UNIVERSAL DIAMETER BOUNDS FOR RANDOM GRAPHS WITH GIVEN DEGREES

LOUIGI ADDARIO-BERRY AND GABRIEL CRUDELE

Abstract. Given a graph G, let $\mathrm{diam}(G)$ be the greatest distance between any two vertices of G which lie in the same connected component, and let $\mathrm{diam}^+(G)$ be the greatest distance between any two vertices of G; so $\mathrm{diam}^+(G) = \infty$ if G is not connected.

Fix a sequence $(d_1, \ldots, d_n)$ of positive integers, and let $\mathbf{G}$ be a uniformly random connected simple graph with $V(\mathbf{G}) = [n] := \{1, \ldots, n\}$ such that $\deg_{\mathbf{G}}(v) = d_v$ for all $v \in [n]$. We show that, unless a $1 - o(1)$ proportion of vertices have degree 2, then $\mathbb{E}[\mathrm{diam}(\mathbf{G})] = O(\sqrt{n})$. It is not hard to see that this bound is best possible for general degree sequences (and in particular in the case of trees, in which $\sum_{v=1}^{n} d_v = 2(n-1)$). We also prove that this bound holds without the connectivity constraint. As a key input to the proofs, we show that graphs with minimum degree 3 are with high probability connected and have logarithmic diameter: if $\min(d_1, \ldots, d_n) \geq 3$ then $\mathrm{diam}^+(\mathbf{G}) = O_{\mathbb{P}}(\log n)$; this bound is also best possible.

Contents

1. Introduction

All graphs considered in this paper are finite and contain no isolated vertices. Given a connected graph G, we write $\mathrm{diam}(G) = \max(\mathrm{dist}_G(u, v) : u, v \in V(G))$, where $\mathrm{dist}_G(u, v)$ is the graph distance from u to v in G. For a not-necessarily-connected graph G, we write $\mathrm{diam}(G)$ for the greatest diameter of any connected component of G and $\mathrm{diam}^+(G) = \max(\mathrm{dist}_G(u, v) : u, v \in V(G))$, so $\mathrm{diam}^+(G) = \infty$ if G is not connected.

A *degree sequence* is a sequence $\mathrm{d} = (d_1, \ldots, d_n)$ of positive integers with even sum. We set $|\mathrm{d}|_0 = n$ and $|\mathrm{d}|_1 = \sum_{i=1}^{n} d_i$. Write $\mathcal{G}_{\mathrm{d}}$ for the set of all simple graphs G with

2020 *Mathematics Subject Classification.* 05C80,60C05.
Key words and phrases. Random graphs, random trees, diameter, switching.

$V(G) = [n]$ such that $\deg_G(i) = d_i$ for all $i \in [n]$, and $\mathcal{C}_{\mathrm{d}} = \{G \in \mathcal{G}_{\mathrm{d}} : G \text{ is connected}\}$. Also, for any $b \in \mathbb{N}$, we write $n_b(\mathrm{d}) := |\{i \in [n] : d_i = b\}|$ for the number of entries of d which equal b. We say d is *b-free* if $n_b(\mathrm{d}) = 0$.

For a finite set $\mathcal{X}$, we write $\mathbf{X} \in_u \mathcal{X}$ to mean that $\mathbf{X}$ is a uniformly random element of $\mathcal{X}$. Whenever we write $\mathbf{X} \in_u \mathcal{X}$, we implicitly assume that $\mathcal{X}$ is nonempty.

Theorem 1.1. *For all $\epsilon > 0$ there exists $C > 0$ such that the following holds. Let $\mathrm{d} = (d_1, \ldots, d_n)$ be a degree sequence such that $n_2(\mathrm{d}) < (1-\epsilon)n$, and let $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}$. Then $\mathbb{E}[\mathrm{diam}(\mathbf{G})] \le C\sqrt{n}$.*

The bound of the theorem can not in general be improved, as the following example illustrates. Fix positive integers k and m, and let d be the sequence

$$(\underbrace{3, \ldots, 3}_{m \text{ times}}, \underbrace{1, \ldots, 1}_{m+2 \text{ times}}, \underbrace{2, \ldots, 2}_{km \text{ times}})$$

The elements of $\mathcal{C}_{\mathrm{d}}$ have $n = (k+2)m+2$ vertices, and are *sub-binary trees* — trees in which the maximum degree is three. Let $\mathbf{T} \in_u \mathcal{C}_{\mathrm{d}}$, and let $\mathbf{T}'$ be the *homeomorphic reduction* of $\mathbf{T}$, obtained from $\mathbf{T}$ by replacing by a single edge each maximal path of $\mathbf{T}$ all of whose internal vertices have exactly one child. Then $\mathbf{T}'$ is a uniformly random labeled binary tree with internal vertices labeled $\{1, \ldots, m\}$ and leaves labeled $\{m+1, \ldots, 2m+2\}$. Such a tree has diameter $\Theta(\sqrt{m})$ in expectation [17]; moreover, the path between two random leaves in $\mathbf{T}'$ also has expected length $\Theta(\sqrt{m})$. Now, conditionally given $\mathbf{T}'$, by symmetry, for each edge uv of $\mathbf{T}'$, the expected length of the path between u and v in $\mathbf{T}$ is $1 + km/(2m+1) = \Theta(k)$. Thus, the path between two random leaves in $\mathbf{T}$ has expected length $\Theta(k\sqrt{m}) = \Theta(\sqrt{kn})$; if almost all vertices have degree 2, then k is large and the diameter is typically much larger than $\sqrt{n}$.

For a connected graph G, we write $V(G)$ and $E(G)$ respectively for the vertex set and edge set of G. We also use the shorthand $v(G) = |V(G)|$ and $e(G) = |E(G)|$. The *surplus* of G is the quantity $s(G) = 1 + e(G) - v(G)$, which is the maximum number of edges that can be deleted from G without disconnecting G. Note that for a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$, writing $s(\mathrm{d}) = 1 + (\sum_{i=1}^n d_i/2) - n = 1 + |\mathrm{d}|_1/2 - |\mathrm{d}|_0$, then for any $G \in \mathcal{C}_{\mathrm{d}}$ we have $s(G) = s(\mathrm{d})$. The preceding example shows that the bound stated in Theorem 1.1 can be saturated by certain ensembles of random trees, which have surplus zero; for any fixed $s > 0$, it is also not very hard to see that degree sequences of the form

$$(\underbrace{3, \ldots, 3}_{m+s \text{ times}}, \underbrace{1, \ldots, 1}_{m+2-s \text{ times}}, \underbrace{2, \ldots, 2}_{km \text{ times}}) \tag{1.1}$$

yield graphs of diameter $\Theta(k\sqrt{m/s}\log(s+1)) = \Theta(\sqrt{kn/s}\log(s+1))$, which likewise saturate the bound. (Many other examples are possible. We write $\log(s+1)$ rather than $\log s$ in order to include the case $s = 1$.)

Our second result establishes the same bound as our first result, but for uniform samples from $\mathcal{G}_{\mathrm{d}}$ rather than from $\mathcal{C}_{\mathrm{d}}$.

Theorem 1.2. *For all $\epsilon > 0$ there exists $C > 0$ such that the following holds. Let $\mathrm{d} = (d_1, \ldots, d_n)$ be a degree sequence such that $n_2(\mathrm{d}) < (1-\epsilon)n$, and let $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$. Then $\mathbb{E}[\mathrm{diam}(\mathbf{G})] \le C\sqrt{n}$.*

An important step in the proofs of the above results is a logarithmic bound on the diameters of random graphs with given degrees, when the minimum degree is at least three, which is of independent interest.

Theorem 1.3. *There exists an absolute constant $C > 0$ such that the following holds. Let $\mathrm{d} = (d_1, \ldots, d_n)$ be a degree sequence such that $n_1(\mathrm{d}) = n_2(\mathrm{d}) = 0$, and let $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$. Then $\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{G}) \ge 1000\log n\right\} \le C\log^8 n/n$.*

1.1. **Three conjectures.** The examples sketched above show that for any fixed integer $s \geq 0$, there are degree sequences with surplus s which saturate the bound of Theorem 1.1, up to a multiplicative factor of order $O((s+2)^{-1/2}\log(s+2))$. However, it seems likely that when s is large, the bound can always be improved.

Conjecture 1. *Fix $\varepsilon > 0$ and let $(\mathrm{d}^n, n \geq 1)$ be a sequence of degree sequences with $|\mathrm{d}^n|_0 = n$ and $n_2(\mathrm{d}^n) \leq (1-\varepsilon)n$, and suppose that $s_n = s(\mathrm{d}^n) \to \infty$ as $n \to \infty$. For $n \geq 1$ let $\mathbf{G}_n \in_u \mathcal{G}_{\mathrm{d}^n}$. Then $\mathbb{E}\left[\mathrm{diam}(\mathbf{G}_n)\right] = O((1+\sqrt{n/s_n})\log s_n)$ as $n \to \infty$.*

It would already be interesting to show that, under the assumptions of the above conjecture, it holds that $\mathbb{E}\left[\mathrm{diam}(\mathbf{G}_n)\right] = o(n^{1/2})$ as $n \to \infty$.

To motivate the second conjecture, it is useful to first present a result from [2]. A *rooted tree* is a tree T with a distinguished node $\rho(T)$ called the root. In a rooted tree, each node aside from the root has a parent; the degree of the root is its number of children; and the degree of each non-root node is its number of children plus one. The *height* $\mathrm{ht}(T)$ is the greatest graph distance of any vertex from the root.

A *child sequence* is a finite sequence $\mathrm{c} = (c_v, v \in V)$ of non-negative integers with $\sum_{v\in V} c_v \leq |V|-1$. For a nonnegative integer b, we write $n_b(\mathrm{c}) = |\{v \in V : c_v = b\}|$. Child sequence c is *sub-binary* if $n_b(c) = 0$ for all $b > 2$. A rooted tree T has child sequence $\mathrm{c} = (c_v, v \in V)$ if for all $v \in V$, vertex v has c_v children in T. (Note that for T to have child sequence c it is necessary that $\sum_{v\in V} c_v = |V|-1$.) The set of all rooted trees with child sequence c is denoted $\mathcal{T}_{\mathrm{c}}$.

The paper [2] showed that random sub-binary rooted trees have stochastically maximal height among trees with a given number of leaves and one-child vertices. More precisely, if b is a sub-binary child sequence and c is any child sequence with $n_0(\mathrm{c}) = n_0(\mathrm{b})$ and $n_1(\mathrm{c}) = n_1(\mathrm{b})$, and $\mathbf{T}_{\mathrm{c}} \in_u \mathcal{T}_{\mathrm{c}}$ and $\mathbf{T}_{\mathrm{b}} \in_u \mathcal{T}_{\mathrm{b}}$, then $\mathrm{ht}(\mathbf{T}_{\mathrm{c}}) \preceq_{\mathrm{st}} \mathrm{ht}(\mathbf{T}_{\mathrm{b}})$. We conjecture that a similar result holds for random graphs with a given degree sequence and surplus.

