The Cameron–Mol edge-addition conjecture for complete bipartite and near-complete graphs

Scott Schubr
Luminary Oddities

November 26, 2026

Abstract

Cameron and Mol conjectured that every connected graph other than a complete graph has two nonadjacent vertices whose joining strictly increases the mean subtree order, the average number of vertices of a subtree. This is the existence form of Conjecture 7.4 of Chin, Gordon, MacPhee and Vincent, that adding any edge to a connected multigraph increases the mean, which Cameron and Mol disproved. We prove the Cameron–Mol conjecture for every complete bipartite graph $K_{m,n}$ with $1 \le m \le n$ and $n \ge 2$: joining two vertices of the larger part increases the mean. The proof sorts subtrees by the numbers of omitted vertices in each part, uses Stein-type identities for these numbers, and shows that a vertex of the larger part is the more likely to be omitted; exact computation covers $n \le 300$. We also prove the conjecture for every near-complete graph, one on N vertices whose complement has maximum degree D with $D + 2 < cN$ for an absolute constant $c > 0$, and there we prove Conjecture 7.4 itself, for new and parallel edges and along every chain of single-edge additions up to K_N. The key estimate holds for subtrees weighted by $\lambda > 0$ per edge: if $D \le N/8$, each missing edge changes the weighted mean by $2/(\mathrm{e}\lambda N^2)$ up to a constant depending only on λ times $(D+1)/N^3$. An exhaustive exact computation confirms the conjecture for all graphs on at most nine vertices, and barbells and circulants show that averaging over missing edges, maximal effective resistance and maximal distance do not always select an improving edge.

Keywords. Mean subtree order, edge addition, complete bipartite graph, random subtree, matrix-tree theorem, effective resistance, Poisson approximation.
MSC 2020. Primary 05C05. Secondary 05C30, 05C50, 05C80, 60C05.

1 Introduction

1.1 The conjecture and the answer

A *subtree* of a graph G is a subgraph of G that is a tree; it may be induced or not and spanning or not, and each single vertex is a subtree. For $\lambda > 0$ give a subtree T the weight $\lambda^{|E(T)|}$, let $Z_\lambda(G)$ be the total weight of all subtrees of G, and let

$$\mu_\lambda(G) = \frac{1}{Z_\lambda(G)} \sum_T |V(T)|\, \lambda^{|E(T)|}$$

be the weighted mean subtree order. Thus $\mu := \mu_1$ is the *mean subtree order*, the expected number of vertices of a subtree chosen uniformly at random. For nonadjacent vertices u, v of G we write $G+uv$ for the graph obtained by joining them, and e denotes Euler's number. The same definitions apply to multigraphs, as in [3]: a subtree is a subgraph that is a tree, so it uses at most one of two parallel edges.

Conjecture (Cameron and Mol) (J. Graph Theory 96(3) (2021), [3, Conjecture 1.2]; also [2, Conjecture 1.6(ii)]). *Let G be a connected graph that is not complete. Then G has two nonadjacent vertices u, v with $\mu(G + uv) > \mu(G)$.*

Cameron and Mol proved this conjecture when G is a tree [3, Theorem 4.1]. We prove it for all complete bipartite graphs and all near-complete graphs. This is edition 2 of the record whose edition 1 is S. Schubr, *The Cameron–Mol edge-addition conjecture for complete bipartite and near-complete graphs*, Luminary Research Archive, 2026, `https://doi.org/10.5281/zenodo.24552819`; it extends case (ii) of Theorem A from $K_{m,n}$ with $n \geq n_0(m)$ to every complete bipartite graph (Theorem 3, Section 3) and adds Section 5.

Theorem A (The Cameron–Mol conjecture for these classes). *Each of the following graphs G has two nonadjacent vertices u, v with $\mu(G + uv) > \mu(G)$:*

(i) $G = K_{2,n}$, *for every $n \geq 2$; here u, v may be any two vertices of the n-vertex part;*

(ii) $G = K_{m,n}$, *for all integers $1 \leq m \leq n$ with $n \geq 2$; here u, v may be any two vertices of the n-vertex part;*

(iii) every graph G on N vertices whose complement has maximum degree D with $1 \leq D$ and $D + 2 < c_1 N$, where $c_1 \in (0, 1/8]$ is the absolute constant of Theorem 4 at $\lambda = 1$; here every pair of nonadjacent vertices u, v has this property.

We call a graph on N vertices *near-complete* if its complement has maximum degree D with $D + 2 < c_1 N$. The conjecture of Cameron and Mol is the existence form of an earlier conjecture of Chin, Gordon, MacPhee and Vincent, which we state as Cameron and Mol restate it [3, Conjecture 1.1].

Conjecture (Chin, Gordon, MacPhee and Vincent) (J. Graph Theory 89(4) (2018), [6, Conjecture 7.4]). *Let G be a connected multigraph, and let H be obtained from G by adding an edge between two distinct vertices of G. Then $\mu(H) > \mu(G)$.*

For a graph G the added edge either joins two nonadjacent vertices or is parallel to an existing edge.

This conjecture is false in general. Cameron and Mol disproved it in [3, §1 and §3]. By computer search they found a unique counterexample of order 7 and 347 counterexamples of order 8, and they stated that for every $N \geq 8$ joining the two vertices of the two-vertex part of $K_{2,N-2}$ lowers μ [3, §1]. Their Theorem 3.1 shows that adding one edge to a graph of order N can lower μ by as much as $N/3$ asymptotically; there the graph is a path with many leaves attached at both ends, and the added edge joins the two ends of the path [3, §3]. Cambie, Jooken and Wagner [2, Theorem 1.5] gave graphs with $o(N^2)$ edges and only one improving missing edge up to isomorphism. In each of these infinite families the complement has maximum degree at least a fixed proportion of the order. Near the complete graph the conjecture of Chin, Gordon, MacPhee and Vincent holds.

Theorem B (Conjecture 7.4 of Chin et al. for near-complete graphs). *Let G be a near-complete graph on N vertices, that is, its complement has maximum degree D with $D + 2 < c_1 N$, where c_1 is as in Theorem A. Then adding an edge between any two distinct vertices of G strictly increases the mean subtree order: $\mu(G + uv) > \mu(G)$ for every pair of nonadjacent vertices u, v, and $\mu(G + f) > \mu(G)$ for every new edge f parallel to an edge of G. In particular the conjecture of Chin, Gordon, MacPhee and Vincent holds for G. Moreover, for every chain $G = G_0 \subset G_1 \subset \cdots \subset G_M = K_N$ of graphs on $V(G)$ in which each G_i has exactly one edge more than G_{i-1},*

$$\mu(G_0) < \mu(G_1) < \cdots < \mu(G_M).$$

For every $\lambda > 0$ the same statements hold for μ_λ, with the constant $c_\lambda \in (0, 1/8]$ of Theorem 4 in place of c_1.

Theorems A and B follow from the detailed results below and are proved in Section 4.6. They concern exactly the stated classes. For $m = 1$ the graph $K_{1,n}$ is a star, a tree, so case (ii) with $m = 1$ is already covered by [3, Theorem 4.1]; what Theorem 2 adds there is the value of the gain at every activity. The constants c_1 and c_λ are given by the proofs as existential constants depending only on the indicated parameters; every constant in Section 3 is explicit. In addition, the conjecture of Cameron and Mol holds for every graph on at most nine vertices (Proposition 7).

1.2 Detailed statements

Throughout, $K_{m,n}$ has parts A and B with $|A| = m$ and $|B| = n$; e_A denotes an edge joining two vertices of A and e_B an edge joining two vertices of B. Up to automorphism these are the only missing edges of $K_{m,n}$.

Theorem 1. *For every $n \geq 2$, $\mu(K_{2,n} + e_B) > \mu(K_{2,n})$. Moreover $\mu(K_{2,n} + e_A) > \mu(K_{2,n})$ for $2 \leq n \leq 5$, and $\mu(K_{2,n} + e_A) < \mu(K_{2,n})$ for every $n \geq 6$.*

Theorem 1 proves the decrease stated by Cameron and Mol in [3, §1], shows that it begins exactly at $n = 6$, and identifies the improving edge e_B.

Theorem 2. *Fix an integer $m \geq 1$ and $\lambda > 0$, and put*

$$p = \frac{m\lambda}{1 + m\lambda}, \qquad \rho = \frac{2p^2}{m}.$$

As $n \to \infty$, with implied constants depending only on m and λ,

$$\mu_\lambda(K_{m,n} + e_B) - \mu_\lambda(K_{m,n}) = \frac{2(1-p)\rho}{1+\rho} + O(n^{-1}), \tag{1}$$

and, if $m \geq 2$,

$$\mu_\lambda(K_{m,n} + e_A) - \mu_\lambda(K_{m,n}) = -\frac{2}{m\lambda n} + O(n^{-2}). \tag{2}$$

For $m = 2$ and $\lambda = 1$ the two changes are $8/39 + O(n^{-1})$ and $-1/n + O(n^{-2})$; the second gives the rate of the decrease in Theorem 1. Since $0 < p < 1$, the constant in (1) is positive. Hence for each $m \geq 2$ and $\lambda > 0$ there is n_0 such that, for $n \geq n_0$, a missing edge of $K_{m,n}$ increases μ_λ exactly when it lies inside the larger part. At $\lambda = 1$ the increase holds for every complete bipartite graph.

Theorem 3. *Let $1 \leq m \leq n$ with $n \geq 2$. Then $\mu(K_{m,n} + e_B) > \mu(K_{m,n})$: joining two vertices of the larger part of a complete bipartite graph strictly increases the mean subtree order.*

For $m = n$ both parts are larger parts, and then every missing edge increases μ. The edge e_A inside a strictly smaller part can lower μ, as in $K_{2,n}$ for $n \geq 6$ (Theorem 1).

Theorem 4. *Fix $\lambda > 0$. There are constants $C_\lambda \geq 16/(\mathrm{e}\lambda)$ and $c_\lambda = 2/(\mathrm{e}\lambda C_\lambda) \in (0, 1/8]$, depending only on λ, with the following property. Let N and D be integers with $N \geq 2$ and $0 \leq D \leq N/8$, let G be a graph on N vertices whose complement has maximum degree at most D, and let u, v be nonadjacent vertices of G. Then*

$$\left| \mu_\lambda(G + uv) - \mu_\lambda(G) - \frac{2}{\mathrm{e}\lambda N^2} \right| \leq C_\lambda \, \frac{D+1}{N^3}. \tag{3}$$

In particular, if $D+1 < c_\lambda N$, then $\mu_\lambda(G+uv) > \mu_\lambda(G)$ for every pair of nonadjacent vertices u, v. If G has M missing edges, then

$$\left|\mu_\lambda(K_N) - \mu_\lambda(G) - \frac{2M}{\mathrm{e}\lambda N^2}\right| \le C_\lambda \frac{M(D+1)}{N^3}. \tag{4}$$

The order of the quantifiers is explicit: C_λ and c_λ are chosen first, and (3) then holds uniformly in N, D, G and the missing edge. Since $c_\lambda \le 1/8$, the condition $D+1 < c_\lambda N$ already implies $D \le N/8$.

Chin, Gordon, MacPhee and Vincent pointed out that their conjecture would make K_N the maximizer of μ among connected graphs of order N, and Cameron and Mol noted that their weaker conjecture would also imply this [3, §1]. That maximality was proved by Cambie, Jooken and Wagner [2, Theorem 1.4] and by Chen, Jin, Li and Ma [4]. The ratio monotonicity of Chen, Jin, Li and Ma gives this strict maximality at every activity $\lambda > 0$ as well. In the near-complete class, Theorem 4 gives it a quantitative form: the new content of Corollary 5 is the explicit lower bound for the gap.

Corollary 5. *Fix $\lambda > 0$. Let $G \neq K_N$ be a graph on N vertices with $M \ge 1$ missing edges whose complement has maximum degree D with $D+1 < c_\lambda N$. Then*

$$\mu_\lambda(K_N) - \mu_\lambda(G) \ge \frac{2M}{\mathrm{e}\lambda N^2}\Big(1 - \frac{D+1}{c_\lambda N}\Big) > 0.$$

In particular K_N is the strict maximizer of μ_λ among graphs on N vertices whose complement has maximum degree D with $D+1 < c_\lambda N$.

1.3 Relation to earlier work

Cameron and Mol also proposed, in the journal version of [3] (the statement is not in its arXiv version 1), that $\mu(K_m + nK_1) < \mu(K_{m,n})$ for n large compared with m (see [1, Conjecture 1.3]), where $K_m + nK_1$ is $K_{m,n}$ with all $\binom{m}{2}$ edges inside the m-part added. Cambie, Chen, Hao and Tokar proved this comparison [1, §3]. They, and independently Chen, Wei and Lian [5], also proved the companion conjecture of Cameron and Mol that for every $k \ge 1$ some addition of k edges lowers the mean subtree order. The second part of Theorem 2 is the single-edge counterpart of the comparison of Cambie, Chen, Hao and Tokar: one edge inside the small part already lowers the mean, at the exact rate $2/(m\lambda n)$, for every activity.

The weight ratios behind Theorem 4 form a Poisson law for the number of omitted vertices. For the complete graph, Chin, Gordon, MacPhee and Vincent proved that the probability that a uniform subtree of K_N is spanning tends to $\mathrm{e}^{-1/\mathrm{e}}$ (see [3, §1]), and Luo, Xu and Tian [8] proved the Poisson law for the number of omitted vertices under the uniform law on K_N; Wagner [11] proved the corresponding Poisson limits for dense Erdős–Rényi graphs, and Wagner and Wang [12, Theorem 2] proved a Poisson approximation, for the uniform law, for every connected graph of minimum degree at least δN, with relative error $O(kN^{-1/3})$ for k omitted vertices. The gain in Theorem 4 has order N^{-2}, so its proof needs two inputs of higher precision in the class of sparse complements: the Poisson law with error $O((D+1)/N)$ at every activity (Lemma 4.5), and an edge-specific estimate for the probability that the subtree contains uv at a given order (Lemma 4.8).

For the complete graph ($D = 0$) both inputs are exact or classical facts. By symmetry the fixed-order edge marginal is $\mathbb{P}(uv \in T \mid |V(T)| = k) = (k-1)/\binom{N}{2} = 2(k-1)/(N(N-1))$, since a tree on k vertices has $k-1$ edges. The number of omitted vertices is asymptotically Poisson with mean $1/(\mathrm{e}\lambda)$; this is the law of Luo, Xu and Tian at $\lambda = 1$, and the ratio computation in the proof

of Lemma 4.4 gives it at every activity. Section 4 extends both facts to complements of maximum degree D, with errors linear in D (Remark 4.9).

For complete bipartite graphs with both parts large, the gain of e_B has order n^{-2} when $m = n$, and the expansions of Theorem 2, in which m is fixed, do not reach it. Section 3 proves Theorem 3 by working with exact identities rather than expansions, with explicit constants throughout.

