Edge correlations in weighted random subtrees of the complete graph: a sharp bound and the activity phase diagram

Scott Schubr
Luminary Oddities

20 November 2026

Abstract

Give each nonempty subtree T of the labelled complete graph K_n weight $\lambda^{|E(T)|}$, where $\lambda > 0$ and single vertices are included. For adjacent edges e, f and disjoint edges e_1, e_2 we study the normalized joint probabilities $R_{n,\lambda} = \mathbb{P}(e, f \subseteq T)/(\mathbb{P}(e \subseteq T)\mathbb{P}(f \subseteq T))$ and $R^{\mathrm{disj}}_{n,\lambda}$; on the uniform spanning tree of K_n they equal $3/4$ and 1. An exact fixed-forest reduction expresses both as ratios of moments of the vertex count K. Theorem A proves $R_{n,\lambda} > 3/4$ for every $n \geq 3$ and every $\lambda > 0$, with an explicit polynomial positivity certificate; the constant is sharp for each fixed n. This settles the finite lower bound conjectured in the author manuscript of Luo, Xu and Tian, for all activities. Theorem B gives the activity phase diagram at $\lambda = c/n$: $R_{n,\lambda}/n \to 1 - T(c)/2$ for $c < 1/e$, where $Te^{-T} = c$; $R_{n,\lambda}/n \to 1/2$ at $c = 1/e$; and $R_{n,\lambda} \to 3ec/(4(ec-1))$ for $c > 1/e$. Adjacent edges are therefore positively correlated for $c < 4/e$ and negatively correlated for $c > 4/e$. At fixed λ, $R_{n,\lambda} = \frac{3}{4}[1 + t/n + (t^2 + 3t/2)/n^2] + o(n^{-2})$ with $t = 1/(e\lambda)$. For disjoint edges we obtain the limits in all four regimes, using a local limit theorem for K above criticality, and the resulting sign pattern with thresholds $c_0 \approx 0.17607$, $1/e$ and $4/(3e)$. Theorem C describes the near-critical crossover $\lambda = e^{-1+s/\sqrt{n}}/n$ uniformly for $s \to \infty$ with $s = o(n^{1/6})$, including two crossover windows, and proves the exact-saddle form $R_{n,\lambda} \sim \frac{3}{4} n/(n - 1/(e\lambda))$ uniformly over all activities above the mass window. Using surplus estimates of Liu and Slade, Theorem D transfers the near-critical asymptotics to random connected subgraphs, for every fixed prescribed forest.

Keywords: random subtrees; connected subgraphs; complete graph; edge correlations; phase diagram; Cayley's formula; local limit theorem.
MSC 2020: 05C80, 05C05, 05C30, 60C05, 60F05.

1 Introduction

Let $\mathcal{T}_n$ be the set of nonempty subtrees of the labelled complete graph K_n: connected acyclic subgraphs, not necessarily induced or spanning, with the n single-vertex trees included. For an activity $\lambda > 0$ put

$$Z_n(\lambda) = \sum_{T \in \mathcal{T}_n} \lambda^{|E(T)|}, \qquad \mathbb{P}_\lambda(T) = \frac{\lambda^{|E(T)|}}{Z_n(\lambda)}.$$

The choice $\lambda = 1$ is the uniform random subtree, and $\lambda \to \infty$ at fixed n gives the uniform spanning tree. Such size-weighted subtree laws are Boltzmann models in the sense of the survey of Fredes

and Marckert [3]. For distinct edges e, f of K_n we study the normalized joint probability

$$R_{n,\lambda} := \frac{\mathbb{P}_\lambda(e, f \subseteq T)}{\mathbb{P}_\lambda(e \subseteq T)\mathbb{P}_\lambda(f \subseteq T)} \quad (e, f \text{ adjacent}), \qquad R^{\mathrm{disj}}_{n,\lambda} \quad (e, f \text{ disjoint}).$$

By symmetry these depend only on n, λ and the type of the pair. The ratio exceeds one exactly when the edges are positively correlated, and $\mathbb{P}_\lambda(e \subseteq T \mid f \subseteq T) = R_{n,\lambda}\mathbb{P}_\lambda(e \subseteq T)$.

Background. Chin, Gordon, MacPhee and Vincent [1] conjectured that, for the uniform random subtree of K_n and adjacent edges, $\mathbb{P}(e \subseteq T \mid f \subseteq T) \le \mathbb{P}(e \subseteq T)$. Luo, Xu and Tian [5] proved this inequality and showed that the ratio of the two sides tends to $3/4$. The strict finite lower bound $\mathbb{P}(e \subseteq T \mid f \subseteq T) > \frac{3}{4}\mathbb{P}(e \subseteq T)$ for every $n \ge 3$ was conjectured in [6, author manuscript, Conjecture 5.1]. The value $3/4$ is the uniform-spanning-tree value on K_n: a uniform spanning tree contains a fixed edge with probability $2/n$ and two fixed adjacent edges with probability $3/n^2$, while two fixed disjoint edges appear with probability exactly $4/n^2 = (2/n)^2$, so disjoint edges are independent there (Remark 2.3). Chin et al. [2] and Wagner [10] studied the probability that a random subtree of K_n is spanning, and Wagner and Wang [11] proved Poisson laws for the number of missing vertices in dense graphs. In the sparse regime $\lambda \asymp 1/n$, Liu and Slade [4] established the finite-size critical scaling profile for trees and connected subgraphs of the complete graph at fixed scaling parameter.

Results. This article merges and extends two earlier Archive papers of the author, [8] and [9]. Section 2 gives the exact fixed-forest reduction: conditional on its vertex count K, the tree is a uniform spanning tree of a uniform vertex set, so every edge observable is a ratio of size moments. Theorem A (Section 3) proves $R_{n,\lambda} > 3/4$ for all $n \ge 3$ and $\lambda > 0$ through an explicit positivity certificate. Theorem B (Section 4) gives the phase diagram in the activity, for adjacent and disjoint edges. Theorem C (Section 5) treats the near-critical crossover and proves the exact-saddle form uniformly above the mass window. Theorem D (Section 6) transfers the near-critical results to random connected subgraphs. Appendix A isolates the input from Liu and Slade, and Appendix B lists the exact and high-precision checks shipped with the article.

Table 1 summarizes the adjacent and disjoint ratios. Here $T(c) \in (0, 1]$ is the smaller solution of $Te^{-T} = c$ for $0 < c \le 1/e$, and $u = 1/(ec)$ for $c > 1/e$.

activity	adjacent edges	disjoint edges
$\lambda = c/n$, $c < 1/e$	$R/n \to 1 - T/2$	$R^{\mathrm{disj}} \to 2T(2-T)/(1-T)$
$\lambda = 1/(en)$	$R/n \to 1/2$	$R^{\mathrm{disj}}/n^{1/4} \to \Gamma(\frac{1}{4})/(2^{1/4}\sqrt{\pi})$
$\lambda = e^{-1+s/\sqrt{n}}/n$, $1 \ll s \ll n^{1/6}$	Theorem C, two windows	Corollary 5.11
above the mass window	$R \sim \frac{3}{4}n/(n - 1/(e\lambda))$ uniformly (Theorem C(b))	
$\lambda = c/n$, $c > 1/e$	$R \to 3/(4(1-u))$	$n(R^{\mathrm{disj}} - 1) \to -u(3-4u)/(1-u)^2$
λ fixed, $t = 1/(e\lambda)$	$R = \frac{3}{4}[1 + \frac{t}{n} + \frac{t^2+3t/2}{n^2}] + o(n^{-2})$	$R^{\mathrm{disj}} = 1 - 3t/n^2 + o(n^{-2})$
$\lambda \to \infty$, n fixed	$R \to 3/4$	$R^{\mathrm{disj}} \to 1$

Table 1: Adjacent and disjoint edge ratios across the activity range.

The phase diagram explains the two classical values. The uniform law $\lambda = 1$ lies deep in the dense regime, where the upper bound $R \le 1$ of [5] holds and R tends to $3/4$ from above. For sparse activities $\lambda = c/n$ with $c < 4/e$, adjacent edges are positively correlated, and below the

critical point $1/(en)$ the ratio is of order n. The constant $3/4$ is thus a spanning-tree effect, and Theorem A shows that it is a lower bound throughout the activity range. The techniques are classical: Cayley counts, tree generating functions, Laplace's method on a log-concave profile, and dominated convergence. The contribution is sharpness of the constant $3/4$ for all activities, together with the limits of both ratios at every fixed sparse and dense activity and the exact-saddle form throughout the supercritical range.

2 The exact fixed-forest reduction

Write $K = |V(T)|$ and, for $1 \le k \le n$,

$$w_{n,1} = n, \qquad w_{n,k} = \binom{n}{k} k^{k-2} \quad (2 \le k \le n), \qquad Z_n(\lambda) = \sum_{k=1}^{n} w_{n,k} \lambda^{k-1}. \tag{1}$$

By Cayley's formula $w_{n,k}$ is the number of subtrees with k vertices; the formula $\binom{n}{k}k^{k-2}$ also gives $w_{n,1} = n$. Hence $\mathbb{P}_\lambda(K = k) = w_{n,k}\lambda^{k-1}/Z_n(\lambda)$. Conditional on $K = k$, the vertex set is uniform among k-sets, and the tree is a uniform spanning tree of that set, since the weight depends only on k.

Lemma 2.1 (Forest extensions)**.** *Let F be a spanning forest of K_k whose components have orders $\sigma_1, \dots, \sigma_j$ (isolated vertices count as components of order one). The number of spanning trees of K_k containing F is $k^{j-2}\prod_i \sigma_i$.*

Proof. Contract each component to a vertex. A spanning tree containing F corresponds to a tree on the j contracted vertices together with a choice of endpoint labels for each of its edges; a contracted vertex of degree d_i contributes $\sigma_i^{d_i}$ choices. The weighted Cayley identity, obtained from the Prüfer code, sums $\prod_i \sigma_i^{d_i}$ over trees on j labelled vertices to $(\sum_i \sigma_i)^{j-2}\prod_i \sigma_i = k^{j-2}\prod_i \sigma_i$. For $j = 1$ the forest is already a spanning tree and the formula gives one. □

For a fixed forest F in K_n without isolated vertices, let r be its number of vertices, m its number of edges, $d = r - m$ its number of components, and c_F the product of its component orders. Write $(x)_r = x(x-1)\cdots(x-r+1)$.