Conjecture 2. *Fix degree sequences* d *and* d′ *such that $s(\mathrm{d}) = s(\mathrm{d}')$. Suppose that $n_1(\mathrm{d}) = n_1(\mathrm{d}')$, that $n_2(\mathrm{d}) = n_2(\mathrm{d}')$, and that $n_b(\mathrm{d}) = 0$ for $b > 3$. Let $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$ and let $\mathbf{G}' \in_u \mathcal{G}_{\mathrm{d}'}$. Then $\mathrm{diam}(\mathbf{G}') \preceq_{\mathrm{st}} \mathrm{diam}(\mathbf{G})$.*

In view of the discussion around (1.1), this conjecture would imply that random graphs with surplus $s > 0$ always have expected diameter at most $O((1+\sqrt{n/s})\log(s+1))$ unless almost all vertices have degree 2, and so would strengthen Theorem 1.1 for such ensembles.

In keeping with the heuristic that as the degrees of random graphs increase, their diameters decrease, we propose the following conjecture. Informally, it states that among random graphs with given degree sequences with minimum degree d, the random d-regular graphs have the largest diameter, at least to first order. For fixed $d \geq 3$, let $\mathcal{S}_{n,d}$ be the set of all degree sequences $(d_1, \ldots, d_n)$ with $\min(d_1, \ldots, d_n) \geq d$.

Conjecture 3. *For all $d \geq 3$ and $\epsilon > 0$, it holds that*

$$\limsup_{n\to\infty} \sup_{\mathrm{d}\in\mathcal{S}_{n,d}} \mathbb{P}_{\mathbf{G}\in_u\mathcal{G}_{\mathrm{d}}}(\mathrm{diam}(\mathbf{G})) > (1+\epsilon)\log_{d-1} n) = 0.$$

It is not hard to see that for all $d \geq 3$, a random d-regular graph with n vertices has diameter $(1+o(1))\log_{d-1} n$ with high probability and in expectation, which justifies the informal statement of the conjecture just before its formal statement. The $d = 3$ case of the conjecture is already interesting, and may be viewed as a slight weakening of Conjecture 2.

1.2. **Related work.** There is a large body of literature on diameters of random graphs. The first result on the subject we could find is due to Burton [8], who proved two-point concentration for the diameter of the Erdős-Rényi random graph $G(n,p)$ when $p = p(n) \gg (\log n)/n$. (A slightly weaker result of the same flavour was independently proved by Bollobás [8], and related results were proved around the same time by Klee and Larman

[26].) Łuczak showed that in the subcritical random graph, when $p = p(n)$ satisfies that $n^{1/3}(1 - np) \to \infty$, the connected component of $G(n,p)$ with maximal diameter is with high probability a tree, and found the asymptotic behaviour of this maximal diameter throughout the subcritical regime. Much later, Chung and Lu [11], and then Riordan and Wormald [30], substantially extended Burtin's result. The latter paper proved two-point concentration of the diameter of $G(n,p)$ whenever $np \to \infty$, and additionally found tight estimates for the diameter throughout the supercritical regime $n^{1/3}(np - 1) \to \infty$. Independently and at roughly the same time, the diameter of $G(n,p)$ in the critical regime $n^{1/3}(np - 1) = O(1)$ was studied in [1, 29]. Combined, the preceding results fully describe all possibilities for the asymptotic behaviour of the diameter of $G(n,p)$.

The diameters of many other random graph ensembles have also been studied, including the cases of random regular graphs [6, 32]; scale-free random graphs [7, 10, 12]; inhomogeneous random graphs [20, 34]; preferential attachment models [5, 13, 25]; the configuration model under various degree assumptions [16, 35]; hyperbolic random graphs [21]; and in directed settings [9, 14, 22].

The setting we consider in this paper – the uniform distribution over simple graphs with a given degree sequence – is an extremely natural reference model that, to date, appears not to have been studied in generality. Past work on the diameters of random graphs with given degree sequences has focussed on asymptotic results, in regimes where the degree sequences are tame enough to be studied via the configuration model [33, Chapter 7], and has largely focussed on settings exhibiting small-world phenomena (and in particular logarithmic diameters). While we believe that Theorem 1.3 is an important contribution to the understanding of the logarithmic-diameter regime, we also hope our paper will spur interest in developing a finer understanding of typical and extreme distances in random graphs with "sparser" degree sequences, where it is necessary to study both the cycle structure (core) of the graph and the structure of the trees which attach to the core. The study of other graph properties, such as mixing times and cutoff phenomena, would also be interesting for such graphs.

1.3. **Notation.** In this section we introduce some notation that will be used throughout the paper, including in the proof overview which immediately follows.

The *height* of a rooted tree T, denoted $\mathrm{ht}(T)$, is the greatest distance of any vertex from the root of T. A *rooted forest* is a set $F = \{T_1, \ldots, T_k\}$ of rooted trees with disjoint vertex sets; the *root set* of F is $\{\rho(T_i) : i \in [k]\}$. The height of F, denoted $\mathrm{ht}(F)$, is the greatest height of any of its constituent trees.

Rooted forest $F = \{T_1, \ldots, T_k\}$ has child sequence $\mathrm{c} = (c_v, v \in V)$ if $\bigcup_{i\in[k]} V(T_i) = V$ and for all $v \in V$, vertex v has c_v children in its tree in F.

The set of all forests with child sequence c and root set R is denoted $\mathcal{F}_{\mathrm{c},R}$, and the set of all forests with child sequence c is denoted $\mathcal{F}_{\mathrm{c}}$. Given a subset A of the vertex set of a child sequence c, we write $\mathrm{c}|_A$ for the restriction of c to A. We use the analogous notation for restrictions of degree sequences.

For positive integers m and h, an *m-composition of h* is a sequence of m nonnegative integers summing to h. We write $\mathcal{P}_{m,h}$ for the set of all m-compositions of h.

We shall encode multigraphs as triples (V, E, ι), where V is the vertex set, E is the edge set, and $\iota : E \to \binom{V}{2} \cup \binom{V}{1}$ is the endpoint map. For all multigraphs considered in this work, the vertex set V is totally ordered, and for $e \in E$ we define $e^- = \min \iota(e)$ and $e^+ = \max \iota(e)$. We say that edges $e, f \in E$ are *parallel* if $\iota(e) = \iota(f)$, and write $\mathrm{mult}(e)$ for the number of edges parallel to e.

Finally, we note that the notation $\mathcal{G}_{\mathrm{d}}$, $\mathcal{C}_{\mathrm{d}}$ and the like makes sense even if $\mathrm{d} = (d_v, v \in S)$ is a degree sequence indexed by a finite set S which is not an initial segment of $\mathbb{N}$; we will require this at a few points in the proof. Relatedly, for a graph $G = (V, E)$ we refer to

$(\deg_G(v), v \in V)$ as the degree sequence of G, where for concreteness we list the vertex degrees according to the total order of V.

1.4. **Proof outline.** Given a graph G, write $C(G)$ for the maximum subgraph of G of minimum degree at least 2, and call $C(G)$ the *core* of G. Thus, if G is a forest, then $C(G)$ is empty. We say a graph C is a core if it has minimum degree 2. For a vertex $v \in V(G)$ not in a tree component of G, we write $\alpha(v)$ for the (unique) core vertex at minimum graph distance from v.

We let $F(G)$ be the rooted forest obtained from G as follows. For each edge rs with $r \notin V(C(G))$ and $s \in V(C(G))$, remove edge rs and let the resulting tree, which contains r, have root r. Second, for each tree component T' of G, delete the leaf of T' with minimum label and root the resulting tree T at the unique neighbour of that leaf. (All the graphs we consider in this paper have $V(G) \subset \mathbb{N}$, so "minimum label" makes sense.) Finally, remove all vertices of $C(G)$ and their remaining incident edges. Note that, for $v \in V(F(G))$, setting $c_v = \deg_G(v) - 1$, then $F(G)$ has child sequence $\mathrm{c} = (c_v, v \in V(F(G))$.

The *kernel* of a connected graph G is the multigraph $K(G)$ constructed from G as follows. If $s(G) \geq 2$, then $K(G)$ is obtained from the core $C = C(G)$ by replacing each maximal path all of whose internal vertices have degree 2 by a single edge. If $s(G) = 1$ and G has a leaf, then we let ℓ be the leaf of minimum label, and define $K(G)$ to be a single loop edge at the core vertex $\alpha(\ell) \in V(C)$ nearest to ℓ. In the remaining cases, where G is either a tree or a cycle, $K(G)$ is empty. Note that if $s(G) \geq 2$, then $K(G) = K(C)$ and $V(K)$ is precisely the set of vertices $v \in V(C)$ with $\deg_C(v) \geq 3$. If G is not connected, but has connected components $G_1, \ldots, G_k$, then we define its kernel $K(G)$ to be the disjoint union of $K(G_1), \ldots, K(G_k)$. For $e \in E(K)$ we write $C(e)$ for the path in C corresponding to edge e, and we write $|C(e)|$ for the number of internal vertices on $C(e)$.

Now assume that $C(G)$ is non-empty. For each vertex $v \in V(G)$, write $\alpha(v)$ for the nearest element of $C(G)$ to v; if $v \in V(C(G))$ then $\alpha(v) = v$. For any vertices $u, v \in V(G)$, a shortest path between u and v in G may be decomposed into three parts: the (unique) shortest paths from u to $\alpha(u)$ and from v to $\alpha(v)$, and a shortest path between vertices $\alpha(u)$ and $\alpha(v)$ in $C(G)$. This shortest path between $\alpha(u)$ and $\alpha(v)$ can be further decomposed into three parts: a path from $\alpha(u)$ to a vertex $u' \in V(K)$, a path from $\alpha(v)$ to a vertex $v' \in V(K)$, and a shortest path between u' and v' in C. It follows that

$$\begin{aligned} \mathrm{diam}(G) &\leq 2(\mathrm{ht}(F(G)) + 1) + \mathrm{diam}(C(G)) && (1.2)\\ &\leq 2(\mathrm{ht}(F(G)) + 1) + (\mathrm{diam}(K) + 2) \max(|C(e)| + 1 : e \in E(K)). && (1.3) \end{aligned}$$

Therefore, in order to bound $\mathrm{diam}(G)$, it suffices to control $\mathrm{ht}(F(G))$, $\max(|C(e)| : e \in E(K)$, and $\mathrm{diam}(K(G))$. We next explain our techniques for each of these, in order.

Fix a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$, let $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$, and write $\mathbf{C} = C(\mathbf{G})$, $\mathbf{K} = K(\mathbf{G})$, and $\mathbf{F} = F(\mathbf{G})$. Writing $\mathbf{R} = \{v \in [n] : d_v = 1, v \notin V(\mathbf{F})\}$, then $V(\mathbf{C}) = [n] \setminus (\mathbf{R} \cup V(\mathbf{F}))$, and $\mathbf{F}$ has $\mathbf{s}$ roots, where

$$\mathbf{s} = |\mathbf{R}| + \sum_{v \in V(\mathbf{C})} (d_v - \deg_{\mathbf{C}}(v)) = \sum_{v \in [n] \setminus V(\mathbf{F})} (d_v - \deg_{\mathbf{C}}(v)) .$$

To reconstruct $\mathbf{G}$ from $\mathbf{F}$ and $\mathbf{C}$, it suffices to specify the partition $(\mathbf{F}^v, v \in \mathbf{R} \cup V(\mathbf{C}))$ of the trees of $\mathbf{F}$, where $\mathbf{F}^v$ is the set of trees of $\mathbf{F}$ whose root is incident to v in $\mathbf{G}$. The number of possible partitions is

$$\frac{\mathbf{s}!}{\prod_{v \in V(\mathbf{C})} (d_v - \deg_{\mathbf{C}}(v))!} = \frac{\mathbf{s}!}{\prod_{[n] \setminus V(\mathbf{F})} (d_v - \deg_{\mathbf{C}}(v))!} .$$

Since this number depends on $\mathbf{F}$ only through its vertex set $V(\mathbf{F})$, it follows that conditionally given $V(\mathbf{F})$, $\mathbf{F}$ is a uniformly random forest with child sequence $\mathrm{c} = (c_v, v \in V(\mathbf{F}))$.