1.4 Route through the argument

Section 2 treats complete bipartite hosts. A subtree of $K_{m,n}$ with at least two vertices is a spanning tree of the complete bipartite graph induced on its vertex set. An added edge changes one Laplacian eigenvalue of that graph, which multiplies its spanning-tree count by an explicit factor (Lemma 2.2). For $K_{2,n}$ this gives exact tables of weights and mean orders, and the sign of the change reduces to explicit inequalities. For fixed m the weights become tilted binomial laws, whose means we expand to the needed order (Lemma 2.3).

Section 3 treats all $K_{m,n}$. Fix $b_1, b_2 \in B$. Every subtree of $K_{m,n}$ is matched with a subtree through b_1 and b_2 by the numbers s, t of its vertices in the two parts, at explicit rates (Lemma 3.1). So the sign of the gain is that of an inequality between moments of the law of (s, t) for a random subtree through b_1 and b_2 (Lemma 3.2). The new subtrees are weighted by $2/s$, and the subtrees missing b_1 or b_2 by twice the omission rate of a vertex of B. The loss caused by the weight $2/s$ is governed by the omission rate of a vertex of A. Stein-type identities for the numbers $\hat{s} = m - s$ and $\hat{t} = n - t$ of omitted vertices (Lemma 3.3) turn both effects into explicit expectations. The comparison (11), which says that a vertex of the larger part is the more likely to be omitted, shows that the gain beats the loss by a factor near 2. The remaining terms are controlled by monotonicity, by shifts of conditional laws (Lemmas 3.4 and 3.5) and by explicit tail bounds. An exact computation covers $n \le 300$ (Proposition 6).

Section 4 treats near-complete hosts. By the matrix-tree theorem the trees on a vertex set U number $|U|^{|U|-2} \det(\mathrm{Id} - L_U/|U|)$, where L_U is the Laplacian of the complement restricted to U (Lemma 4.1); these determinants are stable under removing or exchanging vertices (Lemma 4.3). Comparison with the complete graph gives the Poisson law (Lemma 4.5), and exchanging vertex labels gives the omission probabilities (Lemma 4.6). Kirchhoff's formula expresses edge inclusion as an effective resistance (Lemma 4.7), and a resolvent comparison gives the inclusion probability at each order (Lemma 4.8). The gain is a covariance between the order and that probability (Lemma 2.1), and its leading term is $2\,\mathrm{Var}(J)/N^2$, where J is the number of omitted vertices, whose variance tends to $1/(\mathrm{e}\lambda)$.

2 Complete bipartite hosts

2.1 Two forms of the gain

For a graph G, nonadjacent vertices u, v and $\lambda > 0$, let Z^{uv} be the total weight of the subtrees of $G + uv$ that contain the edge uv, and let ν be their weighted mean order.

Lemma 2.1. *With this notation,*

$$\mu_\lambda(G + uv) - \mu_\lambda(G) = \frac{Z^{uv}}{Z_\lambda(G) + Z^{uv}}\,(\nu - \mu_\lambda(G)). \tag{5}$$

Equivalently, let T be a random subtree of $G + uv$ with $\mathbb{P}(T) \propto \lambda^{|E(T)|}$, let $K = |V(T)|$ and let I be

the indicator of $uv \in E(T)$. Then

$$\mu_\lambda(G+uv) - \mu_\lambda(G) = \frac{\operatorname{Cov}(K, I)}{1 - \mathbb{P}(I=1)}. \tag{6}$$

Proof. The subtrees of $G+uv$ that avoid uv are exactly the subtrees of G, with the same weights. Hence $\mu_\lambda(G+uv)$ is the average of $\mu_\lambda(G)$ and ν with weights $Z_\lambda(G)$ and Z^{uv}, which is (5). In the probabilistic form, conditioning on $I=0$ gives the weighted law on subtrees of G, so $\mu_\lambda(G) = \mathbb{E}[K(1-I)]/(1-\mathbb{P}(I=1))$, and

$$\mathbb{E}K - \frac{\mathbb{E}K - \mathbb{E}[KI]}{1-\mathbb{P}(I=1)} = \frac{\mathbb{E}[KI] - \mathbb{E}K\,\mathbb{P}(I=1)}{1-\mathbb{P}(I=1)},$$

which is (6). □

2.2 Spanning trees of complete bipartite supports

For a graph X let $\tau(X)$ denote its number of spanning trees.

Lemma 2.2. *Let $s, t \geq 1$ and let $K_{s,t}$ have parts A' and B' with $|A'| = s$ and $|B'| = t$. Then $\tau(K_{s,t}) = s^{t-1}t^{s-1}$. If $b_1 \neq b_2$ lie in B', exactly $(2/s)\,\tau(K_{s,t})$ spanning trees of $K_{s,t} + b_1b_2$ contain b_1b_2; if $a_1 \neq a_2$ lie in A', exactly $(2/t)\,\tau(K_{s,t})$ spanning trees of $K_{s,t} + a_1a_2$ contain a_1a_2.*

Proof. Let Λ be the Laplacian of $K_{s,t}$. A vector x supported on B' with zero sum satisfies $\Lambda x = sx$, because each vertex of B' has degree s and all its neighbours lie in A', where x vanishes, while each vertex of A' sees $\sum_{b\in B'} x_b = 0$. Likewise zero-sum vectors on A' have eigenvalue t, the vector equal to t on A' and $-s$ on B' has eigenvalue $s+t$, and constants have eigenvalue 0. These account for all $s+t$ dimensions, so the matrix-tree theorem gives $\tau(K_{s,t}) = (s+t)^{-1}s^{t-1}t^{s-1}(s+t)$; this is the classical count of Scoins [9]. For $x = \mathbf{1}_{b_1} - \mathbf{1}_{b_2}$ the Laplacian of $K_{s,t} + b_1b_2$ is $\Lambda + xx^{\mathsf{T}}$. Since $\Lambda x = sx$ and Λ preserves $x^\perp$, on which xx^{T} vanishes, the eigenvalue s belonging to x becomes $s + \|x\|^2 = s+2$ and all others are unchanged. Hence $\tau(K_{s,t} + b_1b_2) = (1 + 2/s)\,\tau(K_{s,t})$, and the trees through b_1b_2 number $(2/s)\,\tau(K_{s,t})$. The case of a_1a_2 is symmetric. □

A subtree of $K_{m,n}$ with at least two vertices meets A in some $s \geq 1$ vertices and B in some $t \geq 1$ vertices, and it is a spanning tree of the copy of $K_{s,t}$ induced on its vertex set. The subtrees of $K_{m,n} + e_B$ through $e_B = b_1b_2$ are the edge b_1b_2 itself and, for $s \geq 1$ and $t \geq 2$, the spanning trees through b_1b_2 of the copies of $K_{s,t} + b_1b_2$ induced on the $\binom{m}{s}\binom{n-2}{t-2}$ vertex sets that contain b_1, b_2. Since $\binom{n-2}{t-2} = \binom{n}{t}\,t(t-1)/(n(n-1))$, Lemma 2.2 gives

$$Z_\lambda(K_{m,n}) = m + n + \sum_{s=1}^{m}\sum_{t=1}^{n} \binom{m}{s}\binom{n}{t} s^{t-1}t^{s-1}\lambda^{s+t-1}, \tag{7}$$

$$Z^{e_B} = \lambda + \sum_{s=1}^{m}\sum_{t=1}^{n} \binom{m}{s}\binom{n}{t} \frac{2t(t-1)}{s\,n(n-1)}\, s^{t-1}t^{s-1}\lambda^{s+t-1}, \tag{8}$$

$$Z^{e_A} = \lambda + \sum_{s=1}^{m}\sum_{t=1}^{n} \binom{m}{s}\binom{n}{t} \frac{2s(s-1)}{t\,m(m-1)}\, s^{t-1}t^{s-1}\lambda^{s+t-1} \quad (m \geq 2), \tag{9}$$

and the corresponding order sums carry the extra factor $s+t$, and 2 for the single edge. The checker `check_exact.py` compares (7)–(9) with a direct enumeration of acyclic edge sets for small m, n and several activities.

2.3 Proof of Theorem 1

Here $\lambda = 1$, $A = \{a_1, a_2\}$, and we call a_1, a_2 the terminals. A subtree of $K_{2,n}$ with no terminal is a single vertex of B. A subtree with exactly one terminal a is a star centred at a whose leaves form an arbitrary subset of B. A subtree containing both terminals contains a unique path $a_1 b a_2$, and we call $b \in B$ its connector; each of the other $n-1$ vertices of B is absent, a leaf at a_1 or a leaf at a_2. Writing $\alpha = 2^n$ and $\beta = 3^{n-1}$, the subtrees of $K_{2,n}$ are therefore counted as follows.

terminals	number	mean order
0	n	1
1	2α	$(n+2)/2$
2	$n\beta$	$(2n+7)/3$

The edge $e_B = b_1 b_2$. The subtrees through $b_1 b_2$ with no terminal consist of the edge itself. Those with one terminal a contain the triangle ab_1b_2, use b_1b_2 and exactly one of ab_1, ab_2, and have the remaining $n-2$ vertices of B as optional leaves at a: there are $2 \cdot 2 \cdot 2^{n-2} = \alpha$ of them, with mean order $(n+4)/2$. For those with both terminals, Lemma 2.2 with $s = 2$ shows that on every vertex set containing a_1, a_2, b_1, b_2 the trees through b_1b_2 are as numerous as the spanning trees of the corresponding $K_{2,t}$. So they have the same counts by order as the subtrees of $K_{2,n}$ that contain a_1, a_2, b_1, b_2. If the connector is b_1 or b_2, the other one is a leaf at a_1 or a_2: $4 \cdot 3^{n-2}$ subtrees with mean order $(2n+8)/3$. Otherwise the connector is one of the other $n-2$ vertices and b_1, b_2 are both leaves: $4(n-2)3^{n-3}$ subtrees with mean order $(2n+9)/3$ (none when $n = 2$). In total there are $4(n+1)3^{n-3}$ such subtrees, with mean order

$$\frac{3(2n+8) + (n-2)(2n+9)}{3(n+1)} = \frac{2n^2 + 11n + 6}{3(n+1)} = \frac{2n+9}{3} - \frac{1}{n+1}.$$

terminals	number through e_B	mean order
0	1	2
1	α	$(n+4)/2$
2	$4(n+1)\beta/9$	$(2n+9)/3 - 1/(n+1)$

Let $Z = n + 2\alpha + n\beta$ and $\Sigma = n + (n+2)\alpha + n(2n+7)\beta/3$ be the number and the total order of the subtrees of $K_{2,n}$, and let Z^{e_B} and Σ^{e_B} be the same quantities for the subtrees through e_B. By (5), the change has the sign of $\nu - \mu(K_{2,n}) = (\Sigma^{e_B} Z - \Sigma Z^{e_B})/(Z Z^{e_B})$, and expanding the tables gives

$$\begin{aligned}\Sigma^{e_B} Z - \Sigma Z^{e_B} = 2\alpha^2 &+ \Big(-\frac{n^2}{54} + \frac{43n}{27} + \frac{8}{9}\Big)\alpha\beta + \Big(\frac{n^2}{2} + 2\Big)\alpha \\ &+ \frac{8n^2 - 4n}{27}\,\beta^2 + \frac{8n^3 + 14n^2 + 3n}{27}\,\beta + n.\end{aligned}$$

For $n = 2$ every coefficient is positive. For $n \geq 3$ we have $\beta \geq \alpha$, so $\frac{8n^2-4n}{27}\beta^2 \geq \frac{n^2}{54}\alpha\beta$ because $16n^2 - 8n \geq n^2$; the only negative term is dominated and the others are positive. Hence $\mu(K_{2,n} + e_B) > \mu(K_{2,n})$ for every $n \geq 2$.

The edge $e_A = a_1 a_2$. A subtree through a_1a_2 contains both terminals, and each vertex of B is absent or a leaf at exactly one terminal. So there are $3^n = 3\beta$ such subtrees, with mean order $2 + 2n/3$. Then

$$3\big((2 + \tfrac{2n}{3})Z - \Sigma\big) = n(2n+3) + (n+6)2^n - n\,3^{n-1},$$

which equals 40, 72, 96, 12 for $n = 2, 3, 4, 5$ and -600 for $n = 6$. For $n \geq 6$, divide by $n3^{n-1}$: the result is $(2n+3)/3^{n-1} + 3(1+6/n)(2/3)^n - 1$, and both positive summands strictly decrease in n. Hence it stays negative for all $n \geq 6$, and (5) gives the stated signs. □

The checker `check_exact.py` verifies both tables and the two sign expressions exactly for $2 \leq n \leq 44$; the identities themselves are polynomial identities in n, α, β.

2.4 Tilted binomial means

For a real x and an integer $l \geq 0$ write $(x)_l = x(x-1)\cdots(x-l+1)$.

Lemma 2.3. *Fix $0 < p < 1$, an integer $d \geq 0$ and a real polynomial h of degree d with leading coefficient 1. Let $\xi \sim \mathrm{Bin}(n, p)$. As $n \to \infty$, with implied constants depending only on p and h,*

$$\mathbb{E}[h(\xi)\mathbf{1}_{\{\xi \geq 1\}}] = (np)^d(1 + O(n^{-1})), \qquad \frac{\mathbb{E}[\xi h(\xi)]}{\mathbb{E}[h(\xi)\mathbf{1}_{\{\xi \geq 1\}}]} = np + d(1-p) + O(n^{-1}).$$

Proof. Write $h(x) = \sum_{l=0}^{d} h_l (x)_l$ with $h_d = 1$. Since $\mathbb{E}[(\xi)_l] = (n)_l p^l$ and $\xi(\xi)_l = (\xi)_{l+1} + l(\xi)_l$,

$$\mathbb{E}[\xi(\xi)_l] = (n)_l p^l (np + l(1-p)).$$

Put $w_l = h_l (n)_l p^l / ((n)_d p^d)$, so that $w_d = 1$ and $w_l = O(n^{l-d})$. Then $\mathbb{E}[h(\xi)] = (n)_d p^d \sum_l w_l = (np)^d(1 + O(n^{-1}))$, and

$$\frac{\mathbb{E}[\xi h(\xi)]}{\mathbb{E}[h(\xi)]} = np + (1-p)\frac{\sum_l l\, w_l}{\sum_l w_l} = np + d(1-p) - (1-p)\frac{\sum_{l<d}(d-l)w_l}{\sum_l w_l} = np + d(1-p) + O(n^{-1}).$$

Finally $\mathbb{E}[\xi h(\xi)\mathbf{1}_{\{\xi \geq 1\}}] = \mathbb{E}[\xi h(\xi)]$, and restricting $h(\xi)$ to $\{\xi \geq 1\}$ changes its mean by $h(0)(1-p)^n$, which is exponentially small relative to $(np)^d$. □

2.5 Proof of Theorem 2

All constants in this proof depend only on m and λ. For $1 \leq s \leq m$ let $p_s = s\lambda/(1+s\lambda)$, let $\xi_s \sim \mathrm{Bin}(n, p_s)$ and put $q_s = \binom{m}{s}\lambda^{s-1}/s$, so that $p_m = p$. The identity $\binom{n}{t}(s\lambda)^t = (1+s\lambda)^n \mathbb{P}(\xi_s = t)$ turns the (s,t) term of (7) into $q_s(1+s\lambda)^n \mathbb{P}(\xi_s = t)\, t^{s-1}$. Thus the subtrees meeting A in exactly s vertices form a class of total weight

$$Z_s = q_s(1+s\lambda)^n\, \mathbb{E}[\xi_s^{s-1}\mathbf{1}_{\{\xi_s \geq 1\}}]$$

and mean order $s + \mathbb{E}[\xi_s^s]/\mathbb{E}[\xi_s^{s-1}\mathbf{1}_{\{\xi_s \geq 1\}}]$.

