |B \cap A_\ell|)$ has the required distribution. Now we can consider $f(X)$ as a function of B and apply case (i).

Finally, suppose that X has the multinomial distribution described in (iii). Apply case (ii) with $n_j = \lceil p_j t \rceil$ and let $t \to \infty$. □

We remark that by giving tighter bounds on factors of the form $m/(n-m)$ in the above proof, the constants in the error term $|L|$ for (b) can be improved. We do not pursue this here.

3. Moment calculations

Now we prove a lemma that will be used repeatedly in the following sections.

Lemma 3.1. *Suppose $u, v \colon \{1, 2, \dots, n\} \to \mathbb{R}$. Define the function $\Psi = \Psi_{u,v} \colon S_n \to \mathbb{R}$ by*

$$\Psi(\sigma) = \sum_{j=1}^{n} u(j)v(\sigma_j) \quad \textit{for } \sigma \in S_n.$$

Let $X = (X_1, \dots, X_n)$ denote a random permutation uniformly chosen from S_n. Define

$$\bar{u} = \frac{1}{n}\sum_{j=1}^{n} u(j), \qquad \bar{v} = \frac{1}{n}\sum_{j=1}^{n} v(j).$$

Finally, let

$$\alpha = \left(\max_j u(j) - \min_j u(j)\right)\left(\max_j v(j) - \min_j v(j)\right).$$

(i) *Then*

$$\mathbb{E}\,\Psi(X) = n\,\bar{u}\,\bar{v} \quad \textit{and} \quad \mathbb{E}\,e^{\Psi(X)} = e^{\mathbb{E}\,\Psi(X) + \frac{1}{2}\operatorname{Var}\Psi(X) + L}$$

for some $L \in \mathbb{R}$ with $|L| \leqslant \frac{3}{2}n\alpha^3 + 11\,n\alpha^4$.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

(ii) *Now let $u', v' \in \{1, 2, \dots, n\} \to \mathbb{R}$ and let $\bar{u}', \bar{v}'$ be the average value of u', v', respectively. Let $\Psi' = \Psi_{u',v'}$. Then*

$$\mathrm{Cov}\,(\Psi(X), \Psi'(X)) = \frac{1}{n-1}\sum_{j=1}^{n}(u(j)-\bar{u})(u'(j)-\bar{u}')\sum_{k=1}^{n}(v(k)-\bar{v})(v'(k)-\bar{v}').$$

In particular,

$$\mathrm{Var}\,\Psi(X) = \frac{1}{n-1}\sum_{j=1}^{n}(u(j)-\bar{u})^2\sum_{k=1}^{n}(v(k)-\bar{v})^2.$$

(iii) *For distinct $j, k \in \{1, \dots, n\}$, define $E_{jk}\colon S_n \to \mathbb{R}$ by*

$$E_{jk}(\sigma) = (u(j)+v(\sigma_j))(u(k)+v(\sigma_k)).$$

Then

$$\mathbb{E}\,E_{jk}(X) = (u(j)+\bar{v})(u(k)+\bar{v}) - \frac{1}{n(n-1)}\sum_{i=1}^{n}(v(i)-\bar{v})^2.$$

(iv) *For $j, k, \ell, m \in \{1, \dots, n\}$ with j, k distinct and ℓ, m distinct,*

$$\mathrm{Cov}\,(E_{jk}(X), E_{\ell m}(X)) = \begin{cases} O((\|u\|+\|v\|)^4/n) & \text{if } \{j,k\}\cap\{\ell,m\}=\emptyset,\\ O((\|u\|+\|v\|)^4) & \text{otherwise.}\end{cases}$$

(v) *For distinct $j, k \in \{1, \dots, n\}$,*

$$\begin{aligned}&\mathrm{Cov}\,(E_{jk}(X), \Psi'(X))\\ &\quad= \frac{1}{n}\big((u'(j)-\bar{u}')(u(k)+\bar{v}) + (u'(k)-\bar{u}')(u(j)+\bar{v})\big)\sum_{a=1}^{n}(v(a)-\bar{v})(v'(a)-\bar{v}')\\ &\qquad+ O\bigg(\frac{(\|u\|+\|v\|)^2\|u'\|\|v'\|}{n}\bigg)\\ &\quad= O((\|u\|+\|v\|)^2\|u'\|\|v'\|).\end{aligned}$$

Proof. We calculate that

$$\mathbb{E}\,\Psi(X) = \sum_{j=1}^{n} u(j)\;\mathbb{E}\,v(X_j) = \bar{v}\sum_{j=1}^{n}u(j) = n\bar{u}\bar{v}.$$

Next we apply Theorem 2.1(b) and Lemma 2.2 to the Doob martingale for $\Psi\colon S_n \to \mathbb{R}$, as defined in (2.4). Observe that for $1 \leqslant j < a \leqslant n$ we have

$$D^{(j\,a)}\Psi(\sigma) = (u(j)-u(a))(v(\sigma_j)-v(\sigma_a)).$$

Therefore $\|D^{(j\,a)}\Psi\| \leqslant \alpha$ and $\alpha_j[\Psi, S_n] \leqslant \alpha$. When $1 \leqslant j, k, a, b \leqslant n$ are distinct, observe that $D^{(k\,b)}\,D^{(j\,a)}\Psi(\sigma) = 0$. Otherwise we can bound

$$\|D^{(k\,b)}\,D^{(j\,a)}\Psi\| \leqslant 2\|D^{(j\,a)}\Psi\| \leqslant 2\alpha,$$

which leads to the estimate $\Delta_{jk}[\Psi, S_n] \leqslant 4\alpha/(n-j)$. Applying Theorem 2.2 and observing that $\beta_j \leqslant 4\alpha^2$ gives the stated bound on L. This completes the proof of (i).

For part (ii) we may assume without loss of generality that $\bar{u}, \bar{v}, \bar{u}', \bar{v}'$ all equal, by shifting u, v, u', v' if necessary. This shifts the distributions of Ψ and Ψ' but has no effect on their covariance.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Next observe that for $j, k = 1, \ldots, n$,

$$\operatorname{Cov}\,(u(j)v(X_j), u'(k)v'(X_k)) = \biggl(\sum_{i=1}^{n} v(i)v'(i)\biggr)\frac{u(j)u'(k)}{n}\biggl(1 - \frac{(n+1)\,\mathbf{1}_{j\neq k}}{n-1}\biggr),$$

where $\mathbf{1}_{j\neq k}$ is the indicator variable which equals 1 when $j \neq k$ and 0 otherwise. Summing this expression over all pairs (j, k) proves the first statement of (ii), and replacing Ψ' with Ψ completes the proof of (ii).

For part (iii) we calculate that

$$\mathbb{E}\,[v(X_1)v(X_2)] = \bar{v}^2 - \frac{1}{n(n-1)}\sum_{i=1}^{n}(v(i) - \bar{v})^2,$$

from which (iii) follows.

For part (iv), it is not difficult to prove by induction on k that

$$\mathbb{E}\,[v(X_1)v(X_2)\cdots v(X_k)] = \bar{v}^k + O(n^{-1})\,\|v\|^k \quad \text{for } k = O(1). \tag{3.1}$$

This follows using the fact that for $k \geqslant 1$,

$$\mathbb{E}\,[v(X_1)v(X_2)\cdots v(X_k)v(X_{k+1})] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\,[v(X_1)v(X_2)\cdots v(X_k)\mid X_{k+1} = i]\;v(i),$$

after observing that the average of $\{v(j) \mid j \neq i\}$ equals $\bar{v} + O(\|v\|/n)$. It follows that

$$\mathbb{E}\,E_{jk}(X) = (u(j) + \bar{v})(u(k) + \bar{v}) + O\biggl(\frac{\|v\|^2}{n}\biggr). \tag{3.2}$$

Therefore $E_{jk}(X)$, $E_{\ell m}(X)$, $\mathbb{E}\,E_{jk}(X)$ and $\mathbb{E}\,E_{\ell m}(X)$ are all $O((\|u\| + \|v\|)^2)$, from which we conclude that $\operatorname{Cov}\,(E_{jk}, E_{\ell m}) = O((\|u\| + \|v\|)^4)$, for any distinct j, k and distinct ℓ, m. In the case that $\{j, k\} \cap \{\ell, m\} = \emptyset$, it follows from (3.1) that

$$\mathbb{E}\,[E_{jk}(X)E_{\ell m}(X)] = (u(j) + \bar{v})(u(k) + \bar{v})(u(\ell) + \bar{v})(u(m) + \bar{v}) + O\biggl(\frac{(\|u\| + \|v\|)^4}{n}\biggr).$$

The improved bound on $\operatorname{Cov}\,(E_{jk}(X), E_{\ell m}(X))$ follows directly.

Finally we prove part (v). First we calculate $\operatorname{Cov}\,(E_{jk}(X), v'(X_i))$ under the assumption that $i \notin \{j, k\}$. In this case

$$\begin{aligned}
&\mathbb{E}\,[E_{jk}(X)\,v'(X_i)]\\
&= \frac{1}{n}\sum_{a=1}^{n}\mathbb{E}\,[(u(j) + v(X_j))(u(k) + v(X_k))\mid X_i = a]\;v'(a)\\
&= \frac{1}{n}\sum_{a=1}^{n} v'(a)\biggl(\biggl(u(j) + \frac{n\bar{v} - v(a)}{n-1}\biggr)\biggl(u(k) + \frac{n\bar{v} - v(a)}{n-1}\biggr)\\
&\qquad\qquad - \frac{1}{(n-1)(n-2)}\sum_{b:\, b\neq a}(v(b) - \bar{v})^2\biggr)\\
&= \bar{v}'(u(j) + \bar{v})(u(k) + \bar{v}) - \frac{\bar{v}'}{n(n-1)}\sum_{b=1}^{n}(v(b) - \bar{v})^2\\
&\quad - \frac{u(j) + u(k) + 2\bar{v}}{n(n-1)}\sum_{a=1}^{n} v'(a)(v(a) - \bar{v}) + O\biggl(\frac{\|v\|^2\|v'\|}{n^2}\biggr).
\end{aligned}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

The second line follows from applying (iii) to the restriction of u, v to $\{1, \ldots, n\} \setminus \{a\}$. Subtracting $\mathbb{E}\, E_{jk}\ \mathbb{E}\, v'(X_i)$ using (iii), we find that

$$\mathrm{Cov}\,(E_{jk}(X), v'(X_i)) = -\frac{u(j) + u(k) + 2\bar{v}}{n(n-1)} \sum_{a=1}^{n} (v(a) - \bar{v})(v'(a) - \bar{v}') + O\biggl(\frac{\|v\|^2 \|v'\|}{n^2}\biggr).$$

Now suppose that $i \in \{j, k\}$. When $i = j$, using similar calculations as above, we obtain

$$\mathbb{E}\,[E_{jk}(X)v'(X_j)] = \frac{1}{n} \sum_{a=1}^{n} (u(j) + v(a))(u(k) + \bar{v})\, v'(a) + O\biggl(\frac{(\|u\| + \|v\|)^2 \|v'\|}{n}\biggr)$$

and hence

$$\mathrm{Cov}\,(E_{jk}(X), v'(X_j)) = (u(k) + \bar{v})\frac{1}{n} \sum_{a=1}^{n} (v(a) - \bar{v})(v'(a) - \bar{v}') + O\biggl(\frac{(\|u\| + \|v\|)^2 \|v'\|}{n}\biggr).$$

A similar formula holds when $i = k$. Now summing over all i, we obtain the stated formula for $\mathrm{Cov}\,(E_{jk}(X), \Psi'(X))$, completing the proof. □

4. Subgraphs isomorphic to a given graph

If H is a graph on the vertex set $\{1, \ldots, n\}$, let $\mathbb{P}(\boldsymbol{d}, H)$ be the probability that $G \sim \mathcal{G}_{\boldsymbol{d}}$ contains H as a subgraph. The starting point for our arguments is the following result adapted from McKay [9, Theorem 2.1]. We state it using the parameters defined in (1.1) and (1.3).