Next, let $\mathbf{F}'$ be the homeomorphic reduction of $\mathbf{F}$, obtained by replacing by a single edge each maximal path of $\mathbf{F}$ all of whose internal vertices have exactly one child. Write $h = |V(\mathbf{F}) \setminus V(\mathbf{F}')|$ and $m = e(\mathbf{F}') = |E(\mathbf{F}')|$. To recover $\mathbf{F}$ from $\mathbf{F}'$ it suffices to specify, for each edge $e = uv \in E(\mathbf{F}')$, the (possibly empty) ordered sequence of one-child vertices appearing on the $u - v$ path in $\mathbf{F}$. The number of possibilities for this data is

$$h!|\mathcal{P}_{m,h}| = h!\binom{h+m-1}{m-1}.$$

This number only depends on $\mathbf{F}$ and $\mathbf{F}'$ through their vertex sets (since the degrees are specified, $e(\mathbf{F}')$ is determined by $V(\mathbf{F}')$), so conditionally given $V(\mathbf{F}')$, $\mathbf{F}'$ is a uniformly random forest with child sequence $\mathrm{c}' = (c_v, v \in V(\mathbf{F}'))$.

We now appeal to one of the main results of [2], which yields[1] the following theorem.

Theorem 1.4. *Let $\mathrm{c}' = (c_v, v \in S)$ be a 1-free child sequence, and let $\mathbf{F}' \in_u \mathcal{F}_{\mathrm{c}}$. Then for all $x \geq 0$, $\mathbb{P}\left\{\mathrm{ht}(\mathbf{F}') > x\sqrt{|S|}\right\} \leq 4e^{-x^2/2^8}$.*

If $\mathbf{G}$ is 2-free, then $\mathrm{c} = (c_v, v \in V(\mathbf{F}))$ is itself 1-free, so $\mathbf{F}' = \mathbf{F}$; since $\mathrm{diam}(\mathbf{G}) - \mathrm{diam}(\mathbf{C})) \leq 2(\mathrm{ht}(\mathbf{F})+1)$ and $|\mathbf{S}| = |V(\mathbf{F})| \leq n$, in this case the above theorem immediately implies that $\mathbb{E}\,\mathrm{diam}(\mathbf{G}) \leq \mathbb{E}\,\mathrm{diam}(C(\mathbf{G})) + O(\sqrt{n})$.

If we assume only that $n_2(\mathbf{G}) \leq (1-\epsilon)n$, then to bound $\mathrm{diam}(\mathbf{G})$ as above, we also need to control $\mathrm{ht}(\mathbf{F}) - \mathrm{ht}(\mathbf{F}')$. For this, we will essentially show that, on average, each edge in $\mathbf{F}'$ is subdivided $O(1/\epsilon)$ times by one-child vertices in $\mathbf{F}$ and, moreover, the number of times that a long path in $\mathbf{F}'$ is subdivided by one-child vertices is concentrated enough that $\mathbb{E}\,\mathrm{ht}(\mathbf{F}) = O(\mathbb{E}(\mathrm{ht}(\mathbf{F}') + \log n)/\epsilon)$. (In fact, we will also need to control the effect of subdivisions on distances in the core, and we will do this at the same time as we handle their effect on $\mathbf{F}$, so our eventual argument is slightly different from the sketch in this paragraph; but the idea is the same.)

We next explain our approach to bounding $\max(|\mathbf{C}(e)|+1 : e \in E(\mathbf{K}))$. To get the key ideas across, it is useful to assume that d is 2-free, and also to condition on the events that $\mathbf{K}$ is a simple graph and that for all $v \in V(\mathbf{K})$, $\deg_{\mathbf{G}}(v) = \deg_{\mathbf{K}}(v)$, so that v has no neighbours in $\mathbf{F}$. Write $m = |E(\mathbf{K})|$, and list the edges of K in lexicographic order as $e_1, \ldots, e_m$, where $e_i = u_i v_i$ and $u_i < v_i$. If $|\mathbf{C}(e_i)| > 0$ then let T_i be the (possibly empty) subgraph of $\mathbf{G}$ induced by the set of vertices $\{v \in V(\mathbf{G}) : \alpha(v) \in V(\mathbf{C}(e_i)) \setminus \{u_i, v_i\}\}$. If T_i is non-empty then let u_i' and v_i' be the vertices of T_i nearest to u_i and v_i, respectively; it is possible that $u_i = v_i$. Now define a tree $\mathbf{T}$ as follows: for each $i \in [m]$ with T_i non-empty let $j > i$ be minimal such that T_j is non-empty, and add an edge from v_i' to u_j'; if no such j exists then instead add an edge from v_i' to vertex 0. Root the resulting tree at u_{i_0}', where $i_0 = \min(i \in [m] : T_i \text{ is non-empty})$. An example appears in Figure 1.

Write $B = [n] \setminus V(\mathbf{K})$, and let $\mathrm{c} = (c_v, v \in B \cup \{0\})$ be the child sequence with $c_0 = 0$ and $c_v = d_v - 1$ for $n \in B$. Then $\mathbf{T}$ has child sequence c. Moreover, the graph $\mathbf{G}$ can be recovered from $\mathbf{T}$ together with the data $(|C(e_i)|, i \in [m])$. The set of internal vertices of $C(e_1), \ldots, C(e_m)$ are precisely the vertices on the path in $\mathbf{T}$ from its root to the leaf 0 (excluding 0); thus, conditionally, given $\mathbf{T}$, the number of possibilities for this data is $|\mathcal{P}_{m,\mathrm{ht}_{\mathbf{T}}(0)}|$. It follows that conditionally given (simple) $\mathbf{K}$ and given that $(\deg_{\mathbf{G}}(v), v \in V(\mathbf{K})) = (\deg_{\mathbf{K}}(v), v \in V(\mathbf{K}))$,

$$(\mathbf{T}, (|C(e_i)|, i \in [m])) \in_u \{(T, P) : T \in \mathcal{T}_{\mathrm{c}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}.$$

Thus, conditionally given m and given its child sequence c, $\mathbf{T}$ is distributed as a random tree with child sequence c chosen with probability proportional to the number of m-compositions of the root-to-0 path. We build on the techniques of [2] in order to establish

[1] See (2.8) in the proof of Theorem 6 in [2]. That result is stated for random trees with given child sequences, not random forests, but the extension to random forests is an immediate consequence of Lemma 6.6.

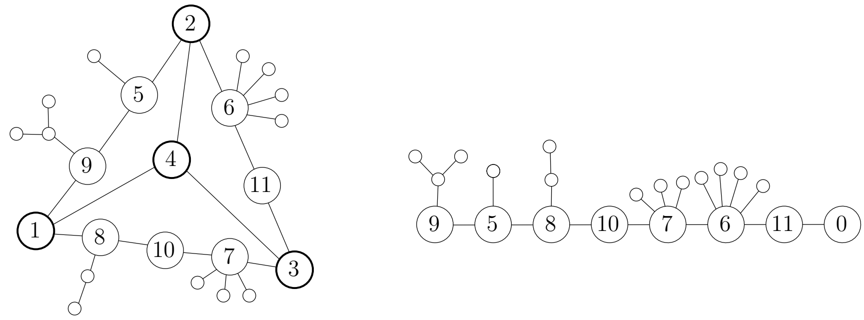

FIGURE 1. Left, a connected graph $\mathbf{G}$, with its kernel vertices circled in bold. Only labels of vertices in the core are shown, for readability. Right: The tree $\mathbf{T}$ constructed from (the trees hanging from the core paths of) $\mathbf{G}$, with root 9. Listing the kernel edges as $e_1, \ldots, e_6$ in lexicographic order, then the numbers of internal vertices of the paths $C(e_1), \ldots, C(e_6)$ are $(2, 3, 0, 2, 0, 0)$, which is a partition of $7 = \mathrm{ht}_{\mathbf{T}}(0)$ with $6 = e(K(\mathbf{G}))$ parts. The graph $\mathbf{G}$ can be recovered from $\mathbf{T}$ together with this partition.

that, up to a parity constraint, among all such composition-biased random rooted trees with a given 1-free child sequence, the height of a fixed leaf is stochastically maximized when the child sequence is binary. More precisely, we prove the following theorem. The *parity* of an integer n is the remainder of n modulo 2.

Theorem 1.5. *Fix positive integers m and n, and let ε be the parity of n. Let $\mathrm{c} = (c_0, \ldots, c_n)$ be a 1-free child sequence with $c_0 = 0$ and let $\mathrm{b} = (b_0, \ldots, b_{n+\varepsilon})$ be a binary child sequence with $b_0 = 0$. Then with $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) : T \in \mathcal{T}_{\mathrm{c}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$ and $(\mathbf{S}, \mathbf{Q}) \in_u \{(S, Q) : S \in \mathcal{T}_{\mathrm{b}}, Q \in \mathcal{P}_{m,\mathrm{ht}_S(0)}\}$,*

$$\mathrm{ht}_{\mathbf{T}}(0) \preceq_{\mathrm{st}} \mathrm{ht}_{\mathbf{S}}(0).$$

The proof of this theorem, which appears in Section 3 is surprisingly challenging, and requires an auxiliary stochastic domination result which is proved by studying the asymptotic behaviour of a Markov chain reminiscent of importance sampling. Once the stochastic domination is proved, however, we are able to establish the following sub-Gaussian tail bound by proving a corresponding bound for the height of a fixed node in a composition-biased binary tree.

Theorem 1.6. *Fix positive integers m and n with $n \geq 64m$. Let $\mathrm{c} = (c_0, \ldots, c_n)$ be a 1-free child sequence with $c_0 = 0$ and let $\mathbf{h} = \mathrm{ht}_{\mathbf{T}}(0)$ where $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) : T \in \mathcal{T}_{\mathrm{c}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$. Then for all $x \geq 0$,*

$$\mathbb{P}\left\{\mathbf{h} \geq 2\sqrt{mn} + x\sqrt{n}\right\} \leq \exp\left(-\frac{x^2}{3} + 4\right).$$

The proof of Theorem 1.6 appears in Section 4. We combine its tail bound with fairly straightforward arguments for the sizes of entries in a uniformly random composition, to show that for any fixed edge e_i, $|\mathbf{C}(e_i)|/(1 + \sqrt{n/m})$ has sub-exponential tails, and so in particular $\mathbb{E}\max(|\mathbf{C}(e_i)| : i \in [m]) = O((1 + \sqrt{n/m})\log m)$ by a union bound. (The arguments become somewhat more complicated when we remove the assumptions that $\mathbf{K}$ is simple and that d is 2-free, but the key ideas are present in the above.)

Given the above, the remaining step in the proof of Theorem 1.1 is to bound the diameter of the kernel. At this point, we stop assuming that the kernel is a simple graph and that d is 2-free.

Write $\mathcal{G}_{\mathrm{d}}^{-}$ for the set of all (not necessarily connected) graphs with degree sequence d and no cycle components, and let $\mathbf{C}^{-}$ be obtained from $\mathbf{C}$ by deleting all cycle components.