Reduction to the class $s = m$. We have $Z_m \geq q_m(1+m\lambda)^n \mathbb{P}(\xi_m \geq 1) \geq q_m(1+m\lambda)^n/2$ for large n, while $Z_s \leq q_s(1+s\lambda)^n n^{s-1}$ for $s < m$, there are $m+n$ singletons, and the single added edge has weight λ. The multipliers in (8) and (9) are at most 2. Hence every contribution to $Z_\lambda(K_{m,n})$, Z^{e_B} or Z^{e_A} from outside the class $s = m$ is at most $\eta_n Z_m$, where $\eta_n = C\, n^m\big((1+(m-1)\lambda)/(1+m\lambda)\big)^n + C(m+n)(1+m\lambda)^{-n}$ decreases exponentially. Since all orders lie in $[1, m+n]$, replacing a family of weighted subtrees by its class-m part changes its mean order by at most $(m+n)$ times the relative weight removed. The class-m parts of Z^{e_B} and Z^{e_A} computed below are at least $Z_m/(Cn)$, so all these replacements cost $O(n^2\eta_n)$ in the means and a factor $1 + O(n\eta_n)$ in the weights, both smaller than any power of n.

The class $s = m$. By Lemma 2.3 with $h(x) = x^{m-1}$,

$$Z_m = q_m(1+m\lambda)^n (np)^{m-1}(1 + O(n^{-1})), \qquad \text{mean order } m + np + (m-1)(1-p) + O(n^{-1}).$$

For e_B the class-m multiplier in (8) is $2t(t-1)/(m\,n(n-1))$, so the new class has weight $2q_m(1+m\lambda)^n\mathbb{E}[\xi_m^m(\xi_m-1)]/(m\,n(n-1))$. Lemma 2.3 with $h(x)=x^m(x-1)$, of degree $m+1$, gives the weight ratio $(2/m)(np)^{m+1}/(n^2(np)^{m-1})\,(1+O(n^{-1}))=\rho+O(n^{-1})$ and the mean order $m+np+(m+1)(1-p)+O(n^{-1})$. Therefore $Z^{e_B}/(Z_\lambda(K_{m,n})+Z^{e_B})=\rho/(1+\rho)+O(n^{-1})$ and $\nu-\mu_\lambda(K_{m,n})=2(1-p)+O(n^{-1})$, and (5) gives (1).

For e_A with $m\ge 2$ the class-m multiplier in (9) is $2/t$, so the new class has weight $2q_m(1+m\lambda)^n\mathbb{E}[\xi_m^{m-2}\mathbf{1}_{\{\xi_m\ge 1\}}]$. Lemma 2.3 with $h(x)=x^{m-2}$ gives the weight ratio $2(np)^{m-2}/(np)^{m-1}\,(1+O(n^{-1}))=2/(np)+O(n^{-2})$ and the mean order $m+np+(m-2)(1-p)+O(n^{-1})$. Therefore $Z^{e_A}/(Z_\lambda(K_{m,n})+Z^{e_A})=2/(np)+O(n^{-2})$ and $\nu-\mu_\lambda(K_{m,n})=-(1-p)+O(n^{-1})$, and (5) gives the change $-2(1-p)/(np)+O(n^{-2})$. Since $(1-p)/p=1/(m\lambda)$, this is (2). □

3 All complete bipartite graphs

This section proves Theorem 3. Throughout, $\lambda=1$, $1\le m\le n$, $n\ge 2$, and b_1,b_2 are two vertices of B.

3.1 Subtrees through two vertices of the larger part

Let $\Omega=\{(s,t):1\le s\le m,\ 2\le t\le n\}$ and

$$w(s,t)=\binom{m}{s}\binom{n-2}{t-2}s^{t-1}t^{s-1},\qquad \mathcal{Z}=\sum_{(s,t)\in\Omega}w(s,t).$$

By Lemma 2.2, $w(s,t)$ is the number of subtrees of $K_{m,n}$ that contain b_1 and b_2 and have s vertices in A and t in B: such a subtree spans the copy of $K_{s,t}$ on its vertex set, and there are $\binom{m}{s}\binom{n-2}{t-2}$ such vertex sets. We write (s,t) for a random element of Ω with $\mathbb{P}(s,t)=w(s,t)/\mathcal{Z}$; thus (s,t) records the part sizes of a uniformly random subtree of $K_{m,n}$ through b_1 and b_2, and $\mathbb{E}$, Var, Cov refer to this law. Put

$$K=s+t,\qquad \hat{s}=m-s,\qquad \hat{t}=n-t,$$

the order of that subtree and the numbers of vertices of A and of B it omits. For integers $x,y\ge 1$ let $\Psi(x,y)=x^{-1}(1-1/y)^{x-1}$, with $0^0=1$, and put

$$\Psi_2(s,t)=s^{-2}(1-2/t)^{s-1},\qquad g=2\Psi(s,t)+\Psi_2(s,t),\qquad \Psi_0=\Psi(m,n).$$

Lemma 3.1. *(i) For $(s,t)\in\Omega$ the subtrees of $K_{m,n}$ that contain exactly one of b_1,b_2 and have s vertices in A and $t-1$ in B number $2\Psi(s,t)w(s,t)$. The other subtrees containing exactly one of b_1,b_2 are $\{b_1\}$ and $\{b_2\}$.*

(ii) For $(s,t)\in\Omega$ the subtrees that contain neither b_1 nor b_2 and have s vertices in A and $t-2$ in B number $\Psi_2(s,t)w(s,t)$. The other subtrees containing neither are the $n-2$ single vertices of $B\setminus\{b_1,b_2\}$.

(iii) The subtrees of $K_{m,n}+b_1b_2$ through b_1b_2 are the edge b_1b_2 and, for $(s,t)\in\Omega$, $(2/s)w(s,t)$ subtrees with s vertices in A and t in B.

Proof. (i) Such a subtree contains one of b_1,b_2, chosen in two ways, and $t-2$ of the other $n-2$ vertices of B, and it spans a copy of $K_{s,t-1}$. So there are $2\binom{m}{s}\binom{n-2}{t-2}s^{t-2}(t-1)^{s-1}=2\Psi(s,t)w(s,t)$ of them. A subtree with no vertex of A is a single vertex. (ii) For $t\ge 3$ there are $\binom{m}{s}\binom{n-2}{t-2}s^{t-3}(t-2)^{s-1}=\Psi_2(s,t)w(s,t)$ such subtrees. For $t=2$ a subtree with no vertex of B is a single vertex of A, and $\Psi_2(s,2)w(s,2)$ is m for $s=1$ and 0 for $s\ge 2$. (iii) is Lemma 2.2. □

Lemma 3.2. *Let $\nu_* = \mathbb{E}[K/s]/\mathbb{E}[1/s]$ and $\tilde{\mu} = (\mathbb{E}[(1+g)K] - 2\mathbb{E}\Psi(s,t) - 2\mathbb{E}\Psi_2(s,t))/(1+\mathbb{E}g)$. If $\nu_* - \tilde{\mu} > m(m+n)/(2\mathcal{Z})$, then $\mu(K_{m,n} + b_1 b_2) > \mu(K_{m,n})$. Moreover*

$$(1+\mathbb{E}g)\,(\nu_* - \tilde{\mu}) = 2\mathbb{E}\Psi(s,t) + 2\mathbb{E}\Psi_2(s,t) + \mathrm{Cov}(g, \hat{s}+\hat{t}) - (1+\mathbb{E}g)\cdot\mathcal{L}, \tag{10}$$

where $\mathcal{L} = \mathbb{E}K - \nu_* = \mathrm{Cov}(\hat{s}+\hat{t}, 1/s)/\mathbb{E}[1/s]$.

Proof. A subtree counted in Lemma 3.1(i) has order $K-1$, one in (ii) has order $K-2$, and a single vertex has order 1. Hence the number of subtrees of $K_{m,n}$ is $\mathcal{Z}\,(1+\mathbb{E}g) + n$ and their total order is $\mathcal{Z}\,(1+\mathbb{E}g)\,\tilde{\mu} + n$, so $\mu(K_{m,n})$ is a weighted average of $\tilde{\mu}$ and 1. Since $\tilde{\mu}$ is itself an average of orders, it is at least 1, and therefore $\mu(K_{m,n}) \le \tilde{\mu}$. By Lemma 3.1(iii) the subtrees through $b_1 b_2$ have mean order

$$\nu = \frac{2\mathcal{Z}\,\mathbb{E}[K/s] + 2}{2\mathcal{Z}\,\mathbb{E}[1/s] + 1} = \nu_* - \frac{\nu_* - 2}{2\mathcal{Z}\,\mathbb{E}[1/s] + 1} \ge \nu_* - \frac{m(m+n)}{2\mathcal{Z}},$$

because $\nu_* \le m+n$ and $\mathbb{E}[1/s] \ge 1/m$. So $\nu - \mu(K_{m,n}) \ge \nu_* - \tilde{\mu} - m(m+n)/(2\mathcal{Z}) > 0$, and (5) gives the claim. For (10), write $\mathbb{E}[(1+g)K] = (1+\mathbb{E}g)\,\mathbb{E}K + \mathrm{Cov}(g,K)$ and $K = m+n-\hat{s}-\hat{t}$, and note that $1+g$ and g have the same covariances. □

3.2 Stein identities and monotonicity

Put

$$\sigma_A(s,t) = s\Psi(t,s) = \frac{s}{t}\Big(1-\frac{1}{s}\Big)^{t-1}, \qquad \sigma_B(s,t) = (t-2)\Psi(s,t) = \frac{t-2}{s}\Big(1-\frac{1}{t}\Big)^{s-1}.$$

Thus $\sigma_A(1,t) = 0$ and $\sigma_B(s,2) = 0$. We also write $\bar{a}(t) = \sigma_A(m,t)$ and $\bar{b}(s) = \sigma_B(s,n)$.

Lemma 3.3. *For every real function f on $\mathbb{Z}^2$,*

$$\mathbb{E}[\hat{s}\,f(s,t)] = \mathbb{E}[\sigma_A(s,t) f(s-1,t)], \qquad \mathbb{E}[\hat{t}\,f(s,t)] = \mathbb{E}[\sigma_B(s,t) f(s,t-1)],$$

and the first identity also holds conditionally on t, the second conditionally on s.

Proof. From $(m-s)\binom{m}{s} = (s+1)\binom{m}{s+1}$ one obtains $(m-s)w(s,t) = \sigma_A(s+1,t)w(s+1,t)$ for $1 \le s < m$. Multiply by $f(s,t)$, sum over s and substitute $s+1 \mapsto s$; the term $s=m$ vanishes on the left and the term $s=1$ on the right. Likewise $(n-t)\binom{n-2}{t-2} = (t-1)\binom{n-2}{t-1}$ gives $(n-t)w(s,t) = \sigma_B(s,t+1)w(s,t+1)$. □

Lemma 3.4. *(i) σ_A is increasing in s and decreasing in t; σ_B is increasing in t and decreasing in s; $\Psi(s,t)$ and $\Psi_2(s,t)$ are decreasing in s and nondecreasing in t; and $\Psi(t,s)$ is increasing in s.*

(ii) $w(s,t)w(s+1,t+1) \ge w(s+1,t)w(s,t+1)$. Hence the law of s given t increases with t in the likelihood-ratio order. Given s, the law of $t-2$ dominates the binomial law $\mathrm{Bin}(n-2, s/(s+1))$ in that order.

(iii) For every t and $k \ge 1$, $\mathbb{E}[(\hat{s})_k \mid t] \le \mathbb{E}[\hat{s} \mid t]^k$, $\mathbb{E}[\hat{s} \mid t] \le \bar{a}(t)$ and $\mathrm{Var}(\hat{s} \mid t) \le \mathbb{E}[\hat{s} \mid t]$. For every s the same holds for $\hat{t}$ given s, with $\bar{b}(s)$ in place of $\bar{a}(t)$.

(iv) Let $z \ge 0$, and let $a \ge \mathbb{E}[\hat{s} \mid t]$. Then $\mathbb{E}[\hat{s}^j (1+z)^{\hat{s}} \mid t] \le M_j(z,a)$ for $j = 0,1,2,3$, where

$$M_0 = \mathrm{e}^{za}, \qquad M_1 = (1+z)a\,\mathrm{e}^{za}, \qquad M_2 = (1+z)a\big(1+(1+z)a\big)\mathrm{e}^{za},$$
$$M_3 = \big((1+z)^3a^3 + 3(1+z)^2a^2 + (1+z)a\big)\mathrm{e}^{za}.$$

The same holds for $\hat{t}$ given s, with any $a \ge \mathbb{E}[\hat{t} \mid s]$.