Theorem 4.1. *For any constant $\eta \in (0, \frac{1}{2})$ there is some $\varepsilon_5(\eta) > 0$ such that the following holds for every fixed $\varepsilon \in (0, \varepsilon_5(\eta))$. Let $\boldsymbol{d}$ be a degree sequence which satisfies (1.2). Suppose that H is a graph on the vertex set $\{1, \ldots, n\}$ with $m \leqslant n^{1+2\varepsilon}$ edges and degree sequence $\boldsymbol{h} = (h_1, \ldots, h_n)$ such that $\|\boldsymbol{h}\| \leqslant n^{1/2+\varepsilon}$. Then*

$$\mathbb{P}(\boldsymbol{d}, H) = \lambda^m\ \exp\,(f(\boldsymbol{d}, \boldsymbol{h}) + g(\boldsymbol{d}, H) + O(n^{-1/2+\eta})), \tag{4.1}$$

where

$$f(\boldsymbol{d}, \boldsymbol{h}) = \frac{(1-\lambda)}{4\lambda}(\mu_1^2 + 2\mu_1 - 2\mu_2) - \frac{(1-\lambda^2)}{6\lambda^2 n}\, \mu_3 + \frac{1}{\lambda n} \sum_{j=1}^{n} (d_j - d)h_j$$
$$+ \frac{1}{2\lambda^2 n^2} \sum_{j=1}^{n} (d_j - d)h_j^2 - \frac{1}{2\lambda^2 n^2} \sum_{j=1}^{n} (d_j - d)^2 h_j,$$
$$g(\boldsymbol{d}, H) = -\frac{1}{\lambda(1-\lambda)n^2} \sum_{jk \in E(H)} (d_j - d - h_j + \lambda h_j)(d_k - d - h_k + \lambda h_k).$$

Proof. Theorem 2.1 in [9] is stated slightly differently. It supposes two constants $a, b > 0$ with $a + b < \frac{1}{2}$ and the second part of (1.2) reads

$$\min\{d, n - d - 1\} \geqslant \frac{n}{3a \log n}.$$

If $\varepsilon = \varepsilon(a, b)$ defined in [9, Theorem 2.1], then (4.1) holds with the error term $O(n^{-b})$. To obtain our formulation, take $a = 2\eta^2$, $b = \frac{1}{2} - \eta$ and $\varepsilon_5(\eta) = \varepsilon(a, b)$. Clearly, for any $\varepsilon < \varepsilon_5(\eta)$ formula (4.1) also holds with the same error term, since all assumptions depending on ε become more strict. □

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Lemma 4.1. *Let $A = \bigcup_{k=1}^{\infty} A_k \subseteq \mathbb{C}$ be such that for each k, $A_k \neq \emptyset$ and $\sup_{a\in A_k} |a| < \infty$. Suppose $a_k = O(1)$ as $k \to \infty$ for every sequence $a_1, a_2, \ldots$ with $a_k \in A_k$ for each k. Then $\sup_{a\in A} |a| < \infty$.*

Proof. For each k there is some $a'_k \in A_k$ such that $|a'_k| \geqslant \sup_{a\in A_k} |a| - 1$. Now we have $\sup_{a\in A} |a| \leqslant \sup_{k\geqslant 1} |a'_k| + 1 < \infty$. □

Remark 4.1. Lemma 4.1 will be useful whenever we need to sum Theorem 4.1, or similar theorems, over many values of the parameters. For fixed η, there are only finitely many degree sequences $\boldsymbol{d}$ and graphs H for each n that satisfy the conditions. Therefore, applying Lemma 4.1 to the sets consisting of the error terms in Theorem 4.1 for these $\boldsymbol{d}$ and H, scaled by a factor $n^{1/2-\eta}$, shows that the error term is uniform over $\boldsymbol{d}$ and H. That is, there is a function $C(\eta)$, not depending on any other parameters, such that the absolute value of the error term is bounded above by $C(\eta) n^{-1/2+\eta}$.

In order to compute the expected number of subgraphs isomorphic to H, we must sum $\mathbb{P}(\boldsymbol{d}, H)$ over all possible locations of H. We will find it convenient to average over permutations of the degree sequence $\boldsymbol{d}$ rather than over labellings of the subgraph H; by symmetry, this is equivalent.

For a permutation $\sigma \in S_n$, let $\boldsymbol{d}^\sigma = (d_{\sigma_1}, \ldots, d_{\sigma_n})$ denote the permuted degree sequence, and let

$$f_{\boldsymbol{h}}(\sigma) = \frac{(1-\lambda)}{4\lambda}(\mu_1^2 + 2\mu_1 - 2\mu_2) - \frac{(1-\lambda^2)}{6\lambda^2 n}\,\mu_3 + \frac{1}{\lambda n}\sum_{j=1}^{n}(d_{\sigma_j} - d)h_j + \frac{1}{2\lambda^2 n^2}\sum_{j=1}^{n}(d_{\sigma_j} - d)h_j^2 - \frac{1}{2\lambda^2 n^2}\sum_{j=1}^{n}(d_{\sigma_j} - d)^2 h_j,$$

$$g_H(\sigma) = -\frac{1}{\lambda(1-\lambda)n^2}\sum_{jk\in E(H)}(d_{\sigma_j} - d - h_j + \lambda h_j)(d_{\sigma_k} - d - h_k + \lambda h_k).$$

Since $f_{\boldsymbol{h}}(\sigma) = f(\boldsymbol{d}^\sigma, \boldsymbol{h})$ and $g_H(\sigma) = g(\boldsymbol{d}^\sigma, H)$, Theorem 4.1 implies that the expected number of subgraphs isomorphic to H in a uniformly random graph with degree sequence $\boldsymbol{d}$ is

$$(1 + O(n^{-1/2+\eta}))\,\frac{n!}{|\,\mathrm{Aut}\,(H)|}\,\lambda^m\,\mathbb{E}\,[\exp\,(f_{\boldsymbol{h}}(X) + g_H(X))], \tag{4.2}$$

where the expectation is taken with respect to a uniformly random element X of S_n. Here we have used the uniformity of the error term $O(n^{-1/2+\eta})$ in Theorem 4.1, as explained in Remark 4.1.

Define

$$\varepsilon_1(\eta) = \min\left\{\varepsilon_5(\eta), \frac{1}{12}\eta\right\}, \tag{4.3}$$

where $\varepsilon_5(\eta)$ is provided by Theorem 4.1. Before proving Theorem 1.1, we apply the results of Section 3 to obtain the following expressions.

Lemma 4.2. *Let $\eta \in (0, \frac{1}{2})$ be constant. If assumptions (1.2) and (1.4) hold with $\varepsilon \in (0, \varepsilon_1(\eta)]$, then*

$$\mathbb{E}\, f_{\boldsymbol{h}}(X) = \frac{(1-\lambda)}{4\lambda}(\mu_1^2 + 2\mu_1 - 2\mu_2) - \frac{(1-\lambda^2)}{6\lambda^2 n}\,\mu_3 - \frac{R}{2\lambda^2 n}\mu_1,$$

$$\mathbb{E}\, g_H(X) = -\frac{1-\lambda}{\lambda n^2}\sum_{jk\in E(H)} h_j h_k + O(n^{-1/2+\eta}),$$

$$\mathrm{Var}\,[f_{\boldsymbol{h}}(X) + g_H(X)] = \frac{R}{\lambda^2 n}(\mu_2 - \mu_1^2) + O(n^{-1/2+\eta}).$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Proof. We repeatedly use the following bounds in our estimates:

$$n\mu_t \leqslant 2\,\|\boldsymbol{h}\|^{t-1}\, m \leqslant 2n^{(t+1)(1/2+\varepsilon)} \quad \text{and} \quad R \leqslant \delta^2 \leqslant n^{1+2\varepsilon}.$$

The expression for $\mathbb{E}\, f_{\boldsymbol{h}}(X)$ follows directly from applying Lemma 3.1(i) to the terms of $f_{\boldsymbol{h}}(X)$. (The first two terms are constants, the third and fourth term both give zero since the average of $d_{\sigma_j} - d$ is zero, and the fifth term of $f_{\boldsymbol{h}}$ provides the final term in the expected value.) Similarly, using Lemma 3.1(iii) and (1.2) we have

$$\mathbb{E}\, g_H(X) = -\frac{1-\lambda}{\lambda n^2} \sum_{jk\in E(H)} h_j h_k + \frac{R}{\lambda(1-\lambda)n^2(n-1)},$$

which matches the given expression after applying the assumptions.

Recall that for real random variables $X_1, \ldots, X_t$ we have

$$\operatorname{Var}\Bigg[\sum_{j=1}^{t} X_j\Bigg] = \sum_{j,k=1}^{t} \operatorname{Cov}(X_j, X_k).$$

For all positive integers i, ℓ, and for all $jk \in E(H)$, define the functions $\Psi^{(i,\ell)}, E_{jk} \colon S_n \to \mathbb{R}$ by

$$\Psi^{(i,\ell)}(\sigma) = \sum_{j=1}^{n} (d_{\sigma_j} - d)^i\, h_j^\ell,$$

$$E_{jk}(\sigma) = ((\lambda-1)h_j + d_{\sigma_j} - d)((\lambda-1)h_k + d_{\sigma_k} - d)$$

for all $\sigma \in S_n$. Using Lemma 3.1(ii) with $u(j) = h_j$ and $v(\sigma_j) = d_{\sigma_j} - d$, we find that

$$\operatorname{Var}\bigg[\frac{1}{\lambda n}\Psi^{(1,1)}(X)\bigg] = \frac{R(\mu_2 - \mu_1^2)}{\lambda^2(n-1)} = \frac{R(\mu_2-\mu_1^2)}{\lambda^2 n} + O(n^{-1/2+\eta}).$$

Applying (1.4) and Lemma 3.1(v) with $u(j) = (\lambda-1)h_j$, $u'(j) = h_j$, and $v(j) = v'(j) = d_j - d$, we obtain

$$\begin{aligned}
&\sum_{jk\in E(H)} \operatorname{Cov}(\Psi^{(1,1)}(X), E_{jk}(X)) \\
&\quad = O(\mu_1(\|u\| + \|v\|)^2 \|u'\|\|v'\|) + R(\lambda-1) \sum_{jk\in E(H)} ((h_j - \mu_1)h_k + (h_k - \mu_1)h_j) \\
&\quad = O(n^{2+6\varepsilon}).
\end{aligned}$$

Consequently

$$\operatorname{Cov}\bigg(\frac{1}{\lambda n}\Psi^{(1,1)}(X), -\frac{1}{\lambda(1-\lambda)n^2} E_{jk}(X)\bigg) = O(n^{-1/2+\eta}).$$

Observe also that

$$\begin{aligned}
&\operatorname{Var}\Bigg[-\frac{1}{\lambda(1-\lambda)n^2} \sum_{jk\in E(H)} E_{jk}(X)\Bigg] \\
&\quad = \frac{1}{\lambda^2(1-\lambda)^2 n^4} \sum_{jk\in E(H)} \sum_{i\ell\in E(H)} \operatorname{Cov}(E_{jk}(X), E_{i\ell}(X)) \\
&\quad = \frac{1}{\lambda^2(1-\lambda)^2 n^4} \bigg(m\|\boldsymbol{h}\| O(((1-\lambda)\|\boldsymbol{h}\| + \delta)^4) + m^2\, O\bigg(\frac{((1-\lambda)\|\boldsymbol{h}\| + \delta)^4}{n}\bigg)\bigg) \\
&\quad = O(n^{-1/2+\eta}).
\end{aligned} \tag{4.4}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

This follows from Lemma 3.1(iv), using the fact that there are at most $m\|\boldsymbol{h}\|$ pairs of adjacent edges of H and at most m^2 pairs of non-adjacent edges of H.