Note that, conditionally given the degree sequence $\mathrm{d}^- = (\deg_{\mathbf{C}^-}(v), v \in V(\mathbf{C}^-))$ of $\mathbf{C}^-$, then $\mathbf{C}^- \in_u \mathcal{G}^-_{\mathrm{d}^-}$, and d^- is a degree sequence with minimum degree 2. Moreover, $\mathbf{K} = K(\mathbf{C}^-)$; so, to prove uniform tail bounds on the diameter of the kernel, it suffices to prove corresponding bounds for kernels $K(\mathbf{C})$ of random graphs $\mathbf{C} \in_u \mathcal{G}^-_{\mathrm{d}}$ that hold uniformly over degree sequences with minimum degree 2. We do so in Section 5, where we prove the following theorem.

Theorem 1.7. *Fix positive integers $n \le \ell$ and a 1-free degree sequence* d *of the form $\mathrm{d} = (d_1, \ldots, d_\ell) = (d_1, \ldots, d_n, 2, \ldots, 2)$ with $d_v \ge 3$ for all $v \in [n]$. Take $\mathbf{C} \in_u \mathcal{G}^-_{\mathrm{d}}$ and let $\mathbf{K} = K(\mathbf{C})$. Write $2m = \sum_{v\in[n]} d_v$ so that if $C \in \mathcal{G}^-_{\mathrm{d}}$, then $K(C)$ has m edges. Then*

$$\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{K}) \ge 500 \log m\right\} = O\left(\frac{\log^8 m}{m}\right).$$

Note that because d is 1-free, graphs $C \in \mathcal{G}^-_{\mathrm{d}}$ are cores, each of whose components have surplus at least 2. Also, observe that Theorem 1.3, follows immediately from the case $\ell = n$ of Theorem 1.7, since in this case $\mathcal{G}_{\mathrm{d}} = \mathcal{G}^-_{\mathrm{d}}$ and $\mathbf{K} = \mathbf{C}$.

To prove Theorem 1.7, we consider a breadth-first exploration of the kernel of the random core. At each time, there will be a queue of vertices which have been discovered but whose neighbourhoods still need to be explored. We will show that with high probability, until the queue is extremely large, its size grows exponentially as a function of the radius of the explored subgraph. We then show that if the queues in two explorations started from different vertices become large enough, they must link up with high probability. We prove that all of these events must occur with high enough probability that a union bound over starting locations for the exploration allows us to conclude.

To implement the above argument, we use the technique of *switching*, independently introduced in [23, 31] (see also [15, 27] for some other early works using the technique), which has found substantial use in the combinatorial analysis of graphs with given degrees, as well as as in developing Markov chains for sampling from random graphs with given degrees (see the ICM article [28] for a fairly recent overview of the uses of the method). In its simplest incarnation, the technique works as follows. Let G be a graph and let $u, v, x, y \in V(G)$ be such that $\{u, v\}, \{x, y\} \in E(G)$ and $\{u, x\}, \{v, y\} \notin E(G)$. Then *switching on uv and xy* consists of removing the edges $\{u, v\}, \{x, y\}$ and adding the edges $\{u, x\}, \{v, y\}$. Note that in the resulting graph G', all vertices have the same degrees as they do in G.

As a very simple example of how switching can be used to prove probability bounds, consider the all-3's degree sequence $\mathrm{d} = (3, \ldots, 3)$ of length n. Let $\mathcal{A} = \{G \in \mathcal{G}_{\mathrm{d}} : G|_{[4]} \text{ is a clique}\}$, and let $\mathcal{B} \subset \mathcal{G}_{\mathrm{d}} \setminus \mathcal{A}$ be the set of graphs G' that can be obtained from A by a switching that uses the edge 12. For each graph in $G \in \mathcal{A}$, there are exactly $3n/2 - 6$ such switchings. On the other hand, for each graph $G' \in \mathcal{B}$, there is exactly one switching which results in a graph in $\mathcal{A}$; so $|\mathcal{B}| \ge (3n/2 - 6)|\mathcal{A}|$. It follows that if $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$ then

$$\mathbb{P}\left\{\mathbf{G}|_{[4]} \text{ is a clique}\right\} = \mathbb{P}\left\{\mathbf{G} \in \mathcal{A}\right\} = \frac{|\mathcal{A}|}{|\mathcal{G}_{\mathrm{d}}|} \le \frac{|\mathcal{A}|}{|\mathcal{A} \cup \mathcal{B}|} \le \frac{1}{3n/2 - 5}.$$

Similarly, let $\mathcal{B}'$ be the set of graphs G'' that can be obtained from a graph G in A by first switching on 12 and ab, then on 13 and cd, where $a, b, c, d \in \{5, \ldots, n\}$ and $\{a, b\}, \{c, d\} \in E(G)$ are non-incident edges. Then the same logic as above shows that $\mathcal{B}'' \ge (3n/2 - 6)(3n/2 - 11)|\mathcal{A}|$, so in fact $\mathbb{P}\left\{\mathbf{G}|_{[4]} \text{ is a clique}\right\} = O\left(n^{-2}\right)$.

The power of switching is that it allows us to compare the probabilities of two events by counting the number of switchings between them. Because switching is a local operation, this is generally much easier than directly counting the sizes of the sets corresponding to the events. We postpone discussion of the details of the specific switching arguments that arise in our proofs, since setting them up carefully is rather involved; but at a high level they all boil down to applications of the following lemma.

Lemma 1.8. *Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint subsets of a set $\mathcal{C} \subset \mathcal{D}$, and let $\mathcal{H}$ be a bipartite multigraph with parts $\mathcal{A}$ and $\mathcal{B}$. Suppose that every $A \in \mathcal{A}$ has degree at least a in $\mathcal{H}$, and every $B \in \mathcal{B}$ has degree at most b in $\mathcal{H}$. Then with $\mathbf{D} \in_u \mathcal{D}$, we have*

$$a\mathbb{P}\{\mathbf{D} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\} \leq b\mathbb{P}\{\mathbf{D} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\}.$$

Proof. It follows from the degree conditions that $\mathcal{H}$ has at least $a|\mathcal{A}|$ edges and at most $b|\mathcal{B}|$ edges, so $a|\mathcal{A}| \leq b|\mathcal{B}|$. The result then follows from the fact that $\mathbf{D}$ is a uniform sample from $\mathcal{D}$, so conditionally given that $\mathcal{D} \in \mathcal{A} \cup \mathcal{B}$ we have $\mathbf{D} \in_u \mathcal{A} \cup \mathcal{B}$. □

When we apply this lemma, $\mathcal{D}$ will frequently be a subset of $\mathcal{G}_{\mathrm{d}}$ for some degree sequence $\mathcal{D}$, and an edge of $\mathcal{H}$ from $A \in \mathcal{A}$ to $B \in \mathcal{B}$ will represent a switching or sequence of switchings that transforms graph A into graph B. Also, sometimes $\mathcal{D}$ will in fact be a set of multigraphs (this is required since since the kernel of a simple graph can be a multigraph); in this case, the representation of graphs as matchings of half-edges is useful for cleanly writing the switching arguments. We also require a more involved version of switching called *path switching*, introduced with a slightly different formalism in [24], which reroutes pairs of core paths between kernel vertices, rather than just pairs of edges.

To summarize, we have the following bounds: the forest $\mathbf{F} = F(\mathbf{G})$, which is essentially the spanned by the set of vertices not lying in cycles, has height $O(\sqrt{n})$ in expectation and has sub-Gaussian tails on this order; the greatest length of a path of the core $\mathbf{C} = C(\mathbf{G})$ between adjacent kernel vertices is $O((1 + \sqrt{n/m}) \log m)$ in expectation, where m is the number of kernel edges, and has and has sub-exponential tails on this order; and the diameter of the kernel $\mathbf{K} = K(\mathbf{G})$ is $O(\log m)$ except with probability $O(\log^8 m/m)$. Combining these bounds with the inequality (1.2), and being careful with the contributions to the expectation from small-probability events, is enough to conclude.

At this point, it seems like a good idea to turn to the details.

2. The simple kernel

Loops and multi-edges complicate analysis of kernels because they have symmetries which must be accounted for (permuting parallel edges and reversing loops) and because they introduce constraints on the number of core vertices which must lie on a given kernel edge (loops must be subdivided at least twice and at most one copy of any edge can be non-subdivided). To account for both problems, we augment the kernel with some additional data which breaks these symmetries and accounts for these constraints.

Fix a graph G, let $C = C(G)$ and let $K = K(G)$. The *simple kernel* of G, denoted $K^* = K^*(G)$, is constructed from C by deleting all cycle and tree components, then replacing $C(e)$ for each $e \in E(K)$ with a short path or cycle in the following way. If e is a loop, then replace the subpath of $C(e)$ between the two neighbours of $e^- = e^+$ with a single edge e^*. If e has distinct endpoints and $\mathrm{mult}(e) = 1$, then replace $C(e)$ with a single edge e^*. If e has distinct endpoints, $\mathrm{mult}(e) \geq 2$, and $|C(e)| \geq 1$, then replace the subpath of $C(e)$ between e^- and the neighbour of e^+ on $C(e)$ with a single edge e^*. In the only remaining case, which is that e has distinct endpoints, $\mathrm{mult}(e) \geq 2$, and $|C(e)| = 0$, we leave $C(e)$ unchanged; in this case e^* is not defined. Note that if K is empty then K^* is empty, and if G has connected components $G_1, \ldots, G_k$ then $K^*(G)$ is the disjoint union of $K^*(G_1), \ldots, K^*(G_k)$.

We call the edges e^* *mutable* and the remaining edges *immutable*. Let $E^*(K^*)$ be the set of all mutable edges, and set $e^*(K^*) = |E^*(K^*)|$. For $e^* \in E^*(K^*)$, we write $C(e^*)$ for the path in C replaced by e^*, and write $|C(e^*)|$ for the number of internal vertices on $C(e^*)$. When reconstructing C from K^*, the mutable edges may be subdivided, but the immutable edges must not be. To justify the notation $E^*(K^*)$, observe that K^* is connected and $s(K^*) \geq 2$, then we can tell which edges in K^* are mutable. Indeed, we have $K(K^*) = K$, and to decide if $e \in E(K^*)$ is mutable or immutable, consider the edge $f \in E(K)$ such that e lies on $K^*(f)$. Then e is immutable if and only if one of its endpoints is not in $V(K)$, and the other endpoint is f^+. When K^* is connected and $s(K^*) = 1$ though, K^* is a triangle and has a unique mutable edge, but from the graph structure of this triangle alone, we cannot tell which of its three edges is mutable. In this case, K^* therefore carries not only the data of this triangle, but also which of its three edges is mutable.

We call K^* the simple kernel because it is simple and homeomorphic to K with each edge being subdivided at most twice. In particular, $e(K) \leq e(K^*) \leq 3e(K)$ and $v(K) \leq v(K^*) \leq v(K) + 2e(K)$. Furthermore, for all $e \in E(K)$, there is exactly one mutable edge on $K^*(e)$ unless $\mathrm{mult}(e) \geq 2$ and $|C(e)| = 0$. Because only one edge $f \in E(K)$ parallel to e can have $|C(f)| = 0$, this implies $e(K)/2 \leq e^*(K^*) \leq e(K)$ and $e(K^*)/3 \leq e^*(K^*) \leq e^*(K)$. We also have $\mathrm{diam}(K) \leq \mathrm{diam}(K^*) \leq 2\,\mathrm{diam}(K) + 2$. Indeed in K^*, every vertex has a neighbour in K, and for every pair of vertices $u, v \in V(K)$ that are adjacent in K, there is a path of length at most 2 between u and v in K^*.