Proof. (i) The logarithmic derivatives in the real variables s, t have the stated signs. For σ_A they are $1/s + (t-1)/(s(s-1))$ and $-1/t + \log(1-1/s)$. For σ_B they are $1/(t-2) + (s-1)/(t(t-1))$ and $-1/s + \log(1-1/t)$. For $\Psi(s,t)$ they are $-1/s + \log(1-1/t)$ and $(s-1)/(t(t-1))$. For Ψ_2 they are $-2/s + \log(1-2/t)$ and $2(s-1)/(t(t-2))$, for $t \ge 3$; at $t = 2$, $\Psi_2(s,2) = \mathbf{1}_{\{s=1\}}$. For $\Psi(t,s)$ the derivative in s is $(t-1)/(s(s-1))$. (ii) The binomial factors cancel, and the ratio $w(s,t)w(s+1,t+1)/(w(s+1,t)w(s,t+1))$ equals $(s+1)(t+1)/(st)$. Given s, the law of $t' = t-2$ is proportional to $\binom{n-2}{t'} s^{t'}(t'+2)^{s-1}$, which is the binomial weight times the nondecreasing factor $(t'+2)^{s-1}$. (iii) Apply Lemma 3.3 conditionally on t with $f = (\hat{s}-1)_{k-1}$, for which $f(s-1,t) = (\hat{s})_{k-1}$:

$$\mathbb{E}[(\hat{s})_k \mid t] = \mathbb{E}[\sigma_A\,(\hat{s})_{k-1} \mid t] \le \mathbb{E}[\sigma_A \mid t]\,\mathbb{E}[(\hat{s})_{k-1} \mid t].$$

The inequality is Chebyshev's, since given t the factor σ_A increases with s while $(\hat{s})_{k-1}$ decreases with s. Also $\mathbb{E}[\sigma_A \mid t] = \mathbb{E}[\hat{s} \mid t]$, by the case $f = 1$, and $\mathbb{E}[\sigma_A \mid t] \le \sigma_A(m,t)$. The bound on the variance is the case $k = 2$. The statements for $\hat{t}$ are proved in the same way. (iv) The power series $\mathbb{E}[(1+z)^{\hat{s}} \mid t] = \sum_k z^k \mathbb{E}[(\hat{s})_k \mid t]/k!$ is dominated coefficientwise by e^{za}, and so are its derivatives by those of e^{za}. Writing $\hat{s}^2 = (\hat{s})_2 + \hat{s}$ and $\hat{s}^3 = (\hat{s})_3 + 3(\hat{s})_2 + \hat{s}$ gives the four bounds. □

3.3 Shifts of the conditional laws

Lemma 3.5. *Let $2 \le t < n$. Let f be a monotone function of $s \in [1,m]$ with $|f| \le C_f$ and with $|f(s) - f(s-1)| \le \ell$ whenever $s - 1 \ge m/2$. Put $a = \mathbb{E}[\hat{s} \mid t]$, $\vartheta = 2/m + 1/(t+1)$ and $d = 1 - (1/m + 1/(t+1))a$, and suppose $d > 0$. Then*

$$\big|\mathbb{E}[f \mid t+1] - \mathbb{E}[f \mid t]\big| \le \frac{\ell\vartheta a + 2C_f\,\mathbb{P}(\hat{s} > m/2 \mid t)}{d}.$$

The same holds for a monotone function of t given s, with $\hat{t}$, n, $\vartheta' = 2/n + 1/(s+1)$ and $d' = 1 - (1/n + 1/(s+1))\mathbb{E}[\hat{t} \mid s]$ in place of $\hat{s}$, m, ϑ and d.

Proof. Given t, the law of $k = \hat{s}$ is $q_t(k) \propto \binom{m}{k}(1-k/m)^{t-1}t^{-k}$ for $0 \le k \le m-1$. Hence $q_{t+1}(k) = q_t(k)r(k)/\mathbb{E}_t r$, where $r(k) = (1-k/m)(t/(t+1))^k$, and

$$\mathbb{E}[f \mid t+1] - \mathbb{E}[f \mid t] = \mathrm{Cov}_t(f,r)/\mathbb{E}_t r.$$

The function r is decreasing, $r(0) = 1$, and $1 - r(k) \le k(1/m + 1/(t+1))$, so $\mathbb{E}_t r \ge d$. For $j < m/2$,

$$0 \le r(j) - r(j+1) \le \frac{1}{m-j} + \frac{1}{t+1} \le \vartheta.$$

Let k, k' be independent with law q_t. Since f and r are monotone, $|\mathrm{Cov}_t(f,r)|$ equals $\frac{1}{2}\mathbb{E}|(f(k) - f(k'))(r(k) - r(k'))|$. Pairs with $k, k' \le m/2$ contribute at most $\frac{1}{2}\ell\vartheta\,\mathbb{E}(k-k')^2 = \ell\vartheta\,\mathrm{Var}_t(\hat{s})$, which is at most $\ell\vartheta a$ by Lemma 3.4(iii). The other pairs contribute at most $2C_f\,\mathbb{P}_t(\hat{s} > m/2)$. For the second statement, the law of $i = \hat{t}$ given s is $\propto \binom{n-2}{i}(1-i/n)^{s-1}s^{-i}$, and the same argument applies with $r'(i) = (1-i/n)(s/(s+1))^i$. □

3.4 Estimates for large parts

In this subsection $3 \le m \le n$ and $n \ge 301$. Put $\varsigma = n/m$, $z_A = 1/(m-1)$, $z_B = 1/(n-1)$, $E_\varsigma = \mathrm{e}^{-(n-2)/(2m)}$ and $c_3 = \frac12 \log(4/3)$, and define

$$a_* = \frac{2m}{n} E_\varsigma = \frac{2E_\varsigma}{\varsigma}, \qquad b_* = \frac{2n}{m}\,\mathrm{e}^{-(m-2)/(2n)}, \qquad \gamma_m = \frac{1}{\lceil m/2\rceil - 1}, \qquad \varrho = \frac{\Psi(n,m)}{\Psi_0},$$

$$\delta_* = m^2\big(\tfrac23\big)^{n/2-1} + (m^2+1)\mathrm{e}^{-c_3(n-2)} + \big(\tfrac{4\mathrm{e}}{n}\big)^{m/2}\mathrm{e}^{-(n-2)/4}.$$

Note that $\bar a(n) = m\Psi(n,m)$ and $\bar b(m) = (n-2)\Psi_0$, and let $\mathcal{G} = \{s \ge m/2,\ t \ge n/2\}$.

Lemma 3.6. *(i) $\bar a(t) \le a_*$ for $t \ge n/2$, and $\bar b(s) \le b_*$ for $s \ge m/2$.*

(ii) $\mathbb{P}(\mathcal{G}^c) \le \delta_*$. Moreover $\mathbb{P}(\hat s > m/2 \mid t) \le \delta_*$ for $t \ge n/2$, and $\mathbb{P}(\hat t > n/2 \mid s) \le \delta_*$ for $s \ge 3$.

(iii) $a_* \le 1.218$, $b_* \le 2\varsigma$, $\Psi_0 \ge 1/(\mathrm{e}m)$, $\delta_* \le 2\cdot 10^{-14}$ and $n^4\delta_* \le 2\cdot 10^{-4}$.

Proof. (i) We have $\bar a(t) = (m/t)(1-1/m)^{t-1} \le (2m/n)\mathrm{e}^{-(t-1)/m}$ and $t - 1 \ge (n-2)/2$. Also $\bar b(s) = ((n-2)/s)(1-1/n)^{s-1} \le (2n/m)\mathrm{e}^{-(s-1)/n}$ and $s - 1 \ge (m-2)/2$.

(ii) Given $s \ge 3$, Lemma 3.4(ii) and the Chernoff bound $\mathbb{P}(\mathrm{Bin}(N,p) \le N/2) \le \mathrm{e}^{-N\,D(1/2\|p)}$ give $\mathbb{P}(t < n/2 \mid s) \le \mathrm{e}^{-c_3(n-2)}$. Here $D(1/2\|3/4) = c_3$ is the relative entropy, and $\{t < n/2\} = \{\hat t > n/2\}$. For $s \le 2$ and every t, $w(s,t)/w(m,t) = \binom{m}{s}(s/m)^{t-1}t^{s-m} \le \binom{m}{s}(2/3)^{t-1}$, and $\binom{m}{1} + \binom{m}{2} \le m^2$. Hence

$$\mathbb{P}(s \le 2) \le m^2 \sum_t (2/3)^{t-1}\mathbb{P}(s = m, t) \le m^2\big((2/3)^{n/2-1} + \mathrm{e}^{-c_3(n-2)}\big).$$

For $t \ge n/2$ put $k = \lfloor m/2\rfloor + 1$. Then Lemma 3.4(iii) and $k! \ge (k/\mathrm{e})^k$ give

$$\mathbb{P}(\hat s \ge k \mid t) \le \frac{\bar a(t)^k}{k!} \le \Big(\frac{2\mathrm{e}a_*}{m}\Big)^k \le \Big(\frac{4\mathrm{e}}{n}\Big)^{m/2}\mathrm{e}^{-(n-2)/4},$$

because $2\mathrm{e}a_*/m = (4\mathrm{e}/n)E_\varsigma < 1$ and $k \ge m/2$. Finally $\mathcal{G}^c \subseteq \{s \le 2\} \cup \{t < n/2,\ s \ge 3\} \cup \{\hat s > m/2,\ t \ge n/2\}$.

(iii) $a_* \le 2\mathrm{e}^{-(n-2)/(2n)} \le 1.218$ since $m \le n$, and $b_* \le 2n/m$. Next, $\Psi_0 \ge m^{-1}(1-1/n)^{n-1} > 1/(\mathrm{e}m)$. Bounding $m \le n$ in the first two terms and $m \ge 3$ in the third, each term of $n^4\delta_*$ is at most its value at $n = 301$, because $n^6\mathrm{e}^{-c_3 n}$, $n^6(2/3)^{n/2}$ and $n^4(4\mathrm{e}/n)^{3/2}\mathrm{e}^{-n/4}$ decrease for $n \ge 301$. At $n = 301$, $n^4\delta_* \le 1.6\cdot 10^{-4}$ and $\delta_* \le 1.9\cdot 10^{-14}$. □

Put

$$\chi_A = \mathrm{e}^{z_A b_*}\Big(1 + \frac{2(1+z_A)b_*}{n}\Big), \qquad \chi_B = \mathrm{e}^{z_B a_*}\Big(1 + \frac{2(1+z_B)a_*}{m}\Big), \qquad \bar g = 1 + \frac4m + \frac{4}{m^2} + 3\delta_*,$$

and $\Xi = \bar g\,\frac{m}{m-1}(1+\gamma_m a_*)\chi_A\varrho$.

Lemma 3.7. *(i) $\mathbb{E}\hat s \le \bar e_A := \min(\bar a(n)\chi_A, a_*) + (m/2 + a_*)\delta_*$ and $\mathbb{E}\hat t \le \bar e_B := \bar b(m)\chi_B + 2n\delta_*$.*

(ii) $\mathbb{E}\Psi(s,t) \ge \Psi_0(1 - \frac{m-1}{n-1}(2\bar e_B/n + n\delta_*/2))$, and $(1 + \mathbb{E}g) \le \bar g$.

(iii) $m\bar g\,\mathbb{E}[\Psi(t,s)/(s-1)] \le \Xi\Psi_0 + m\bar g\delta_*$, where the term with $s = 1$ is read as 0.

Proof. (i) By Lemma 3.3 with $f = 1$ and Lemma 3.4, $\mathbb{E}\hat{s} = \mathbb{E}\sigma_A \le \mathbb{E}\bar{a}(t)$. On $\mathcal{G}$, $\bar{a}(t) = \bar{a}(n)(n/t)(1+z_A)^{\hat{t}} \le \bar{a}(n)(1+2\hat{t}/n)(1+z_A)^{\hat{t}}$. By Lemma 3.4(iv) with $a = \bar{b}(s) \le b_*$, its conditional mean given $s \ge m/2$ is at most $\bar{a}(n)\chi_A$; it is also at most a_*. Off $\mathcal{G}$, $\bar{a}(t) \le m/2$, and $\bar{a}(t) \le a_*$ when $t \ge n/2$. For the second bound, $\mathbb{E}\hat{t} = \mathbb{E}\sigma_B \le \mathbb{E}\bar{b}(s)$, and $\bar{b}(s) = \bar{b}(m)(m/s)(1+z_B)^{\hat{s}} \le \bar{b}(m)(1+2\hat{s}/m)(1+z_B)^{\hat{s}}$ on $\mathcal{G}$, whose conditional mean given $t \ge n/2$ is at most $\bar{b}(m)\chi_B$. Off $\mathcal{G}$ we have $\bar{b}(s) \le n$.

(ii) Since Ψ decreases in s,

$$\Psi(s,t) \ge \Psi(m,t) = \Psi_0\Big(1 - \frac{\hat{t}}{t(n-1)}\Big)^{m-1} \ge \Psi_0\Big(1 - \frac{(m-1)\hat{t}}{t(n-1)}\Big).$$

Moreover $\mathbb{E}[\hat{t}/t] \le 2\mathbb{E}\hat{t}/n + (n/2)\mathbb{P}(t < n/2)$. Also $g \le 2/s + 1/s^2$, which is at most $4/m + 4/m^2$ when $s \ge m/2$, and $g \le 3$ always.

(iii) Use $\mathbb{E}[1/s] \ge 1/m$. On $\mathcal{G}$, $\Psi(t,s) \le \Psi(t,m) = \bar{a}(t)/m$ and $(m-1)/(s-1) \le 1 + \gamma_m\hat{s}$. By Lemma 3.3 with $f = \bar{a}(t)\mathbf{1}_{\{t \ge n/2\}}$, $\mathbb{E}[\bar{a}(t)\hat{s}; t \ge n/2] = \mathbb{E}[\bar{a}(t)\sigma_A; t \ge n/2] \le a_*\mathbb{E}[\bar{a}(t); t \ge n/2]$, and $\mathbb{E}[\bar{a}(t); t \ge n/2] \le \bar{a}(n)\chi_A + a_*\delta_*$ as in (i). Hence $\mathbb{E}[\Psi(t,s)/(s-1); \mathcal{G}] \le (1+\gamma_m a_*)(\bar{a}(n)\chi_A + a_*\delta_*)/(m(m-1))$. Off $\mathcal{G}$, $\Psi(t,s)/(s-1) \le 1/t \le 1/2$. Finally $\bar{a}(n) = m\varrho\Psi_0$, and $(1+\gamma_m a_*)a_*/(m-1) \le m/2$. □

Lemma 3.8. *With $e_f(t) = \mathbb{E}[f \mid t]$,*

$$(1 + \mathbb{E}g)\,(\nu_* - \tilde{\mu}) \ge 2\mathbb{E}\Psi(s,t) - \Xi\Psi_0 - m\bar{g}\delta_* - R_1 - R_2 - R_3 - R_4 - R_5,$$

where

$$R_1 = \mathbb{E}\big[\sigma_B\,(g(s,t) - g(s,t-1))\big] - 2\mathbb{E}\Psi_2, \qquad R_2 = \mathbb{E}\big[(\sigma_A(m,t) - \sigma_A)(g - g(m,t))\big],$$
$$R_3 = \big|\mathrm{Cov}(e_{\sigma_A+\sigma_B}(t), e_g(t))\big|, \qquad R_4 = m\bar{g}\,\mathbb{E}\big[(\sigma_B - \sigma_B(m,t))(1/s - 1/m)\big],$$
$$R_5 = m\bar{g}\,\big|\mathrm{Cov}(e_{\sigma_A+\sigma_B}(t), e_{1/s}(t))\big|.$$

Proof. We use (10). By Lemma 3.3,

$$\mathrm{Cov}(g,\hat{s}) = \mathbb{E}[\sigma_A(g(s-1,t) - g)] + \mathrm{Cov}(\sigma_A, g), \qquad \mathrm{Cov}(g,\hat{t}) = \mathbb{E}[\sigma_B(g(s,t-1) - g)] + \mathrm{Cov}(\sigma_B, g).$$

The first expectation is nonnegative, because g decreases in s. Write $\mathrm{Cov}(\sigma_A + \sigma_B, g)$ as the mean of the conditional covariances given t plus the covariance of the conditional means. Given t, both σ_B and g decrease in s, so $\mathrm{Cov}(\sigma_B, g \mid t) \ge 0$. For $X \le x_0$ and $Y \ge y_0$ we have $\mathrm{Cov}(X,Y) \ge -\mathbb{E}[(x_0 - X)(Y - y_0)]$. With $X = \sigma_A$, $x_0 = \sigma_A(m,t)$, $Y = g$ and $y_0 = g(m,t)$, this gives $\mathbb{E}[\mathrm{Cov}(\sigma_A, g \mid t)] \ge -R_2$. Hence $\mathrm{Cov}(g, \hat{s} + \hat{t}) \ge -R_1 - 2\mathbb{E}\Psi_2 - R_2 - R_3$.