Proposition 2.2 (Exact reduction)**.** *For every $n \ge r$, $\lambda > 0$ and $1 \le k \le n$,*

$$\mathbb{P}_\lambda(F \subseteq T \mid K = k) = \frac{c_F}{(n)_r}\frac{(k)_r}{k^m}, \tag{2}$$

which is zero for $k < r$. In particular, for a fixed edge e, adjacent edges e, f and disjoint edges e_1, e_2,

$$\mathbb{P}_\lambda(e \subseteq T \mid K = k) = \frac{2(k-1)}{n(n-1)}, \tag{3}$$

$$\mathbb{P}_\lambda(e, f \subseteq T \mid K = k) = \frac{3h(k)}{(n)_3}, \qquad h(k) = \frac{(k-1)(k-2)}{k} = k - 3 + \frac{2}{k}, \tag{4}$$

$$\mathbb{P}_\lambda(e_1, e_2 \subseteq T \mid K = k) = \frac{4g(k)}{(n)_4}, \qquad g(k) = \frac{(k)_4}{k^2} = k^2 - 6k + 11 - \frac{6}{k}. \tag{5}$$

Proof. The selected k-set contains the r prescribed vertices with probability $(k)_r/(n)_r$. Given this, Lemma 2.1 applied to F together with the other $k-r$ vertices as singletons ($j = k-m$ components) shows that a uniform spanning tree of the k-set contains F with probability $k^{k-m-2}c_F/k^{k-2} = c_F/k^m$. The three special cases have $(r, m, c_F) = (2, 1, 2)$, $(3, 2, 3)$ and $(4, 2, 4)$. □

Averaging over K gives the exact moment identities

$$R_{n,\lambda} = \frac{3n(n-1)}{4(n-2)} \frac{\mathbb{E}_\lambda h(K)}{(\mathbb{E}_\lambda(K-1))^2} = \frac{3}{4} \frac{n(n-1)}{n-2} \frac{\mathbb{E}_\lambda K - 3 + 2\mathbb{E}_\lambda(1/K)}{(\mathbb{E}_\lambda K - 1)^2}, \tag{6}$$

$$R^{\mathrm{disj}}_{n,\lambda} = \frac{n(n-1)}{(n-2)(n-3)} \frac{\mathbb{E}_\lambda g(K)}{(\mathbb{E}_\lambda(K-1))^2} = \Pi_n \Big[1 + \frac{\mathrm{Var}_\lambda\, K - 4\mathbb{E}_\lambda K + 10 - 6\mathbb{E}_\lambda(1/K)}{(\mathbb{E}_\lambda K - 1)^2}\Big], \tag{7}$$

with $\Pi_n = n(n-1)/((n-2)(n-3))$ and $n \geq 4$ in (7). The second form of (7) uses $\mathbb{E}g(K) - (\mathbb{E}K - 1)^2 = \mathrm{Var}\, K - 4\mathbb{E}K + 10 - 6\mathbb{E}(1/K)$. In polynomial form, with

$$N = \sum_{k=1}^{n} w_{n,k} \lambda^{k-1}, \quad B = \sum_{k=1}^{n} (k-1) w_{n,k} \lambda^{k-1}, \quad A = \sum_{k=1}^{n} h(k) w_{n,k} \lambda^{k-1},$$

we have $R_{n,\lambda} = \frac{3n(n-1)}{4(n-2)} NA/B^2$.

Remark 2.3 (Uniform spanning tree values). At $K = n$, (3)–(5) give $\mathbb{P}(e) = 2/n$, $\mathbb{P}(e,f) = 3/n^2$ and $\mathbb{P}(e_1, e_2) = 4/n^2$ exactly. On the uniform spanning tree of K_n adjacent edges therefore have ratio $3/4$, and disjoint edges are independent, with ratio 1. For fixed n, the top-degree term of $Z_n(\lambda)$ comes from spanning trees, so $K \to n$ in probability as $\lambda \to \infty$ and both ratios converge to these values.

Size weights at $\lambda = c/n$. For $\lambda = c/n$ define the normalized size weights

$$t_{n,k} := \frac{w_{n,k} \lambda^{k-1}}{n} = \frac{c^{k-1} k^{k-2}}{k!} \prod_{j=0}^{k-1} \Big(1 - \frac{j}{n}\Big), \qquad 1 \leq k \leq n, \tag{8}$$

so that $t_{n,1} = 1$ and $\mathbb{P}(K = k) = t_{n,k} / \sum_j t_{n,j}$. Their successive ratios are

$$\frac{t_{n,k+1}}{t_{n,k}} = c\Big(1 - \frac{k}{n}\Big)\Big(1 + \frac{1}{k}\Big)^{k-2}, \qquad 1 \leq k \leq n-1. \tag{9}$$

Put $\eta = \log(ec) \in \mathbb{R}$ and

$$\Psi(k) = \eta(k-1) + \sum_{j=0}^{k-1} \log\Big(1 - \frac{j}{n}\Big), \qquad 1 \leq k \leq n. \tag{10}$$

Robbins' form of Stirling's formula [7], $k! = \sqrt{2\pi k}\, k^k e^{-k} e^{\theta_k/(12k)}$ with $\theta_k \in (0,1)$, gives for every $k \geq 1$

$$t_{n,k} = \kappa_k\, k^{-5/2} e^{\Psi(k)}, \qquad \kappa_k = \frac{e}{\sqrt{2\pi}} e^{-\theta_k/(12k)} \in \Big[\frac{e^{11/12}}{\sqrt{2\pi}}, \frac{e}{\sqrt{2\pi}}\Big]. \tag{11}$$

The increments $\Psi(k+1) - \Psi(k) = \eta + \log(1 - k/n)$ decrease in k, and the increments at $k-1$ and k differ by $-\log(1 - 1/(n-k+1)) \geq 1/n$. Thus Ψ is concave with curvature at least $1/n$. If $\eta > 0$, the increment is nonnegative exactly for $k \leq k_*$, where

$$k_* = k_*(\lambda) = n(1 - e^{-\eta}) = n - \frac{1}{e\lambda}. \tag{12}$$

Let k° be a maximizer of Ψ on $\{1, \ldots, n\}$; then $|k^\circ - k_*| \leq 1$, and concavity gives, for every k,

$$\Psi(k) \leq \Psi(k^\circ) - \frac{|k - k^\circ|(|k - k^\circ| - 1)}{2n}. \tag{13}$$

Finally $\log \prod_{j<k}(1 - j/n) \leq -k(k-1)/(2n)$ for all $k \leq n$.

Tree series. The rooted and unrooted labelled-tree series

$$T(z)=\sum_{k\ge1}\frac{k^{k-1}}{k!}z^k=ze^{T(z)},\qquad U(z)=\sum_{k\ge1}\frac{k^{k-2}}{k!}z^k=T-\frac{T^2}{2}$$

satisfy, on $0\le z\le 1/e$ and with $T=T(z)\in[0,1]$,

$$V(z):=\sum_{k\ge1}\frac{k^{k-3}}{k!}z^k=T-\frac{3}{4}T^2+\frac{1}{6}T^3,\qquad W(z):=\sum_{k\ge1}\frac{k^k}{k!}z^k=\frac{T}{1-T}\quad(z<1/e).\tag{14}$$

Indeed $dz/z=(1-T)\,dT/T$, so integrating T/z and then U/z gives U and V, while $W=zT'(z)=T/(1-T)$ by differentiating $T=ze^T$. At $z=1/e$ one has $T=1$, $U=1/2$ and $V=5/12$ by Abel's theorem.

3 Theorem A: the sharp lower bound

Theorem A. *For every $n\ge3$, every $\lambda>0$ and distinct adjacent edges e,f of K_n,*

$$\mathbb{P}_\lambda(e\subseteq T\mid f\subseteq T)>\frac{3}{4}\,\mathbb{P}_\lambda(e\subseteq T).$$

More precisely, the polynomial $P_n(\lambda):=n(n-1)NA-(n-2)B^2$ satisfies

$$P_n(\lambda)=\lambda^2\sum_{m=0}^{2n-4}\pi_{n,m}\lambda^m,\qquad \pi_{n,m}\ge0,\qquad \pi_{n,0}=\frac{n^2(n-1)^2(n-2)(4n-3)}{12}>0,\tag{15}$$

and $R_{n,\lambda}-\frac{3}{4}=\dfrac{3P_n(\lambda)}{4(n-2)B(\lambda)^2}$. For each fixed n, $R_{n,\lambda}\to3/4$ as $\lambda\to\infty$, so $3/4$ is the largest constant valid for all $\lambda>0$, already for each fixed n.

The representation (15) is a positivity certificate: a polynomial with nonnegative coefficients and a positive lowest coefficient, divided by the positive quantity B^2. The size of the excess $R_{n,\lambda}-3/4$ is given by Theorem B: it is $3/(4e\lambda n)+O(n^{-2})$ at fixed λ, and of order one or larger when $\lambda=O(1/n)$.

Proof. The identity for $R_{n,\lambda}-3/4$ is (6) in polynomial form. Symmetrize

$$P_n(\lambda)=\frac{1}{2}\sum_{i,j=1}^{n}w_{n,i}w_{n,j}Q_{ij}\lambda^{i+j-2},\qquad Q_{ij}=n(n-1)[h(i)+h(j)]-2(n-2)(i-1)(j-1).$$

Let $n\ge4$. We check that $Q_{ij}\ge0$ for every pair other than $(2,2)$. If $i=1$ or $j=1$, this follows from $h(1)=h(2)=0$ and $h(k)\ge0$ for integers $k\ge1$. If $i=2$ and $j\ge3$, then

$$Q_{2j}=(j-1)\Big[\frac{n(n-1)(j-2)}{j}-2(n-2)\Big]\ge\frac{(j-1)(n-3)(n-4)}{3}\ge0,$$

because $(j-2)/j$ increases in j. Symmetry treats $j=2<i$. If $i,j\ge3$, put $m=(i+j)/2\in[3,n]$. Since $h''(x)=4/x^3>0$ and $(i-1)(j-1)\le(m-1)^2$,

$$Q_{ij}\ge2\big[n(n-1)h(m)-(n-2)(m-1)^2\big]=\frac{2(m-1)}{m}(n-2)(n-m)\Big(m-1-\frac{n}{n-2}\Big)\ge0,$$

where the last step uses $n/(n-2) \le 2$. The only negative entry is $Q_{22} = -2(n-2)$, at total degree two.