Fix a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$, let K^* be the simple kernel of a graph in $\mathcal{C}_{\mathrm{d}}$, and suppose that K^* is nonempty. Define $\mathcal{C}_{\mathrm{d}}(K^*) = \{G \in \mathcal{C}_{\mathrm{d}} : K^*(G) = K^*\}$. We next describe a bijective encoding of $\mathcal{C}_{\mathrm{d}}(K^*)$ which we use for analyzing the structure of a typical element. Let $\mathrm{c} = (c_0, \ldots, c_n)$ be the child sequence given by

$$c_v = \begin{cases} 0 & v = 0 \\ d_v - 1 & v \in [n] \setminus V(K^*) \\ d_v - \deg_{K^*}(v) & v \in V(K^*) \end{cases}, \tag{2.1}$$

and note that if $\mathrm{d}|_{[n]\setminus V(K^*)}$ is 2-free then $\mathrm{c}|_{[0,n]\setminus V(K^*)}$ is 1-free. Here and below, we write $[0, n] = \{0, 1, \ldots, n\}$. Let $m = e^*(K^*)$ and define

$$\mathcal{X}'_{\mathrm{d}}(K^*) = \bigcup_{(A,B)} \{(F, T, P) : F \in \mathcal{F}_{\mathrm{c}|_A, V(K^*)}, T \in \mathcal{T}_{\mathrm{c}|_B}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\},$$

where the union runs over all partitions (A, B) of $[0, n]$ such that $V(K^*) \subseteq A$ and $0 \in B$. If $s(K^*) \geq 2$, then set $\mathcal{X}_{\mathrm{d}}(K^*) = \mathcal{X}'_{\mathrm{d}}(K^*)$. If $s(K^*) = 1$, then K^* is a triangle with a unique mutable edge e^*. Let v^* be the vertex of K^* not incident to e^*. In this case we let $\mathcal{X}_{\mathrm{d}}(K^*)$ be the set of triples $(F, T, P) \in \mathcal{X}'_{\mathrm{d}}(K^*)$ such that the component tree of F rooted at v^* contains the leaf $\ell = \min(v \in [n] : d_v = 1)$ of minimum label in d.

We now provide a bijection $\mathcal{X}_{\mathrm{d}}(K^*) \to \mathcal{C}_{\mathrm{d}}(K^*)$. Given a triple $(F, T, P) \in \mathcal{X}_{\mathrm{d}}(K^*)$, we need to construct a graph $G \in \mathcal{C}_{\mathrm{d}}(K^*)$. Write $P = (P_1, \ldots, P_m)$ and let $(w_1, \ldots, w_h, w_{h+1})$ be the sequence of vertices on the path in T from its root to the leaf 0, so w_1 is the root of T, $w_{h+1} = 0$, and $h = \mathrm{ht}_T(0)$. List the mutable edges of K^* in lexicographic order as $e_1^*, \ldots, e_m^*$, and for $i \in [m]$, let $u_i = (e_i^*)^-$ and $v_i = (e_i^*)^+$. We divide T into a sequence of smaller trees $T_1, \ldots, T_m$, where for $i \in [m]$, the tree T_i is obtained from T as follows. If $i = 1$, then give w_1 a new leaf neighbour with label u_1^*, and subdivide the edge between w_{P_1+1} and its parent w_{P_1} with a vertex labeled v_1^*. Then remove the edge between v_1^* and its child. There are two resulting components, and T_1 is the component containing u_1^* and v_1^*. If $i \geq 2$, then subdivide the edge between $w_{P_1+\cdots+P_{i-1}+1}$ and its parent $w_{P_1+\cdots+P_{i-1}}$ with a vertex labeled u_i^*, and subdivide the edge between $w_{P_1+\cdots+P_i+1}$ and its parent $w_{P_1+\cdots+P_i}$ with a vertex labeled v_i^*. Then remove the edges in the resulting tree between u_i^* and its parent and between v_i^* and its child. There are three resulting components, and T_i is the component containing u_i^* and v_i^*. Let G be the graph on $[n]$ with edge set

$$E(G) = (E(K^*) \setminus E^*(K^*)) \cup E(F) \cup E(T_1) \cup \cdots \cup E(T_m).$$

It is not hard to see that $G \in \mathcal{C}_{\mathrm{d}}(K^*)$, so this map is well-defined. We argue that it is bijective by describing its inverse. Fix a graph $G \in \mathcal{C}_{\mathrm{d}}(K^*)$, and let $C = C(G)$. Let F be the forest obtained G by deleting all edges in C, rooting each tree at its unique vertex in C, then deleting all trees in the resulting forest not rooted at a vertex of K^*. Let $P = (P_1, \ldots, P_m)$ be the m-composition where $P_i = |C(e_i^*)|$ for $i \in [m]$. For $i \in [m]$, let T_i be the tree consisting of the path $C(e_i^*)$, together with all trees in $F(G)$ rooted at an internal vertex of $C(e_i^*)$. For each $i \in [m-1]$, identify $v_i^* \in V(T_i)$ with $u_{i+1}^* \in V(T_{i+1})$, then remove this vertex and join its two neighbours with an edge. Finally, remove the leaf u_1^* and root T at its child, and relabel the leaf v_m^* with 0. Let T be the resulting tree. One can check that this construction of (F, T, P) inverts the map $\mathcal{X}_{\mathrm{d}}(K^*) \to \mathcal{C}_{\mathrm{d}}(K^*)$.

Due to this bijection, a uniformly random graph $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}(K^*)$, can by encoded by a uniformly random triple $(\mathbf{F}, \mathbf{T}, \mathbf{P}) \in_u \mathcal{X}_{\mathrm{d}}(K^*)$. Under this bijection, writing $\mathbf{C} = C(\mathbf{G})$, the number of internal vertices on $\mathbf{C}(e_i^*)$ is given by $\mathbf{P}_i$. In particular, the total number of vertices in $\mathbf{C}$ but outside the simple kernel is equal to $\sum_{i=1}^m \mathbf{P}_i = \mathrm{ht}_{\mathbf{T}}(0)$. This implies the following proposition. It is not hard to see that $G \in \mathcal{C}_{\mathrm{d}}(K^*)$, so this map is well-defined. We argue that it is bijective by describing its inverse. Fix a graph $G \in \mathcal{C}_{\mathrm{d}}(K^*)$, and let $C = C(G)$. Let F be the forest obtained G by deleting all edges in C, rooting each tree at its unique vertex in C, then deleting all trees in the resulting forest not rooted at a vertex of K^*. Let $P = (P_1, \ldots, P_m)$ be the m-composition where $P_i = |C(e_i^*)|$ for $i \in [m]$. For $i \in [m]$, let T_i be the tree consisting of the path $C(e_i^*)$, together with all trees in $F(G)$ rooted at an internal vertex of $C(e_i^*)$. For each $i \in [m-1]$, identify $v_i^* \in V(T_i)$ with $u_{i+1}^* \in V(T_{i+1})$, then remove this vertex and join its two neighbours with an edge. Finally, remove the leaf u_1^* and root T at its child, and relabel the leaf v_m^* with 0. Let T be the resulting tree. One can check that this construction of (F, T, P) inverts the map $\mathcal{X}_{\mathrm{d}}(K^*) \to \mathcal{C}_{\mathrm{d}}(K^*)$.

Due to this bijection, a uniformly random graph $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}(K^*)$, can by encoded by a uniformly random triple $(\mathbf{F}, \mathbf{T}, \mathbf{P}) \in_u \mathcal{X}_{\mathrm{d}}(K^*)$. Under this bijection, writing $\mathbf{C} = C(\mathbf{G})$, the number of internal vertices on $\mathbf{C}(e_i^*)$ is given by $\mathbf{P}_i$. In particular, the total number of vertices in $\mathbf{C}$ but outside the simple kernel is equal to $\sum_{i=1}^m \mathbf{P}_i = \mathrm{ht}_{\mathbf{T}}(0)$. This implies the following proposition.

Proposition 2.1. *Let* $\mathrm{d} = (d_1, \ldots, d_n)$ *be a degree sequence, let* K^* *be the simple kernel of a graph in* $\mathcal{C}_{\mathrm{d}}$*, and suppose that* K^* *is nonempty. Let* $(\mathbf{F}, \mathbf{T}, \mathbf{P}) \in_u \mathcal{X}_{\mathrm{d}}(K^*)$*, let* $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}(K^*)$*, and set* $\mathbf{C} = C(\mathbf{G})$*. Write* $m = e^*(K^*)$ *and list the mutable edges of* K^* *in lexicographic order as* $e_1^*, \ldots, e_m^*$*. Then*

$$(|\mathbf{C}(e_1^*)|, \ldots, |\mathbf{C}(e_m^*)|) \stackrel{d}{=} (\mathbf{P}_1, \ldots, \mathbf{P}_m).$$

In particular, $v(\mathbf{C}) - v(K^*) \stackrel{d}{=} \mathrm{ht}_{\mathbf{T}}(0)$.

For any partition (A, B) of $[0, n]$ with $V(K^*) \subseteq A$ and $0 \in B$, given that $(V(\mathbf{F}), V(\mathbf{T})) = (A, B)$, we have $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) \mid T \in \mathcal{T}_{\mathrm{c}}|_B, P \in \mathcal{P}_{m, \mathrm{ht}_T(0)}\}$. Since B is a subset of $[0, n] \setminus V(K^*)$ containing 0, $\mathrm{c}|_B$ is a child sequence in which 0 is a leaf. Moreover if $\mathrm{d}|_{[n] \setminus V(K^*)}$ is 2-free then $\mathrm{c}|_B$ is 1-free. Therefore if we can prove a tail bound on $\mathrm{ht}_{\mathbf{S}}(0)$ where $(\mathbf{S}, \mathbf{Q}) \in_u \{(S, Q) \mid S \in \mathcal{T}_{\mathrm{c}}, Q \in \mathcal{P}_{m, \mathrm{ht}_S(0)}\}$ which is uniform over all 1-free child sequences $\mathrm{c} = (c_0, \ldots, c_n)$ with $c_0 = 0$ and which is increasing in n, then it will apply without change to $v(\mathbf{C}) - v(K^*)$ provided that $d_v \neq 2$ for all $v \in [n] \setminus V(K^*)$. In order to understand $v(\mathbf{C})$, like in the proof outline, it therefore suffices to study the height of 0 in a random rooted tree with a given 1-free child sequence $\mathrm{c} = (c_0, \ldots, c_n)$ such that $c_0 = 0$, where the probability of selecting a given tree $T \in \mathcal{T}_{\mathrm{c}}$ is proportional to the number of m-compositions of $\mathrm{ht}_T(0)$.

3. Stochastic domination of biased trees

This section is devoted to the proof of Theorem 1.5. For a random variable $\mathbf{X}$ and an event E with $\mathbb{P}\{E\} > 0$, we use the notation $\mathbf{Y} \stackrel{d}{=} (\mathbf{X} \mid E)$ to mean that $\mathbb{P}\{\mathbf{Y} \in B\} = \mathbb{P}\{\mathbf{X} \in B \mid E\}$ for every measurable set B. Whenever we refer to a random variable of the form $(\mathbf{X} \mid E)$, we implicitly assume that $\mathbb{P}\{E\} > 0$. Our proof of Theorem 1.5 hinges on Proposition 3.2, which in turn makes repeated use of the following easily verified fact.