For the loss, Lemma 3.3 gives

$$\mathrm{Cov}(\hat{s} + \hat{t}, 1/s) = \mathbb{E}\Big[\frac{\Psi(t,s)}{s-1}\Big] + \mathrm{Cov}(\sigma_A, 1/s) + \mathrm{Cov}(\sigma_B, 1/s),$$

because $\sigma_A/(s-1) - \sigma_A/s = \Psi(t,s)/(s-1)$. Given t, σ_A increases and $1/s$ decreases in s, so the conditional covariance is at most 0. Both σ_B and $1/s$ decrease, and $X \ge x_0$, $Y \ge y_0$ imply $\mathrm{Cov}(X,Y) \le \mathbb{E}[(X - x_0)(Y - y_0)]$. Therefore $\mathrm{Cov}(\hat{s} + \hat{t}, 1/s) \le \mathbb{E}[\Psi(t,s)/(s-1)] + (R_4 + R_5)/(m\bar{g})$. Since $(1 + \mathbb{E}g) \le \bar{g}$ and $1/\mathbb{E}[1/s] \le m$, it follows that $(1 + \mathbb{E}g) \cdot \mathcal{L} \le m\bar{g}\,\mathbb{E}[\Psi(t,s)/(s-1)] + R_4 + R_5$, and if $\mathcal{L} < 0$ this bound holds trivially. Lemma 3.7(iii) completes the proof. □

Lemma 3.9. *Put $z' = 2z_B + z_B^2$, $\vartheta = 2/m + 2/n$ and $d_A = 1 - (1/m + 2/n)a_*$, and define*

$$\chi_2 = M_0(z', a_*) + \frac{4}{m}M_1(z', a_*) + \frac{4}{m^2}M_2(z', a_*), \qquad \mathcal{X} = M_2(z_B, a_*) + \frac{2}{m}M_3(z_B, a_*),$$

$$\Delta_\sigma = a_*\Big(\frac{2}{n}+\frac{2}{m}\Big) + b_*\Big(\frac{2.014}{n}+\frac{4m}{n^2}\Big) + 3n\delta_* + \frac{(\ell_A+\ell_B)\vartheta a_* + 3n\delta_*}{d_A},$$
$$\Delta_g = \Psi_0\Big(26.5\Big(1+\frac{2.02}{m}\Big)\frac{m}{n^2} + 6.61\Big(1+\frac{2}{m}\Big)\frac{\vartheta^2 a_*}{d_A}\Big) + 3\delta_* + \frac{6\delta_*}{d_A}, \qquad \Delta_q = \frac{4\vartheta a_*/m^2 + 2\delta_*}{d_A},$$

with $\ell_A = a_*(2/m + 4n/m^2)$ *and* $\ell_B = b_*\vartheta$. *Let* $\mathcal{V} = \bar e_B + b_*^2\bar e_A + n^2\delta_*$. *Then*

$$R_1 \le 8.12\frac{m}{n^2}\bar b(m)\chi_2\Psi_0 + 3n\delta_*, \quad R_2 \le 2\Big(1+\frac{2}{m}\Big)\Big(\frac{1}{m}+\frac{2n}{m(m-1)}\Big)\Big(\frac{1}{m}+\frac{2}{n}\Big)a_*\mathcal{X}\Psi_0 + \frac{3m}{2}\delta_*,$$
$$R_3 \le \Delta_\sigma\Delta_g\mathcal{V} + 24n\delta_*, \qquad R_4 \le \frac{2\bar g}{m}\Big(\frac{1}{m}+\frac{2}{n}\Big)\bar b(m)\mathcal{X} + mn\bar g\delta_*,$$
$$R_5 \le m\bar g\,(\Delta_\sigma\Delta_q\mathcal{V} + 8n\delta_*).$$

Proof. On $\mathcal{G}$ we use $s \ge m/2$, $t \ge n/2$, $(n/2-1)^2 \ge n^2/4.06$ and $(n/2-1)(n/2-2) \ge n^2/4.09$, together with $\Psi(s,t) \le \Psi(s,n) = \Psi_0(m/s)(1+z_B)^{\hat s}$, so that $\Psi(s,t) \le \Psi_0(1+2\hat s/m)(1+z_B)^{\hat s} \le 3.303\Psi_0$ there. We also use $\sigma_B \le \bar b(s) = \bar b(m)(m/s)(1+z_B)^{\hat s}$ and $\Psi_2 \le \Psi(s,t)/s$. Off $\mathcal{G}$ every integrand below is bounded by the number written next to δ_*, by Lemma 3.6(ii).

R_1: from $\Psi(s,t-1)/\Psi(s,t) = (1-1/(t-1)^2)^{s-1}$ and $\Psi_2(s,t-1)/\Psi_2(s,t) = (1-2/((t-1)(t-2)))^{s-1}$,

$$g(s,t) - g(s,t-1) \le 2\Psi(s,t)\frac{s-1}{(t-1)^2} + \Psi_2(s,t)\frac{2(s-1)}{(t-1)(t-2)}.$$

On $\mathcal{G}$ the first term contributes at most $8.12(m/n^2)\mathbb{E}[\sigma_B\Psi]$. Also $\sigma_B\Psi \le \bar b(m)\Psi_0(1+2\hat s/m)^2(1+z')^{\hat s}$, whose conditional mean is at most $\bar b(m)\Psi_0\chi_2$ by Lemma 3.4(iv). The second term contributes at most $b_* \cdot 8.18(m/n^2)\mathbb{E}\Psi_2 \le 2\mathbb{E}\Psi_2$, since $b_* \le 2n/m$ and $n \ge 301$.

R_2: from $(1-1/s)/(1-1/m) = 1-\hat s/(s(m-1))$ we get $\sigma_A(m,t)-\sigma_A \le \bar a(t)\hat s(1/m+2n/(m(m-1)))$ on $\mathcal{G}$. From $\Psi(m,t)/\Psi(s,t) \ge 1-\hat s/m-\hat s/t$ and the analogous inequality for Ψ_2, $g-g(m,t) \le 2(1+2/m)\Psi(s,t)\hat s(1/m+2/n)$. Then bound $\mathbb{E}[\bar a(t)\hat s^2(1+2\hat s/m)(1+z_B)^{\hat s} \mid t]$ by $a_*\mathcal{X}$, using Lemma 3.4(iv) with $a = \bar a(t) \le a_*$.

R_4: $\sigma_B - \sigma_B(m,t) \le \sigma_B\hat s(1/m+2/n)$ and $1/s - 1/m = \hat s/(ms) \le 2\hat s/m^2$ on $\mathcal{G}$, and the conditional mean of $\hat s^2(1+2\hat s/m)(1+z_B)^{\hat s}$ is at most $\mathcal{X}$.

R_3 and R_5: for $\lceil n/2\rceil \le t < n$ we bound $|e_f(t+1) - e_f(t)|$ for $f = \sigma_A+\sigma_B$, g and $1/s$. Write it as $\mathbb{E}[f(s,t+1)-f(s,t) \mid t+1]$ plus the change in $\mathbb{E}[f(\cdot,t) \mid \cdot]$ from t to $t+1$, and bound the second part with Lemma 3.5. For $f = \sigma_A+\sigma_B$, apply that lemma to σ_A and σ_B separately; each is monotone in s.

On $s \ge m/2$:

- $|\Delta_t\sigma_A| \le \sigma_A(1/(t+1)+1/s) \le a_*(2/n+2/m)$;
- $|\Delta_t\sigma_B| \le \sigma_B(s,t+1)(1/(t-1)+(s-1)/t^2) \le b_*(2.014/n+4m/n^2)$;
- the increments in s of σ_A and σ_B are at most ℓ_A and ℓ_B;
- $|\Delta_t g| \le 3.303\Psi_0(8m/n^2)(1+2.015/m)$;
- the increments of g in s are at most $6.61\Psi_0\vartheta(1+2/m)$;
- the increments of $1/s$ are at most $4/m^2$.

With $a = \mathbb{E}[\hat s \mid t] \le a_*$, Lemma 3.5 gives the bounds Δ_σ, Δ_g and Δ_q for the increments of $e_{\sigma_A+\sigma_B}$, e_g and $e_{1/s}$ on $[\lceil n/2\rceil, n]$. Extend each e_f to $t < \lceil n/2\rceil$ by its value at $\lceil n/2\rceil$. This changes a covariance by at most $6\|e_f\|\|e_{f'}\|\mathbb{P}(t < n/2)$, by the triangle inequality applied to $\mathbb{E}[\phi\psi]$ and to

$\mathbb{E}\phi\,\mathbb{E}\psi$; here $\|e_{\sigma_A+\sigma_B}\| < n$, $\|e_g\| \le 3$ and $\|e_{1/s}\| \le 1$. For a function ϕ whose increments are at most c, $\operatorname{Var}\phi(t) \le c^2 \operatorname{Var} t$, so the covariance of the extended functions is at most $\Delta_f\Delta_{f'}\operatorname{Var} t$.

It remains to show $\operatorname{Var} t \le \mathcal{V}$. We have $\operatorname{Var} t = \mathbb{E}\operatorname{Var}(\hat{t} \mid s) + \operatorname{Var} e_{\hat{t}}(s)$. The first term is at most $\mathbb{E}\hat{t} \le \bar{e}_B$ by Lemma 3.4(iii). For the second, $\operatorname{Var} e_{\hat{t}}(s) \le \mathbb{E}(e_{\hat{t}}(s) - e_{\hat{t}}(m))^2$. When $m/2 \le s \le m$, both values lie in $[0, b_*]$ and differ only if $s < m$; so the second term is at most $b_*^2\mathbb{P}(\hat{s} \ge 1) + n^2\delta_*$, and $\mathbb{P}(\hat{s} \ge 1) \le \bar{e}_A$. □

The comparison of $\Psi(n,m)$ with $\Psi_0 = \Psi(m,n)$ is the elementary inequality

$$\log\frac{\Psi(n,m)}{\Psi(m,n)} \le -(\varsigma - 1) - \log\varsigma \le 0 \qquad (2 \le m \le n,\ \varsigma = n/m). \tag{11}$$

To prove it, let $\mathcal{F}(x,y) = \log\Psi(y,x) - \log\Psi(x,y)$ for real $y \ge x \ge 2$. Then $\mathcal{F}(x,x) = 0$ and

$$\partial_y\mathcal{F} = \log(1 - 1/x) - \frac{x-1}{y(y-1)} - \frac{1}{y} \le -\frac{1}{x} - \frac{1}{y}.$$

Integrating from $y = m$ to $y = n$ at $x = m$ gives (11). In words, a vertex of the larger part is the more likely to be left out of a random subtree. Together with the factor 2 in Lemma 3.1(iii), this inequality produces the margin in Lemma 3.10.

Lemma 3.10. *Let $3 \le m \le n$ and $n \ge 301$. Then $(1 + \mathbb{E}g)\,(\nu_* - \tilde{\mu}) \ge 0.6\,\Psi_0 > m(m+n)(1 + \mathbb{E}g)/(2\mathcal{Z})$.*

Proof. The terms of Lemmas 3.8 and 3.9 that contain δ_* are the δ_*-terms displayed in Lemmas 3.7–3.9 and their products with the bounded factors Δ_σ, Δ_g/Ψ_0, $m\Delta_q/\Psi_0$ and $\mathcal{V}/n$; there are fewer than fifteen, and each is at most $30n^2\delta_*$ (the largest is $8mn\bar{g}\delta_* \le 22n^2\delta_*$). Since $\Psi_0 \ge 1/(em)$, each is at most $30\mathrm{e}\,n^3\delta_*\Psi_0 \le 82\,(n^4\delta_*/n)\Psi_0 \le 6\cdot 10^{-5}\Psi_0$ by Lemma 3.6(iii), so together they are less than $10^{-3}\Psi_0$. We use $\sup_{y>0} y^j\mathrm{e}^{-y/2} = (2j/\mathrm{e})^j$ and $E_\varsigma = \mathrm{e}^{1/m}\mathrm{e}^{-\varsigma/2}$.

Case $\varsigma \le 2$. Then $m \ge 151$, $\bar{g} \le 1.0267$, $\gamma_m \le 1/75$, $z_A \le 1/150$ and $b_* = 2\varsigma\,\mathrm{e}^{-1/(2\varsigma)+1/n} \le 3.126$. Hence $\chi_A \le 1.0425$ and, as $\varrho \le 1$ by (11), $\Xi \le 1.095$. Also $\bar{b}(m) \le \varsigma\,\mathrm{e}^{-1/\varsigma+1/n} \le 1.218$ and $\chi_B \le 1.021$, so $2\mathbb{E}\Psi(s,t) \ge 1.983\Psi_0$. The bounds of Lemma 3.9 give the following, using

- $m\bar{b}(m)/n^2 \le 1/n$;
- $(n-2)a_*/m \le 2E_\varsigma \le 1.222$;
- $\mathcal{X} \le 2.321a_*$;
- $\chi_2 \le 1.0415$;
- $\Delta_\sigma \le 0.0695$;
- $\Delta_g \le 0.0926\Psi_0$;
- $\mathcal{V} \le 4.87$, since $8\varsigma\mathrm{e}^{-1/\varsigma-\varsigma/2}$ increases on $[1,2]$;
- $m(\Delta_q - 2\delta_*/d_A)/\Psi_0 \le 4\mathrm{e}\vartheta a_*/d_A \le 0.268$, since $\Psi_0 \ge 1/(em)$; the term $2\delta_*/d_A$ is part of the δ_*-budget.

The bounds are

$$R_1 \le 0.0281\Psi_0, \quad R_2 \le 0.0031\Psi_0, \quad R_3 \le 0.0314\Psi_0, \quad R_4 \le 0.0773\Psi_0, \quad R_5 \le 0.0932\Psi_0.$$

Thus $(1+\mathbb{E}g)\,(\nu_*-\tilde{\mu}) \geq (1.983-1.095-0.234-0.001)\Psi_0 > 0.6\Psi_0$.

Case $\varsigma > 2$. Here $a_* \leq E_\varsigma \leq 0.514$, $a_*/m = 2E_\varsigma/n \leq 1.027/n$, and $b_* \leq 2\varsigma$. Since $E_\varsigma = \mathrm{e}^{1/m}\mathrm{e}^{-\varsigma/2}$ and $\varsigma > 2$, we have $E_\varsigma \varsigma^j \leq \mathrm{e}^{1/3}\sup_{y\geq 2} y^j \mathrm{e}^{-y/2}$, which gives $E_\varsigma\varsigma^j \leq 0.514$, 1.027, 3.022, 15.01 and 104.7 for $j = 0,\dots,4$.