We can now verify that all the remaining contributions to $\operatorname{Var}[f_{\boldsymbol{h}}(X)+g_H(X)]$ are $O(n^{-1/2+\eta})$. Using Lemma 3.1(ii), we have

$$\operatorname{Var}\biggl[\frac{1}{2\lambda^2 n^2}\Psi^{(1,2)}(X)\biggr] = \frac{(\mu_4-\mu_2^2)R}{4\lambda^4 n^2(n-1)} = O(n^{-1/2+\eta}), \tag{4.5}$$

$$\operatorname{Var}\biggl[-\frac{1}{2\lambda^2 n^2}\Psi^{(2,1)}(X)\biggr] = \frac{(\mu_2-\mu_1^2)(R_4-R^2)}{4\lambda^4 n^2(n-1)} = O(n^{-1/2+\eta}), \tag{4.6}$$

where $R_4=\sum_{j=1}^n (d_j-d)^4$. For any two real random variables Z, Z' we have $|\operatorname{Cov}(Z,Z')| \leqslant \max\{\operatorname{Var} Z, \operatorname{Var} Z'\}$, so the three covariances involving two of the quantities in (4.4)–(4.6) are also $O(n^{-1/2+\eta})$. The only remaining covariances are

$$\operatorname{Cov}\biggl(\frac{1}{\lambda n}\Psi^{(1,1)}(X), \frac{1}{2\lambda^2 n^2}\Psi^{(1,2)}(X)\biggr) = \frac{R(\mu_3-\mu_1\mu_2)}{2\lambda^3 n(n-1)} = O(n^{-1/2+\eta}) \quad \text{and}$$

$$\operatorname{Cov}\biggl(\frac{1}{\lambda n}\Psi^{(1,1)}(X), -\frac{1}{2\lambda^2 n^2}\Psi^{(2,1)}(X)\biggr) = -\frac{(\mu_2-\mu_1^2)\sum_{j=1}^n (d_j-d)^3}{2\lambda^3 n^2(n-1)} = O(n^{-1/2+\eta}),$$

using (1.4), since $R \leqslant \delta^2$ and $\mu_1^2 \leqslant \mu_2 \leqslant \mu_3$. This completes the proof. □

Proof of Theorem 1.1. Define $\varepsilon_1(\eta)$ as in (4.3). We will apply Theorem 2.2 to estimate (4.2). The expected value and variance of $f_{\boldsymbol{h}}+g_H$ are given in Lemma 4.2. It remains to prove that

$$\sum_{j=1}^{n-1}\biggl(\frac{1}{6}\alpha_j^3+\frac{1}{3}\alpha_j\beta_j+\frac{5}{8}\alpha_j^4+\frac{5}{8}\beta_j^2\biggr) = O(n^{-1/2+\eta}),$$

where $\alpha_j=\alpha_j[f_{\boldsymbol{h}}+g_H, S_n]$, $\beta_j=\sum_{k=j+1}^{n-1}\alpha_k\Delta_{jk}$ and $\Delta_{jk}=\Delta_{jk}[f_{\boldsymbol{h}}+g_H, S_n]$.

Without loss of generality, we can assume that

$$h_1 \geqslant h_2 \geqslant \cdots \geqslant h_n.$$

We calculate that for $1 \leqslant j < a \leqslant n$,

$$\|D^{(j\,a)} f_{\boldsymbol{h}}\| = O\biggl(\frac{\delta h_j}{\lambda n}\biggr).$$

Therefore

$$\alpha_j[f_{\boldsymbol{h}}, S_n] = O\biggl(\frac{\delta h_j}{\lambda n}\biggr).$$

Observe also that $\|D^{(k\,b)}\,D^{(j\,a)} f_{\boldsymbol{h}}\| = 0$ whenever j,k,a,b are distinct. Otherwise, for $1 \leqslant j < a \leqslant n$ and $1 \leqslant k < b \leqslant n$ with j,k,a,b not all distinct, we use the bound

$$\|D^{(k\,b)}\,D^{(j\,a)} f_{\boldsymbol{h}}\| \leqslant 2\|D^{(j\,a)} f_{\boldsymbol{h}}\| = O\biggl(\frac{\delta h_j}{\lambda n}\biggr).$$

Thus

$$\Delta_{jk}[f_{\boldsymbol{h}}, S_n] = O\biggl(\frac{\delta h_j}{\lambda n(n-j)}\biggr).$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Next we need to consider g_H, and calculate that

$$\|D^{(j\,a)}g_H\| = O\biggl(\frac{(\delta+h_j)^2 h_j}{\lambda(1-\lambda)n^2}\biggr),$$

since only the terms of g_H corresponding to edges incident with j or a can contribute. This gives us

$$\alpha_j[g_H, S_n] = O\biggl(\frac{(\delta+h_j)^2 h_j}{\lambda(1-\lambda)n^2}\biggr) = O(n^{-1/2+4\varepsilon}).$$

Suppose that j, k, a, b are all distinct, with $j < a$ and $j < k < b$. If $\{jk,\ jb,\ ab,\ kb\} \cap E(H) = \emptyset$, then

$$\|D^{(k\,b)}\,D^{(j\,a)}g_H\| = 0,$$

and otherwise

$$\|D^{(k\,b)}\,D^{(j\,a)}g_H\| = O\biggl(\frac{(\delta+h_j)^2}{\lambda(1-\lambda)n^2}\biggr).$$

Therefore

$$\Delta_{jk}[g_H, S_n)] = O\biggl(\frac{(\delta+h_j)^2}{\lambda(1-\lambda)n^2}\biggr)\biggl(\mathbf{1}_{jk\in E(H)} + \frac{h_j}{n-k}\biggr).$$

To see this, we recall that

$$\Delta_{jk}[g_H, S_n] = \frac{1}{(n-j)(n-k)} \sum_{a=j+1}^{n} \sum_{b=k+1}^{n} \|D^{(k\,b)}D^{(j\,a)}g_H\|$$

and observe that there are at most h_j choices for $b > k$ such that $jb \in E(H)$. Also, there are at most $n-j$ choices for a: dividing the product of these by $(n-k)(n-j)$ leads to the term $h_j/(n-k)$. Similarly, there are at most h_k choices for $a > j$ such that $ka \in E(H)$, and then at most $n-k$ choices for $b > k$. If $jk \in E(H)$ then there are at most $(n-k)(n-j)$ choices for a, b, and there are at most $(n-k)h_k$ edges with at least one end-vertex in $\{k+1, \ldots, n\}$ (this counts the choices for $ab \in E(H)$). The 'diagonal' terms (where $a = b$ or $a = k$) satisfy

$$\|D^{(k\,b)}\,D^{(j\,a)}g_H\| \leqslant 2\|D^{(j\,a)}g_H\| = O\biggl(\frac{(\delta+h_j)^2 h_j}{\lambda(1-\lambda)n^2}\biggr).$$

So their contribution to $\Delta_{jk}(g_H, S_n)$ is bounded by

$$O\biggl(\frac{(\delta+h_j)^2 h_j}{\lambda(1-\lambda)n^2(n-j)}\biggr).$$

Combining estimates above and recalling (2.5), we conclude

$$\alpha_j = O\biggl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\biggr), \qquad 1 \leqslant j \leqslant n,$$

$$\Delta_{jk} = O\biggl(\frac{\delta h_j}{\lambda n(n-j)} + \frac{(\delta+h_j)^2}{\lambda(1-\lambda)n^2}\biggl(\mathbf{1}_{jk\in E(H)} + \frac{h_j}{n-k}\biggr)\biggr), \qquad 1 \leqslant j < k \leqslant n.$$

Using the inequality $(|x| + |y|)^3 \leqslant 4(|x|^3 + |y|^3)$ for each term of the sum, we get

$$\sum_{j=1}^{n-1} \alpha_j^3 = \sum_{j=1}^{n-1} O\biggl(\biggl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\biggr)^3\biggr) = O\biggl(\frac{\delta^3 \mu_3}{\lambda^3 n^2} + n^{-1/2+12\varepsilon}\biggr). \tag{4.7}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

This is $O(n^{-1/2+\eta})$ by (1.4) and our assumption $\varepsilon \leqslant \frac{1}{12}\eta$. We now want to show that the other error terms from Theorem 2.2 all fit inside this bound too.

Now

$$\sum_{k=j+1}^{n-1} \mathbf{1}_{jk\in E(H)} \leqslant h_j, \qquad \sum_{k=j+1}^{n-1} \frac{1}{n-k} \leqslant 1+\log n,$$

so Lemma 2.2 gives

$$\begin{aligned}\beta_j &= 2\sum_{k=j+1}^{n-1} \alpha_k \Delta_{jk}\\ &= O\biggl(\frac{\delta^2 h_j^2}{\lambda^2 n^2} + \frac{(\delta+h_j)^2\, \delta\, h_j^2\, \log n}{\lambda^2(1-\lambda)n^3}\biggr)\\ &= O\biggl(\frac{\delta^2 h_j^2}{\lambda^2 n^2} + \frac{\delta h_j^4 \log n}{\lambda^2(1-\lambda)n^3}\biggr)\\ &= O\biggl(\Bigl(\frac{\delta h_j}{\lambda n} + \frac{h_j^3 \log n}{\lambda(1-\lambda)n^2}\Bigr)^2\biggr)\\ &= O\biggl(\Bigl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\Bigr)^2\biggr).\end{aligned}$$

From this we find that

$$\sum_{j=1}^{n-1} \alpha_j\beta_j = \sum_{j=1}^{n-1} O\biggl(\Bigl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\Bigr)^3\biggr),$$

which is $O(n^{-1/2+\eta})$, similarly to (4.7).

For the two remaining terms, note that

$$\sum_{j=1}^{n-1} O\biggl(\Bigl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\Bigr)^4\biggr) = \sum_{j=1}^{n-1} O\biggl(\Bigl(\frac{\delta h_j}{\lambda n} + n^{-1/2+4\varepsilon}\Bigr)^3\biggr)$$

whenever the right-hand side is $O(1)$, and furthermore both $\sum_{j=1}^{n-1} \alpha_j^4$ and $\sum_{j=1}^{n-1} \beta_j^2$ are covered by $O(n^{-1/2+\eta})$. Applying Theorem 2.2 and using Lemma 4.2 completes the proof. □

Proof of Corollary 1.1. To show (1.5), we recall that, by assumption, $\mu_3 \leqslant \lambda^2 n^{1/2+\eta}$. Hence

$$\frac{1-\lambda}{\lambda n^2} \sum_{jk\in E(H)} h_j h_k \leqslant \frac{1-\lambda}{2\lambda n^2} \sum_{jk\in E(H)} (h_j^2 + h_k^2) = \frac{1-\lambda}{2\lambda n}\mu_3 = O(n^{-1/2+\eta}).$$

To prove the second part of the corollary, we can use the bound $R \leqslant \delta^2 \leqslant n^{1+2\varepsilon}$ and observe that $\mu_1 \leqslant \mu_2$ since each h_j is a natural number, and $\mu_1^2 \leqslant \mu_2$ by the power mean inequality. □

Proof of Corollary 1.2. Part (a) is a simple application of Corollary 1.1.

To prove part (b), note that the argument of the exponential in part (a) depends only on the degree h, and not on the finer structure of H, within the error term. Therefore, using the logic in Remark 4.1, the expected number of spanning h-regular subgraphs is

$$\mathrm{RG}\,(n,h)\,\lambda^m \exp\biggl(-\frac{1-\lambda}{4\lambda}h(h-2) - \frac{Rh}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr),$$

where RG (n, h) is the number of labelled regular graphs of order n and degree h. From [8] we know that

$$\mathrm{RG}\,(n,h)=\frac{(nh)!}{(nh/2)!\,2^{nh/2}(h!)^n}\exp\left(-\frac{h^2-1}{4}+O(h^3/n)\right).$$

Now apply Stirling's formula and observe that $h^3/n \leqslant n^{-1+6\varepsilon} \leqslant n^{-1/2+\eta}$, by (4.3). □

5. Spanning trees

As another application of our results, we calculate the expected number of spanning trees of $G \sim \mathcal{G}_{\boldsymbol{d}}$ where the degree sequence $\boldsymbol{d} = (d_1, \ldots, d_n)$ satisfies (1.2). Recall the parameters defined in (1.1).