Lemma 3.1. *For any random variable $\mathbf{X}$ and any $a \le b$,*

$$(\mathbf{X} \mid \mathbf{X} \ge a) \preceq_{\mathrm{st}} (\mathbf{X} \mid \mathbf{X} \ge b), \text{ and}$$
$$(\mathbf{X} \mid \mathbf{X} \le a) \preceq_{\mathrm{st}} (\mathbf{X} \mid \mathbf{X} \le b).$$

Proposition 3.2. *Let $\mathbf{X}_1, \mathbf{Y}_1, \mathbf{X}_2, \mathbf{Y}_2$ be discrete random variables such that $\mathbf{X}_i$ and $\mathbf{Y}_i$ are independent for $i \in \{1, 2\}$. Suppose that for all x, $(\mathbf{X}_1 \mid \mathbf{X}_1 \le x) \preceq_{\mathrm{st}} (\mathbf{X}_2 \mid \mathbf{X}_2 \le x)$ and for all y, $(\mathbf{Y}_1 \mid \mathbf{Y}_1 \ge y) \preceq_{\mathrm{st}} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \ge y)$. Then*

$$(\mathbf{Y}_1 \mid \mathbf{Y}_1 \ge \mathbf{X}_1) \preceq_{\mathrm{st}} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \ge \mathbf{X}_2).$$

Proof. We first prove the special case of the lemma when $\mathbf{X}_1 = \mathbf{X}_2$.

Claim 3.3. *Let $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$ be discrete random variables such that $\mathbf{W}$ is independent of $\mathbf{U}$ and $\mathbf{V}$. Suppose that for all w, $(\mathbf{U} \mid \mathbf{U} \ge w) \preceq_{\mathrm{st}} (\mathbf{V} \mid \mathbf{V} \ge w)$. Then*

$$(\mathbf{U} \mid \mathbf{U} \ge \mathbf{W}) \preceq_{\mathrm{st}} (\mathbf{V} \mid \mathbf{V} \ge \mathbf{W}).$$

Proof. Choose u_0, v_0, and w_0 in the supports of $\mathbf{U}$, $\mathbf{V}$, and $\mathbf{W}$ respectively such that $v_0 \ge u_0 \ge w_0$. Consider the Markov chain $((\mathbf{U}_t, \mathbf{W}_t), t \ge 0)$ started at $(\mathbf{U}_0, \mathbf{W}_0) = (u_0, w_0)$ specified by the following transition rules: for $t \ge 1$, and for $u \ge w$,

$$(\mathbf{U}_t \mid \mathbf{W}_{t-1} = w) \stackrel{\mathrm{d}}{=} (\mathbf{U} \mid \mathbf{U} \ge w),$$
$$(\mathbf{W}_t \mid \mathbf{U}_t = u) \stackrel{\mathrm{d}}{=} (\mathbf{W} \mid \mathbf{W} \le u),$$

$\mathbf{U}_t$ is conditionally independent of $\mathbf{U}_{t-1}$ given that $\mathbf{W}_{t-1} = w$, and $\mathbf{W}_t$ is conditionally independent of $\mathbf{W}_{t-1}$ given that $\mathbf{U}_t = u$. It is not hard to check that this chain is irreducible and aperiodic. A routine calculation using the independence of $\mathbf{U}$ and $\mathbf{W}$ shows that its stationary measure, and therefore its limit in total variation distance, is $((\mathbf{U}, \mathbf{W}) \mid \mathbf{U} \ge \mathbf{W})$. The analogous chain $((\mathbf{V}_t, \mathbf{W}'_t), t \ge 0)$ started at $(\mathbf{V}_0, \mathbf{W}'_0) = (v_0, w_0)$ converges in total variation distance to its stationary distribution $((\mathbf{V}, \mathbf{W}) \mid \mathbf{V} \ge \mathbf{W})$.

We will inductively construct a coupling of the two chains such that almost surely $\mathbf{U}_t \le \mathbf{V}_t$ and $\mathbf{W}_t \le \mathbf{W}'_t$ for all $t \ge 0$. The base case is automatic since $\mathbf{U}_0 = u_0 \le v_0 = \mathbf{V}_0$ and $\mathbf{W}_0 = w_0 = \mathbf{W}'_0$. Suppose then that $\mathbf{W}_{t-1}$ and $\mathbf{W}'_{t-1}$ are coupled so that $\mathbf{W}_{t-1} \le \mathbf{W}'_{t-1}$ almost surely. For any $w \le w'$, using Lemma 3.1 and our assumption,

$$(\mathbf{U}_t \mid \mathbf{W}_{t-1} = w) \stackrel{\mathrm{d}}{=} (\mathbf{U} \mid \mathbf{U} \ge w) \preceq_{\mathrm{st}} (\mathbf{U} \mid \mathbf{U} \ge w') \preceq_{\mathrm{st}} (\mathbf{V} \mid \mathbf{V} \ge w') \stackrel{\mathrm{d}}{=} (\mathbf{V}_t \mid \mathbf{W}'_{t-1} = w').$$

Since $\mathbf{W}_{t-1} \le \mathbf{W}'_{t-1}$ almost surely, it follows that $\mathbf{U}_t \preceq_{\mathrm{st}} \mathbf{V}_t$, so there is a coupling $(\mathbf{U}_t, \mathbf{V}_t)$ with $\mathbf{U}_t \le \mathbf{V}_t$ almost surely. It remains to couple $\mathbf{W}_t$ and $\mathbf{W}'_t$. For any $u \le v$, using Lemma 3.1,

$$(\mathbf{W}_t \mid \mathbf{U}_t = u) \stackrel{\mathrm{d}}{=} (\mathbf{W} \mid \mathbf{W} \le u) \preceq_{\mathrm{st}} (\mathbf{W} \mid \mathbf{W} \le v) \stackrel{\mathrm{d}}{=} (\mathbf{W}'_t \mid \mathbf{V}_t = v).$$

Since $\mathbf{V}_t \le \mathbf{U}_t$ almost surely, it follows that there is a coupling $(\mathbf{W}_t, \mathbf{W}'_t)$ with $\mathbf{W}_t \le \mathbf{W}'_t$ almost surely.

Let $\widehat{\mathbf{U}} \stackrel{\mathrm{d}}{=} (\mathbf{U} \mid \mathbf{U} \ge \mathbf{W})$ and $\widehat{\mathbf{V}} \stackrel{\mathrm{d}}{=} (\mathbf{V} \mid \mathbf{V} \ge \mathbf{W})$, so $\mathbf{U}_t \to \widehat{\mathbf{U}}$ and $\mathbf{V}_t \to \widehat{\mathbf{V}}$ in total variation distance. Then the coupled sequence $((\mathbf{U}_t, \mathbf{V}_t), t \ge 0)$ yields a coupling $(\widehat{\mathbf{U}}, \widehat{\mathbf{V}})$ of the limits, and because $\mathbf{U}_t \le \mathbf{V}_t$ for all $t \ge 0$ almost surely, $\widehat{\mathbf{U}} \le \widehat{\mathbf{V}}$ almost surely. ■

With two applications of the claim above, we can deduce the lemma. We may assume that $\mathbf{X}_1$, $\mathbf{Y}_1$, $\mathbf{X}_2$, $\mathbf{Y}_2$ share a common probability space in which $\mathbf{X}_1$ and $\mathbf{Y}_2$ are independent. By Claim 3.3 with $\mathbf{U} = \mathbf{Y}_1$, $\mathbf{V} = \mathbf{Y}_2$, and $\mathbf{W} = \mathbf{X}_1$,

$$(\mathbf{Y}_1 \mid \mathbf{Y}_1 \geq \mathbf{X}_1) \preceq_{\mathrm{st}} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_1).$$

By Claim 3.3 with $\mathbf{U} = -\mathbf{X}_2$, $\mathbf{V} = -\mathbf{X}_1$, and $\mathbf{W} = -\mathbf{Y}_2$, there is a coupling $(\widehat{\mathbf{X}}_1, \widetilde{\mathbf{X}}_2)$ of $(\mathbf{X}_1 \mid \mathbf{X}_1 \leq \mathbf{Y}_2)$ and $(\mathbf{X}_2 \mid \mathbf{X}_2 \leq \mathbf{Y}_2)$ such that $\widehat{\mathbf{X}}_1 \leq \widetilde{\mathbf{X}}_2$ almost surely. Now for each $a \leq b$, Lemma 3.1 yields a coupling of $(\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq a)$ and $(\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq b)$ such that the former is at most the latter almost surely. Obtain $(\widehat{\mathbf{Y}}_2, \widetilde{\mathbf{Y}}_2)$ by first sampling $(\widehat{\mathbf{X}}_1, \widetilde{\mathbf{X}}_2)$, and then given that $(\widehat{\mathbf{X}}_1, \widetilde{\mathbf{X}}_2) = (a, b)$, sampling $(\widehat{\mathbf{Y}}_2, \widetilde{\mathbf{Y}}_2)$ according the coupling of $(\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq a)$ and $(\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq b)$. Then $\widehat{\mathbf{Y}}_2 \leq \widetilde{\mathbf{Y}}_2$ almost surely and it is straightforward to check that $\widehat{\mathbf{Y}}_2 \overset{\mathrm{d}}{=} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_1)$ and $\widetilde{\mathbf{Y}}_2 \overset{\mathrm{d}}{=} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_2)$. We have thus shown

$$(\mathbf{Y}_1 \mid \mathbf{Y}_1 \geq \mathbf{X}_1) \preceq_{\mathrm{st}} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_1) \preceq_{\mathrm{st}} (\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_2),$$

as required. □

Having established this general lemma, we turn to realizing $\mathrm{ht}_{\mathbf{T}}(0)$ and $\mathrm{ht}_{\mathbf{S}}(0)$ as random variables of the form $(\mathbf{Y}_1 \mid \mathbf{Y}_1 \geq \mathbf{X}_1)$ and $(\mathbf{Y}_2 \mid \mathbf{Y}_2 \geq \mathbf{X}_2)$ for some choice of $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2$ satisfying the necessary hypotheses. For this we will require a bijection due to Foata and Fuchs introduced first in [19], which we refer to as the (discrete) line-breaking construction and describe shortly. The present work adds to the list of probabilistic uses of the discrete line-breaking construction, also called the Foata-Fuchs bijection. For example, the proof of Theorem 1.4 ([2, Theorem 6]) relies heavily on the line-breaking construction, and the article [3] gives several variations and consequences of it. Given a child sequence $\mathrm{c} = (c_0, \ldots, c_n)$ with $\sum_{v \in [0,n]} c_v = n$, we write $\mathcal{V}_{\mathrm{c}}$ for the set of all sequences $V = (V_1, \ldots, V_n)$ of length n with alphabet $[0, n]$ such that $|\{i \in [n] : V_i = v\}| = c_v$ for all $v \in [0, n]$.

