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 $K_{2,n}$ (all $n \geq 2$), for $K_{m,n}$ with m fixed and n large, and for every near-complete graph, meaning one on N vertices whose complement has maximum degree D with $D + 2 < cN$ for an absolute constant $c > 0$. For near-complete graphs we prove Conjecture 7.4 itself: adding any edge, new or parallel to an existing one, increases the mean, and the mean increases at each step of every chain of single-edge additions up to the complete graph K_N. This gives a quantitative form, in this class, of the strict maximality of K_N proved by Cambie, Jooken and Wagner and by Chen, Jin, Li and Ma, with an explicit lower bound for the gap. The key estimate holds for subtrees weighted by $\lambda > 0$ per edge: if $D \leq N/8$, each missing edge changes the weighted mean by $2/(\mathrm{e}\lambda N^2)$ up to a constant depending only on λ times $(D + 1)/N^3$. In $K_{2,n}$ the edge inside the two-vertex part lowers the mean exactly when $n \geq 6$, proving a decrease stated by Cameron and Mol.

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 three further classes.

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 each $m \geq 1$ and every $n \geq n_0(m)$, where $n_0(m)$ depends only on m; 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 3 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 3 in place of c_1.

Theorems A and B follow from the detailed results below and are proved in Section 3.6. They concern exactly the three 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 $n_0(m)$, c_1 and c_λ are given by the proofs as existential constants depending only on the indicated parameters.

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.

Theorem 3. *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

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

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 $c_\lambda \leq 1/8$, the condition $D + 1 < c_\lambda N$ already implies $D \leq N/8$.

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 3 gives it a quantitative form: the new content of Corollary 4 is the explicit lower bound for the gap.

Corollary 4. *Fix $\lambda > 0$. Let $G \neq K_N$ be a graph on N vertices with $M \geq 1$ missing edges whose complement has maximum degree D with $D + 1 < c_\lambda N$. Then*

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

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, in the journal version of [3], also proposed 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 $k \geq 1$ 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 3 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 [12] proved the corresponding Poisson limits for dense Erdős–Rényi graphs, and Wagner and Wang [13, 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 3 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 3.5), and an edge-specific estimate for the probability that the subtree contains uv at a given order (Lemma 3.8).

For the complete graph ($D = 0$) both inputs are exact or classical facts, and both are used in our earlier articles. The fixed-order edge marginal $\mathbb{P}(uv \in T \mid |V(T)| = k) = 2(k-1)/(N(N-1))$ is the edge case of the fixed-forest reduction in [10, Proposition 2.2, equation (3)], and the Poisson law with mean $1/(\mathrm{e}\lambda)$ for the number of omitted vertices, due to Luo, Xu and Tian at $\lambda = 1$, is used at every fixed activity in [10, Lemma 4.6] and in the proof of Theorem 2 in [9, §3]. Section 3 extends both to complements of maximum degree D, with errors linear in D (Remark 3.9).

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 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 3.1); these determinants are stable under removing or exchanging vertices (Lemma 3.3). Comparison with the complete graph gives the Poisson law (Lemma 3.5), and exchanging vertex labels gives the omission probabilities (Lemma 3.6). Kirchhoff's formula expresses edge inclusion as an effective resistance (Lemma 3.7), and a resolvent comparison gives the inclusion probability at each order (Lemma 3.8). 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

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

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 [11]. 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_1ba_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 $e_B = b_1b_2$. The subtrees through b_1b_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

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

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 $a_1 a_2$ 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,*

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

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 $\xi(\xi)_l = (\xi)_{l+1} + l(\xi)_l$,

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

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 $1 \le s \le m$ 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

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

and mean order $s + \mathbb{E}[\xi_s^s]/\mathbb{E}[\xi_s^{s-1}\mathbf{1}_{\{\xi_s\ge1\}}]$.

Reduction to the class $s = m$. We have $Z_m \ge q_m(1+m\lambda)^n \mathbb{P}(\xi_m \ge 1) \ge q_m(1+m\lambda)^n/2$ for large n, while $Z_s \le q_s(1+s\lambda)^n n^{s-1}$ 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 $h(x) = x^{m-1}$,

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

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 $2q_m(1+m\lambda)^n \mathbb{E}[\xi_m^{m-2}\mathbf{1}_{\{\xi_m\ge1\}}]$. 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 Near-complete hosts

3.1 Setting

Throughout this section $\lambda > 0$ is fixed, $\theta = 1/(\mathrm{e}\lambda)$, $N \ge 8$, and H is a graph on a vertex set V of size N whose complement F has maximum degree at most D, where $D \le N/8$. 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 $2D$,

$$\|L_U\| \le 2D, \qquad \operatorname{tr} L_{\partial U} \le 2D\,|V \setminus U|, \qquad \operatorname{tr} L \le ND. \tag{10}$$

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 $k \ge 1$ put

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

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

3.2 Determinants and their comparison

Lemma 3.1. *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 3.2. *Let X_0, X_1 be positive semidefinite with $\|X_0\|, \|X_1\| \leq 1/2$. Then*