First, $\Xi \leq 0.221$. If $m \geq 76$, then $\log(\chi_A\varrho) \leq 2\varsigma/(m-1)+4(1+z_A)/m-(\varsigma-1)-\log\varsigma \leq -1.586$ by (11), and the prefactor of Ξ is at most 1.073, because $\varsigma \geq 301/m$ gives $\gamma_m a_* = 2\gamma_m \mathrm{e}^{1/m}\mathrm{e}^{-\varsigma/2}/\varsigma \leq 0.005$ for $m \geq 76$. If $m < 76$, then $\varsigma > 4$. In that case $\log\varrho \leq -\log\varsigma - (n-1)(1/m + 1/(2m^2)) + (m-1)/(n-1)$ and $z_A b_* \leq 2n/(m(m-1))$, so

$$\log(\chi_A\varrho) \leq -\log\varsigma - \varsigma\Big(1+\frac{1}{2m}-\frac{2}{m-1}\Big)+\frac{1}{m}+\frac{1}{2m^2}+\frac{4(1+z_A)}{m}+\frac{1}{4}.$$

For $m \geq 6$ this gives $\chi_A\varrho \leq 0.077$; the prefactor of Ξ is at most 2.14, its value at $m = 6$ (where $\varsigma \geq 50$ and $a_* < 10^{-10}$) being the largest, so $\Xi \leq 0.18$. For $m \leq 5$ we have $\varsigma \geq 60$ and $\Xi < 10^{-4}$.

Next, $\chi_B \leq 1.0086$, so $2\mathbb{E}\Psi(s,t) \geq 1.986\Psi_0$. Using $1/m = \varsigma/n$, $\bar{b}(m) \leq \varsigma$, $\mathcal{X} \leq 3.406 a_*$ and $\chi_2 \leq 1.0242$, we get

- $R_1 \leq 0.0277\Psi_0$;
- $R_2 \leq 90.9(1+3\varsigma)E_\varsigma^2\Psi_0/n^2 \leq 0.0019\Psi_0$;
- $R_4 \leq 8\cdot 3.406\,\bar{g}E_\varsigma\varsigma\Psi_0/n \leq 0.097\Psi_0$, since $\bar{g}E_\varsigma\varsigma \leq E_\varsigma\varsigma + 4E_\varsigma\varsigma^2/n + 4E_\varsigma\varsigma^3/n^2 \leq 1.068$.

Moreover:

- $\Delta_\sigma \leq (4.028\varsigma + 11.552)/n$;
- $\Delta_g \leq \Psi_0(44.35/\varsigma + 200.4E_\varsigma\varsigma/n)/n$;
- $\mathcal{V} \leq 1.0086\varsigma + 8.215$, since $b_*^2 a_* \leq 8\varsigma E_\varsigma \leq 8.214$;
- $m(\Delta_q - 2\delta_*/d_A)/\Psi_0 \leq 65.9E_\varsigma/n$; the term $2\delta_*/d_A$ is part of the δ_*-budget above.

With $\varsigma \leq n/3$ these give $R_3 \leq 0.247\Psi_0$ and $R_5 \leq 0.081\Psi_0$. Hence

$$(1+\mathbb{E}g)\,(\nu_*-\tilde{\mu}) \geq (1.986-0.221-0.455-0.001)\Psi_0 > 1.3\Psi_0.$$

In both cases the last inequality of the lemma holds, because $\mathcal{Z} \geq w(m,n) = m^{n-1}n^{m-1} \geq 3^{300}$. □

3.5 Proof of Theorem 3

Proposition 6. *For all integers* $1 \leq m \leq n \leq 300$ *with* $n \geq 2$, $\mu(K_{m,n}+e_B) > \mu(K_{m,n})$.

Proof. By (5), the change has the sign of $\Sigma^{e_B}Z - \Sigma Z^{e_B}$, where Z, Σ, Z^{e_B} and Σ^{e_B} are the integers given by (7), (8) and the corresponding order sums at $\lambda = 1$. The script `check_kmn_exact.py` evaluates these integers exactly for all 45,149 pairs (m,n) and finds the sign positive in every case. □

Proof of Theorem 3. For $n \leq 300$ this is Proposition 6. Let $n \geq 301$.

Case $m = 1$. Put $x = 2^{n-2}$. The star $K_{1,n}$ has $Z = 4x + n$ subtrees of total order $\Sigma = (2n+4)x + n$, and the subtrees of $K_{1,n}+e_B$ through e_B number $Z^{e_B} = 2x+1$ with total order $\Sigma^{e_B} = (n+4)x+2$. Hence $\Sigma^{e_B}Z - \Sigma Z^{e_B} = 8x^2 + (n^2+4)x + n > 0$, and (5) gives the claim. (This is also a case of [3, Theorem 4.1].)

Case $m = 2$. This is Theorem 1.

Case $m \geq 3$. By Lemma 3.10, $\nu_* - \tilde{\mu} > m(m+n)/(2\mathcal{Z})$, and Lemma 3.2 applies. □

4 Near-complete hosts

4.1 Setting

Throughout this section $\lambda > 0$ is fixed, $\theta = 1/(\mathrm{e}\lambda)$, $N \geq 8$, and H is a graph on a vertex set V of size N whose complement F has maximum degree at most D, where $D \leq N/8$. Let L be the Laplacian of F. For $U \subseteq V$ let L_U be the Laplacian of the induced subgraph $F[U]$, extended by zero to an $N \times N$ matrix, and let $L_{\partial U} = L - L_U$, the Laplacian of the edges of F that meet $V \setminus U$. All three matrices are positive semidefinite. We write $\|\cdot\|$ for the operator norm and $\|\cdot\|_1$ for the trace norm, which equals the trace on positive semidefinite matrices. Since the Laplacian of a graph of maximum degree at most D has norm at most $2D$,

$$\|L_U\| \leq 2D, \qquad \operatorname{tr} L_{\partial U} \leq 2D\,|V \setminus U|, \qquad \operatorname{tr} L \leq ND. \tag{12}$$

Let T be a random subtree of H with $\mathbb{P}(T) \propto \lambda^{|E(T)|}$, and let $S = V(T)$, $K = |S|$ and $J = N - K$, the number of omitted vertices. For a k-set U with $k \geq 1$ put

$$\Phi(U) = \det\bigl(\mathrm{Id} - L_U/k\bigr),$$

the determinant of an $N \times N$ matrix.

4.2 Determinants and their comparison

Lemma 4.1. *For every $U \subseteq V$ with $|U| = k \geq 1$, $\tau(H[U]) = k^{k-2}\Phi(U)$. Consequently the subtrees of H with k vertices have total weight $w_k = \lambda^{k-1}k^{k-2}\sum_{|U|=k}\Phi(U)$; conditionally on $K = k$, the vertex set satisfies $\mathbb{P}(S = U \mid K = k) = \Phi(U)/\sum_{|U'|=k}\Phi(U')$, and given $S = U$ the subtree T is a uniform spanning tree of $H[U]$.*

Proof. For $k = 1$ both sides equal 1. For $k \geq 2$, the Laplacian of $H[U]$ on $\mathbb{R}^U$ is $k\mathrm{Id}_U - \mathbf{1}\mathbf{1}^{\mathsf{T}} - L_{F[U]}$. It vanishes on the constant vector, and on the orthogonal complement W of the constants, which $L_{F[U]}$ preserves, it acts as $k\mathrm{Id} - L_{F[U]}$. By the matrix-tree theorem [7], $\tau(H[U])$ is $1/k$ times the product of the eigenvalues on W, that is, $k^{-1}\det(k\mathrm{Id}_U - L_{F[U]})/k = k^{k-2}\det(\mathrm{Id}_U - L_{F[U]}/k)$, since $k\mathrm{Id}_U - L_{F[U]}$ acts as k on the constants. Extending by zero outside U adds unit eigenvalues, so this is $k^{k-2}\Phi(U)$. Every tree with vertex set U has weight λ^{k-1}, which gives the remaining statements. □

Lemma 4.2. *Let X_0, X_1 be positive semidefinite with $\|X_0\|, \|X_1\| \leq 1/2$. Then*

$$\bigl|\log\det(\mathrm{Id} - X_1) - \log\det(\mathrm{Id} - X_0)\bigr| \leq 2\|X_1 - X_0\|_1, \qquad \log\det(\mathrm{Id} - X_0) \geq -2\operatorname{tr} X_0.$$

Proof. The matrices $X_t = (1-t)X_0 + tX_1$, $0 \leq t \leq 1$, are positive semidefinite with $\|X_t\| \leq 1/2$, so $\|(\mathrm{Id} - X_t)^{-1}\| \leq 2$. The derivative of $\log\det(\mathrm{Id} - X_t)$ is $-\operatorname{tr}\bigl((\mathrm{Id} - X_t)^{-1}(X_1 - X_0)\bigr)$, of absolute value at most $2\|X_1 - X_0\|_1$; integrate over t. The second inequality follows from $\log(1-x) \geq -2x$ on $[0, 1/2]$, applied to the eigenvalues of X_0. □

Lemma 4.3. *Let $U \subseteq V$ with $|U| = k \geq N/2$, and $j = N - k$. Then $\Phi(U) > 0$,*

$$\bigl|\log\Phi(U) - \log\Phi(V)\bigr| \leq \frac{6Dj}{k} \leq \frac{12Dj}{N}, \qquad \Phi(V) \geq \mathrm{e}^{-2D}, \tag{13}$$

and, for every transposition σ of V,

$$\bigl|\log\Phi(\sigma U) - \log\Phi(U)\bigr| \leq \frac{8D}{k} \leq \frac{16D}{N}. \tag{14}$$

Proof. By (12) and $D \le N/8 \le k/4$, the matrices L_U/k, L/N and $L_{\sigma U}/k$ are positive semidefinite of norm at most $1/2$, so $\Phi(U) \ge 2^{-N} > 0$ and Lemma 4.2 applies. For (13),

$$\Big\| \frac{L_U}{k} - \frac{L}{N} \Big\|_1 \le \frac{\operatorname{tr} L_{\partial U}}{k} + \operatorname{tr} L\Big(\frac{1}{k} - \frac{1}{N}\Big) \le \frac{2Dj}{k} + \frac{N D\, j}{kN} = \frac{3Dj}{k},$$

and $\log \Phi(V) \ge -2 \operatorname{tr} L/N \ge -2D$. For (14), if $\sigma = (w\, w')$ with $w \in U$ and $w' \notin U$ (otherwise $\sigma U = U$), then $L_U - L_{U\setminus\{w\}}$ and $L_{\sigma U} - L_{U\setminus\{w\}}$ are the Laplacians of at most D edges each, so $\|L_{\sigma U} - L_U\|_1 \le 4D$. □

4.3 The number of omitted vertices

For $0 \le j \le N-1$ let $Q_j = w_{N-j}/w_N = \mathbb{P}(J = j)/\mathbb{P}(J = 0)$.

Lemma 4.4. *For $0 \le j \le N/2$,*

$$\Big|\log \frac{j!\, Q_j}{\theta^j}\Big| \le \frac{3j + 3j^2 + 12Dj}{N}, \qquad Q_j \le \frac{(\mathrm{e}^{5/2}\theta)^j}{j!}. \tag{15}$$

For $N/2 < j \le N-1$,

$$Q_j \le \mathrm{e}^{2D} \lambda^{-j} \frac{j+1}{j!}. \tag{16}$$

Proof. Let $k = N - j$. By Lemma 4.1, Q_j equals $\lambda^{-j}\binom{N}{j}(N-j)^{N-j-2}/N^{N-2}$, the same ratio for the complete graph, multiplied by the average of $\Phi(U)/\Phi(V)$ over the k-sets U. The complete-graph ratio equals $(\theta^j/j!)\,\mathrm{e}^j \prod_{i=0}^{j-1}(1 - i/N)\,(1 - j/N)^{N-j-2}$. For $j \le N/2$ use $-x - x^2 \le \log(1-x) \le -x$ on $[0, 1/2]$ and $N - j - 2 \ge 0$: the logarithm of $\mathrm{e}^j \prod_{i<j}(1 - i/N)(1 - j/N)^{N-j-2}$ lies between $(j^2 + 5j)/(2N) - j^3/(3N^2) - j^2/N \ge -7j^2/(6N)$ and $(j^2 + 5j)/(2N)$, so its absolute value is at most $3(j + j^2)/N$ and it is at most $j/4 + 5j/16$. By (13) each ratio $\Phi(U)/\Phi(V)$, and hence their average, lies in $[\mathrm{e}^{-12Dj/N}, \mathrm{e}^{12Dj/N}]$, and $12Dj/N \le 3j/2$. This proves (15). For $j > N/2$, bound $\tau(H[U])$ by Cayley's count k^{k-2} for the complete graph on U and use $\Phi(V) \ge \mathrm{e}^{-2D}$: $Q_j \le \mathrm{e}^{2D}\lambda^{-j}\binom{N}{j}k^{k-2}/N^{N-2}$. For $k \ge 2$ this is at most $\mathrm{e}^{2D}\lambda^{-j}/j!$, since $\binom{N}{j} \le N^j/j!$ and $k^{k-2} \le N^{k-2}$. For $k = 1$ it is $\mathrm{e}^{2D}\lambda^{-j}N/N^{N-2} \le \mathrm{e}^{2D}\lambda^{-j}N/(N-1)!$, which is (16) with $j = N-1$. □

Lemma 4.5. *Let ζ be a Poisson random variable with mean θ. For every integer $r \ge 0$ there is a constant $C_{\lambda,r}$, depending only on λ and r, such that for all N, D and H as in the setting,*

$$\big|\mathbb{E}[J^r] - \mathbb{E}[\zeta^r]\big| \le C_{\lambda,r}\frac{D+1}{N}, \qquad \mathbb{P}(J > N/2) \le C_{\lambda,r} N^{-r}. \tag{17}$$

In particular

$$\big|\operatorname{Var}(J) - \theta\big| \le C'_\lambda \frac{D+1}{N} \tag{18}$$

for a constant C'_λ depending only on λ.