5.1 Plan of attack

For some $\varepsilon > 0$, to be chosen later, define

$$\begin{aligned} D &= \{(h_1,\ldots,h_n)\in\{1,2,\ldots,n-1\}^n\colon h_1+\cdots+h_n=2n-2\},\\ D_{\text{good}} &= \{(h_1,\ldots,h_n)\in D\colon h_1,\ldots,h_n\leqslant n^{3\varepsilon}\},\\ D_{\text{bad}} &= D\setminus D_{\text{good}}. \end{aligned}$$

Let $\mathcal{T}$ be the set of all labelled trees with n vertices. For $\boldsymbol{h} \in D$, let $\mathcal{T}_{\boldsymbol{h}}$ be the set of all $T \in \mathcal{T}$ with degree sequence $\boldsymbol{h}$, and note that every $T \in \mathcal{T}$ belongs to $\mathcal{T}_{\boldsymbol{h}}$ for some $\boldsymbol{h} \in D$. It is well known that the number of trees with degree sequence $\boldsymbol{h}$ is

$$|\mathcal{T}_{\boldsymbol{h}}| = \binom{n-2}{h_1-1,\ldots,h_n-1} \tag{5.1}$$

(see [11, Theorem 3.1]). Also define

$$\mathcal{T}_{\text{good}} = \bigcup_{\boldsymbol{h}\in D_{\text{good}}} \mathcal{T}_{\boldsymbol{h}} \quad\text{and}\quad \mathcal{T}_{\text{bad}} = \bigcup_{\boldsymbol{h}\in D_{\text{bad}}} \mathcal{T}_{\boldsymbol{h}}.$$

A tree is called *good* if it belongs to $\mathcal{T}_{\text{good}}$, and otherwise it is bad.

Our approach will be to write the expected number of spanning trees in a uniformly random graph with degree sequence $\boldsymbol{d}$ as

$$\sum_{T\in\mathcal{T}} \mathbb{P}(\boldsymbol{d},T) = \sum_{T\in\mathcal{T}_{\text{good}}} \mathbb{P}(\boldsymbol{d},T) + \sum_{T\in\mathcal{T}_{\text{bad}}} \mathbb{P}(\boldsymbol{d},T). \tag{5.2}$$

Theorem 1.2 follows immediately from Lemma 5.3, which counts good spanning trees, and Lemma 5.4, which counts bad spanning trees. In fact bad spanning trees will turn out to be rare, so the second sum will contribute a negligible amount relative to the first sum.

5.2 The expected number of good spanning trees

It will be useful to define a random variable related to the degree sequence of a tree uniformly chosen from $\mathcal{T}$, or from $\mathcal{T}_{\boldsymbol{h}}$. Since we are only interested in good trees, we will also consider the truncation of these random vectors, where any entry larger than $\lfloor n^{3\varepsilon}\rfloor$ is replaced by $\lfloor n^{3\varepsilon}\rfloor$.

Lemma 5.1. *Let $X = (X_1, \ldots, X_n)$ be the degree sequence of a random tree uniformly chosen from $\mathcal{T}$. Then, for any fixed $\varepsilon > 0$, the following hold.*

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

(i) *The random vector* $(X_1 - 1, \dots, X_n - 1)$ *has a multinomial distribution with parameters* $m = n-2$, $k = n$, $\lambda_1 = \cdots = \lambda_k = 1$, *in the notation of* (2.20).

(ii) *Next, consider a random variable* $Y \in \{0, 1, 2, \dots\}^n$, *whose components are i.i.d. Poisson variables with mean* 1. *For each* $\boldsymbol{y}$, *we have that*

$$\mathbb{P}(X_1 = y_1{+}1, \dots, X_n = y_n{+}1) = \mathbb{P}(Y_1 = y_1, \dots, Y_n = y_n \mid Y_1 + \cdots + Y_n = n-2).$$

(iii) *Define the random variable* $Z = (Z_1, \dots, Z_n)$, *where* $Z_j = \min\{X_j, \lfloor n^{3\varepsilon} \rfloor\}$ *for each j. Then, uniformly over* $j \neq k$,

$$\mathbb{E}\, Z_j = 2 + O(n^{-1}), \quad \mathbb{E}\, Z_j^2 = 5 + O(n^{-1}),$$
$$\operatorname{Var} Z_j = 1 + O(n^{-1}), \quad \operatorname{Var} Z_j^2 = 27 + O(n^{-1}), \quad \operatorname{Cov}(Z_j, Z_k) = -n^{-1} + O(n^{-2}).$$

Proof. Statements (i) and (ii) are well known and follow easily from (5.1).

For (iii), note that the probability generating function of X is

$$p(\boldsymbol{x}) = \sum_{\boldsymbol{h}} |\mathcal{T}_{\boldsymbol{h}}|\, \boldsymbol{x}^{\boldsymbol{h}} = n^{-n+2}\, x_1 \cdots x_n (x_1 + \cdots + x_n)^{n-2}.$$

This allows computation of small moments of X, for example

$$\mathbb{E}\, X_1 = \frac{\partial}{\partial x_1} p(\boldsymbol{x})|_{(1,\dots,1)} = 2 - \frac{2}{n}.$$

The differences between small moments of Z and the corresponding moments of X are within the given error terms. To see this, we can set all but one of the arguments of $p(\boldsymbol{x})$ equal to 1 to find the distribution of the degree of one vertex. Thus we find that $\mathbb{P}(X_1 = t) \leqslant 1/(t-1)!$ for $t \geqslant 1$ and so

$$\mathbb{P}(Z \neq X) \leqslant n\, \mathbb{P}(X_1 > \lfloor n^{3\varepsilon} \rfloor) = o(e^{-n^{3\varepsilon}}).$$

Statement (iii) follows. □

The following result from [3, Section 3] will be useful.

Lemma 5.2. *Let* $\phi_1, \dots, \phi_n \in \mathbb{R}$ *and let* $\boldsymbol{h}$ *be a sequence such that* $\mathcal{T}_{\boldsymbol{h}} \neq \emptyset$. *Then*

$$\frac{1}{|\mathcal{T}_{\boldsymbol{h}}|} \sum_{T \in \mathcal{T}_{\boldsymbol{h}}} \sum_{jk \in E(T)} \phi_j \phi_k = \frac{1}{n-2} \left(\left(\sum_{k=1}^n \phi_k \right) \left(\sum_{j=1}^n (h_j{-}1)\phi_j \right) - \left(\sum_{j=1}^n (h_j{-}1)\phi_j^2 \right) \right)$$

and

$$\frac{1}{|\mathcal{T}_{\boldsymbol{h}}|} \sum_{T \in \mathcal{T}_{\boldsymbol{h}}} \exp\left(- \sum_{jk \in E(T)} \phi_j \phi_k \right) = \exp\left(K - \frac{1}{|\mathcal{T}_{\boldsymbol{h}}|} \sum_{T \in \mathcal{T}_{\boldsymbol{h}}} \sum_{jk \in E(T)} \phi_j \phi_k \right)$$

for some K with

$$|K| \leqslant \frac{1}{8} n \left(\max_j |\phi_j| - \min_j |\phi_j| \right)^4.$$

Define

$$\varepsilon_2(\eta) = \min\left\{ \varepsilon_5(\eta), \frac{1}{8}\eta \right\}, \tag{5.3}$$

where $\varepsilon_5(\eta)$ is provided by Theorem 4.1.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Lemma 5.3. *Let $\eta \in (0, \frac{1}{2})$ be constant. If assumption (1.2) holds with $\varepsilon \in (0, \varepsilon_2(\eta)]$, then the expected number of good spanning trees is*

$$\sum_{T\in\mathcal{T}_{\text{good}}} \mathbb{P}(\boldsymbol{d}, T) = n^{n-2}\, \lambda^{n-1} \exp\biggl(-\frac{1-\lambda}{2\lambda} - \frac{R}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr).$$

Proof. For a good tree T, we can apply Theorem 4.1 to estimate $\mathbb{P}(\boldsymbol{d}, T)$. The function $f(\boldsymbol{d}, \boldsymbol{h})$ in Theorem 4.1 depends only on $\boldsymbol{h}$, not on the tree T itself. We can 'average out' the contribution of $g(\boldsymbol{d}, T)$, which depends on the structure of T, using the function $\bar{g}(\boldsymbol{d}, \boldsymbol{h})$ defined by

$$e^{\bar{g}(\boldsymbol{d},\boldsymbol{h})} = \frac{1}{|\mathcal{T}_{\boldsymbol{h}}|} \sum_{T\in\mathcal{T}_{\boldsymbol{h}}} e^{g(\boldsymbol{d},T)}.$$

Then we can write

$$\begin{aligned}\sum_{T\in\mathcal{T}_{\boldsymbol{h}}} \mathbb{P}(\boldsymbol{d}, T) &= \lambda^{n-1}\, \exp\,(f(\boldsymbol{d}, \boldsymbol{h}) + O(n^{-1/2+\eta})) \sum_{T\in\mathcal{T}_{\boldsymbol{h}}} e^{g(\boldsymbol{d},T)}\\ &= \lambda^{n-1}\, |\mathcal{T}_{\boldsymbol{h}}|\, \exp\,(f(\boldsymbol{d}, \boldsymbol{h}) + \bar{g}(\boldsymbol{d}, \boldsymbol{h}) + O(n^{-1/2+\eta})).\end{aligned}$$

Define

$$\phi_j = \frac{d_j - d - h_j + \lambda h_j}{n\sqrt{\lambda(1-\lambda)}}$$

for $j = 1, \ldots, n$. Using $\sum_{j=1}^n (d_j - d) = 0$, $\|\boldsymbol{h}\| \leqslant n^{3\varepsilon}$ and $|\sum_{j=1}^n h_j a_j| = O(n \max_j |a_j|)$ for any $a_1, \ldots, a_n$, the first part of Lemma 5.2 gives

$$\frac{1}{|\mathcal{T}_{\boldsymbol{h}}|} \sum_{T\in\mathcal{T}_{\boldsymbol{h}}} \sum_{jk\in E(T)} \phi_j\phi_k = O(n^{-1/2+2\varepsilon}).$$

Combining this with the second part of Lemma 5.2, we find that $\bar{g}(\boldsymbol{d}, \boldsymbol{h}) = O(n^{-1/2+2\varepsilon})$. The assumption $\|\boldsymbol{h}\| \leqslant n^{3\varepsilon}$ and the fact that $\mu_1 = 2 - 2/n$ for a tree imply that

$$\begin{aligned}&\frac{1-\lambda}{4\lambda}(\mu_1^2 + 2\mu_1) - \frac{(1-\lambda^2)}{6\lambda^2 n}\mu_3 + \frac{1}{2\lambda^2 n^2} \sum_{j=1}^{n} (d_j - d) h_j^2\\ &\quad = \frac{2(1-\lambda)}{\lambda} + O(n^{-1/2+8\varepsilon} + n^{-1+10\varepsilon}).\end{aligned}$$

Applying Theorem 4.1, we have, since $\varepsilon \leqslant \frac{1}{8}\eta$,

$$\sum_{T\in\mathcal{T}_{\text{good}}} \mathbb{P}(\boldsymbol{d}, T) = (1 + O(n^{-1/2+\eta}))\lambda^{n-1} \sum_{\boldsymbol{h}\in D_{\text{good}}} \binom{n-2}{h_1-1, \ldots, h_n-1} e^{f^*(\boldsymbol{d},\boldsymbol{h})}, \tag{5.4}$$

where

$$f^*(\boldsymbol{d}, \boldsymbol{h}) = \frac{2(1-\lambda)}{\lambda} - \frac{1-\lambda}{2\lambda}\mu_2 - \frac{1}{2\lambda^2 n^2} \sum_{j=1}^{n} (d_j - d)^2 h_j + \frac{1}{\lambda n} \sum_{j=1}^{n} (d_j - d) h_j.$$