Degrees zero and one carry only the pairs $(1,1)$, $(1,2)$, $(2,1)$, and $Q_{11} = Q_{12} = 0$; hence P_n is divisible by λ^2. Every coefficient of degree at least three is a sum of terms $w_{n,i}w_{n,j}Q_{ij}/2$ with $(i,j) \neq (2,2)$ and is nonnegative. At degree two the pairs are $(1,3)$, $(3,1)$ and $(2,2)$, and, with $w_{n,3} = 3\binom{n}{3}$ and $h(3) = 2/3$,

$$\begin{aligned}\pi_{n,0} &= n(n-1)w_{n,1}w_{n,3}h(3) - (n-2)w_{n,2}^2 \\ &= \frac{n^3(n-1)^2(n-2)}{3} - \frac{n^2(n-1)^2(n-2)}{4} = \frac{n^2(n-1)^2(n-2)(4n-3)}{12}.\end{aligned}$$

For $n = 3$, direct substitution gives $N = 3 + 3\lambda + 3\lambda^2$, $B = 3\lambda + 6\lambda^2$, $A = 2\lambda^2$ and $P_3(\lambda) = 27\lambda^2$, which is (15); equivalently $R_{3,\lambda} = 3(1 + \lambda + \lambda^2)/(1 + 2\lambda)^2$. Since $B > 0$ for $\lambda > 0$, $R_{n,\lambda} > 3/4$. Sharpness is Remark 2.3. □

4 Theorem B: the activity phase diagram

Throughout this section $T(c)$ denotes the smaller root of $Te^{-T} = c$ for $0 < c \le 1/e$, and $u = 1/(ec) \in (0,1)$ for $c > 1/e$. Put

$$\Phi(c) = \frac{3ec}{4(ec-1)} = \frac{3}{4(1-u)} \qquad (c > 1/e).$$

Theorem B. *Let e, f be adjacent and e_1, e_2 disjoint edges of K_n.*

(i) (Adjacent edges at $\lambda = c/n$.) *For fixed $c > 0$, as $n \to \infty$,*

$$\frac{R_{n,c/n}}{n} \to 1 - \frac{T(c)}{2} \quad (c < 1/e), \qquad \frac{R_{n,1/(en)}}{n} \to \frac{1}{2}, \qquad R_{n,c/n} \to \Phi(c) \quad (c > 1/e).$$

(ii) (Sign change.) *If $0 < c < 4/e$, then $R_{n,c/n} > 1$ for all large n, so e and f are positively correlated. If $c > 4/e$, then $3/4 < R_{n,c/n} < 1$ for all large n, so they are negatively correlated.*

(iii) (Fixed activity.) *For fixed $\lambda > 0$ and $t = 1/(e\lambda)$,*

$$R_{n,\lambda} = \frac{3}{4}\Big[1 + \frac{t}{n} + \frac{t^2 + 3t/2}{n^2}\Big] + o(n^{-2}).$$

(iv) (Disjoint edges.) *For fixed $\lambda > 0$, $R^{\mathrm{disj}}_{n,\lambda} = 1 - 3t/n^2 + o(n^{-2})$. For fixed $c > 0$,*

$$R^{\mathrm{disj}}_{n,c/n} \to \frac{2T(2-T)}{1-T} \quad (c < 1/e), \qquad \frac{R^{\mathrm{disj}}_{n,1/(en)}}{n^{1/4}} \to \frac{\Gamma(1/4)}{2^{1/4}\sqrt{\pi}},$$

$$n(R^{\mathrm{disj}}_{n,c/n} - 1) \to -\frac{u(3-4u)}{(1-u)^2} \quad (c > 1/e).$$

Corollary 4.1 (Sign pattern for disjoint edges)**.** *Let $c_0 = T_0 e^{-T_0} = 0.1760677\ldots$ with $T_0 = (5 - \sqrt{17})/4$. For fixed c and all large n, disjoint edges are negatively correlated if $c < c_0$, positively correlated if $c_0 < c < 4/(3e)$, and negatively correlated if $c > 4/(3e)$; they are negatively correlated at every fixed λ.*

Proof. For $c < 1/e$, $2T(2-T)/(1-T) = 1$ is $2T^2 - 5T + 1 = 0$, whose root in $(0,1)$ is T_0; the limit increases in T, which increases in c. At $c = 1/e$ the ratio tends to infinity. For $c > 1/e$ the limit $-u(3-4u)/(1-u)^2$ is positive for $3/4 < u < 1$, that is $1/e < c < 4/(3e)$, and negative for $u < 3/4$. □

Remark 4.2. The regimes match. As $c \uparrow 1/e$, $1 - T(c)/2 \to 1/2$; as $c \downarrow 1/e$, $\Phi(c) \to \infty$, consistent with order-n growth; and as $c \to \infty$, $\Phi(c) = \frac{3}{4}(1 + u + u^2 + \cdots)$ with $u = t/n$ when $c = \lambda n$, which agrees with (iii) to first order and with the t^2 term at second order. The additional $3t/(2n^2)$ in (iii) is a finite-c correction. Part (ii) shows that the upper bound $R \le 1$ of [5] is a dense-activity phenomenon, valid on the side $c > 4/e$ of the sign change.

4.1 Subcritical and critical activities

Lemma 4.3. *Fix $0 < c \le 1/e$. If $c < 1/e$, then for every function ϕ with $|\phi(k)| \le C(1+k)^m$,*

$$\sum_{k=1}^{n} \phi(k) t_{n,k} \to \sum_{k \ge 1} \phi(k) \frac{c^{k-1} k^{k-2}}{k!}.$$

If $c = 1/e$, the same holds for $|\phi(k)| \le Ck$.

Proof. Each $t_{n,k}$ increases to $c^{k-1}k^{k-2}/k!$ as $n \to \infty$, and by (11) the limit is at most $(e/\sqrt{2\pi})k^{-5/2}(ec)^{k-1}$. Dominated convergence applies: the majorant times $(1+k)^m$ is summable if $ec < 1$, and times k is summable if $ec = 1$. □

Proof of Theorem B(i) *for $c \le 1/e$ and* (iv) *for $c < 1/e$.* By (14) and Lemma 4.3, with $T = T(c)$,

$$\sum_k t_{n,k} \to \frac{U}{c} = \frac{T - T^2/2}{c}, \qquad \sum_k (k-1) t_{n,k} \to \frac{T-U}{c} = \frac{T^2}{2c},$$
$$\sum_k h(k) t_{n,k} \to \frac{T - 3U + 2V}{c} = \frac{T^3}{3c}, \qquad \sum_k g(k) t_{n,k} \to \frac{W - 6T + 11U - 6V}{c} = \frac{T^4}{c(1-T)},$$

the last for $c < 1/e$; it uses $W = T + T^2 + T^3 + T^4/(1-T)$. Since $(n-1)/(n-2) \to 1$ and $\Pi_n \to 1$, (6) and (7) give

$$\frac{R_{n,c/n}}{n} \to \frac{3}{4} \cdot \frac{(U/c)(T^3/(3c))}{T^4/(4c^2)} = \frac{U}{T} = 1 - \frac{T}{2}, \qquad R^{\mathrm{disj}}_{n,c/n} \to \frac{(U/c)\, T^4/(c(1-T))}{T^4/(4c^2)} = \frac{2T(2-T)}{1-T}.$$

At $c = 1/e$, the first three limits are valid with $T = 1$ (the functions 1, $k-1$ and $h(k)$ are at most k), giving $e/2$, $e/2$ and $e/3$, so $R_{n,1/(en)}/n \to \frac{3}{4} \cdot \frac{(e/2)(e/3)}{(e/2)^2} = \frac{1}{2}$. □

Lemma 4.4. *At $c = 1/e$, $n^{-1/4} \sum_{k=1}^{n} k^2 t_{n,k} \to \dfrac{e}{\sqrt{2\pi}} \displaystyle\int_0^\infty y^{-1/2} e^{-y^2/2}\, dy = \frac{e\,\Gamma(1/4)}{2^{3/4}\sqrt{2\pi}}$.*

Proof. Here $\eta = 0$ and $k^2 t_{n,k} = \kappa_k k^{-1/2} \exp\{\sum_{j<k} \log(1 - j/n)\}$ by (11). Fix $0 < \epsilon < M$. For $k < \epsilon\sqrt{n}$, the sum is at most $(e/\sqrt{2\pi}) \sum_{k < \epsilon\sqrt{n}} k^{-1/2} \le C\epsilon^{1/2} n^{1/4}$. For $k > M\sqrt{n}$, the bound $\exp\{-k(k-1)/(2n)\}$ gives at most $C(M\sqrt{n})^{-1/2} \sum_{k > M\sqrt{n}} e^{-(k-1)^2/(2n)} \le Cn^{1/4} \int_{M-1}^\infty e^{-y^2/2} dy$ for large n. For $\epsilon\sqrt{n} \le k \le M\sqrt{n}$, the inequality $|\log(1-x) + x| \le x^2$ for $0 \le x \le 1/2$ gives $\sum_{j<k} \log(1 - j/n) = -k^2/(2n) + O(M/\sqrt{n} + M^3/\sqrt{n})$, and $\kappa_k = (e/\sqrt{2\pi})(1 + O(1/(\epsilon\sqrt{n})))$. With $y = k/\sqrt{n}$, the middle sum is $n^{1/4}$ times a Riemann sum of mesh $n^{-1/2}$ for $(e/\sqrt{2\pi}) \int_\epsilon^M y^{-1/2} e^{-y^2/2} dy$, up to a factor $1 + o(1)$. Let $n \to \infty$, then $\epsilon \to 0$ and $M \to \infty$. Finally $\int_0^\infty y^{-1/2} e^{-y^2/2} dy = 2^{-3/4}\Gamma(1/4)$ by the substitution $y^2 = 2v$. □

Proof of Theorem B(iv) *at $c = 1/e$.* Since $g(k) = k^2 - 6k + 11 - 6/k$ and $\sum_k k t_{n,k} = O(1)$ by Lemma 4.3, Lemma 4.4 gives $\sum_k g(k) t_{n,k} \sim n^{1/4} e\Gamma(1/4)/(2^{3/4}\sqrt{2\pi})$. With $\sum_k t_{n,k} \to e/2$ and $\sum_k (k-1) t_{n,k} \to e/2$,

$$\frac{R^{\mathrm{disj}}_{n,1/(en)}}{n^{1/4}} \to \frac{e\Gamma(1/4)}{2^{3/4}\sqrt{2\pi}} \cdot \frac{e/2}{(e/2)^2} = \frac{2^{1/4}\Gamma(1/4)}{\sqrt{2\pi}} = \frac{\Gamma(1/4)}{2^{1/4}\sqrt{\pi}}. \qquad \square$$

4.2 Supercritical activities: a local limit theorem

Proposition 4.5 (Local limit theorem for K)**.** *Fix $c > 1/e$, let $u = 1/(ec)$ and $k_* = n(1-u)$. For every fixed $M > 0$,*

$$\sup_{|k-k_*| \le M\sqrt{n}} \Big| \sqrt{2\pi n u}\, \mathbb{P}_{c/n}(K = k)\, e^{(k-k_*)^2/(2nu)} - 1 \Big| \longrightarrow 0.$$

Moreover $\mathbb{E}(K - k_*)/\sqrt{n} \to 0$, $\mathbb{E}(K-k_*)^2/n \to u$, hence $\operatorname{Var} K = un(1+o(1))$; also $K/n \to 1-u$ in probability and $\mathbb{E}(1/K) = O(1/n)$.

Proof. Here $\eta = \log(ec) = -\log u > 0$ is fixed, and k_* is (12). Let k° maximize Ψ, so $|k^\circ - k_*| \le 1$.