The variant of the line-breaking construction that we make use of is the bijection $\mathcal{T}_{\mathrm{c}} \to \mathcal{V}_{\mathrm{c}}$, which constructs a sequence $V = (V_1, \ldots, V_n) \in \mathcal{V}_{\mathrm{c}}$ from a rooted tree $T \in \mathcal{T}_{\mathrm{c}}$ in the following way. List the leaves of T in increasing order of label as $\ell_1, \ldots, \ell_k$. Define a growing sequence $(S_i, i \in [0, k])$ of subsets of $[0, n]$ and a sequence of vertex-sequences $(W_i, i \in [k])$ as follows. Let $S_0 = \{\rho(T)\}$. For $i \in [k]$, let W_i be the sequence of vertices on the S_{i-1}-to-ℓ_i path in T, excluding the final vertex ℓ_i, and let S_i be the union of S_{i-1} and the vertices in W_i. Let V be the concatenation $(W_1, \ldots, W_k)$.

It is not hard to check that the resulting sequence V lies in $\mathcal{V}_{\mathrm{c}}$, nor is it hard to construct the inverse map $\mathcal{V}_{\mathrm{c}} \to \mathcal{T}_{\mathrm{c}}$ (both of which are done in [3]). Thus, the line-breaking construction is a well-defined bijection which encodes trees with a given child sequence as sequences with specified multiplicities. For us, the crucial property of this bijection is that if $T \in \mathcal{T}_{\mathrm{c}}$ maps to $V = (V_1, \ldots, V_n) \in \mathcal{T}_{\mathrm{c}}$, then the first repetition in V encodes the height of the leaf ℓ_1 of minimum label. Specifically, $\mathrm{ht}_T(\ell_1) = r(V) - 1$, where $r(V) = \min(i \in [n] : V_i \in \{V_1, \ldots, V_{i-1}\})$. For the rest of this section, we will work with child sequences $\mathrm{c} = (c_0, \ldots, c_n)$ such that $c_0 = 0$, which implies that $\ell_1 = 0$. Thus, if $\mathbf{T} \in_u \mathcal{T}_{\mathrm{c}}$ and $\mathbf{V} \in_u \mathcal{V}_{\mathrm{c}}$ then $\mathrm{ht}_{\mathbf{T}}(0) \overset{\mathrm{d}}{=} r(\mathbf{V}) - 1$.

It is an easy exercise to deduce the following formulae for the probability mass function of $\mathrm{ht}_{\mathbf{T}}(0)$ when $\mathbf{T}$ is either a uniform or composition-biased binary tree from the line-breaking construction. We use the falling factorial notation $(x)_k = x(x-1)\cdots(x-k+1)$, so for fixed m, as h varies, $|\mathcal{P}_{m,h}| = \binom{h+m-1}{m-1} \propto (h+m-1)_{m-1}$.

Lemma 3.4. *Let m be a positive integer, and let $\mathrm{b} = (b_0, \ldots, b_n)$ be a binary child sequence with $b_0 = 0$. Let $\mathbf{S} \in_u \mathcal{T}_{\mathrm{b}}$ and let $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) : T \in \mathcal{T}_{\mathrm{b}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$. Then for all*

$h \in [n/2]$,

$$\mathbb{P}\{\mathrm{ht}_{\mathbf{S}}(0) = h\} = \frac{h}{n-h} \prod_{i=1}^{h-1} \left(1 - \frac{i}{n-i}\right), \text{ and hence}$$

$$\mathbb{P}\{\mathrm{ht}_{\mathbf{T}}(0) = h\} \propto (h+m-1)_{m-1} \frac{h}{n-h} \prod_{i=1}^{h-1} \left(1 - \frac{i}{n-i}\right).$$

We will therefore take $\mathbf{Y}_1 = r(\mathbf{V})$ and $\mathbf{Y}_2 = r(\mathbf{W})$ where $\mathbf{V} \in_u \mathcal{V}_{\mathrm{c}}$ and $\mathbf{W} \in_u \mathcal{V}_{\mathrm{b}}$ for some 1-free child sequence c and some binary child sequence b. This way, for $i \in \{1, 2\}$, $\mathbf{Y}_i$ encodes the height of 0 in a random tree. Soon we will see how to choose $\mathbf{X}_i$ so that conditioning on $\mathbf{Y}_i \geq \mathbf{X}_i$ precisely accounts for the fact that the trees in Theorem 1.5 are composition-biased. Before describing $\mathbf{X}_i$ though, we verify that the hypotheses of Proposition 3.2 hold with these choices of $\mathbf{Y}_i$. We start with an easy enumeration formula.

Lemma 3.5. *Let* $\mathrm{c} = (c_0, \ldots, c_n)$ *be a child sequence on* $[0, n]$ *and write* $I = \{v \in [0, n] \mid c_v \neq 0\}$. *Then for all* $h \geq 0$,

$$|\{V \in \mathcal{V}_{\mathrm{c}} : r(V) > h\}| = \frac{h!(n-h)!}{\prod_{v \in I} c_v!} \sum_{A \in \binom{I}{h}} \prod_{v \in A} c_v.$$

Proof. To construct any such sequence, first choose a set A of size h from I to be the first h entries of V. Then let $(V_1, \ldots, V_h)$ be one of the $h!$ permutations of A. Finally, choose a permutation of the remaining multiset, which contains $c_v - 1$ copies of each $v \in A$ and c_v copies of each $v \in I \setminus A$ to form $(V_{h+1}, \ldots, V_n)$. There are

$$\frac{(n-h)!}{\prod_{v \in A}(c_v - 1)! \prod_{v \in I \setminus A} c_v!} = \frac{(n-h)!}{\prod_{v \in I} c_v!} \prod_{v \in A} c_v$$

ways to do so. ☐

Lemma 3.6. *Fix a positive integer* n *and let* ε *be the parity of* n. *Let* $\mathrm{c} = (c_0, \ldots, c_n)$ *be a* 1*-free child sequence and let* $\mathrm{b} = (b_0, \ldots, b_{n+\varepsilon})$ *be a binary child sequence. Take* $\mathbf{V} \in_u \mathcal{V}_{\mathrm{c}}$ *and* $\mathbf{W} \in_u \mathcal{V}_{\mathrm{b}}$. *Then for all* y,

$$(r(\mathbf{V}) \mid r(\mathbf{V}) \geq y) \preceq_{\mathrm{st}} (r(\mathbf{W}) \mid r(\mathbf{W}) \geq y).$$

Proof. We need to prove that for any positive integers y and h,

$$\frac{|\{V \in \mathcal{V}_{\mathrm{c}} : r(V) \geq \max(h, y)\}|}{|\{V \in \mathcal{V}_{\mathrm{c}} : r(V) \geq y\}|} \leq \frac{|\{W \in \mathcal{V}_{\mathrm{b}} : r(W) \geq \max(h, y)\}|}{|\{W \in \mathcal{V}_{\mathrm{b}} : r(W) \geq y\}|}.$$

When $h \leq y$, both of the expressions above are equal to 1 and the claim holds. By induction then, this reduces to showing for all $h \geq 1$ that

$$\frac{|\{V \in \mathcal{V}_{\mathrm{c}} : r(V) \geq h+1\}|}{|\{W \in \mathcal{V}_{\mathrm{b}} : r(W) \geq h+1\}|} \leq \frac{|\{V \in \mathcal{V}_{\mathrm{c}} : r(V) \geq h\}|}{|\{W \in \mathcal{V}_{\mathrm{b}} : r(W) \geq h\}|}. \tag{3.1}$$

Since b is binary, either $b_v = 0$ and $b_v! = 1$ or $b_v! = b_v = 2$. Also, $|\{i \in [0, n+\varepsilon] \mid b_v \neq 0\}| = (n+\varepsilon)/2$. Applying Lemma 3.5 to b then, we obtain

$$|\{W \in \mathcal{V}_{\mathrm{b}} : r(W) \geq h+1\}| = \frac{h!(n+\varepsilon-h)!}{2^{\frac{n+\varepsilon}{2}-h}} \binom{\frac{n+\varepsilon}{2}}{h}.$$

This allows us to replace the denominators in the desired inequality (3.1). Applying the general form of Lemma 3.5 to the child sequence c and writing $I = \{i \in [0, n] \mid c_v \neq 0\}$, after some cancellation we find that (3.1) is equivalent to

$$h(n+\varepsilon-h+1) \sum_{A \in \binom{I}{h}} \prod_{v \in A} c_v \leq (n+\varepsilon-2h+2)(n-h+1) \sum_{B \in \binom{I}{h-1}} \prod_{u \in B} c_u. \tag{3.2}$$

First we show that

$$h \sum_{A\in\binom{I}{h}} \prod_{v\in A} c_v \le (n-2h+2) \sum_{B\in\binom{I}{h-1}} \prod_{u\in B} c_u, \tag{3.3}$$

or in other words that (3.2) holds when $\varepsilon = 0$. For each $A \in \binom{I}{h}$, there are exactly h different ways to write $A = B \cup \{v\}$ where $B \in \binom{I}{h-1}$ and $v \in I \setminus B$, so

$$\sum_{B\in\binom{I}{h-1}} \sum_{v\in I\setminus B} \prod_{u\in B\cup\{v\}} c_u = h \sum_{A\in\binom{I}{h}} \prod_{v\in A} c_v.$$

As c is 1-free, $c_v \ge 2$ for all $v \in I$. But also $\sum_{v\in I} c_v = n$. Thus, for each $B \in \binom{I}{h-1}$,

$$\sum_{v\in I\setminus B} c_v \le n - 2|B| = n - 2h + 2.$$

Therefore

$$\begin{aligned} h \sum_{A\in\binom{I}{h}} \prod_{v\in A} c_v &= \sum_{B\in\binom{I}{h-1}} \sum_{v\in I\setminus B} \prod_{u\in B\cup\{i\}} c_u \\ &= \sum_{B\in\binom{I}{h-1}} \sum_{v\in I\setminus B} c_v \prod_{u\in B} c_u \\ &= \sum_{B\in\binom{I}{h}} \left(\sum_{v\in I\setminus B} c_v \right) \prod_{u\in B} c_u \\ &\le (n-2h+2) \sum_{B\in\binom{I}{h}} \prod_{u\in B} c_u, \end{aligned}$$

so (3.3) holds.

To verify (3.2) when $\varepsilon = 1$, note that $n - 2h + 2 \le n - h + 1$ since $h \ge 1$, and hence $(n-2h+2)(n+\varepsilon-h+1) \le (n+\varepsilon-2h+2)(n-h+1)$. Combining this with (3.3),

$$h \sum_{A\in\binom{I}{h}} \prod_{v\in A} c_v \le (n-2h+2) \sum_{B\in\binom{I}{h}} \prod_{u\in B} c_u \le \frac{(n+\varepsilon-2h+2)(n-h+1)}{n+\varepsilon-h+1} \sum_{B\in\binom{I}{h}} \prod_{u\in B} c_u,$$

which is equivalent to the required inequality (3.2). □

For $i \in \{1, 2\}$, we will let $\mathbf{X}_i$ be (a translate of) the maximum of a random set of indices. The following proposition certifies that this choice of $\mathbf{X}_i$ satisfies the assumptions of Proposition 3.2.