We now rewrite (5.4) using the random variable Z defined in Lemma 5.1, by extending the sum over D_{good} to all of D, as follows:

$$\sum_{T\in\mathcal{T}_{\text{good}}} \mathbb{P}(\boldsymbol{d}, T) = n^{n-2}\lambda^{n-1} e^{O(n^{-1/2+\eta})} (\,\mathbb{E}\, e^{f^*(\boldsymbol{d},Z)} - \mathbb{E}\,[\mathbf{1}_{D_{\text{bad}}} e^{f^*(\boldsymbol{d},Z)}]). \tag{5.5}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Applying the estimates from Lemma 5.1(iii) to $f^*(\boldsymbol{d}, Z)$, we have

$$\begin{aligned}\mathbb{E} f^*(\boldsymbol{d}, Z) &= -\frac{1-\lambda}{2\lambda} - \frac{R}{\lambda^2 n} + O(n^{-1/2+\eta}),\\ \operatorname{Var} f^*(\boldsymbol{d}, Z) &= \frac{R}{\lambda^2 n} + O(n^{-1/2+\eta}).\end{aligned} \tag{5.6}$$

Observe that by Lemma 5.1(i), $f^*(\boldsymbol{d}, Z)$ can be written as a function of a multinomial distribution:

$$f^*(\boldsymbol{d}, Z) = \widetilde{f}(X_1 - 1, \dots, X_n - 1).$$

We will apply Theorem 2.3(iii) to the function $\widetilde{f}$. Recalling (2.18) and (2.19), we calculate that $\alpha_{\max} = O(n^{-1/2+2\varepsilon})$ and $\Delta_{\max} = 0$. Therefore, using (5.6),

$$\mathbb{E}\, e^{f^*(\boldsymbol{d},Z)} = \exp\biggl(-\frac{1-\lambda}{2\lambda} - \frac{R}{2\lambda^2 n} + O(n^{-1/2+\eta})\biggr). \tag{5.7}$$

Next we bound $\mathbb{E}\,[\mathbf{1}_{D_{\text{bad}}} e^{f^*(\boldsymbol{d},Z)}]$. From the definition of f we have $f^*(\boldsymbol{z}) \leqslant \hat{f}(\boldsymbol{z})$ for all $\boldsymbol{z} \in \{1, 2, \dots\}^n$, where

$$\hat{f}(\boldsymbol{z}) = \frac{2}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^{n} (d_j - d) z_j.$$

For $\sigma \in S_n$, define

$$\hat{f}_\sigma(\boldsymbol{z}) = \frac{2}{\lambda} + \frac{1}{\lambda n} \sum_{j=1}^{n} (d_{\sigma_j} - d) z_j.$$

We now apply Lemma 3.1 to estimate

$$\frac{1}{n!} \sum_{\sigma \in S_n} e^{\hat{f}_\sigma(\boldsymbol{z})} \quad \text{for } \boldsymbol{z} \in \{1, 2, \dots\}^n.$$

Defining

$$u_j = \frac{d_j - d}{\lambda n} \quad \text{and} \quad v_j = z_j$$

for $j = 1, \dots n$, we find with respect to a uniformly random permutation $\sigma \in S_n$ that $\hat{f}_\sigma(\boldsymbol{z})$ has expectation $2/\lambda$ and variance at most

$$\frac{R}{\lambda^2 n(n-1)} \sum_{j=1}^{n} z_j^2.$$

The parameter α required by Lemma 3.1 satisfies $\alpha = O(n^{-1/2+4\varepsilon}/\lambda)$. Consequently, by Lemma 3.1(i),

$$\frac{1}{n!} \sum_{\sigma \in S_n} e^{\hat{f}_\sigma(\boldsymbol{z})} = O(e^{2/\lambda})\, e^{C \sum_{j=1}^{n} z_j^2}, \tag{5.8}$$

where

$$C = \frac{R}{2\lambda^2 n(n-1)}.$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Since $D_{\rm good}$ is invariant under permutations of the components, $\mathbb{E}\, e^{\hat f(Z)}$ is a symmetric function of $d_1, \ldots, d_n$. Therefore

$$\mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{f(Z)}] \leqslant \mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{\hat f(Z)}] = \mathbb{E}\left[\mathbf{1}_{D_{\rm bad}} \frac{1}{n!} \sum_{\sigma\in S_n} e^{\hat f_\sigma(Z)}\right] = O(e^{2/\lambda})\,\mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{C\sum_{j=1}^n Z_j^2}].$$

Next, note that $Y_1 + \cdots + Y_n$ has a Poisson distribution with mean n, and hence, by Stirling's approximation,

$$\mathbb{P}(Y_1 + \cdots + Y_n = n-2) = \frac{e^{-n}\, n^{n-2}}{(n-2)!} = \Theta(n^{-1/2}).$$

Applying Lemma 5.1(ii), we obtain

$$\mathbb{P}(X_1 = y_1 + 1, \ldots, X_n = y_n + 1) = O(n^{1/2})\,\mathbb{P}(Y_1 = y_1, \ldots, Y_n = y_n). \tag{5.9}$$

Therefore

$$\mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{f(Z)}] = O(e^{2/\lambda} n^{1/2}) \sum_{y_1,\ldots,y_n} \mathbb{P}(Y = (y_1, \ldots, y_n))\, e^{C\sum_{j=1}^n \min\{y_j+1, \lfloor n^{3\varepsilon}\rfloor\}^2},$$

where the sum is restricted to sequences $(y_1, \ldots, y_n)$ of non-negative integers such that $(y_1 + 1, \ldots, y_n + 1) \notin D_{\rm good}$. Recalling that the components of Y are independent, we can separate the sum and use the union bound on the constraint. This gives

$$\mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{f(Z)}] \leqslant O(e^{2/\lambda} n^{3/2})(\Sigma_1 + \Sigma_2)^{n-1}\Sigma_2,$$

where

$$\Sigma_1 = \sum_{y=0}^{\lfloor n^{3\varepsilon}\rfloor} \frac{e^{-1}}{y!}\, e^{C(y+1)^2} \quad \text{and} \quad \Sigma_2 = \sum_{y=\lfloor n^{3\varepsilon}\rfloor+1}^{\infty} \frac{e^{-1}}{y!}\, e^{Cn^{6\varepsilon}}.$$

Since

$$C = O(n^{-1+2\varepsilon}/\lambda^2) \quad \text{and} \quad \sum_{j=0}^{\infty} \frac{1}{j!}(j+1)^2 = 5e,$$

we conclude that

$$\Sigma_1 = \sum_{y=0}^{\lfloor n^{3\varepsilon}\rfloor-1} \frac{e^{-1}}{y!}\,(1 + C(y+1)^2 + O(n^{-2+17\varepsilon})) = 1 + 5C + O(n^{-2+17\varepsilon}), \tag{5.10}$$

$$\Sigma_2 = O(e^{-n^{3\varepsilon}}).$$

Therefore

$$\mathbb{E}\,[\mathbf{1}_{D_{\rm bad}} e^{f(Z)}] = O(n^{3/2})\exp\left(2/\lambda - n^{3\varepsilon} + O(n^{2\varepsilon}/\lambda^2)\right) = O(e^{-n^{3\varepsilon}/2}). \tag{5.11}$$

Combining (5.5), (5.7) and (5.11) completes the proof. □

5.3 The expected number of bad spanning trees

To complete the proof of Theorem 1.2, it remains for us to bound $\sum_{T\in\mathcal{T}_{\rm bad}} \mathbb{P}(\boldsymbol{d}, T)$. Note that we cannot use Theorem 4.1 directly since T fails the required degree bound. However, we can choose a subgraph $F \subseteq T$ to which Theorem 4.1 applies and use the fact that $\mathbb{P}(\boldsymbol{d}, T) \leqslant \mathbb{P}(\boldsymbol{d}, F)$.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Lemma 5.4. *Let $\eta \in (0, \frac{1}{2})$ be constant. If assumption (1.2) holds with $\varepsilon \in (0, \varepsilon_2(\eta)]$, where $\varepsilon_2(\eta)$ is defined in (5.3), then the expected number of bad spanning trees is*

$$\sum_{T \in \mathcal{T}_{\mathrm{bad}}} \mathbb{P}(\boldsymbol{d}, T) = n^{n-2}\, \lambda^{n-1}\, O(e^{-n^{3\varepsilon}/2}).$$

Proof. Let T be a bad tree. Define $F(T)$ to be the set of all subgraphs of T that have maximum degree at most $n^{3\varepsilon}$ and at least $n-1-\sum_{j=1}^n \max\{0, x_j - n^{3\varepsilon}\}$ edges. Since one such subgraph is obtained by deleting $\max\{0, x_j - \lfloor n^{3\varepsilon} \rfloor\}$ arbitrary edges incident with each vertex j, we have $F(T) \neq \emptyset$. We also have that for any permutation σ of the vertices, $F(\sigma(T)) = \sigma(F(T))$. Since $g(\boldsymbol{d}, F') = O(n^{2\varepsilon}/\lambda(1-\lambda))$ for all $F' \in F(T)$, using Theorem 4.1 we can write

$$\sum_{T \in \mathcal{T}_{\mathrm{bad}}} \mathbb{P}(\boldsymbol{d}, T) \leqslant e^{\lambda^{-1}(1-\lambda)^{-1} O(n^{2\varepsilon})} \sum_{T \in \mathcal{T}_{\mathrm{bad}}} \lambda^{n-1-\sum_{j=1}^n \max\{0, x_j - n^{3\varepsilon}\}} \frac{1}{|F(T)|} \sum_{F' \in F(T)} e^{\hat{f}(\boldsymbol{z}(F'))},$$

where $\boldsymbol{z}(F')$ is the degree sequence of F' and $\hat{f}$ is defined as before. The expression on the right is a symmetric function of $\boldsymbol{d}$, so we can average it over all permutations of the elements of $\boldsymbol{d}$. The same calculations that led to (5.8) show that

$$\frac{1}{n!} \sum_{\sigma \in S_n} e^{\hat{f}_\sigma(\boldsymbol{z}(F'))} = O(e^{2/\lambda}) e^{C \sum_{j=1}^n \min\{x_j, n^{3\varepsilon}\}^2}.$$

Therefore

$$\begin{aligned}
&\sum_{T \in \mathcal{T}_{\mathrm{bad}}} \mathbb{P}(\boldsymbol{d}, T) \\
&\quad\leqslant e^{\lambda^{-1}(1-\lambda)^{-1} O(n^{2\varepsilon})}\, \lambda^{n-1} \sum_{T \in \mathcal{T}_{\mathrm{bad}}} \lambda^{-\sum_{j=1}^n \max\{0, x_j - n^{3\varepsilon}\}} e^{C \sum_{j=1}^n \min\{x_j, n^{3\varepsilon}\}^2} \\
&\quad\leqslant e^{\lambda^{-1}(1-\lambda)^{-1} O(n^{2\varepsilon})}\, \lambda^{n-1} n^{n-2} \\
&\qquad\times \sum_{y_1, \dots, y_n} \mathbb{P}(Y = (y_1, \dots, y_n))\, \lambda^{-\sum_{j=1}^n \max\{0, y_j + 1 - n^{3\varepsilon}\}}\, e^{C \sum_{j=1}^n \min\{y_j + 1, n^{3\varepsilon}\}^2},
\end{aligned}$$

using (5.9). As before, the sum is restricted to those sequences $(y_1, \dots, y_n)$ of non-negative integers such that $(y_1 + 1, \dots, y_n + 1) \notin D_{\mathrm{good}}$.