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

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 $-\,\mathrm{tr}\bigl((\mathrm{Id} - X_t)^{-1}(X_1 - X_0)\bigr)$, 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 3.3. *Let $U \subseteq V$ with $|U| = k \geq N/2$, and $j = N - k$. Then $\Phi(U) > 0$,*

$$\bigl|\log\Phi(U) - \log\Phi(V)\bigr| \leq \frac{6Dj}{k} \leq \frac{12Dj}{N}, \qquad \Phi(V) \geq \mathrm{e}^{-2D}, \tag{11}$$

and, for every transposition σ of V,

$$\bigl|\log\Phi(\sigma U) - \log\Phi(U)\bigr| \leq \frac{8D}{k} \leq \frac{16D}{N}. \tag{12}$$

Proof. By (10) and $D \leq N/8 \leq k/4$, the matrices L_U/k, L/N and $L_{\sigma U}/k$ are positive semidefinite of norm at most $1/2$, so $\Phi(U) \geq 2^{-N} > 0$ and Lemma 3.2 applies. For (11),

$$\Bigl\|\frac{L_U}{k} - \frac{L}{N}\Bigr\|_1 \leq \frac{\mathrm{tr}\,L_{\partial U}}{k} + \mathrm{tr}\,L\Bigl(\frac{1}{k} - \frac{1}{N}\Bigr) \leq \frac{2Dj}{k} + \frac{ND\,j}{kN} = \frac{3Dj}{k},$$

and $\log\Phi(V) \geq -2\,\mathrm{tr}\,L/N \geq -2D$. For (12), 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 $\|L_{\sigma U} - L_U\|_1 \leq 4D$. □

3.3 The number of omitted vertices

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

Lemma 3.4. *For $0 \le j \le N/2$,*

$$\Big|\log \frac{j!\,Q_j}{\theta^j}\Big| \le \frac{3j + 3j^2 + 12Dj}{N}, \qquad Q_j \le \frac{(\mathrm{e}^{5/2}\theta)^j}{j!}. \tag{13}$$

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

$$Q_j \le \mathrm{e}^{2D}\lambda^{-j}\frac{j+1}{j!}. \tag{14}$$

Proof. Let $k = N - j$. By Lemma 3.1, 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 (11) 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 (13). 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 (14) with $j = N-1$. □

Lemma 3.5. *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,*

$$\big|\mathbb{E}[J^r] - \mathbb{E}[\zeta^r]\big| \le C_{\lambda,r}\frac{D+1}{N}, \qquad \mathbb{P}(J > N/2) \le C_{\lambda,r}N^{-r}. \tag{15}$$

In particular

$$\big|\mathrm{Var}(J) - \theta\big| \le C'_\lambda\frac{D+1}{N} \tag{16}$$

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 $j > N/2$, (14), $j! \ge (j/\mathrm{e})^j$ 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 (15). For $j \le N/2$, write $t_j = \log(j!Q_j/\theta^j)$; by (13), $|t_j| \le (3j + 3j^2 + 12Dj)/N$ 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 $\sum_j Q_j \ge 1$,

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

Finally $\mathrm{Var}(\zeta) = \theta$, and (16) follows from the cases $r = 1, 2$ because $\mathbb{E}J$ and $\mathbb{E}\zeta$ are bounded. □

3.4 Omission probabilities

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

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

Lemma 3.6. *For $N/2 \le k \le N$, every vertex w and every pair $x \neq y$,*

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

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

$$\Big|\psi_k - \frac{k(k-1)}{N(N-1)}\Big| \le C_\omega \frac{Dj}{N^2}, \qquad 1 - \psi_k \le 2\mathrm{e}^2\frac{j}{N}, \tag{17}$$

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

Proof. Let $w \neq w'$ and $\sigma = (w\,w')$. By Lemma 3.1 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 (12), $\omega_{w'} = \sum_{U \not\ni w} \Phi(\sigma U)/\sum_{|U|=k}\Phi(U)$ 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 (17), and $1 - \psi_k \le \omega_u + \omega_v \le 2\mathrm{e}^2 j/N$. □

3.5 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 3.7. *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 3.1), 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 3.8. *Let $\kappa = R_N(L) - 2/N$. Then $0 \le \kappa \le 16D/(3N^2)$, and for $0 \le j \le N/2$,*

$$\pi_j := \mathbb{P}(uv \in T \mid J = j) = \frac{2}{N} + \kappa - \frac{2j}{N(N-1)} + \varepsilon_j, \qquad |\varepsilon_j| \le C_\pi\frac{Dj}{N^3}, \tag{18}$$

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

Proof. Let $k = N - j \geq N/2$ and $\Gamma_k = k\mathrm{Id} - L$, whose eigenvalues are at least $k - 2D \geq k/2$. By Lemmas 3.1 and 3.7,