Local shape. For an integer i with $|i - k_*| \le (M+2)\sqrt{n}$, the increment is $\eta + \log(1 - i/n) = \log(1 - (i-k_*)/(nu)) = -(i-k_*)/(nu) + O(M^2/n)$ uniformly. Summing from k° to $k-1$ (or from k to $k^\circ - 1$) for $|k - k_*| \le M\sqrt{n}$,

$$\Psi(k) - \Psi(k^\circ) = -\frac{(k-k_*)^2 - (k^\circ - k_*)^2}{2nu} + O\Big(\frac{M}{\sqrt{n}}\Big) + O\Big(\frac{M^3}{\sqrt{n}}\Big) = -\frac{(k-k_*)^2}{2nu} + o(1).$$

On this range $k, k^\circ \ge n(1-u)/2$, so $(k^\circ/k)^{5/2} = 1 + O(M/\sqrt{n})$ and $\kappa_k/\kappa_{k^\circ} = 1 + O(1/n)$. By (11),

$$\frac{t_{n,k}}{t_{n,k^\circ}} = e^{-(k-k_*)^2/(2nu)}(1+o(1)) \qquad \text{uniformly for } |k - k_*| \le M\sqrt{n}. \tag{16}$$

Tails. By (11) and (13), with $m = |k - k^\circ|$,

$$\frac{t_{n,k}}{t_{n,k^\circ}} \le e^{1/12}\Big(\frac{k^\circ}{k}\Big)^{5/2} e^{-m(m-1)/(2n)} \le \begin{cases} Ce^{-m(m-1)/(2n)}, & k \ge k_*/2, \\ Cn^{5/2} e^{-(k_*/2-2)^2/(2n)}, & k < k_*/2. \end{cases}$$

The second line is at most $e^{-n(1-u)^2/9}$ for large n.

Summation. Let ϕ be continuous with $|\phi(z)| \le 1 + z^2$. By (16), the sum of $\phi((k - k_*)/\sqrt{n}) t_{n,k}/t_{n,k^\circ}$ over $|k - k_*| \le M\sqrt{n}$ is $\sqrt{n}\big(\int_{-M}^{M} \phi(z) e^{-z^2/(2u)} dz + o(1)\big)$, a Riemann sum of mesh $n^{-1/2}$. Since $m(m-1)/(2n) \ge z^2/2 - |z|$ for $m = |z|\sqrt{n}$ and large n, the sum over $k \ge k_*/2$, $|k - k_*| > M\sqrt{n}$, is at most $C\sqrt{n}\int_{|z|>M-1} (1+z^2) e^{-z^2/2+|z|} dz$. The sum over $k < k_*/2$ is at most $n(1+n)e^{-n(1-u)^2/9} = o(\sqrt{n})$. Letting $M \to \infty$,

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \phi\Big(\frac{k-k_*}{\sqrt{n}}\Big) \frac{t_{n,k}}{t_{n,k^\circ}} \to \int_{\mathbb{R}} \phi(z) e^{-z^2/(2u)} dz.$$

Taking $\phi = 1$ gives $\sum_k t_{n,k} = t_{n,k^\circ}\sqrt{2\pi n u}(1+o(1))$, which with (16) proves the local limit theorem. Taking $\phi(z) = z$ and $\phi(z) = z^2$ gives the two moment limits, and $\operatorname{Var} K/n = \mathbb{E}(K-k_*)^2/n - (\mathbb{E}(K - k_*)/\sqrt{n})^2 \to u$. Chebyshev's inequality gives $K/n \to 1-u$. Finally $\mathbb{E}(1/K) \le 2/k_* + \mathbb{P}(K < k_*/2) = O(1/n)$. $\square$

Proof of Theorem B(i) *for* $c > 1/e$, (ii), *and* (iv) *for* $c > 1/e$. By Proposition 4.5, $\mathbb{E}K = n(1-u) + o(\sqrt{n})$ and $\mathbb{E}(1/K) = O(1/n)$, so (6) gives $R_{n,c/n} \to \frac{3}{4} \cdot \frac{n \cdot n(1-u)}{n^2(1-u)^2} = \Phi(c)$.

For (ii): if $c \le 1/e$, $R_{n,c/n}/n$ has a positive limit, so $R_{n,c/n} \to \infty$. If $1/e < c < 4/e$, then $\Phi(c) > 1$, since $\Phi(c) = 1$ exactly when $ec = 4$ and Φ decreases in c. If $c > 4/e$, then $3/4 < \Phi(c) < 1$. In each case $R_{n,c/n} - 1$ eventually has the sign of $4/e - c$; the inequality $R_{n,c/n} > 3/4$ is Theorem A.

For (iv) with $c > 1/e$, use the second form of (7). Here $n(\Pi_n - 1) \to 4$, $\Pi_n \to 1$, and by Proposition 4.5

$$n \cdot \frac{\operatorname{Var} K - 4\mathbb{E}K + 10 - 6\mathbb{E}(1/K)}{(\mathbb{E}K - 1)^2} \to \frac{u - 4(1-u)}{(1-u)^2}.$$

Hence $n(R^{\mathrm{disj}}_{n,c/n} - 1) \to 4 + \dfrac{5u-4}{(1-u)^2} = \dfrac{4u^2 - 3u}{(1-u)^2} = -\dfrac{u(3-4u)}{(1-u)^2}$. □

4.3 Fixed activity: the Poisson deficiency to second order

Fix $\lambda > 0$, let $t = 1/(e\lambda)$ and $D = n - K$. Relative to the spanning-tree weight put

$$\rho_{n,j} = \frac{w_{n,n-j}\lambda^{n-j-1}}{w_{n,n}\lambda^{n-1}} = \frac{\lambda^{-j}}{j!}\prod_{r=0}^{j-1}\Big(1 - \frac{r}{n}\Big)\Big(1 - \frac{j}{n}\Big)^{n-j-2}, \qquad 0 \le j \le n-1,$$

and $\rho_{n,j} = 0$ for $j \ge n$, so that $\mathbb{P}_\lambda(D = j) = \rho_{n,j} / \sum_i \rho_{n,i}$.

Lemma 4.6. *For every function* ϕ *with* $|\phi(j)| \le C(1+j)^m$,

$$\sum_{j \ge 0} \phi(j)\rho_{n,j} = \sum_{j \ge 0} \phi(j)\frac{t^j}{j!}\Big(1 + \frac{5j}{2n}\Big) + o(n^{-1}).$$

Consequently $\mathbb{E}_\lambda D = t + 5t/(2n) + o(n^{-1})$, $\mathbb{E}_\lambda D^2 \to t + t^2$, $\operatorname{Var}_\lambda D \to t$, *and* $\mathbb{E}_\lambda(1/K) = 1/n + O(n^{-2})$.

Proof. For $3 \le k \le n$, $w_{n,k-1}/w_{n,k} = (1-1/k)^{k-3}/(n-k+1) \le 1/(n-k+1)$, and $w_{n,1}/w_{n,2} = 2/(n-1)$. Multiplying these ratios for $k = n, n-1, \dots, n-j+1$ gives $\rho_{n,j} \le 2\lambda^{-j}/j!$ for all j. Hence the parts of both sums with $j > n^{1/4}$ are $O(\sum_{j > n^{1/4}} (1+j)^{m+1} \max(\lambda^{-1}, t)^j / j!)$, which is smaller than every power of n.

For $j \le n^{1/4}$, write $\rho_{n,j} = (t^j/j!) f_n(j)$ with $f_n(j) = e^j \prod_{r<j}(1 - r/n)(1 - j/n)^{n-j-2}$. Using $\log(1-x) = -x - x^2/2 + O(x^3)$,

$$\sum_{r<j} \log\Big(1 - \frac{r}{n}\Big) = -\frac{j(j-1)}{2n} + O\Big(\frac{j^3}{n^2}\Big), \qquad (n-j-2)\log\Big(1 - \frac{j}{n}\Big) = -j + \frac{j^2}{2n} + \frac{2j}{n} + O\Big(\frac{j^3}{n^2}\Big),$$

so $\log f_n(j) = 5j/(2n) + O(j^3/n^2)$ and $f_n(j) = 1 + 5j/(2n) + O(j^3/n^2)$ uniformly for $1 \le j \le n^{1/4}$, with $f_n(0) = 1$. The resulting error is $O(n^{-2} \sum_j (1+j)^{m+3} t^j / j!) = O(n^{-2})$.

For the consequences, $\sum_j t^j/j! = e^t$, $\sum_j j t^j/j! = te^t$ and $\sum_j j^2 t^j/j! = (t+t^2)e^t$ give $\mathbb{E}D = (t + \frac{5}{2n}(t+t^2))/(1 + \frac{5t}{2n}) + o(n^{-1}) = t + \frac{5t}{2n} + o(n^{-1})$, and similarly $\mathbb{E}D^2 \to t + t^2$. Finally $1/(n-D) - 1/n = D/(n(n-D)) \le 2D/n^2$ on $\{D \le n/2\}$, while $\mathbb{P}(D > n/2)$ is smaller than every power of n; hence $\mathbb{E}(1/K) = 1/n + O(n^{-2})$. □

Proof of Theorem B(iii) *and* (iv) *at fixed* λ. Let $x = 1/n$, $d = \mathbb{E}D$ and $\varepsilon = \mathbb{E}(1/K)$. By (6),

$$\frac{4}{3}R_{n,\lambda} = \frac{1-x}{1-2x} \cdot \frac{1 - (3+d)x + 2\varepsilon x}{(1-(1+d)x)^2}.$$

Insert $d = t + \delta x + o(x)$ with $\delta = 5t/2$ and $\varepsilon x = x^2 + O(x^3)$ from Lemma 4.6. The first two factors multiply to $1 - (2+t)x + (1 - t - \delta)x^2 + o(x^2)$, and $(1-(1+d)x)^{-2} = 1 + 2(1+t)x + (2\delta + 3(1+t)^2)x^2 + o(x^2)$. The product is $1 + tx + (t^2 - t + \delta)x^2 + o(x^2) = 1 + tx + (t^2 + \frac{3}{2}t)x^2 + o(x^2)$, which is (iii).

For disjoint edges, use the second form of (7), with $\operatorname{Var} K = \operatorname{Var} D \to t$ and $\mathbb{E}K = n - d$, $d \to t$:

$$\begin{aligned}\frac{\operatorname{Var} K - 4\mathbb{E}K + 10 - 6\varepsilon}{(\mathbb{E}K - 1)^2} &= (-4n + \operatorname{Var} D + 4d + 10 + o(1))x^2(1 + 2(1+d)x + O(x^2)) \\ &= -4x + (2 - 3t)x^2 + o(x^2).\end{aligned}$$

Since $\Pi_n = 1 + 4x + 14x^2 + O(x^3)$, the product is $1 + (14 - 16 + 2 - 3t)x^2 + o(x^2) = 1 - 3tx^2 + o(x^2)$. □

5 Theorem C: the near-critical crossover

For $s > 0$ set

$$\eta = \frac{s}{\sqrt{n}}, \quad c = e^{-1+\eta}, \quad \lambda = \frac{c}{n}, \quad \hat{k} = \sqrt{n}\, s, \qquad q = 2n^{-1/4}s^{-3/2}e^{s^2/2}, \quad p = \frac{q}{\hat{k}}. \tag{17}$$

Here $\hat{k} = n\eta$ is the linearization of the saddle $k_*(\lambda) = n(1 - e^{-\eta})$ of (12); thresholds involving it are rounded when used for integer sums. Write $H_r = e\sqrt{n}\,\hat{k}^{r-5/2}e^{s^2/2}$, so that $H_0 = (e/2)p$ and $H_1 = (e/2)q$. In the parametrization $\lambda = p_{\mathrm{LS}}/(en)$ of Liu and Slade [4], $p_{\mathrm{LS}} = e^{s/\sqrt{n}}$, and $\sqrt{n}(p_{\mathrm{LS}} - 1) = s + O(s^2/\sqrt{n})$; their scaling theorems hold at fixed s.