Separating the sum and applying the union bound, we have

$$\sum_{y_1, \dots, y_n} \mathbb{P}(Y = (y_1, \dots, y_n))\, \lambda^{-\sum_{j=1}^n \max\{0, y_j + 1 - n^{3\varepsilon}\}}\, e^{C \sum_{j=1}^n \min\{y_j + 1, n^{3\varepsilon}\}^2} \leqslant n\, (\Sigma_1 + \Sigma_2')^{n-1} \Sigma_2',$$

where Σ_1 was defined earlier and

$$\Sigma_2' = \sum_{y = \lfloor n^{3\varepsilon} \rfloor}^{\infty} \frac{e^{-1} \lambda^{-(y - n^{3\varepsilon})}}{y!} e^{C n^{6\varepsilon}} = O(e^{-n^{3\varepsilon}}).$$

Therefore, using (5.10),

$$\begin{aligned}
\sum_{T \in \mathcal{T}_{\mathrm{bad}}} \mathbb{P}(\boldsymbol{d}, T) &\leqslant \lambda^{n-1}\, n^{n-2} \exp\biggl(-n^{3\varepsilon} + \lambda^{-2} n^{2\varepsilon} + \frac{1}{\lambda(1-\lambda)} O(n^{2\varepsilon})\biggr) \\
&= \lambda^{n-1}\, n^{n-2}\, O(e^{-n^{3\varepsilon}/2}).
\end{aligned}$$

□

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

6. Counting induced subgraphs

Recall the parameters defined in (1.1). In this section our starting point is the following result adapted from McKay [9, Theorem 2.4] in the same way as Theorem 4.1 was adapted from [9, Theorem 2.1]. We require notation that generalizes (1.6):

$$\omega_{s,t} = \sum_{j=1}^{r} (d_j - d)^s (h_j - \lambda(r-1))^t \quad \text{for } s, t \geqslant 0.$$

Theorem 6.1. *Let $\eta \in (0, \frac{1}{2})$ be constant. Then there is a constant $\varepsilon_6(\eta) > 0$ such that the following holds for every fixed $\varepsilon \in (0, \varepsilon_6(\eta)]$. Let $\boldsymbol{d}$ be a degree sequence which satisfies (1.2). Suppose that $H^{[r]}$ is a graph on the vertex set $\{1, \ldots, r\}$ with degree sequence $\boldsymbol{h}^{[r]} = (h_1, \ldots, h_r)$ such that $r \leqslant n^{1/2+\varepsilon}$. Then the probability that $G \sim \mathcal{G}_{\boldsymbol{d}}$ has $H^{[r]}$ as an induced subgraph is*

$$\begin{aligned} &\lambda^m (1-\lambda)^{\binom{r}{2}-m} \\ &\quad \times \exp\biggl(\frac{2\omega_{1,1} - \omega_{0,2}}{2\lambda(1-\lambda)n} + \frac{r^2}{2n} + \frac{(1-2\lambda)\omega_{0,1}}{2\lambda(1-\lambda)n} + \frac{4\omega_{1,0}\omega_{0,1} - \omega_{0,1}^2 - 2\omega_{1,0}^2}{4\lambda(1-\lambda)n^2} \\ &\quad + \frac{r(2\omega_{1,1} - \omega_{2,0} - \omega_{0,2})}{2\lambda(1-\lambda)n^2} - \frac{(1-2\lambda)(\omega_{0,3} + 3\omega_{2,1} - 3\omega_{1,2})}{6\lambda^2(1-\lambda)^2 n^2} + O(n^{-1/2+\eta}) \biggr). \end{aligned}$$

Now, for a given permutation $\sigma \in S_n$, let

$$\omega_{s,t}(\sigma) = \sum_{j=1}^{r} (d_{\sigma_j} - d)^s \, (h_j - \lambda(r-1))^t.$$

Note that $\omega_{0,t}(\sigma)$ is independent of σ and equals ω_t from (1.6). Let $f_{\text{ind}} : S_n \to \mathbb{R}$ be defined as

$$\begin{aligned} f_{\text{ind}}(\sigma) = &\frac{2\omega_{1,1}(\sigma) - \omega_2}{2\lambda(1-\lambda)n} + \frac{r^2}{2n} + \frac{(1-2\lambda)\omega_1}{2\lambda(1-\lambda)n} + \frac{4\omega_{1,0}(\sigma)\omega_1 - \omega_1^2 - 2\omega_{1,0}(\sigma)^2}{4\lambda(1-\lambda)n^2} \\ &+ \frac{r(2\omega_{1,1}(\sigma) - \omega_{2,0}(\sigma) - \omega_2)}{2\lambda(1-\lambda)n^2} - \frac{(1-2\lambda)(\omega_3 + 3\omega_{2,1}(\sigma) - 3\omega_{1,2}(\sigma))}{6\lambda^2(1-\lambda)^2 n^2}. \end{aligned}$$

Observe that the considerations of Remark 4.1 apply to Theorem 6.1. Thus we find (under the assumptions of Theorem 6.1) that the expected number of induced copies of $H^{[r]}$ in a uniformly random graph with degree sequence $\boldsymbol{d}$ is

$$(1 + O(n^{-1/2+\eta})) \frac{r!}{|\operatorname{Aut}(H^{[r]})|} \binom{n}{r} \lambda^m (1-\lambda)^{\binom{r}{2}-m} \, \mathbb{E}\,[e^{f_{\text{ind}}(X)}], \tag{6.1}$$

where the expectation is taken with respect to a uniformly random element X of S_n.

In the proof of Theorem 1.3 we will use the following bounds given by the power mean inequality:

$$\begin{aligned} \frac{\delta}{\lambda(1-\lambda)n} \sum_{j=1}^{r} |h_j - \lambda(r-1)| &\leqslant r^{2/3} \biggl(\frac{\delta^3}{\lambda^3(1-\lambda)^3 n^3} \sum_{j=1}^{r} |h_j - \lambda(r-1)|^3 \biggr)^{1/3}, \\ \frac{\delta^2}{\lambda^2(1-\lambda)^2 n^2} \sum_{j=1}^{r} |h_j - \lambda(r-1)|^2 &\leqslant r^{1/3} \biggl(\frac{\delta^3}{\lambda^3(1-\lambda)^3 n^3} \sum_{j=1}^{r} |h_j - \lambda(r-1)|^3 \biggr)^{2/3}. \end{aligned} \tag{6.2}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Before proving Theorem 1.3, we apply the results of Section 3 to obtain the following expressions. Define

$$\varepsilon_3(\eta) = \min\left\{\varepsilon_6(\eta), \frac{1}{8}\eta\right\}. \tag{6.3}$$

Lemma 6.1. *Let $\eta \in (0, \frac{1}{2})$ be constant. If assumptions (1.2) and (1.7) hold with $\varepsilon \in (0, \varepsilon_3(\eta)]$, then*

$$\begin{aligned}\mathbb{E} f_{\mathrm{ind}}(X) = &-\frac{\omega_2}{2\lambda(1-\lambda)n} + \frac{r^2}{2n} + \frac{(1-2\lambda)\omega_1}{2\lambda(1-\lambda)n} - \frac{\omega_1^2}{4\lambda(1-\lambda)n^2} - \frac{r^2R}{2\lambda(1-\lambda)n^2}\\ &- \frac{r\,\omega_2}{2\lambda(1-\lambda)n^2} - \frac{(1-2\lambda)\omega_3}{6\lambda^2(1-\lambda)^2n^2} - \frac{(1-2\lambda)R\,\omega_1}{2\lambda^2(1-\lambda)^2n^2} + O(n^{-1/2+\eta}),\\ \operatorname{Var} f_{\mathrm{ind}}(X) = &\frac{R\,\omega_2}{\lambda^2(1-\lambda)^2n^2} - \frac{r\,\omega_1\sum_{j=1}^n (d_j-d)^3}{\lambda^2(1-\lambda)^2n^4} + O(n^{-1/2+\eta}).\end{aligned}$$

Proof. We will often employ the bounds $R \leqslant \delta^2$ and $|h_j - \lambda(r-1)| \leqslant r$.

In order to easily apply Lemma 3.1, we extend the sum defining $\omega_{s,t}(\sigma)$ to n terms by appending zeros:

$$\omega_{s,t}(\sigma) = \sum_{j=1}^{n} u_j v_{\sigma_j}, \tag{6.4}$$

where, for $j = 1, \ldots n$,

$$u_j = \begin{cases}(h_j - \lambda(r-1))^t & \text{if } j \leqslant r,\\ 0 & \text{if } j \geqslant r+1,\end{cases} \quad \text{and} \quad v_j = (d_j - d)^s.$$

When applying Lemma 3.1, it will be convenient to use the identity

$$\sum_{j=1}^{n} (u(j) - \bar{u})(u'(j) - \bar{u}') = \sum_{j=1}^{n} u(j)u'(j) - n\bar{u}\bar{u}'.$$

If $u(j) = q_j^k$ and $u'(j) = q_j^t$ for $j = 1, \ldots, n$, for some sequence $(q_1, \ldots, q_n) \in \mathbb{R}^n$ and $k, t \geqslant 0$, then we can apply the power mean inequality to bound this expression by $O(1)\sum_{j=1}^n |q_j|^{k+t}$.

By Lemma 3.1(i),

$$\mathbb{E}\,[\omega_{s,t}(X)] = \frac{1}{n}\left(\sum_{j=1}^{n} (d_j - d)^s\right)\left(\sum_{j=1}^{r} (h_j - \lambda(r-1))^t\right).$$

This implies that $\mathbb{E}\,[\omega_{1,t}(X)] = 0$ for any $t \geqslant 0$, and that

$$\mathbb{E}\left[-\frac{r\,\omega_{2,0}(X)}{2\lambda(1-\lambda)n^2}\right] = -\frac{r^2R}{2\lambda(1-\lambda)n^2}, \quad \mathbb{E}\left[-\frac{(1-2\lambda)\omega_{2,1}(X)}{2\lambda^2(1-\lambda)^2n^2}\right] = -\frac{(1-2\lambda)R\,\omega_1}{2\lambda^2(1-\lambda)^2n^2}.$$

Finally, applying Lemma 3.1(ii) shows that

$$\begin{aligned}\mathbb{E}\,[\omega_{1,0}(X)^2] &= \operatorname{Var}\,[\omega_{1,0}(X)]\\ &= \frac{1}{n-1}\sum_{j=1}^{n} (d_j - d)^2\left(r\left(1 - \frac{r}{n}\right)^2 + (n-r)\left(\frac{r}{n}\right)^2\right)\\ &= O(Rr),\end{aligned}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

again using the fact that $\mathbb{E}\,[\omega_{1,0}(X)] = 0$ for the first equality. Therefore

$$\mathbb{E}\left[-\frac{\omega_{1,0}(X)^2}{2\lambda(1-\lambda)n^2}\right] = O\left(\frac{Rr}{\lambda(1-\lambda)n^2}\right) = O(n^{-1/2+4\varepsilon}) = O(n^{-1/2+\eta}).$$

Combining the above expressions and estimates leads to the expression for $\mathbb{E}\, f_{\rm ind}(X)$.