Proof. Let $N_\lambda = \max(8, \lceil 4\mathrm{e}^{3/2}/\lambda \rceil)$. For $N < N_\lambda$ both bounds hold with a large constant, since $J \le N$. Let $N \ge N_\lambda$. For $j > N/2$, (16), $j! \ge (j/\mathrm{e})^j$ and $\mathrm{e}^{2D} \le \mathrm{e}^{N/4} \le \mathrm{e}^{j/2}$ give

$$Q_j \le (j+1)\Big(\frac{\mathrm{e}^{3/2}}{\lambda j}\Big)^j \le N\Big(\frac{2\mathrm{e}^{3/2}}{\lambda N}\Big)^j \le N\, 2^{-N/2},$$

because the base is at most $1/2$. Hence, for every $r \ge 0$, $\sum_{j>N/2} j^r Q_j \le N^{r+2} 2^{-N/2}$, which is smaller than any fixed power of N. Since $\mathbb{P}(J = j) = Q_j / \sum_i Q_i$ and $\sum_i Q_i \ge Q_0 = 1$, the

case $r = 0$ proves the second bound in (17). For $j \le N/2$, write $t_j = \log(j!Q_j/\theta^j)$; by (15), $|t_j| \le (3j + 3j^2 + 12Dj)/N$ and $t_j \le 5j/2$, so $|\mathrm{e}^{t_j} - 1| \le |t_j|\mathrm{e}^{5j/2}$ and

$$\sum_{j \le N/2} j^r \Big| Q_j - \frac{\theta^j}{j!} \Big| \le \frac{1}{N} \sum_{j \ge 0} j^r (3j + 3j^2 + 12Dj) \frac{(\mathrm{e}^{5/2}\theta)^j}{j!} \le C \frac{D+1}{N}.$$

Together with the Poisson tail $\sum_{j > N/2} j^r \theta^j/j!$, this gives $\sum_j j^r Q_j = \mathrm{e}^\theta \mathbb{E}[\zeta^r] + O((D+1)/N)$ for each r, and for $r = 0$ it gives $\sum_j Q_j = \mathrm{e}^\theta + O((D+1)/N)$. Since $\sum_j Q_j \ge 1$,

$$\big|\mathbb{E}[J^r] - \mathbb{E}[\zeta^r]\big| = \frac{\big|\sum_j j^r Q_j - \mathrm{e}^\theta \mathbb{E}[\zeta^r] - \mathbb{E}[\zeta^r](\sum_j Q_j - \mathrm{e}^\theta)\big|}{\sum_j Q_j} \le C_{\lambda,r} \frac{D+1}{N}.$$

Finally $\mathrm{Var}(\zeta) = \theta$, and (18) follows from the cases $r = 1, 2$ because $\mathbb{E}J$ and $\mathbb{E}\zeta$ are bounded. □

4.4 Omission probabilities

For $N/2 \le k \le N$ and $j = N - k$ write, for vertices w and $x \ne y$,

$$\omega_w = \mathbb{P}(w \notin S \mid K = k), \qquad \omega_{xy} = \mathbb{P}(x \notin S,\ y \notin S \mid K = k).$$

Lemma 4.6. *For $N/2 \le k \le N$, every vertex w and every pair $x \ne y$,*

$$\mathrm{e}^{-16D/N} \frac{j}{N} \le \omega_w \le \mathrm{e}^{16D/N} \frac{j}{N}, \qquad \mathrm{e}^{-32D/N} \frac{j(j-1)}{N(N-1)} \le \omega_{xy} \le \mathrm{e}^{32D/N} \frac{j(j-1)}{N(N-1)}.$$

Consequently, for distinct vertices u, v, the probability $\psi_k = \mathbb{P}(u, v \in S \mid K = k)$ satisfies

$$\Big|\psi_k - \frac{k(k-1)}{N(N-1)}\Big| \le C_\omega \frac{Dj}{N^2}, \qquad 1 - \psi_k \le 2\mathrm{e}^2 \frac{j}{N}, \tag{19}$$

with the absolute constant $C_\omega = 32\mathrm{e}^2 + 16\mathrm{e}^4$.

Proof. Let $w \ne w'$ and $\sigma = (w\, w')$. By Lemma 4.1 the conditional law of S given $K = k$ is proportional to Φ, and σ maps the k-sets avoiding w bijectively onto those avoiding w'. By (14), $\omega_{w'} = \sum_{U \not\ni w} \Phi(\sigma U) / \sum_{|U|=k} \Phi(U)$ lies between $\mathrm{e}^{-16D/N}\omega_w$ and $\mathrm{e}^{16D/N}\omega_w$. Summing over w' and using $\sum_{w'} \omega_{w'} = \mathbb{E}[J \mid K = k] = j$ gives the first pair of bounds. For pairs, at most two transpositions carry $\{x, y\}$ to any other pair $\{x', y'\}$, and their product maps the k-sets avoiding x, y onto those avoiding x', y', changing each Φ by a factor in $[\mathrm{e}^{-32D/N}, \mathrm{e}^{32D/N}]$; summing over pairs and using $\sum_{x<y} \omega_{xy} = \binom{j}{2}$ gives the second pair of bounds. Now $\psi_k = 1 - \omega_u - \omega_v + \omega_{uv}$, while $1 - 2j/N + j(j-1)/(N(N-1)) = k(k-1)/(N(N-1))$. Since $16D/N \le 2$ and $32D/N \le 4$, we have $|\mathrm{e}^{\pm x} - 1| \le x\mathrm{e}^x$, so $|\omega_w - j/N| \le 16\mathrm{e}^2 Dj/N^2$ and, using $j(j-1)/(N(N-1)) \le j/(2N)$, $|\omega_{uv} - j(j-1)/(N(N-1))| \le 16\mathrm{e}^4 Dj/N^2$. This gives the first bound in (19), and $1 - \psi_k \le \omega_u + \omega_v \le 2\mathrm{e}^2 j/N$. □

4.5 Edge inclusion at fixed order

Fix an edge uv of H and let $b = \mathbf{1}_u - \mathbf{1}_v$. For $k > 0$ and a symmetric matrix Y with $\|Y\| < k$ put

$$R_k(Y) = b^{\mathsf{T}} (k\mathrm{Id} - Y)^{-1} b.$$

Lemma 4.7. *Let $U \ni u, v$ with $|U| = k > 2D$. The proportion of spanning trees of $H[U]$ that contain uv equals $R_k(L_U)$.*

Proof. On the space W of zero-sum vectors on U the Laplacian of $H[U]$ acts as the positive definite operator $\Gamma = k\mathrm{Id} - L_{F[U]}$ (proof of Lemma 4.1), and the Laplacian of $H[U] - uv$ acts as $\Gamma - bb^{\mathsf{T}}$, with $b \in W$. By the matrix-tree theorem and the matrix determinant lemma,

$$\frac{\tau(H[U]-uv)}{\tau(H[U])} = \frac{\det(\Gamma - bb^{\mathsf{T}})}{\det\Gamma} = 1 - b^{\mathsf{T}}\Gamma^{-1}b,$$

so the proportion of trees through uv is $b^{\mathsf{T}}\Gamma^{-1}b$; this is Kirchhoff's identification of edge inclusion with effective resistance [7]. On $\mathbb{R}^V$ the operator $(k\mathrm{Id} - L_U)^{-1}$ agrees with Γ^{-1} on W, and b lies in W, so $b^{\mathsf{T}}\Gamma^{-1}b = R_k(L_U)$. □

Lemma 4.8. *Let $\kappa = R_N(L) - 2/N$. Then $0 \le \kappa \le 16D/(3N^2)$, and for $0 \le j \le N/2$,*

$$\pi_j := \mathbb{P}(uv \in T \mid J = j) = \frac{2}{N} + \kappa - \frac{2j}{N(N-1)} + \varepsilon_j, \qquad |\varepsilon_j| \le C_\pi \frac{Dj}{N^3}, \tag{20}$$

where C_π is an absolute constant. In particular $\varepsilon_0 = 0$.

Proof. Let $k = N - j \ge N/2$ and $\Gamma_k = k\mathrm{Id} - L$, whose eigenvalues are at least $k - 2D \ge k/2$. By Lemmas 4.1 and 4.7,

$$\pi_j = \mathbb{E}\big[\mathbf{1}_{\{u,v\in S\}} R_k(L_S) \mid K = k\big].$$

Removing the boundary. Since $k\mathrm{Id} - L_S = \Gamma_k + L_{\partial S}$ with $L_{\partial S} \succeq 0$, conjugating the inequalities $0 \preceq \mathrm{Id} - (\mathrm{Id}+Y)^{-1} \preceq Y$ for $Y = \Gamma_k^{-1/2} L_{\partial S} \Gamma_k^{-1/2}$ by $\Gamma_k^{-1/2}$ gives

$$0 \preceq \Gamma_k^{-1} - (\Gamma_k + L_{\partial S})^{-1} \preceq \Gamma_k^{-1} L_{\partial S} \Gamma_k^{-1}, \qquad \text{so} \qquad 0 \le R_k(L) - R_k(L_S) \le b^{\mathsf{T}}\Gamma_k^{-1} L_{\partial S} \Gamma_k^{-1} b.$$

An edge xy of F contributes to $L_{\partial S}$ exactly when $x \notin S$ or $y \notin S$, an event of conditional probability at most $\omega_x + \omega_y \le 2\mathrm{e}^2 j/N$ by Lemma 4.6. Hence $\mathbb{E}[L_{\partial S} \mid K = k] \preceq (2\mathrm{e}^2 j/N) L$, and with $\|b\|^2 = 2$, $\|L\| \le 2D$ and $\|\Gamma_k^{-1}\| \le 2/k \le 4/N$,

$$0 \le \psi_k R_k(L) - \pi_j \le \frac{2\mathrm{e}^2 j}{N}\, b^{\mathsf{T}}\Gamma_k^{-1} L \Gamma_k^{-1} b \le \frac{2\mathrm{e}^2 j}{N} \cdot 2 \cdot 2D \cdot \frac{16}{N^2} = 128\mathrm{e}^2 \frac{Dj}{N^3}.$$

The resistance on the full vertex set. Let L have eigenvalues $\eta_i \in [0, 2D]$ with orthonormal eigenvectors φ_i, and let $g_i = \langle b, \varphi_i\rangle^2$, so that $g_i \ge 0$ and $\sum_i g_i = \|b\|^2 = 2$. Then

$$R_x(L) - \frac{2}{x} = \sum_i g_i \frac{\eta_i}{x(x-\eta_i)} \qquad (x \ge N/2).$$

At $x = N$ this gives $0 \le \kappa \le 2\cdot 2D/(N\cdot 3N/4) = 16D/(3N^2)$. For $x \ge N/2$ we have $\eta_i \le N/4 \le x/2$, so the derivative of $\eta_i/(x(x-\eta_i))$ has absolute value $\eta_i(2x-\eta_i)/(x^2(x-\eta_i)^2) \le 8\eta_i/x^3 \le 64\eta_i/N^3$. Hence

$$\Big|R_k(L) - \frac{2}{k} - \kappa\Big| \le \sum_i g_i \cdot 64\eta_i \frac{j}{N^3} \le 256\frac{Dj}{N^3}.$$

Combination. Write $\psi_k = k(k-1)/(N(N-1)) + \Delta_k$ with $|\Delta_k| \le C_\omega Dj/N^2$ by (19). Since $\frac{k(k-1)}{N(N-1)} \cdot \frac{2}{k} = \frac{2}{N} - \frac{2j}{N(N-1)}$,

$$\psi_k R_k(L) - \Big(\frac{2}{N} + \kappa - \frac{2j}{N(N-1)}\Big) = \Delta_k R_k(L) - \kappa\Big(1 - \frac{k(k-1)}{N(N-1)}\Big) + \frac{k(k-1)}{N(N-1)}\Big(R_k(L) - \frac{2}{k} - \kappa\Big).$$

Here $0 \le R_k(L) \le \|b\|^2\|\Gamma_k^{-1}\| \le 8/N$, $0 \le 1 - k(k-1)/(N(N-1)) \le 3j/N$ and $\kappa \le 16D/(3N^2)$, so the three terms are at most $8C_\omega Dj/N^3$, $16Dj/N^3$ and $256Dj/N^3$ in absolute value. Together with the boundary estimate this proves (20) with $C_\pi = 8C_\omega + 272 + 128\mathrm{e}^2$. At $j = 0$ we have $S = V$, $L_{\partial S} = 0$ and $\psi_N = 1$, so $\pi_0 = R_N(L)$ exactly. □

4.6 Proofs of Theorems 4, A and B and of Corollary 5

Proof of Theorem 4. If $N < 8$, then $D = 0$, G is complete and there is nothing to prove. We first prove (3) for $N \geq N'_\lambda$, where N'_λ depends only on λ and is fixed below; all constants C depend only on λ. Apply the setting to $H = G + uv$, whose complement is a subgraph of the complement of G and so has maximum degree at most D. Let T, K, J be as there and $I = \mathbf{1}_{\{uv \in T\}}$.

Since $\mathbb{E}[I \mid J] = \pi_J$, we have $\operatorname{Cov}(K, I) = -\operatorname{Cov}(J, \pi_J)$. For $J \leq N/2$, Lemma 4.8 gives $\pi_J = 2/N + \kappa - 2J/(N(N-1)) + \varepsilon_J$. For $J > N/2$ define ε_J by the same equation; then $|\varepsilon_J| \leq 1$, because $\pi_J \in [0, 1]$ and $2/N + \kappa - 2J/(N(N-1)) \in [\kappa, 2/N + \kappa] \subset [0, 1]$. Therefore

$$\operatorname{Cov}(K, I) = \frac{2\operatorname{Var}(J)}{N(N-1)} - \operatorname{Cov}(J, \varepsilon_J). \tag{21}$$

The remainder satisfies

$$\begin{aligned} |\operatorname{Cov}(J, \varepsilon_J)| = \big|\mathbb{E}[(J - \mathbb{E}J)\varepsilon_J]\big| &\leq \frac{C_\pi D}{N^3}\mathbb{E}\big[J\,|J - \mathbb{E}J|\big] + \mathbb{E}\big[|J - \mathbb{E}J|\mathbf{1}_{\{J > N/2\}}\big] \\ &\leq \frac{2C_\pi D}{N^3}\mathbb{E}[J^2] + N\,\mathbb{P}(J > N/2). \end{aligned}$$

By Lemma 4.5, $\mathbb{E}[J^2] \leq C$ and $N\mathbb{P}(J > N/2) \leq C/N^3$, so $|\operatorname{Cov}(J, \varepsilon_J)| \leq C(D+1)/N^3$. By (18) and $0 \leq 1/(N(N-1)) - 1/N^2 \leq 2/N^3$,

$$\frac{2\operatorname{Var}(J)}{N(N-1)} = \frac{2\theta}{N^2} + O\Big(\frac{D+1}{N^3}\Big), \qquad \text{hence} \qquad \operatorname{Cov}(K, I) = \frac{2\theta}{N^2} + O\Big(\frac{D+1}{N^3}\Big).$$

Moreover $\mathbb{P}(I = 1) = \mathbb{E}[\pi_J] \leq 2/N + \kappa + C_\pi D/(2N^2) + \mathbb{P}(J > N/2) \leq C/N$. Choose $N'_\lambda \geq 8$, depending only on λ, so that $\mathbb{P}(I = 1) \leq 1/2$ for $N \geq N'_\lambda$; then $|1/(1 - \mathbb{P}(I = 1)) - 1| \leq 2\mathbb{P}(I = 1) \leq C/N$, and (6) gives

$$\mu_\lambda(G + uv) - \mu_\lambda(G) = \Big(\frac{2\theta}{N^2} + O\Big(\frac{D+1}{N^3}\Big)\Big)(1 + O(N^{-1})) = \frac{2}{\mathrm{e}\lambda N^2} + O\Big(\frac{D+1}{N^3}\Big).$$

For $8 \leq N < N'_\lambda$ both mean orders lie in $[1, N]$, so the left side of (3) is at most $N + 1/\lambda$, which is at most $(N'_\lambda + 1/\lambda)(N'_\lambda)^3(D+1)/N^3$. This proves (3) for all N with a constant C_λ depending only on λ. Enlarging C_λ if necessary, we take $C_\lambda \geq 16/(\mathrm{e}\lambda)$, so that $c_\lambda = 2/(\mathrm{e}\lambda C_\lambda) \leq 1/8$.