Theorem C. (a) (Uniform composite law.) *Let $a_n \to \infty$ and $b_n = o(n^{1/6})$, with $a_n \le b_n$ eventually. Uniformly for $a_n \le s \le b_n$, with* (17)*,*

$$\frac{Z_n(\lambda)}{n} = \frac{e}{2}(1+p)(1+o(1)), \tag{18}$$

$$\frac{B}{n} = \frac{e}{2}(1+q)(1+o(1)), \tag{19}$$

$$\frac{A}{n} = \frac{e}{2}\Big(\frac{2}{3} + q\Big)(1+o(1)), \tag{20}$$

and consequently

$$\frac{R_{n,\lambda}}{n} = \frac{3}{4}\frac{(1+p)(2/3+q)}{(1+q)^2}(1+o(1)). \tag{21}$$

Every error is relative and uniform on the interval; p and q may be unbounded.

(b) (Exact-saddle form above the mass window.) *Let $\sigma_n \to \infty$ with $p(\sigma_n) \to \infty$, equivalently $\sigma_n^2 - \frac{3}{2}\log n - \frac{5}{2}\log\log n \to \infty$ when $\sigma_n^2 = O(\log n)$. Then*

$$\sup_{\lambda \ge e^{-1+\sigma_n/\sqrt{n}}/n} \Big| \frac{4k_*(\lambda)}{3n} R_{n,\lambda} - 1 \Big| \longrightarrow 0, \qquad k_*(\lambda) = n - \frac{1}{e\lambda}.$$

Part (b) covers every activity above the mass window, from $s \to \infty$ through fixed $c > 1/e$ and fixed λ to $\lambda \to \infty$. It is consistent with Theorem B: for fixed $c > 1/e$, $3n/(4k_*) = \Phi(c)$, and for fixed λ, $3n/(4k_*) = \frac{3}{4}(1 + t/n + t^2/n^2 + \cdots)$.

5.1 The composite law

The convergent critical sums are, by (14) at $z = 1/e$,

$$\sum_{k\ge1} \frac{e^{1-k}k^{k-2}}{k!} = \frac{e}{2}, \quad \sum_{k\ge1}(k-1)\frac{e^{1-k}k^{k-2}}{k!} = \frac{e}{2}, \quad \sum_{k\ge1} h(k)\frac{e^{1-k}k^{k-2}}{k!} = \frac{e}{3}, \tag{22}$$

and the corresponding sum with factor k is e.

Lemma 5.1 (Three-range decomposition)**.** *Under the hypotheses of Theorem C*(a)*, call $k \le 1/\eta$ the low range, $1/\eta < k < \hat{k}/2$ the intermediate range and $k \ge \hat{k}/2$ the high range. Uniformly for $a_n \le s \le b_n$:*

1. *for every fixed function ϕ with $0 \le \phi(k) \le k$, the low-range sum of $\phi(k)t_{n,k}$ equals $\sum_{k\ge1}\phi(k)e^{1-k}k^{k-2}/k! + o(1)$;*
2. *for $r = 0, 1$, the intermediate-range sum of $k^r t_{n,k}$ is $o(1 + H_r)$, and $\sum_{k=1}^{n} k^r t_{n,k} = O(1 + H_r)$;*
3. *for $r = 0, 1$, the high-range sum of $k^r t_{n,k}$ is $H_r(1 + o(1))$, and for every bounded continuous ϕ on $\mathbb{R}$ and every choice of $s = s_n \in [a_n, b_n]$,*

$$\frac{\sum_{k\ge\hat{k}/2} t_{n,k}\,\phi((k-\hat{k})/\sqrt{n})}{\sum_{k\ge\hat{k}/2} t_{n,k}} \longrightarrow \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} \phi(z)e^{-z^2/2}\,dz.$$

Proof of Lemma 5.1 and Theorem C(a). For $r = 0, 1$ put $S_r = \sum_{k=1}^{n} k^r t_{n,k}$. We prove

$$S_0 = \frac{e}{2} + H_0 + o(1 + H_0), \qquad S_1 = e + H_1 + o(1 + H_1), \tag{23}$$

and keep track of the low and high pieces for the other two moments. All estimates below are uniform for $a_n \le s \le b_n$; in particular $\eta \to 0$, $s \to \infty$ and $\hat{k}^3/n^2 \le b_n^3/\sqrt{n} \to 0$.

For the low range $k \le 1/\eta$, the factor $e^{\eta k}$ is at most e and the product in (8) is at most one, so (11) dominates $k^r t_{n,k}$ by $Ck^{r-5/2}$, which is summable. At each fixed k the weight tends uniformly to $e^{1-k}k^{k-2}/k!$, and the cutoff tends uniformly to infinity. Dominated summation shows that the low range contributes its full critical sum, with uniform additive error $o(1)$. The same holds with any fixed factor ϕ satisfying $0 \le \phi(k) \le k$, in particular with $k - 1$ and $h(k)$; this is part (i).

Next consider $1/\eta < k < \hat{k}/2$. By (11) and $\log\prod_{j<k}(1 - j/n) \le -k(k-1)/(2n)$,

$$k^r t_{n,k} \le Ck^{r-5/2}\exp\Big\{\eta k - \frac{k^2}{2n} + \frac{k}{2n}\Big\}. \tag{24}$$

The last term in the exponent is at most $1/2$. On $[1/\eta, 5/\eta]$ the sum is $O(\eta^{3/2-r}) = o(1)$. On $[5/\eta, \hat{k}/2]$ the function $F_r(k) = k^{r-5/2}\exp\{\eta k - k^2/(2n)\}$ is increasing for large n: its logarithmic derivative times k is the concave quadratic $\eta k - k^2/n - (5/2 - r)$, positive at both endpoints for large s. The sum is therefore at most the number of terms times the value at the upper end:

$$\sum_{1/\eta<k<\hat{k}/2} k^r t_{n,k} \le o(1) + C\hat{k}^{r-3/2}e^{3s^2/8} = o(1) + O\big(H_r s e^{-s^2/8}\big) = o(1 + H_r).$$

For the high range $k \ge \hat{k}/2$, let $z = (k - \hat{k})/\sqrt{n}$. By (11), $t_{n,k} = (e/\sqrt{2\pi})(1 + O(\eta) + O(k^{-1}))k^{-5/2}e^{\eta k}\prod_{j<k}(1 - j/n)$. On every bounded z-interval the logarithm of the product is $-k^2/(2n) + o(1)$ uniformly, the remainder being $O(k/n + k^3/n^2) = o(1)$; also $k/\hat{k} = 1 + z/s \to 1$

uniformly. Completing the square gives $\eta k - k^2/(2n) = s^2/2 - z^2/2$. After division of $k^r t_{n,k}$ by $\hat{k}^{r-5/2} e^{s^2/2}$, the bound (24) and $k/\hat{k} \geq 1/2$ give the majorant $Ce^{-z^2/2}$ for $r = 0, 1$. On bounded z-intervals the normalized summand tends uniformly to $(e/\sqrt{2\pi})e^{-z^2/2}$. The mesh is $1/\sqrt{n}$, the lower boundary $-s/2$ tends uniformly to $-\infty$, and the upper boundary tends uniformly to $+\infty$. Gaussian tail bounds remove $|z| > M$ uniformly; local Riemann sums followed by $M \to \infty$ give

$$\sum_{k \geq \hat{k}/2} k^r t_{n,k} = H_r(1 + o(1)), \qquad r = 0, 1,$$

a relative estimate even when $H_r \to 0$. Inserting the bounded factor $\phi(z)$ and dividing by the case $\phi = 1$ proves part (iii); part (ii) is the intermediate bound together with the other two ranges.

The three ranges prove (23). In the high range, replacing k by $k-1$ changes the sum by its zeroth mass, and replacing it by $h(k)$ changes it by at most three times that mass; since $H_0/H_1 = 1/\hat{k} \to 0$ these changes are $o(H_1)$. The intermediate ranges are bounded by the $r = 1$ estimate, and the low ranges have the constants (22). Hence

$$\frac{B}{n} = \frac{e}{2} + H_1 + o(1 + H_1), \qquad \frac{A}{n} = \frac{e}{3} + H_1 + o(1 + H_1).$$

With $H_0 = (e/2)p$ and $H_1 = (e/2)q$, the positive background constants convert these additive estimates into the relative forms (18)–(20). Substitution into (6), with $(n-1)/(n-2) \to 1$, proves (21). □

Remark 5.2. The first and second size moments S_1 and S_2 are the one-point function and the susceptibility of Liu and Slade [4]. The high-range term H_2 coincides with $n^{1/4}$ times the large-s form $e\, s^{-1/2} e^{s^2/2}$ of their profile, while $H_1 = H_2/\hat{k}$ is a correction of order $n^{-1/4}$ at fixed s, below the resolution of their one-point limit e, which matches the background of S_1. Their parameter $s_{\mathrm{LS}} = \sqrt{n}(e^{s/\sqrt{n}} - 1)$ satisfies $s_{\mathrm{LS}}^2 - s^2 = O(s^3/\sqrt{n}) = o(1)$ in the present range. The increments here are uniformity as $s \to \infty$ with $s = o(n^{1/6})$, and the unrooted mass S_0, whose crossover is delayed by the factor $\hat{k}$.

5.2 The two windows and conditional Gaussian fluctuations

Corollary 5.3 (First-moment window)**.** *If* $s^2 = \frac{1}{2}\log n + \frac{3}{2}\log\log n + x_n$ *with* $x_n \to x$, *and* $Q = 2^{7/4} e^{x/2}$, *then* $q \to Q$, $p \to 0$ *and*

$$\frac{R_{n,\lambda}}{n} \longrightarrow \frac{3}{4}\frac{2/3 + Q}{(1+Q)^2}.$$

Moreover K *converges weakly to the probability law*

$$\pi_k = \frac{2e^{-k} k^{k-2}}{k!}, \qquad k \geq 1, \tag{25}$$

the Borel law with parameter one reweighted by $1/k$, *whose mean is* 2, *whereas* $\mathbb{E}K \to 2 + Q$.