Now for the variance. From Lemma 3.1(ii) we have

$$\mathrm{Var}\,[\omega_{1,1}(X)] = \frac{nR}{n-1}\left(\omega_2 - \frac{\omega_1^2}{n}\right) = R\,\omega_2 + O\left(\frac{R(\omega_2+\omega_1^2)}{n}\right). \tag{6.5}$$

Combining (1.7) and (6.2) gives

$$\frac{R\,\omega_1^2}{\lambda^2(1-\lambda)^2n^3} = O\left(\frac{r^{4/3}n^{-1/3+2\eta/3}}{n}\right) = O(n^{-1/2+\eta}),$$

$$\frac{R\,\omega_2}{\lambda^2(1-\lambda)^2n^3} = O\left(\frac{r^{1/3}n^{-1/3+2\eta/3}}{n}\right) = O(n^{-1/2+\eta}).$$

Therefore the first term of $f_{\rm ind}$ has variance

$$\mathrm{Var}\left[\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}\right] = \frac{R\omega_2}{\lambda^2(1-\lambda)^2n^2} + O(n^{-1/2+\eta}).$$

Also

$$\mathrm{Var}\left[\frac{r\,\omega_{1,1}(X)}{\lambda(1-\lambda)n^2}\right] = O\left(\frac{r^2R\,\omega_2}{\lambda^2(1-\lambda)^2\,n^4}\right) = O\left(\frac{\delta^2r^5}{\lambda^2(1-\lambda)^2\,n^4}\right) = O(n^{-1/2+\eta}).$$

Using Lemma 3.1(ii), we have the following rough bound:

$$\mathrm{Var}\,[\omega_{k,t}(X)] = O(\delta^{2k}\,r^{2t+1}). \tag{6.6}$$

Hence

$$\mathrm{Var}\left[\frac{\omega_{1,0}(X)\,\omega_1}{4\lambda(1-\lambda)n^2}\right] = O\left(\frac{r^5\,\delta^2}{\lambda^2(1-\lambda)^2n^4}\right) = O(n^{-1/2+\eta}),$$

$$\mathrm{Var}\left[\frac{r\,\omega_{2,0}(X)}{\lambda(1-\lambda)n^2}\right] = O\left(\frac{r^3\delta^4}{\lambda^2(1-\lambda)^2n^4}\right) = O(n^{-1/2+\eta}),$$

$$\mathrm{Var}\left[\frac{\omega_{2,1}(X)}{\lambda^2(1-\lambda)^2n^2}\right] = O\left(\frac{\delta^4r^3}{\lambda^4(1-\lambda)^4n^4}\right) = O(n^{-1/2+\eta}),$$

$$\mathrm{Var}\left[\frac{\omega_{1,2}(X)}{\lambda^2(1-\lambda)^2n^2}\right] = O\left(\frac{\delta^2r^5}{\lambda^4(1-\lambda)^4n^4}\right) = O(n^{-1/2+\eta}).$$

The final variance that we must calculate is $\mathrm{Var}\,[\omega_{1,0}(X)^2]$. We have

$$\mathrm{Var}\,[\omega_{1,0}(X)^2] \leqslant \mathbb{E}\,[\omega_{1,0}(X)^4] = \frac{1}{n!}\sum_{\sigma\in S_n}\left(\sum_{j=1}^{r}(d_{\sigma_j}-d)\right)^4 - \frac{(r)_4}{(n)_4}\left(\sum_{k=1}^{n}(d_k-d)\right)^4. \tag{6.7}$$

Note that the second term is 0 since $\sum_{k=1}^{n}(d_k-d)=0$. Expanding the right-hand side of (6.7) as

$$\mathbb{E}\left[\sum_{i_1,i_2,i_3,i_4=1}^{n} c(i_1,i_2,i_3,i_4)(d_{i_1}-d)(d_{i_2}-d)(d_{i_3}-d)(d_{i_4}-d)\right],$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

we find that $c(i_1, i_2, i_3, i_4) = 0$ if i_1, i_2, i_3, i_4 are distinct. The part

$$\left(\sum_{j=1}^{r} (d_{\sigma_j} - d)\right)^4$$

has $r^4 - (r)_4 = O(r^3)$ other terms while the part

$$\frac{(r)_4}{(n)_4}\left(\sum_{k=1}^{n} (d_k - d)\right)^4$$

has $n^4 - (n)_4 = O(n^3)$ other terms. Therefore

$$\mathrm{Var}\,[\omega_{1,0}(X)^2] = O(r^3\delta^4 + r^4\delta^4/n) = O(r^3\delta^4),$$

which proves that

$$\mathrm{Var}\left[\frac{\omega_{1,0}(X)^2}{2\lambda(1-\lambda)n^2}\right] = O(n^{-1/2+\eta}).$$

Hence we see that only the first term of $f_{\mathrm{ind}}(X)$ has non-negligible variance. It follows that any covariance which does not involve the first term will automatically fit within the $O(n^{-1/2+\eta})$ error term. We now compute the remaining covariances. Lemma 3.1(ii) implies that

$$\mathrm{Cov}\,(\omega_{j,k}(X), \omega_{s,t}(X)) = O(\delta^{j+s}\, r^{k+t+1}),$$

which shows that

$$\mathrm{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{\omega_{1,0}(X)}{\lambda(1-\lambda)n^3}\right) = O\left(\frac{\delta^2 r^2}{\lambda^2(1-\lambda)^2 n^3}\right) = O(n^{-1/2+\eta}).$$

Using assumption (1.7) and bounds (6.2), (6.5), we find that

$$\begin{aligned}\mathrm{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{r\,\omega_{1,1}(X)}{\lambda(1-\lambda)n^2}\right) &= \frac{r}{\lambda^2(1-\lambda)^2 n^3}\ \mathrm{Var}\,[\omega_{1,1}(X)]\\ &= O\left(\frac{\delta^2 r\omega_2}{\lambda^2(1-\lambda)^2 n^3}\right)\\ &= O\left(\frac{r^{4/3}n^{-1/3+2\eta/3}}{n}\right)\\ &= O(n^{-1/2+\eta}).\end{aligned}$$

Applying Lemma 3.1(ii) and using the same kind of argument, we find that

$$\begin{aligned}\mathrm{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{\omega_{2,1}(X)}{\lambda^2(1-\lambda)^2 n^2}\right) &= O\left(\frac{\delta^3\omega_2}{\lambda^3(1-\lambda)^3 n^3}\right)\\ &= O\left(\frac{\delta r^{1/3}n^{-1/3+2\eta/3}}{\lambda(1-\lambda)n}\right)\\ &= O(n^{-1/2+\eta}),\\ \mathrm{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{\omega_{1,2}(X)}{\lambda^2(1-\lambda)^2 n^2}\right) &= O\left(\frac{\delta^2\sum_{j=1}^{r}|h_j - \lambda(r-1)|^3}{\lambda^3(1-\lambda)^3 n^3}\right)\\ &= O(n^{-1/2+\eta}).\end{aligned}$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

The following term will contribute to the answer, so we calculate it precisely. Writing $\omega_{1,1}(\sigma)) = \sum_{j=1}^{n} u_j v_{\sigma_j}$ using (6.4) with $s=t=1$, we have $\bar{u}=\omega_1/n$ and $\bar{v}=0$. Similarly, write $\omega_{2,0}(\sigma) = \sum_{j=1}^{n} u'_j v'_{\sigma_j}$ using (6.4) with $s=2$ and $t=0$, giving $\bar{u}'=r/n$ and $\bar{v}'=R$. Applying Lemma 3.1(ii) gives

$$\begin{aligned}
&\operatorname{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{-r\,\omega_{2,0}(X)}{2\lambda(1-\lambda)n^2}\right) \\
&= -\frac{r}{2\lambda^2(1-\lambda)^2 n^3(n-1)} \sum_{j=1}^{n} (d_j-d)((d_j-d)^2-R) \\
&\quad \times \left(\sum_{j=1}^{r} \left(h_j - \lambda(r-1) - \frac{\omega_1}{n}\right)\left(1-\frac{r}{n}\right) + \frac{\omega_1\, r(n-r)}{n^2}\right) \\
&= -\frac{r\,\omega_1 \sum_{j=1}^{n} (d_j-d)^3}{2\lambda^2(1-\lambda)^2 n^4} + O(n^{-1/2+\eta}). \qquad (6.8)
\end{aligned}$$

Finally, we need

$$\operatorname{Cov}\left(\frac{\omega_{1,1}(X)}{\lambda(1-\lambda)n}, \frac{\omega_{1,0}(X)^2}{\lambda(1-\lambda)n^2}\right) = \frac{1}{\lambda^2(1-\lambda)^2 n^3} \sum_{j=1}^{r}\sum_{k=1}^{r} \operatorname{Cov}(\hat{E}_{jk}(X), \omega_{1,1}(X)), \qquad (6.9)$$

where $\hat{E}_{jk}(\sigma) = (d_{\sigma_j}-d)(d_{\sigma_k}-d)$. If $j \neq k$ then each covariance in the sum on the right-hand side matches the setting of Lemma 3.1(v) with $u_j=0$ for all j, and $v_j = d_j - d$. Note $\bar{u}=\bar{v}=0$. Write $\Psi'(X) = \omega_{1,1}(X)$ using (6.4) with $s=t=1$, giving $\bar{u}'=\omega_1/n$ and $\bar{v}'=0$. By Lemma 3.1(iv), if $j \neq k$ then

$$\operatorname{Cov}(\hat{E}_{jk}(X), \omega_{1,1}(X)) = O\left(\frac{\delta^3 r}{n}\right)$$

since $u(k)+\bar{v} = u(j)+\bar{v} = 0$ for all j,k. Therefore the terms in (6.9) with $j \neq k$ contribute

$$2\binom{r}{2} O\left(\frac{\delta^3 r}{\lambda^2 n^4}\right) = O(n^{-1/2+\eta}).$$

The terms in (6.9) with $j=k$ contribute

$$\frac{1}{\lambda^2(1-\lambda)^2 n^3} \operatorname{Cov}(\omega_{1,1}(X),\ \omega_{2,0}(X)) = O\left(\frac{\delta^3 \omega_1}{\lambda^2(1-\lambda)^2 n^3}\right) = O(n^{-1/2+\eta}),$$

using the earlier expression for this covariance (6.8) and the bound $\omega_1 = O(r^2)$. Thus we see that (6.9) does not contribute significantly.

Combining all the estimates above (and multiplying (6.8) by 2) gives the stated expression for the variance of $f_{\text{ind}}(X)$. This completes the proof. □

We can now prove our main result about induced subgraphs.

Proof of Theorem 1.3. Define $\varepsilon_3(\eta)$ as in (6.3). We will apply Theorem 2.2 to estimate (6.1). The expected value and variance of f_{ind} are given in Lemma 6.1. It remains to prove that

$$\sum_{j=1}^{n-1} \left(\frac{1}{6}\alpha_j^3 + \frac{1}{3}\alpha_j\beta_j + \frac{5}{8}\alpha_j^4 + \frac{5}{8}\beta_j^2\right) = O(n^{-1/2+\eta}),$$

where $\alpha_j = \alpha_j[f_{\text{ind}}, S_n]$, $\beta_j = \sum_{k=j+1}^{n-1} \alpha_k \Delta_{jk}$ and $\Delta_{jk} = \Delta_{jk}[f_{\text{ind}}, S_n]$.

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Without loss of generality, we can assume that

$$|h_1 - \lambda(r-1)| \geqslant |h_2 - \lambda(r-1)| \geqslant \cdots \geqslant |h_r - \lambda(r-1)|.$$

For any s, t, and $1 \leqslant j < a \leqslant n$, the function $\omega_{s,t}$ satisfies

$$\|D^{(ja)}\omega_{s,t}\| = \begin{cases} O(\delta^s|h_j - \lambda(r-1)|^t) & \text{for } j \leqslant r, \\ 0 & \text{otherwise.} \end{cases}$$

Also, we have

$$\|D^{(ja)}\omega_{1,0}^2\| \leqslant 2\|\omega_{1,0}\|\,\|D^{(ja)}\omega_{1,0}\| = \begin{cases} O(\delta^2 r) & \text{for } j \leqslant r, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\alpha_j = \alpha_j[f_{\text{ind}}, S_n]$. Observe that $\alpha_j = 0$ for $j > r$. By our assumptions, we have

$$r, \delta, h_j = O(n^{1/2+\varepsilon}), \qquad \omega_1 = 2m - \lambda\binom{r}{2} = O(n^{1+2\varepsilon}).$$

Thus, using the above bounds, we find that $\alpha_j = O(\gamma_j)$ for $1 \leqslant j \leqslant r$, where

$$\gamma_j = \frac{\delta|h_j - \lambda(r-1)|}{\lambda(1-\lambda)n} + n^{-1/2+4\varepsilon}.$$

Note that for any s, t, and distinct $1 \leqslant j, k, a, b \leqslant n$ with $j < a$ and $j < k < b$, we have

$$\|D^{(k\,b)}D^{(j\,a)}\omega_{s,t}\| = 0,$$

$$\|D^{(k\,b)}D^{(j\,a)}\omega_{1,0}^2\| = \begin{cases} O(\delta^2) & \text{for } k \leqslant r, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\Delta_{jk} = \Delta_{jk}[f_{\text{ind}}, S_n]$. We have that $\Delta_{jk} = 0$ for $k > r$. Observe also that

$$\|D^{(k\,a)}D^{(j\,a)}f_{\text{ind}}\| \leqslant 2\|D^{(k\,a)}f_{\text{ind}}\| = O(n^{3\varepsilon}),$$

$$\|D^{(k\,b)}D^{(j\,k)}f_{\text{ind}}\| \leqslant 2\|D^{(j\,k)}f_{\text{ind}}\| = O(n^{3\varepsilon}).$$

Thus, using the bounds above, we find that $\Delta_{jk} = O(n^{-1+3\varepsilon})$ for $1 \leqslant j < k \leqslant r$.