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

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

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

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 $\omega_x + \omega_y \leq 2\mathrm{e}^2 j/N$ by Lemma 3.6. Hence $\mathbb{E}[L_{\partial S} \mid K = k] \preceq (2\mathrm{e}^2 j/N) L$, and with $\|b\|^2 = 2$, $\|L\| \leq 2D$ and $\|\Gamma_k^{-1}\| \leq 2/k \leq 4/N$,

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

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 $g_i \geq 0$ and $\sum_i g_i = \|b\|^2 = 2$. Then

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

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

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

Combination. Write $\psi_k = k(k-1)/(N(N-1)) + \Delta_k$ with $|\Delta_k| \leq C_\omega Dj/N^2$ by (17). 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 $0 \leq R_k(L) \leq \|b\|^2 \|\Gamma_k^{-1}\| \leq 8/N$, $0 \leq 1 - k(k-1)/(N(N-1)) \leq 3j/N$ and $\kappa \leq 16D/(3N^2)$, 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 (18) 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. □

3.6 Proofs of Theorems 3, A and B and of Corollary 4

Proof of Theorem 3. 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 $\mathrm{Cov}(K, I) = -\mathrm{Cov}(J, \pi_J)$. For $J \leq N/2$, Lemma 3.8 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

$$\mathrm{Cov}(K, I) = \frac{2\,\mathrm{Var}(J)}{N(N-1)} - \mathrm{Cov}(J, \varepsilon_J). \tag{19}$$

The remainder satisfies

$$|\operatorname{Cov}(J,\varepsilon_J)| = \bigl|\mathbb{E}[(J-\mathbb{E}J)\varepsilon_J]\bigr| \le \frac{C_\pi D}{N^3}\mathbb{E}\bigl[J\,|J-\mathbb{E}J|\bigr] + \mathbb{E}\bigl[|J-\mathbb{E}J|\mathbf{1}_{\{J>N/2\}}\bigr]$$
$$\le \frac{2C_\pi D}{N^3}\mathbb{E}[J^2] + N\,\mathbb{P}(J>N/2).$$

By Lemma 3.5, $\mathbb{E}[J^2]\le C$ and $N\mathbb{P}(J>N/2)\le C/N^3$, so $|\operatorname{Cov}(J,\varepsilon_J)|\le C(D+1)/N^3$. By (16) and $0\le 1/(N(N-1))-1/N^2\le 2/N^3$,

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

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

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

For $8\le N<N'_\lambda$ 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 $C_\lambda\ge 16/(\mathrm{e}\lambda)$, so that $c_\lambda=2/(\mathrm{e}\lambda C_\lambda)\le 1/8$.

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. Part (ii) is Theorem 2 at $\lambda=1$: the main term in (1) is a positive number depending only on m, so the change is positive for $n\ge n_0(m)$. For part (iii), $D+1<D+2<c_1N\le N/8$, so G is connected because its minimum degree is at least $N-1-D>N/2$, it is not complete because $D\ge 1$, and Theorem 3 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 3 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 $\mathrm{Cov}(K, I) = (1 - \mathbb{P}(I = 1))(\mu_\lambda(G) - \mu_\lambda(G - uv))$. 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 3 gives $\mu_\lambda(G) > \mu_\lambda(G - uv)$. Hence $\mathrm{Cov}(K, I) > 0$ and $\mu_\lambda(G + f) > \mu_\lambda(G)$.

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

Proof of Corollary 4. 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 $M \ge 1$. □

3.7 Remarks

Remark 3.9 (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 3.8 reduces to the exact identity $\pi_j = 2(N - 1 - j)/(N(N - 1))$, the edge case of the fixed-forest reduction in [10, Proposition 2.2]. The first factor in the proof of Lemma 3.4 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 is used at every fixed activity in [10, Lemma 4.6] and [9, §3]; 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 3 is the covariance of the order with this exact complete-graph slope, namely $2\,\mathrm{Var}(\zeta)/N^2$. Lemmas 3.3–3.8 carry these complete-graph facts over to complements of maximum degree D, with errors linear in D.

Remark 3.10 (The density scale). Theorems B and 3 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 3), and at these orders it fails for complements of maximum degree about 0.41 times the order.

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. Anthropic Claude reviewed and completed the proofs, restructured and cross-referenced the manuscript and checked computations. Scott Schubr directed the research and is responsible for the content.

Data availability. No datasets were used or generated. The package contains three Python scripts. `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 3.2 on random matrices and Lemmas 3.1, 3.3–3.8 and (6) on explicit hosts by enumeration of vertex sets, and evaluates the error terms of Theorem 2. The proofs are independent of these 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] S. Schubr. *A sharp lower bound for adjacent-edge correlations in random subtrees*, edition 4. Luminary Research Archive, November 2026. https://doi.org/10.5281/zenodo.24472625.

[10] S. Schubr. *Edge correlations in weighted random subtrees of the complete graph: a sharp bound and the activity phase diagram*, edition 1. Luminary Research Archive, November 2026. https://doi.org/10.5281/zenodo.24404894.

[11] 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.

[12] 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.

[13] 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.