Proof. Substitution into (17) gives $q \to Q$; the powers of $\log n$ cancel because $s \sim \sqrt{(\log n)/2}$. Also $p = q/\hat{k} \to 0$, and the ratio limit follows from Theorem C(a). For each fixed k, dividing (8) by $Z_n/n \to e/2$ gives (25). Its probabilities sum to one by $U(1/e) = 1/2$, so pointwise convergence implies weak convergence on the positive integers. Its mean is $2T(1/e) = 2$. Finally (19) and (18) give $\mathbb{E}(K-1) \to 1 + Q$. For each fixed integer M, $\lim_n \mathbb{E}[K \mathbf{1}_{\{K>M\}}] = 2 + Q - \sum_{k \leq M} k\pi_k$; letting $M \to \infty$ leaves $Q > 0$, so the size law fails to be uniformly integrable. A vanishing-probability high-size contribution carries the extra mean. □

Corollary 5.4 (Mass window)**.** *If* $s^2 = \frac{3}{2}\log n + \frac{5}{2}\log\log n + y_n$ *with* $y_n \to y$*, and* $P = 2(3/2)^{-5/4}e^{y/2}$*, then* $p \to P$*,* $q \to \infty$*, and*

$$\frac{\hat{k}}{n}R_{n,\lambda} \longrightarrow \frac{3}{4}\frac{1+P}{P}, \qquad \frac{K}{\hat{k}} \;\Rightarrow\; \frac{1}{1+P}\delta_0 + \frac{P}{1+P}\delta_1.$$

Proof. Direct substitution gives $p \to P$; here $s \sim \sqrt{(3\log n)/2}$ and $q = \hat{k}p \to \infty$. Multiplying (21) by $\hat{k}$ gives the ratio limit. By Lemma 5.1 the low zeroth mass tends to $e/2$, the intermediate mass is $o(1) + o(H_0) = o(1)$, and the high mass tends to $eP/2$. For bounded continuous f on $[0,\infty)$, the low range contributes $(e/2)f(0)+o(1)$ to the unnormalized sum with test factor $f(k/\hat{k})$: dominated summation reduces it to fixed k, where $k/\hat{k} \to 0$, and on the low range $k/\hat{k} \le 1/s^2$. In the high range $k/\hat{k} = 1 + z/s \to 1$ uniformly on bounded z-intervals, and z is tight under the high-range weights by Lemma 5.1(iii), so the high range contributes $(eP/2)f(1)+o(1)$. The intermediate part vanishes by boundedness of f. Dividing by the total mass $e(1+P)/2 + o(1)$ proves the mixture law. □

Corollary 5.5 (Conditional Gaussian fluctuations)**.** *For every choice of* $s = s_n \in [a_n, b_n]$ *in Theorem C*(a)*, conditional on* $K \ge \hat{k}/2$*,*

$$\frac{K - \hat{k}}{\sqrt{n}} \;\Rightarrow\; \mathcal{N}(0,1),$$

whether or not the conditioning event has positive limiting probability.

Proof. The conditional law of K given $K \ge \hat{k}/2$ has weights proportional to $t_{n,k}$ on the high range, so this is Lemma 5.1(iii), whose estimate is relative. □

Both windows lie in $s \to \infty$, $s = o(n^{1/6})$. Their distinct locations reflect the factor $\hat{k}$ between the saddle's first moment and its mass. Between them, where $q \to \infty$ and $p \to 0$, Theorem C(a) gives $R_{n,\lambda} \sim \frac{3}{4}n/q$.

5.3 The exact-saddle form

Lemma 5.6 (Concentration at the saddle)**.** *Fix* $\epsilon \in (0, 1/2)$*. Uniformly over all* λ *with* $k_*(\lambda) \ge n^{9/14}/2$*, the window* $W = \{k : |k - k_*| \le \epsilon k_*\}$ *satisfies*

$$\frac{\sum_{k\notin W} k t_{n,k}}{\sum_{k\in W} t_{n,k}} \le Cn^{7/2}\exp\Big\{-\frac{(\epsilon k_* - 2)^2}{2n}\Big\} \longrightarrow 0.$$

Proof. Write $\lambda = c/n$ with $\eta = \log(ec) > 0$; the representation (11) and (13) hold for every such c. Let k° maximize Ψ; since $|k^\circ - k_*| \le 1$ and $\epsilon k_* \ge 1$ for large n, $k^\circ \in W$, and $\sum_{k\in W} t_{n,k} \ge t_{n,k^\circ} \ge (e^{11/12}/\sqrt{2\pi})(k_*+1)^{-5/2}e^{\Psi(k^\circ)}$. For $k \notin W$, $|k - k^\circ| \ge \epsilon k_* - 1$, so (11) and (13) give $kt_{n,k} \le (e/\sqrt{2\pi})\, n\, k^{-5/2}e^{\Psi(k^\circ)}e^{-(\epsilon k_* - 2)^2/(2n)}$. Summing $k^{-5/2}$ over $k \ge 1$ and using $k_* + 1 \le 2n$ gives the bound. Since $k_*^2/n \ge n^{2/7}/4$, the right side tends to zero uniformly. □

Proof of Theorem C(b). Split the range of $s = \sqrt{n}\log(e\lambda n)$ into $\sigma_n \le s \le n^{1/7}$ and $s \ge n^{1/7}$.

Range $s \ge n^{1/7}$. Here $\eta = s/\sqrt{n} \ge n^{-5/14}$. Since $(1 - e^{-\eta})/\eta$ decreases and equals $1 - e^{-1}$ at $\eta = 1$, $k_* = n(1 - e^{-\eta}) \ge (1 - e^{-1})\min(s\sqrt{n}, n) \ge n^{9/14}/2$. Fix ϵ and let δ_n be the bound of Lemma 5.6. Then $\sum_k t_{n,k} \le (1+\delta_n)\sum_W t_{n,k}$ and $\sum_k kt_{n,k} \le (1+\delta_n)\sum_W kt_{n,k}$. On W,

$(1-\epsilon)k_* \le k \le (1+\epsilon)k_*$, and for integers $k \ge 1$ one has $k-3 \le h(k) \le k-1$ and $h(k) \ge 0$. Therefore

$$\frac{(1-\epsilon)k_*-3}{1+\delta_n} \le \mathbb{E}h(K) \le \mathbb{E}(K-1) \le (1+\epsilon)(1+\delta_n)k_*, \qquad \mathbb{E}(K-1) \ge \frac{(1-\epsilon)k_*-1}{1+\delta_n}.$$

Inserting these bounds in $\frac{4k_*}{3n}R_{n,\lambda} = \frac{n-1}{n-2}\,\frac{k_*\mathbb{E}h(K)}{(\mathbb{E}(K-1))^2}$ shows that, for all large n and uniformly in this range, $|4k_*R_{n,\lambda}/(3n)-1| \le C\epsilon$. As ϵ is arbitrary, the supremum over this range tends to zero.

Range $\sigma_n \le s \le n^{1/7}$. Apply Theorem C(a) with $a_n = \sigma_n$ and $b_n = n^{1/7}$. Since $\frac{d}{ds}\log p = s - \frac{5}{2s} > 0$ for $s > 2$, $p \ge p(\sigma_n) \to \infty$ uniformly, and $q = \hat{k}p \to \infty$ as well. Hence

$$\frac{(1+p)(2/3+q)}{(1+q)^2} = \frac{1}{\hat{k}}\Big(1+\frac{1}{p}\Big)\frac{(2/3+q)\,q}{(1+q)^2} = \frac{1+o(1)}{\hat{k}}$$

uniformly, and (21) gives $R_{n,\lambda} = \frac{3}{4}(n/\hat{k})(1+o(1))$. Finally $k_*/\hat{k} = (1-e^{-\eta})/\eta \in [1-\eta/2, 1]$ with $\eta \le n^{-5/14}$. The two ranges together prove (b). When $\sigma_n^2 = O(\log n)$, the identity $\log p = \frac{s^2}{2} - \frac{3}{4}\log n - \frac{5}{2}\log s + \log 2$ shows that $p(\sigma_n) \to \infty$ is equivalent to $\sigma_n^2 - \frac{3}{2}\log n - \frac{5}{2}\log\log n \to \infty$. □

Remark 5.7 (Numerical accuracy). Above the mass window the exact-saddle form is the more accurate one, increasingly so as s grows. At $n = 10^8$, the ratio $4k_*R_{n,\lambda}/(3n)$ equals 1.0433, 1.0181, 1.0063 and 1.0016 at $s = 8, 12, 20, 40$, while the ratio of $R_{n,\lambda}$ to the composite form (21) equals 1.0437, 1.0187, 1.0074 and 1.0036; at $n = 10^6$ and $s = 300$ the two ratios are 1.00002 and 1.158 (Appendix B).

5.4 Higher moments and fixed forest markings

Lemma 5.8 (Higher size moments). *Under the hypotheses of Theorem C*(a)*, for each fixed integer* $d \ge 2$*,*

$$S_d := \sum_{k=1}^{n} k^d t_{n,k} = H_d(1+o(1)), \qquad H_d = e\sqrt{n}\,\hat{k}^{d-5/2}e^{s^2/2} = \hat{k}^d H_0,$$

uniformly for $a_n \le s \le b_n$*.*

Proof. The bound (24) holds with r replaced by d. For $k \le 5/\eta$ summation gives $O(\eta^{3/2-d})$, since $d - 5/2 > -1$; relative to H_d this is $O(s^{4-2d}e^{-s^2/2}) = o(1)$. On $[5/\eta, \hat{k}/2]$ the function $k^{d-5/2}\exp\{\eta k - k^2/(2n)\}$ is increasing: for $d = 2$ its logarithmic derivative times k is a concave quadratic positive at both endpoints for large s, and for $d \ge 3$ its logarithmic derivative is at least $\eta/2 > 0$. The intermediate sum is at most $C\hat{k}^{d-3/2}e^{3s^2/8}$, whose ratio to H_d is $O(se^{-s^2/8}) = o(1)$. In the high range, the bounded-z expansion of the previous proof is unchanged, and the normalized summand tends to $(e/\sqrt{2\pi})e^{-z^2/2}$ on bounded z-intervals. For $z > 0$, $k/\hat{k} = 1 + z/s \le 1 + z$ eventually, and for $z < 0$ in this range $1/2 \le k/\hat{k} \le 1$, so $C(1+|z|^d)e^{-z^2/2}$ is a uniform integrable majorant. Riemann sums show that the high sum is $H_d(1+o(1))$. □

Corollary 5.9 (Fixed disconnected forests). *For every fixed forest* F *with* $d \ge 2$ *components and no isolated vertices, uniformly in the range of Theorem C*(a)*,*

$$\mathbb{P}_\lambda(F \subseteq T) = \frac{c_F}{n^r}\frac{H_d}{e/2+H_0}(1+o(1)) = \frac{c_F\hat{k}^d}{n^r}\frac{p}{1+p}(1+o(1)).$$

Proof. Set $\gamma_F(k) = (k)_r/k^m$ for $k \geq r$ and $\gamma_F(k) = 0$ otherwise. For fixed F, $|\gamma_F(k) - k^d| \leq C_F k^{d-1}$ for all $k \geq 1$. If $d \geq 3$, Lemma 5.8 makes the error in the weighted sum $O(H_{d-1}) = o(H_d)$. If $d = 2$, Lemma 5.1(ii) gives the error $O(1 + H_1) = o(H_2)$, since $H_1/H_2 = 1/\hat{k} \to 0$ and $H_2 = en^{1/4}s^{-1/2}e^{s^2/2} \to \infty$ uniformly. Average (2), use $(n)_r = n^r(1 + o(1))$ and (18), and use $H_d = \hat{k}^d H_0$ and $H_0 = (e/2)p$. □

Corollary 5.10 (Fixed connected tree markings)**.** *Let F be a fixed tree on $r \geq 2$ prescribed vertices, and set $\gamma_r(k) = (k)_r/k^{r-1}$ for $k \geq r$, $\gamma_r(k) = 0$ for $k < r$. Then $\Lambda_r := \sum_{k \geq r} e^{1-k}k^{k-2}\gamma_r(k)/k! = e/r$, and uniformly in the range of Theorem C*(a)*,*

$$\mathbb{P}_\lambda(F \subseteq T) = \frac{1}{n^r}\frac{2 + rq}{1 + p}(1 + o(1)).$$

In particular $\Lambda_2 = e/2$ and $\Lambda_3 = e/3$.