Using the inequality $(|x| + |y|)^3 \leqslant 4(|x|^3 + |y|^3)$ for each term of the sum, we find that

$$\begin{aligned} \sum_{j=1}^{n-1} \alpha_j^3 &= O(1)\sum_{j=1}^{r} \gamma_j^3 \\ &= \sum_{j=1}^{r} O\biggl(\biggl(\frac{\delta|h_j - \lambda(r-1)|}{\lambda(1-\lambda)n} + n^{-1/2+4\varepsilon}\biggr)^3\biggr) \\ &= O\biggl(\frac{\delta^3 \sum_{j=1}^{r}|h_j - \lambda(r-1)|^3}{\lambda^3(1-\lambda)^3 n^3} + rn^{-3/2+12\varepsilon}\biggr) \\ &= O(n^{-1/2+\eta}) \end{aligned}$$

by (1.7) and the bound $\varepsilon \leqslant \frac{1}{8}\eta$. Observe that $\beta_j = 0$ for $j > r$ and

$$\beta_j = O(r\alpha_j n^{-1+3\varepsilon}) = O(\gamma_j\, n^{-1/2+4\varepsilon}) \quad \text{for } j \leqslant r.$$

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Using the power mean inequality, we bound

$$\begin{aligned}\sum_{j=1}^{n}\alpha_j\beta_j &= O(n^{-1/2+4\varepsilon})\sum_{j=1}^{r}\gamma_j^2\\ &\leqslant O(n^{-1/2+4\varepsilon})r^{1/3}\left(\sum_{j=1}^{r}\gamma_j^3\right)^{2/3}\\ &= O(n^{-1/3+4\varepsilon+\varepsilon/3})\left(\sum_{j=1}^{r}\gamma_j^3\right)^{2/3}\\ &= O(n^{-1/2+\eta})\end{aligned}$$

as before. The two remaining terms also have negligible contribution:

$$\sum_{j=1}^{n-1}\alpha_j^4 \leqslant \sum_{j=1}^{r}\gamma_j^4 = O(1)\sum_{j=1}^{r}\gamma_j^3 = O(n^{-1/2+\eta}),$$

$$\sum_{j=1}^{n-1}\beta_j^2 = O(n^{-1+8\varepsilon})\sum_{j=1}^{r}\gamma_j^2 = O(n^{-1/2+\eta}).$$

Applying Theorem 2.2 and using Lemma 6.1, we complete the proof. The bound $\Lambda_2 = O(n^{-1/3+4\varepsilon+\eta/3})$ in the theorem statement follows directly from (1.7) and (6.2). □

Proof of Corollary 1.3. To show (1.8), observe that $\omega_t = O(r^{t+1})$. Therefore the assumption $r^2(1+\delta^2/n) = O(\lambda^2(1-\lambda)^2 n^{1/2+\eta})$ implies that

$$\frac{r^2}{2n} + \frac{(1-2\lambda)\omega_1}{2\lambda(1-\lambda)n} - \frac{r^2R}{2\lambda(1-\lambda)n^2} - \frac{(1-2\lambda)R\omega_1}{2\lambda^2(1-\lambda)^2n^2} = O\left(\frac{r^2(1+\delta^2/n)}{\lambda^2(1-\lambda)^2n}\right) = O(n^{-1/2+\eta}),$$

$$\frac{\omega_1^2}{4\lambda(1-\lambda)n^2} + \frac{r\omega_2}{2\lambda(1-\lambda)n^2} + \frac{(1-2\lambda)\omega_3}{6\lambda^2(1-\lambda)^2n^2} = O\left(\frac{r^4}{\lambda^2(1-\lambda)^2n^2}\right) = O(n^{-1/2+\eta}),$$

$$\frac{r\omega_1\sum_{j=1}^{n}(d_j-d)^3}{2\lambda^2(1-\lambda)^2n^4} = O\left(\frac{r^3\delta^3}{\lambda^2(1-\lambda)^2n^3}\right) = O(n^{-1/2+\eta}).$$

For the second statement, observe that the assumption $r = O(n^{1/3-\varepsilon})$ implies that $r^2(1+\delta^2/n) = O(n^{2/3})$ and so

$$\begin{aligned}-\frac{\omega_2}{2\lambda(1-\lambda)n} + \frac{R\omega_2}{2\lambda^2(1-\lambda)^2n^2} &= O\left(\frac{r^3(1+\delta^2/n)}{\lambda^2(1-\lambda)^2n}\right)\\ &= O\left(\frac{r}{n^{1/3}\lambda^2(1-\lambda)^2}\right)\\ &= O\left(\frac{n^{-\varepsilon}}{\lambda^2(1-\lambda)^2}\right)\\ &= O(n^{-\varepsilon/2}).\end{aligned}$$

Applying (1.8) completes the proof. □

https://doi.org/10.1017/S0963548320000498 Published online by Cambridge University Press

Proof of Corollary 1.4. The bound on r in the corollary statement implies that $r = O(\log n)$ and is equivalent to $\lambda_{\min}^r \geqslant n^{-2+\varepsilon}$. The fact that $\mathbb{E}\, Y_n \to \infty$ thus follows from Corollary 1.3. In order to prove the concentration, we use the second moment method in a standard fashion. Define

$$N = \binom{n}{r} \frac{r!}{|\mathrm{Aut}\,(H^{[r]})|},$$

and let $H_1, \ldots, H_N$ be a list of all the potential induced copies of $H^{[r]}$. Let $p_{j,k}$ be the probability that both H_j and H_k occur simultaneously as induced subgraphs, and define

$$E_t = \sum_{\substack{1 \leqslant j,k \leqslant N \\ |V(H_j) \cap V(H_k)| = t}} p_{j,k}, \quad 0 \leqslant t \leqslant r.$$

We know that $\mathbb{E}\, Y_n^2 = \sum_{t=0}^r E_t$ and now we compare $\mathbb{E}\, Y_n^2$ to $(\mathbb{E}\, Y_n)^2$. The probability $p_{j,k}$ is not provided directly by either Theorem 1.1 or Theorem 1.3 but we can infer it from the second part of Corollary 1.3. By summing over all the possible subgraphs induced by $V(H_j) \cup V(H_k)$, we find that E_t asymptotically matches the corresponding expectation for the binomial random graph model $\mathcal{G}(n, \lambda)$ to relative error $O(n^{-\varepsilon/2} + n^{-1/2+\eta})$. Therefore we have

$$\begin{aligned} E_0 &= \binom{n}{r}\binom{n-r}{r}\left(\frac{r!}{|\mathrm{Aut}\,(H^{[r]})|}\right)^2 \lambda^{2m}(1-\lambda)^{r(r-1)-2m}(1 + O(n^{-\varepsilon/2} + n^{-1/2+\eta})) \\ &= (\mathbb{E}\, Y_n)^2(1 + O(n^{-\varepsilon/2} + n^{-1/2+\eta})). \end{aligned}$$

To bound E_t from above for $t \geqslant 1$, we can assume that two induced copies of $H^{[r]}$ always overlap correctly. This gives

$$\begin{aligned} E_t &\leqslant \binom{n}{r}\binom{r}{t}\binom{n-r}{r-t}\left(\frac{r!}{|\mathrm{Aut}\,(H^{[r]})|}\right)^2 \lambda^{2m}(1-\lambda)^{r(r-1)-2m}\lambda_{\min}^{-\binom{t}{2}} \\ &\leqslant (\mathbb{E}\, Y_n)^2 (2n^{-1} r^2 \lambda_{\min}^{-(t-1)/2})^t (1 + o(1)). \end{aligned}$$

Using the condition $\lambda_{\min}^r \geqslant n^{-2+\varepsilon}$, we have that

$$(n^{-1} 2r^2 \lambda_{\min}^{-(t-1)/2})^t = O(n^{-1/2+\eta}) \quad \text{for } t = 1$$

and

$$(n^{-1} 2r^2 \lambda_{\min}^{-(t-1)/2})^t = O(n^{-t\varepsilon/3}) \quad \text{for } 2 \leqslant t \leqslant r.$$

Therefore

$$\mathbb{E}\, Y_n^2 = (\mathbb{E}\, Y_n)^2 (1 + O(n^{-\varepsilon/2} + n^{-1/2+\eta})),$$

which implies that

$$\mathrm{Var}\, Y_n = (\mathbb{E}\, Y_n)^2\, O(n^{-\varepsilon/2} + n^{-1/2+\eta}).$$

The desired result now follows from Chebyshev's inequality. □

References

[1] Barron, E. N., Cardaliaguet, P. and Jensen, R. (2003) Conditional essential suprema with applications. *Appl. Math. Optim.* **48** 229–253.

[2] Erdős, P. and Gallai, T. (1960) Graphs with prescribed degrees of vertices (Hungarian). *Mat. Lapok* **11** 264–274.

[3] Greenhill, C., Isaev, M., Kwan, M. and McKay, B. D. (2017) The average number of spanning trees in sparse graphs with given degrees. *European J. Combin* **63** 6–25.

[4] Isaev, M. and McKay, B. D. (2018) Complex martingales and asymptotic enumeration. *Random Struct. Algorithms* **52** 617–661.
[5] Kim, J. H., Sudakov, B. and Vu, V. (2007) Small subgraphs of random regular graphs. *Discrete Math.* **307** 1961–1967.
[6] Krivelevich, M., Sudakov, B., Vu, V. H. and Wormald, N. C. (2001) Random regular graphs of high degree. *Random Struct. Algorithms* **13** 346–363.
[7] McDiarmid, C. (1998) Concentration. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, Vol. 16 of Algorithms and Combinatorics, pp. 195–248. Springer.
[8] McKay, B. D. (1985) Asymptotics for symmetric 0–1 matrices with prescribed row sums. *Ars Combin.* **19A** 15–26.
[9] McKay, B. D. (2011) Subgraphs of dense random graphs with specified degrees. *Combin. Probab. Comput.* **20** 413–433.
[10] McKay, B. D. and Wormald, N. C. (1990) Asymptotic enumeration by degree sequence of graphs of high degree. *European J. Combin.* **11** 565–580.
[11] Moon, J. W. (1970) *Counting Labelled Trees*, Vol. 1 of Canadian Mathematical Monographs. Canadian Mathematical Congress.
[12] Xiao, L., Yan, G., Wu, Y. and Ren, W. (2008) Induced subgraph in random regular graph. *J. Systems Sci. Complexity* **21** 645–650.

Cite this article: Greenhill C, Isaev M and McKay BD (2021). Subgraph counts for dense random graphs with specified degrees. *Combinatorics, Probability and Computing* **30**, 460–497. https://doi.org/10.1017/S0963548320000498