If $D + 1 < c_\lambda N$, the right side of (3) is smaller than $C_\lambda c_\lambda / N^2 = 2/(\mathrm{e}\lambda N^2)$, so the change is positive. For (4), add the missing edges of G one at a time in any order, obtaining $G = G_0 \subset G_1 \subset \cdots \subset G_M = K_N$. The complement of each G_{i-1} is a subgraph of the complement of G, so it has maximum degree at most D, and (3) applies to each of the M steps; summing the M estimates gives (4). □

Proof of Theorem A. Part (i) is Theorem 1, and part (ii) is Theorem 3. For part (iii), $D + 1 < D + 2 < c_1 N \leq N/8$, so G is connected because its minimum degree is at least $N - 1 - D > N/2$, it is not complete because $D \geq 1$, and Theorem 4 at $\lambda = 1$ shows that every missing edge increases μ. □

Proof of Theorem B. Fix $\lambda > 0$ and suppose $D + 2 < c_\lambda N$; then $D + 1 < N/8$.

New edges and chains. In the chain, the complement of each G_{i-1} is a subgraph of the complement of G, so it has maximum degree at most D, and $D + 1 < c_\lambda N$. Theorem 4 applied to G_{i-1} and the edge of G_i not in G_{i-1} gives $\mu_\lambda(G_i) > \mu_\lambda(G_{i-1})$. Taking $G_1 = G + uv$ for an arbitrary pair of nonadjacent vertices u, v gives $\mu_\lambda(G + uv) > \mu_\lambda(G)$.

Parallel edges. Let uv be an edge of G and let $G+f$ be the multigraph obtained by adding a new edge f with the same endpoints. Two parallel edges form a cycle, so no subtree of $G+f$ contains both uv and f. The subtrees of $G+f$ that avoid f are the subtrees of G, and exchanging f with uv is a bijection, preserving vertex sets and weights, from the subtrees of $G+f$ that contain f onto the subtrees of G that contain uv. This is the reduction in the proof of [3, Proposition 4.3]. Let T be a random subtree of G with $\mathbb{P}(T) \propto \lambda^{|E(T)|}$, let $K = |V(T)|$, and let I be the indicator of $uv \in E(T)$; then $0 < \mathbb{P}(I=1) < 1$. By the bijection, $\mu_\lambda(G+f)$ is the average of $\mu_\lambda(G)$ and $\mathbb{E}[K \mid I=1]$ with positive weights, so $\mu_\lambda(G+f) - \mu_\lambda(G)$ has the sign of

$$\mathbb{E}[K \mid I=1] - \mathbb{E}K = \frac{\operatorname{Cov}(K,I)}{\mathbb{P}(I=1)}.$$

Now $G = (G-uv)+uv$, and (6), applied to the graph $G-uv$ and its missing edge uv, gives $\operatorname{Cov}(K,I) = (1-\mathbb{P}(I=1))(\mu_\lambda(G)-\mu_\lambda(G-uv))$. The complement of $G-uv$ has maximum degree at most $D+1$, with $D+1 \le N/8$ and $(D+1)+1 < c_\lambda N$, so Theorem 4 gives $\mu_\lambda(G) > \mu_\lambda(G-uv)$. Hence $\operatorname{Cov}(K,I) > 0$ and $\mu_\lambda(G+f) > \mu_\lambda(G)$.

The case $\lambda = 1$ is the first part of the theorem. □

Proof of Corollary 5. Since $D+1 < c_\lambda N$ we have $D \le N/8$, and (4) gives $\mu_\lambda(K_N) - \mu_\lambda(G) \ge 2M/(\mathrm{e}\lambda N^2) - C_\lambda M(D+1)/N^3$, which equals the stated bound because $C_\lambda = 2/(\mathrm{e}\lambda c_\lambda)$. The bound is positive because $D+1 < c_\lambda N$ and $M \ge 1$. □

4.7 Remarks

Remark 4.9 (The complete graph). When $D=0$ the host is K_N: then $L=0$, $\Phi \equiv 1$, $\kappa = 0$, $\varepsilon_j = 0$ for every j, and Lemma 4.8 reduces to the exact identity $\pi_j = 2(N-1-j)/(N(N-1))$, which also follows from symmetry because a tree on $N-j$ vertices has $N-j-1$ edges. The first factor in the proof of Lemma 4.4 is the weight ratio of the complete graph, whose Poisson limit with mean θ is due to Luo, Xu and Tian [8] at $\lambda = 1$ and follows from that proof at every activity; at $\lambda = 1$ the resulting spanning probability $\mathrm{e}^{-1/\mathrm{e}}$ is the limit of Chin, Gordon, MacPhee and Vincent. The main term $2\theta/N^2$ of Theorem 4 is the covariance of the order with this exact complete-graph slope, namely $2\operatorname{Var}(\zeta)/N^2$. Lemmas 4.3–4.8 carry these complete-graph facts over to complements of maximum degree D, with errors linear in D.

Remark 4.10 (The density scale). Theorems B and 4 bound the maximum degree of the complement by a fixed proportion of the order, and this linear scale is the natural one. Cambie, Jooken and Wagner [2, §4.2] consider $K_n \vee mK_1$, the graph K_{n+m} with the edges of a K_m removed, and an edge f inside its K_n part. For $m = 7n/10+2$ they verified by computer, for every multiple n of 10 in $[10, 400]$ and for $n = 800$, that deleting f increases μ. Equivalently, the graph $(K_n \vee mK_1) - f$, whose complement consists of a K_m and the disjoint edge f and so has maximum degree $m-1$, has the missing edge f whose addition decreases μ. In these examples $(m-1)/(n+m)$ is close to $7/17 \approx 0.41$. Thus the conclusion that every missing edge improves holds below a fixed proportion of the order (Theorem 4), and at these orders it fails for complements of maximum degree about 0.41 times the order.

5 Small orders and edge rules

For a graph G and a missing edge uv, let Z^{uv} and Σ^{uv} be the number and the total order of the subtrees of $G+uv$ through uv, and let $\mathcal{D}(uv) = \Sigma^{uv} - \mu(G)Z^{uv}$. By (5), $\mu(G+uv) > \mu(G)$ exactly when $\mathcal{D}(uv) > 0$.

Proposition 7. *Let G be a connected graph that is not complete and has at most nine vertices. Then G has two nonadjacent vertices u, v with $\mu(G+uv) > \mu(G)$, and $\mathcal{D}(uv) \neq 0$ for every missing edge uv. Moreover, each of the following holds for G:*

(a) $\sum_{uv} \mathcal{D}(uv) > 0$ *over all missing edges uv;*

(b) the same sum over the missing edges uv at distance 2 is positive, if there is such an edge;

(c) every missing edge uv of maximal effective resistance increases μ;

(d) some missing edge uv of maximal distance increases μ;

(e) every missing edge uv minimising $\deg u + \deg v$ *increases μ;*

(f) every missing edge uv minimising $\deg u \deg v$ *increases μ.*

Proof. The noncomplete connected graphs of orders $3, \dots, 9$ form 1, 5, 20, 111, 852, 11,116 and 261,079 isomorphism classes. These are the numbers of connected graphs [10] minus one for the complete graph. The program `gen.mjs` lists one graph per class: it extends every class of order $k-1$ by a vertex joined to a nonempty set of vertices, which reaches every class because every connected graph has a vertex whose removal leaves it connected, and it identifies isomorphic graphs by a canonical labelling. Its counts agree with [10] for every order.

For each listed graph and each missing edge uv, the program `verify_small.mjs` computes $Z(G)$, $\Sigma(G)$, Z^{uv} and Σ^{uv} in exact integer arithmetic. It sums spanning-tree counts over vertex sets by the matrix-tree theorem, and it counts the trees of $G[S]+uv$ through uv as the spanning trees of $G[S]+uv$ with uv contracted. It also obtains the effective resistance from the same count on the full vertex set. In every graph some $\mathcal{D}(uv)$ is positive, no $\mathcal{D}(uv)$ is zero, and (a)–(f) hold. As a check of the program, the graphs with a decreasing missing edge number 1 at order 7 and 347 at order 8, as found by Cameron and Mol [3, §1]; at order 9 there are 39,308. □

Remark 5.1 (Edge rules beyond order nine). The rules (a)–(f) of Proposition 7 do not extend to all graphs. Exact computation (`rule_counterex.py` and `g_exact.py`) gives the following.

Let $\mathcal{B}_{12,5}$ be the barbell formed by two copies of K_5 joined by a path with two internal vertices; it has 12 vertices and 23 edges. Only 8 of its 43 missing edges increase μ. Each of them joins a vertex of one K_5, other than its attachment vertex, to the internal path vertex next to that attachment vertex. This is the pattern that Cambie, Jooken and Wagner established for long barbells [2, Proposition 4.1]. In $\mathcal{B}_{12,5}$:

- $\sum_{uv} \mathcal{D}(uv) < 0$, so (a) fails;
- all 16 missing edges of maximal effective resistance $19/5$ decrease μ, so (c) fails;
- all 16 missing edges at maximal distance 5 decrease μ, so (d) fails;
- 8 of the 16 missing edges minimising $\deg u + \deg v$, which are also the ones minimising $\deg u \deg v$, decrease μ, so (e) and (f) fail.

In the circulant $C_{13}(1,2)$, with vertex set $\mathbb{Z}_{13}$ and $i \sim j$ when $i - j \equiv \pm 1, \pm 2$, the missing edges at circular distance 3 or 4 increase μ, and the 26 at circular distance 5 or 6 decrease it. There:

- $\sum_{uv} \mathcal{D}(uv) < 0$;
- the 13 missing edges of maximal effective resistance $2498/3029$ decrease μ;

- the 26 missing edges at maximal distance 3 decrease μ.

Finally, join a vertex of K_6 and a vertex of K_5 by a path with one internal vertex, giving 12 vertices and 27 edges. Its 10 missing edges at distance 2 have $\sum \mathcal{D}(uv) < 0$, so (b) fails, although 9 of them increase μ.

The first decreasing example of Cameron and Mol has order 7. The rules (a)–(f) thus hold at every order up to nine and fail at orders 12 and 13. In each of these three graphs, some missing edge between two vertices at distance 2 increases μ.

Statements and Declarations

Funding. This research was self-funded.

Competing interests. The author owns Luminary Oddities and has no other relevant competing interests.

AI assistance. OpenAI ChatGPT assisted with literature comparison, derivations, proof development and the exact checks of edition 1. Anthropic Claude reviewed and completed the proofs, restructured and cross-referenced the manuscript and checked computations; for edition 2 it developed and wrote the proof of Theorem 3 and wrote the programs for Propositions 6 and 7 and for Remark 5.1. Scott Schubr directed the research and is responsible for the content.

Data availability. No datasets were used or generated. The package contains the following programs.

- `check_exact.py` compares (7)–(9) with an exact enumeration of small complete bipartite hosts and verifies the certificates of Theorem 1.
- `check_dense.py` computes exact gains for K_N minus a matching.
- `check_lemmas.py` checks the certificates symbolically, tests Lemma 4.2 on random matrices, tests Lemmas 4.1, 4.3–4.8 and (6) on explicit hosts by enumeration of vertex sets, and evaluates the error terms of Theorem 2.
- `check_kmn_exact.py` carries out the integer computation of Proposition 6.
- `check_kmn_analytic.py` evaluates the explicit bounds of Lemmas 3.7–3.9 on a grid of (m, n) and compares them with the exact moments of the law of (s, t).
- `gen.mjs` and `verify_small.mjs` carry out Proposition 7.
- `rule_counterex.py` and `g_exact.py` carry out Remark 5.1.

Propositions 6 and 7 and Remark 5.1 rest on these exact computations. The other proofs are independent of the computations.

References

[1] S. Cambie, G. Chen, Y. Hao and N. Tokar. Decreasing the mean subtree order by adding k edges. *Journal of Graph Theory* **105**(3):357–366, 2024. https://doi.org/10.1002/jgt.23043; arXiv:2308.12808.

[2] S. Cambie, J. Jooken and S. Wagner. On the extrema of the mean subtree order of graphs. arXiv:2508.20593v2, 2026. https://arxiv.org/abs/2508.20593v2.

[3] B. Cameron and L. Mol. On the mean subtree order of graphs under edge addition. *Journal of Graph Theory* **96**(3):403–413, 2021. https://doi.org/10.1002/jgt.22621. Section and theorem numbers refer to arXiv:1911.05794v1, https://arxiv.org/abs/1911.05794.

[4] J. Chen, X. Jin, Z. Li and H. Ma. Subtree counts and the maximum mean subtree order. arXiv:2609.14355, 2026. https://arxiv.org/abs/2609.14355.

[5] X. Chen, G. Wei and H. Lian. Solution to a conjecture on the mean subtree order of graphs under edge addition. *Journal of Graph Theory* **104**(3):645–658, 2023. https://doi.org/10.1002/jgt.22993.

[6] A. J. Chin, G. Gordon, K. J. MacPhee and C. Vincent. Subtrees of graphs. *Journal of Graph Theory* **89**(4):413–438, 2018. https://doi.org/10.1002/jgt.22359.

[7] G. Kirchhoff. Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. *Annalen der Physik* **148**(12):497–508, 1847. https://doi.org/10.1002/andp.18471481202.

[8] Z. Luo, K. Xu and J. Tian. Random subtrees and unimodal sequences in graphs. *Discrete Mathematics* **347**(1):113654, 2024. https://doi.org/10.1016/j.disc.2023.113654.

[9] H. I. Scoins. The number of trees with nodes of alternate parity. *Proceedings of the Cambridge Philosophical Society* **58**(1):12–16, 1962. https://doi.org/10.1017/S0305004100036173.

[10] N. J. A. Sloane, ed. *The On-Line Encyclopedia of Integer Sequences*, sequence A001349, number of simple connected graphs on n unlabeled nodes. https://oeis.org/A001349, accessed 26 November 2026.

[11] S. Wagner. On the probability that a random subtree is spanning. *Journal of Graph Theory* **98**(2):195–215, 2021. https://doi.org/10.1002/jgt.22688; arXiv:1910.07349.

[12] S. Wagner and R. Wang. The distribution of subtrees in dense graphs and the roots of the subtree polynomial. arXiv:2605.03583, 2026. https://arxiv.org/abs/2605.03583.