Proof. Since $0 \leq \gamma_r(k) \leq k$, the series for Λ_r converges. Putting $k = j + r$ and cancelling the falling factorial gives $\Lambda_r = e^{1-r}\sum_{j \geq 0}(j + r)^{j-1}e^{-j}/j!$. Lagrange inversion for $T = ze^T$ gives $e^{rT(z)} = \sum_{j \geq 0} r(j + r)^{j-1}z^j/j!$; letting $z \uparrow 1/e$ with $T(1/e) = 1$ and nonnegative coefficients gives $\Lambda_r = e^{1-r}e^r/r = e/r$. Lemma 5.1(i) with factor γ_r gives $\Lambda_r + o(1)$ for the low range, and part (ii) bounds the intermediate range by $o(1 + H_1)$. Since $|\gamma_r(k) - k| \leq C_r$ for all $k \geq 1$, the high sum differs from the high first moment by $O(H_0) = o(H_1)$, so the total is $\Lambda_r + H_1 + o(1 + H_1)$. Apply (2) with $m = r - 1$ and $c_F = r$, and divide by (18). The identities $\gamma_2(k) = k - 1$ and $\gamma_3(k) = h(k)$ give the explicit constants (22). □

Corollary 5.11 (Disjoint edges near criticality)**.** *Uniformly in the range of Theorem C*(a)*,*

$$R^{\mathrm{disj}}_{n,\lambda} = \frac{\hat{k}q(1 + p)}{(1 + q)^2}(1 + o(1)).$$

At the first-moment window $R^{\mathrm{disj}}_{n,\lambda}/\hat{k} \to Q/(1 + Q)^2$, and at the mass window $R^{\mathrm{disj}}_{n,\lambda} \to (1 + P)/P$.

Proof. Apply Corollary 5.9 with $r = 4$, $d = 2$, $c_F = 4$, and Corollary 5.10 with $r = 2$, and use $\hat{k}^2 p = \hat{k}q$. □

6 Theorem D: transfer to connected subgraphs

Give every nonempty connected simple labelled subgraph G of K_n weight $\lambda^{|E(G)|}$, retaining the singleton graphs. Write N^{c} for its partition function and $\mathbb{P}_{\mathrm{c}}$ for the resulting probability measure; the superscript t denotes the tree model of the previous sections. The size variable is again $K = |V(G)|$. In this section η, λ and $\hat{k}$ are as in (17) and $a_n \leq s \leq b_n$ with $a_n \to \infty$, $b_n = o(n^{1/6})$. Every asymptotic statement is uniform over this interval, whose endpoints may depend on n.

6.1 Surplus enumeration and size tails

Let $C_k(\lambda)$ be the total weight of connected graphs on one fixed labelled k-set, and put

$$t_k(\lambda) = k^{k-2}\lambda^{k-1}, \qquad t_1(\lambda) = 1, \qquad D_k(\lambda) = \frac{C_k(\lambda)}{t_k(\lambda)} - 1.$$

For $k = 1, 2$, $D_k = 0$. The connected-subgraph surplus enumeration of Liu and Slade enters through the following lemma, whose derivation is recorded in Appendix A.

Lemma 6.1 (Surplus bound)**.** *There is a constant C such that, for $k \geq 3$, $z > 0$, $zk^{3/2} \leq 1$ and $kz \leq 3/e$,*

$$0 \leq D_k(z) \leq Czk^{3/2} + Cz^2.$$

Let $t_{n,k} = n^{-1}\binom{n}{k} t_k(\lambda)$ be the normalized tree size weight, and define the connected size weight

$$a_{n,k} = \frac{1}{n}\binom{n}{k} C_k(\lambda) = t_{n,k}(1 + D_k(\lambda)). \tag{26}$$

Both singleton weights equal one.

Lemma 6.2 (Local surplus and global tails)**.** *For each fixed $L > 0$, $\sup_{k \leq L\hat{k}} D_k(\lambda) = o(1)$. For $r = 0, 1$,*

$$\sum_{k > 8\hat{k}} k^r a_{n,k} = o(H_r). \tag{27}$$

Moreover, given fixed $\beta > 0$ and an integer $J \geq 0$, take $C_0 > 8(\beta + J/2 + 2)$ and $M = \lceil \max\{8\hat{k}, \sqrt{C_0 n \log n}\} \rceil$. Then

$$\sup_{k \leq M} D_k(\lambda) = o(1), \qquad \lambda M = o(1), \qquad \sum_{k > M} k^J a_{n,k} = o(n^{-\beta}). \tag{28}$$

Proof. For $k \leq L\hat{k}$, $\lambda k^{3/2} \leq Cn^{-1/4} b_n^{3/2} = o(1)$ and $k\lambda = o(1)$, so Lemma 6.1 applies uniformly. The same holds up to M: the additional cutoff contributes at most $Cn^{-1/4}(\log n)^{3/4} = o(1)$ to $\lambda M^{3/2}$, and $\lambda M = o(1)$.

For larger sizes, sum over a spanning tree and then over all choices of further edges. Every connected graph is counted at least once, so for $k \geq 2$

$$1 + D_k(\lambda) \leq (1 + \lambda)^{\binom{k}{2} - k + 1} \leq \exp(\lambda k^2/2).$$

With (11) and $\prod_{j<k}(1 - j/n) \leq \exp[-k(k-1)/(2n)]$,

$$k^r a_{n,k} \leq Ck^{r-5/2} \exp\Big\{\eta k - \frac{(1-c)k^2}{2n} + \frac{k}{2n}\Big\}, \qquad c = e^{-1+\eta} \text{ as in (17)}.$$

Here $c \leq 1/2$ eventually. For $k \geq 8\hat{k}$, $\eta k \leq k^2/(8n)$ and $\exp(k/(2n)) \leq e^{1/2}$, whence

$$k^r a_{n,k} \leq Ck^{r-5/2} e^{-k^2/(8n)}.$$

For $r = 0, 1$ the tail sum is at most $Cn^{r/2-3/4} \int_{7s}^{\infty} u^{r-5/2} e^{-u^2/8} du$, where $7s$ in place of $8s$ absorbs the lattice endpoint. Dividing by $H_r = en^{r/2-3/4} s^{r-5/2} e^{s^2/2}$ proves (27), since $s \geq a_n \to \infty$. The same Gaussian bound holds for any fixed power $r = J$; for $J \geq 3$ the integrand $u^{J-5/2} e^{-u^2/8}$ decreases for $u \geq 2\sqrt{J}$, which holds beyond the cutoff for large n. Beyond M the integral starts at least at $\sqrt{C_0 \log n} - n^{-1/2}$, and the bound is $n^{J/2-3/4-C_0/8}$ times a fixed power of $\log n$, which is $o(n^{-\beta})$ by the choice of C_0. □

Proposition 6.3 (Transfer of size laws)**.** *The connected model has the same relative zeroth and first size-moment asymptotics as the tree model. Its size law and the tree size law have total variation distance $o(1)$, and so do their conditional size laws given $K \geq \hat{k}/2$, even when this event has probability tending to zero. Consequently Corollaries 5.3–5.5 hold for connected graphs with the same constants, including the first-moment limit of $\mathbb{E}K$ and the escaped-mean identity.*

Proof. For $r = 0, 1$, the excess normalized moment $\Delta_r = \sum_k k^r t_{n,k} D_k(\lambda) \geq 0$ is $o(1 + H_r)$ up to $8\hat{k}$ by the local surplus bound and $\sum_k k^r t_{n,k} = O(1 + H_r)$, and $o(H_r)$ beyond $8\hat{k}$ by Lemma 6.2 and $t_{n,k} \leq a_{n,k}$. Thus $\Delta_r = o(1 + H_r)$. For $k - 1$ and $h(k)$ the nonnegative excesses are bounded by Δ_1; their tree expansions have strictly positive critical backgrounds, so the error is relative $o(1)$, and likewise for the zeroth moment. The excess partition function is the total weight of cyclic graphs, a fraction $o(1)$ of the total, and conditional on being a tree the connected law is exactly the tree law; this gives the unconditional total-variation comparison. On $\hat{k}/2 \leq k \leq 8\hat{k}$ the excess zeroth weight is $o(1)$ times the tree weight there, which is $H_0(1 + o(1))$, and the far tail is $o(H_0)$ by Lemma 6.2; this gives the conditional comparison. Total variation transfers bounded-test weak limits, including the mixture and the conditional Gaussian; the moment comparison for $r = 1$ transfers the mean, and with it the escaped-mean identity. □

6.2 Deleting a fixed marked forest

Probabilities of prescribed edges require a finer comparison than the size law; the following deletion argument supplies it for every fixed forest. Let F have $m \geq 1$ edges and r incident labelled vertices, with no isolated vertices, and let c_F be the product of its component orders. On a fixed k-set containing its vertices, let $M_F(k, \lambda)$ be the weight of connected graphs containing its edges. By Lemma 2.1 its tree counterpart is

$$T_F(k, \lambda) = c_F k^{k-m-2} \lambda^{k-1}, \qquad k \geq r. \tag{29}$$

Lemma 6.4 (Fixed-forest comparison)**.** *Suppose $C_j(\lambda) \leq (1+\epsilon) t_j(\lambda)$ for $1 \leq j \leq \kappa$. For $r \leq k \leq \kappa$,*

$$1 \leq \frac{M_F(k, \lambda)}{T_F(k, \lambda)} \leq (1 + \epsilon)^{m+1} + \sum_{j=1}^{m} \frac{\binom{r}{j} j}{c_F} (1 + \epsilon)^j (\lambda k)^{m+1-j}. \tag{30}$$

In particular the ratio tends uniformly to one if $\epsilon = o(1)$ and $\lambda\kappa = o(1)$.

Proof. We use the rooted spanning-forest identity: for a specified j-set of roots in a k-set, the number of spanning forests with one root per component is jk^{k-j-1}. By the rooted matrix-tree theorem this is the determinant of the complete-graph Laplacian with the root rows and columns deleted, the $(k - j) \times (k - j)$ matrix $kI - J$, whose eigenvalues are k with multiplicity $k - j - 1$ and j with multiplicity one; for $j = k$ the empty determinant equals one.

Delete the m prescribed edges from a connected graph containing F. Every remaining component contains a marked vertex, and there are $j \leq m+1$ components; reinserting the edges of F gives a connected quotient multigraph on them. If $j = m+1$, the quotient has j vertices and $j-1$ edges, hence is a tree. All prescribed edges then run between components, so the component weights are unrestricted connected weights; replacing each by its tree weight gives exactly the trees containing F, since deletion and reinsertion are inverse. This contribution is at most $(1+\epsilon)^{m+1} T_F$. For $j \leq m$ we drop quotient connectivity and the internal restrictions, which only enlarges the count. After replacing each component weight by at most $(1 + \epsilon)$ times its tree weight, we count tree forests whose components each contain one of the r marks. Choosing one marked root per component and taking a union over the $\binom{r}{j}$ root sets, the rooted forest identity bounds their number by $\binom{r}{j} j k^{k-j-1}$. Multiplying by λ^{k-j} for the forest edges and λ^m for reinsertion, and dividing by (29), gives the jth term. The lower bound counts the trees alone. □

Theorem D. *For every fixed nonempty labelled forest F as above, uniformly for $a_n \leq s \leq b_n$,*

$$\frac{\mathbb{P}_{\mathrm{c}}(F \subseteq G)}{\mathbb{P}_{\mathrm{t}}(F \subseteq T)} = 1 + o(1).$$

Inclusion prescribes edges, and F may be disconnected. Consequently, for adjacent edges e, f,

$$\frac{R^{\mathrm{c}}_{n,\lambda}}{n} = \frac{3}{4}\frac{(1+p)(2/3+q)}{(1+q)^2}(1+o(1)), \qquad R^{\mathrm{c}}_{n,\lambda} = \frac{\mathbb{P}_{\mathrm{c}}(e,f)}{\mathbb{P}_{\mathrm{c}}(e)\mathbb{P}_{\mathrm{c}}(f)}, \tag{31}$$

with $\mathbb{P}_{\mathrm{c}}(e) = \frac{2(1+q)}{n(n-1)(1+p)}(1+o(1))$ and $\mathbb{P}_{\mathrm{c}}(e,f) = \frac{3(2/3+q)}{n(n-1)(n-2)(1+p)}(1+o(1))$, and the disjoint-edge profile of Corollary 5.11 holds for connected graphs.

Proof. Use Lemma 6.2 with $J = 0$, $\beta = r+2$ and $C_0 = 16(r+4)$. At its cutoff M, $\epsilon = \sup_{j\le M} D_j(\lambda) = o(1)$ and $\lambda M = o(1)$, so Lemma 6.4 compares all marked weights below M with relative error $o(1)$; multiplication by $\binom{n-r}{k-r}$ and summation preserve this comparison. The raw connected partition tail is $\sum_{k>M}\binom{n}{k}C_k(\lambda) = n\sum_{k>M} a_{n,k} = o(n^{-r-1})$, and each marked tail is at most this quantity; tree tails are no larger. Joining the components of F by fixed edges on its r vertices gives a tree containing F, which contributes λ^{r-1} to the marked tree numerator. Since $\lambda \sim 1/(en)$ uniformly, both tails are negligible relative to this lower bound, and the full marked numerators agree up to a factor $1+o(1)$. Proposition 6.3 gives a partition-function ratio $1+o(1)$. The edge, wedge and disjoint-edge statements follow by applying this to a single edge, a path with two edges and two disjoint edges, together with Theorem C(a), Corollaries 5.10 and 5.11, and (3)–(4). □

At the first-moment window of Corollary 5.3, $R^{\mathrm{c}}_{n,\lambda}/n \to \frac{3}{4}(2/3+Q)/(1+Q)^2$; at the mass window of Corollary 5.4, $(\hat{k}/n)R^{\mathrm{c}}_{n,\lambda} \to \frac{3}{4}(1+P)/P$.

Remark 6.5 (Dense-regime contrast, numerical). The constant $3/4$ is a spanning-tree effect. In the connected-subgraph model at fixed n, the law tends as $\lambda \to \infty$ to the uniform law on all graphs with vertex set $[n]$, where edges are independent, so $R^{\mathrm{c}}_{n,\lambda} \to 1$. Exact recursions for C_k and for the marked weights (Appendix B), checked against brute-force enumeration for $n = 4, 5$, give $\inf_\lambda R^{\mathrm{c}}_{n,\lambda} = 0.99713$, 0.99494, 0.99483, 0.99512 and 0.99572 for $n = 6, 10, 14, 20, 40$ on the grid $10^{-3} \le \lambda \le 10^2$, compared with 0.75007 for the tree model at $n = 40$. These figures are numerical.

A The Liu–Slade surplus input

Let $C(k,m)$ be the number of connected labelled graphs on k vertices with m edges, and $S(k,z) = \sum_{\ell\ge 1} C(k, k-1+\ell)z^\ell$, so that $D_k(z) = k^{-(k-2)}S(k,z)$ in the notation of Section 6. Liu and Slade split $S(k,z) \le k^{k-2}(A(k,z)+B(k,z))$ into sparse and dense parts [4, Section 3.2, (3.15)–(3.16)], with their vertex count n in the role of our k. Their Lemma 3.3(i), with $b = 1$, gives $A(k,z) \le Czk^{3/2}$ when $zk^{3/2} \le 1$, and their Lemma 3.4 gives $B(k,z) \le Cz^2$ when $kz \le 3/e$. Together these give Lemma 6.1; nonnegativity holds because every term of S is nonnegative. This is the only external enumeration input of the article; Theorems A–C use only Cayley counts and Stirling's formula.

B Exact and high-precision checks

The package contains Python scripts (standard library and NumPy) that reproduce every constant in the article. `core.py` evaluates the size law exactly in rational arithmetic, and in floating point through the exact ratio (9); the two agree to 9×10^{-16} relative error for $n \le 120$.

- `check_exact.py`: brute-force enumeration of all subtrees for $n = 4, 5, 6$ reproduces $R_{n,\lambda}$ and $R^{\mathrm{disj}}_{n,\lambda}$ exactly at $\lambda = 1/3, 1, 5/2$; the spanning-tree values $3/4$ and 1; the certificate (15) with all coefficients nonnegative and the stated $\pi_{n,0}$ for $3 \le n \le 80$; $Q_{ij} \ge 0$ for $(i,j) \ne (2,2)$ for $4 \le n \le 200$; and $R_{3,\lambda} = 3(1+\lambda+\lambda^2)/(1+2\lambda)^2$.

- `check_phase.py`: Theorem B. For example, at $n = 10^6$, $R/n = 0.944084$, 0.870414, 0.755298 at $c = 0.1, 0.2, 0.3$ (limits 0.944084, 0.870414, 0.755299); $R = 2.838384$, 1.186485, 1.000001, 0.919050, 0.809565 at $c = 0.5, 1, 4/e, 2, 5$ (limits 2.838317, 1.186483, 1, 0.919050, 0.809564). At $\lambda = 1$, $n = 5 \times 10^4$, $n^2(4R/3 - 1 - t/n) = 0.68716$ against $t^2 + 3t/2 = 0.68715$, and $n^2(R^{\mathrm{disj}} - 1) = -1.10374$ against $-3t = -1.10364$. At $c = 1$, $n = 10^7$, $n(R^{\mathrm{disj}} - 1) = -1.40723$ (limit -1.40723) and $\operatorname{Var} K/n = 0.367880$ ($u = 0.367879$). At $c = 1/e$, $R^{\mathrm{disj}}/n^{1/4} = 1.70332$ at $n = 10^8$, increasing towards 1.72008. The constants are $c_0 = 0.17606770185\ldots$, $4/e = 1.47151776\ldots$ and $4/(3e) = 0.49050592\ldots$.

- `check_saddle.py`: Theorem C at $n = 10^6$ and 10^8, the comparison in Remark 5.7, and the sweep of $|4k_*R/(3n) - 1|$ over $\sigma_n \le s \le 6\sqrt{n}$ (activities up to $\lambda \approx e^5/n$) with $p(\sigma_n) = \log n$, which gives 0.266, 0.201, 0.164, 0.141 for $n = 10^4, \ldots, 10^7$, attained at $s = \sigma_n$ where $1/p = 1/\log n$ dominates.

- `check_connected.py`: the connected-model recursions, checked against brute force, and the figures of Remark 6.5.

In the same computations, $R_{n,\lambda}$ decreases in λ on a logarithmic grid $10^{-4} \le \lambda \le 10^4$ for thirteen values $3 \le n \le 400$, and the solution λ_n of $R_{n,\lambda_n} = 1$ satisfies $n\lambda_n = 1.981$, 1.506, 1.4750, 1.47186, 1.471552, 1.471521 for $n = 10, 10^2, \ldots, 10^6$, approaching $4/e$, consistent with Theorem B(ii).

Funding, AI assistance and data availability

No external funding was received for this work. OpenAI and Anthropic AI tools assisted literature comparison, proof exploration, exact and high-precision computation, drafting and consistency checks. The arguments supporting the conclusions are contained in this article, with the external enumeration input identified in Appendix A; the scripts of Appendix B are distributed with it. The author takes responsibility for the statements, attribution and presentation.

References

[1] A. J. Chin, G. Gordon, K. J. MacPhee and C. Vincent, Subtrees of graphs, *Journal of Graph Theory* **89** (2018), no. 4, 413–438. doi:10.1002/jgt.22359.

[2] A. J. Chin, G. Gordon, K. J. MacPhee and C. Vincent, *Random subtrees of complete graphs*, arXiv:1308.4613.

[3] L. Fredes and J.-F. Marckert, Models of random subtrees of a graph, *Probability Surveys* **20** (2023), 722–801. doi:10.1214/23-PS22.

[4] Y. Liu and G. Slade, Critical scaling profile for trees and connected subgraphs on the complete graph, *Canadian Mathematical Bulletin* **69** (2026), no. 2, 550–565. doi:10.4153/S0008439525100763; arXiv:2412.05503.

[5] Z. Luo, K. Xu and J. Tian, Random subtrees and unimodal sequences in graphs, *Discrete Mathematics* **347** (2024), no. 1, 113654. doi:10.1016/j.disc.2023.113654.

[6] Z. Luo, K. Xu and J. Tian, *Random subtrees and unimodal sequences in graphs*, 20-page author manuscript, publicly posted 2 August 2023; Conjecture 5.1, p. 18. Public manuscript record.

[7] H. Robbins, A remark on Stirling's formula, *American Mathematical Monthly* **62** (1955), 26–29.

[8] S. Schubr, *A sharp lower bound for adjacent-edge correlations in random subtrees*, Luminary Research Archive, edition 4, 2026. `https://research.luminaryoddities.org/sharp-lower-bound-adjacent-edge-correlations-random-subtrees/v/4`.

[9] S. Schubr, *Two crossover scales above the critical window for weighted random connected subgraphs*, Luminary Research Archive, version 1, 15 November 2026. `https://research.luminaryoddities.org/two-crossover-scales-above-critical-window-weighted-random-connected-subgraphs/v/1`.

[10] S. Wagner, On the probability that a random subtree is spanning, *Journal of Graph Theory* **98** (2021), no. 2, 195–215. doi:10.1002/jgt.22688.

[11] S. Wagner and R. Wang, *The distribution of subtrees in dense graphs and the roots of the subtree polynomial*, arXiv:2605.03583, 2026.