Lemma 3.7. *Fix $1 \le j \le k \le \ell$ and take $\mathbf{A} \in_u \binom{[k]}{j}$ and $\mathbf{B} \in_u \binom{[\ell]}{j}$. Then for all x,*

$$(\max \mathbf{A} \mid \max \mathbf{A} \le x) \preceq_{\mathrm{st}} (\max \mathbf{B} \mid \max \mathbf{B} \le x).$$

Proof. We may assume that x is a positive integer. Let $k' = \min(k, x)$ and $\ell' = \min(\ell, x)$, and take $\mathbf{A}' \in_u \binom{[k']}{j}$ and $\mathbf{B}' \in_u \binom{[\ell']}{j}$. Then $\mathbf{A}' \overset{d}{=} (\mathbf{A} \mid \max \mathbf{A} \le x)$ and $\mathbf{B}' \overset{d}{=} (\mathbf{B} \mid \max \mathbf{B} \le x)$, so we just need $\max \mathbf{A}' \preceq_{\mathrm{st}} \max \mathbf{B}'$. This holds because for $x' \le k'$,

$$\mathbb{P}\left\{\max \mathbf{A}' \le x'\right\} = \frac{\binom{x'}{j}}{\binom{k'}{j}} \ge \frac{\binom{x'}{j}}{\binom{\ell'}{j}} = \mathbb{P}\left\{\max \mathbf{B}' \le x'\right\},$$

and for $x' > k'$,

$$\mathbb{P}\left\{\max \mathbf{A}' \le x'\right\} = 1 \ge \mathbb{P}\left\{\max \mathbf{B}' \le x'\right\}.$$

□

For a child sequence $\mathrm{c} = (c_0, \ldots, c_n)$ with $c_0 = 0$, let $\mathcal{T}^*_{\mathrm{c}}$ and $\mathcal{V}^*_{\mathrm{c}}$ be the set of trees $T \in \mathcal{T}_{\mathrm{c}}$ and sequences $V \in \mathcal{V}^*_{\mathrm{c}}$ respectively such that each $i \in [0, n]$ with $c_v = 1$ lies on the root-to-0 path in T and appears before the first repetition in V. Note that $\mathcal{T}^*_{\mathrm{c}}$ is the image of $\mathcal{V}^*_{\mathrm{c}}$ under the line-breaking construction. The lemma below follows from a natural correspondence between pairs (T, P) where $T \in \mathcal{T}_{\mathrm{c}}$ and $P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}$ and trees $T \in \mathcal{T}^*_{\mathrm{c}^+}$ for a certain simple modification c^+ of c.

Lemma 3.8. *Let* $\mathrm{c} = (c_0, \ldots, c_n)$ *be a* 1*-free child sequence with* $c_0 = 0$ *and write* c^+ *for the child sequence on* $[0, n+m-1]$ *of the form* $\mathrm{c}^+ = (c_0, \ldots, c_n, 1, \ldots, 1)$*. Take* $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) \mid T \in \mathcal{T}_{\mathrm{c}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$ *and* $\mathbf{T}^* \in_u \mathcal{T}^*_{\mathrm{c}}$*. Then*

$$\mathrm{ht}_{\mathbf{T}}(0) \overset{d}{=} \mathrm{ht}_{\mathbf{T}^*}(0) - (m-1).$$

Proof. We will exhibit an $(m-1)!$-to-1 function which maps trees $T^* \in \mathcal{T}^*_{\mathrm{c}^+}$ to pairs (T, P) with $T \in \mathcal{T}_{\mathrm{c}}$ and $P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}$ in such a way that $\mathrm{ht}_T(0) = \mathrm{ht}_{T^*}(0) - (m-1)$. Let T be the homeomorphic reduction of T^*, and define $P = (P_1, \ldots, P_m)$ as follows. List the one-child vertices in the order they appear on the root-to-0 path in T^* as $(v_1, \ldots, v_{m-1})$; also define $v_m = 0$. Set $P_1 = \mathrm{ht}_{T^*}(v_1)$, and for $i \in [2, m]$, set $P_i = \mathrm{ht}_{T^*}(v_i) - \mathrm{ht}_{T^*}(v_{i-1}) - 1$. It is clear that $T \in \mathcal{T}_{\mathrm{c}}$. Furthermore, $\mathrm{ht}_{T^*}(v_i) \geq \mathrm{ht}_{T^*}(v_{i-1}) + 1$ and so $P_i \geq 0$ for all $i \in [m]$. Finally, the root-to-0 path in T is obtained from a path of length $\mathrm{ht}_{T^*}(0)$ containing $m-1$ one-child vertices by homeomorphic reduction, and thus it has length $\mathrm{ht}_T(0) = \mathrm{ht}_{T^*}(0) - (m-1)$. Hence

$$\begin{aligned}\sum_{i=1}^{m} P_i &= \mathrm{ht}_{T^*}(v_1) + \sum_{i=2}^{m} (\mathrm{ht}_{T^*}(v_i) - \mathrm{ht}_{T^*}(v_{i-1}) - 1) \\ &= \mathrm{ht}_{T^*}(v_m) - (m-1) \\ &= \mathrm{ht}_{T^*}(0) - (m-1) \\ &= \mathrm{ht}_T(0),\end{aligned}$$

and so $P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}$ and our map is well-defined. The $(m-1)!$ preimages of (T, P) are the trees obtained from T^* by permuting the labels of the $m-1$ one-child vertices. $\square$

Given a sequence $V \in \mathcal{V}_{\mathrm{c}}$ for some child sequence $\mathrm{c} = (c_0, \ldots, c_n)$ containing a 1, let $f(V) = \max(i \in [n] : c_{V_i} = 1)$ denote the final index of an element of V which appears exactly once. Note that $r(V) \geq f(V)$ if and only if $V \in \mathcal{V}^*_{\mathrm{c}}$. We are now prepared to prove Theorem 1.5.

Proof of Theorem 1.5. Let c^+ be the child sequence on $[0, n+m-1]$ extending c of the form $\mathrm{c}^+ = (c_0, \ldots, c_n, 1, \ldots, 1)$. Take $\mathbf{A} \in_u \binom{[n+m-1]}{m-1}$ independent of $\mathbf{V}$, and generate $\mathbf{V}^+ \in_u \mathcal{V}_{\mathrm{c}^+}$ from $\mathbf{V}$ and $\mathbf{A}$ as follows. Let the entries of $\mathbf{V}^+$ at the indices in $\mathbf{A}$ be the elements of $[n+1, n+m-1]$ ordered by a uniformly random permutation independent of $\mathbf{A}$ and $\mathbf{V}$. Let the entries of $\mathbf{V}^+$ at the indices outside $\mathbf{A}$ be those of $\mathbf{V}$, with the same order as in $\mathbf{V}$. Take $\mathbf{V}^* \in_u \mathcal{V}^*_{\mathrm{c}^+}$ and $\mathbf{T}^* \in_u \mathcal{T}^*_{\mathrm{c}^+}$.

By Lemma 3.8, $\mathrm{ht}_{\mathbf{T}}(0) \overset{d}{=} \mathrm{ht}_{\mathbf{T}^*}(0) - m + 1$, and since the height of 0 corresponds to the index preceding the first repetition under the line-breaking construction, $\mathrm{ht}_{\mathbf{T}^*}(0) \overset{d}{=} r(\mathbf{V}^*) - 1$. Note that $\mathbf{V}^+ \in \mathcal{V}^*_{\mathrm{c}^+}$ if and only if $r(\mathbf{V}^+) \geq f(\mathbf{V}^+)$, so $\mathbf{V}^* \overset{d}{=} (\mathbf{V}^+ \mid r(\mathbf{V}^+) \geq f(\mathbf{V}^+))$. This establishes

$$\mathrm{ht}_{\mathbf{T}}(0) \overset{d}{=} \mathrm{ht}_{\mathbf{T}^*}(0) - (m-1) \overset{d}{=} r(\mathbf{V}^*) - m \overset{d}{=} (r(\mathbf{V}^+) \mid r(\mathbf{V}^+) \geq f(\mathbf{V}^+)) - m.$$

Now observe that $r(\mathbf{V}^+) \geq f(\mathbf{V}^+)$ if and only if $r(\mathbf{V}^+) = r(\mathbf{V}) + m - 1$, because c is 1-free and no repetition in $\mathbf{V}^+$ occurs at an index in $\mathbf{A}$. Therefore $r(\mathbf{V}^+) \geq f(\mathbf{V}^+)$ if and only if $r(\mathbf{V}) + m - 1 \geq f(\mathbf{V}^+)$, and on this event, $r(\mathbf{V}^+) = r(\mathbf{V}) + m - 1$. Hence

$$\begin{aligned}(r(\mathbf{V}^+) \mid r(\mathbf{V}^+) \geq f(\mathbf{V}^+)) &= (r(\mathbf{V}) + m - 1 \mid r(\mathbf{V}) + m - 1 \geq f(\mathbf{V}^+)) \\ &= (r(\mathbf{V}) \mid r(\mathbf{V}) \geq f(\mathbf{V}^+) - m + 1) + m - 1 \\ &= (r(\mathbf{V}) \mid r(\mathbf{V}) \geq \max \mathbf{A} - m + 1) + m - 1.\end{aligned}$$

The same reasoning applies to $\mathbf{W}$ and an independent index set $\mathbf{B} \in_u \binom{[n+\varepsilon+m-1]}{m-1}$. We conclude that

$$\begin{aligned}\mathrm{ht}_{\mathbf{T}}(0) &\stackrel{d}{=} (r(\mathbf{V}) \mid r(\mathbf{V}) \geq \max \mathbf{A} - m + 1) - 1, \text{ and} \\ \mathrm{ht}_{\mathbf{S}}(0) &\stackrel{d}{=} (r(\mathbf{W}) \mid r(\mathbf{W}) \geq \max \mathbf{B} - m + 1) - 1.\end{aligned}$$

Now $(r(\mathbf{V}) \mid r(\mathbf{V}) \geq y) \preceq_{\mathrm{st}} (r(\mathbf{W}) \mid r(\mathbf{W}) \geq y)$ for all y by Lemma 3.6 and if follows from Lemma 3.7 that $(\max \mathbf{A} - m + 1 \mid \max \mathbf{A} - m + 1 \leq x) \preceq_{\mathrm{st}} (\max \mathbf{B} - m + 1 \mid \max \mathbf{B} - m + 1 \leq x)$ for all x. Since $\mathbf{A}$ and $\mathbf{V}$ are independent and $\mathbf{B}$ and $\mathbf{W}$ are independent, Proposition 3.2 yields

$$\begin{aligned}\mathrm{ht}_{\mathbf{T}}(0) &\stackrel{d}{=} (r(\mathbf{V}) \mid r(\mathbf{V}) \geq \max \mathbf{A} - m + 1) - 1 \\ &\preceq_{\mathrm{st}} (r(\mathbf{W}) \mid r(\mathbf{W}) \geq \max \mathbf{B} - m + 1) - 1 \\ &\stackrel{d}{=} \mathrm{ht}_{\mathbf{S}}(0),\end{aligned}$$

as required. □

4. A tail bound for biased binary trees

In this section we prove Theorem 1.6, which establishes a sub-Gaussian tail bound on the centered and rescaled height of a fixed leaf in a composition-biased random tree. Using the stochastic domination established in Section 3, we can focus on the case of a binary child sequence. Only at the end of our proof will we incorporate the parity adjustment.

Before proceeding with the proof, we note that the following considerations suggest that the form of the bound in Theorem 1.6 cannot be improved. For positive integers m and n with n even, let $\mathbf{h}_{m,n} \stackrel{\mathrm{d}}{=} \mathrm{ht}_{\mathbf{T}}(0)$, where $(\mathbf{T}, \mathbf{P}) \in_u \{(T, P) : T \in \mathcal{T}_{\mathrm{b}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$ for some binary child sequence $\mathrm{b} = (b_0, \ldots, b_n)$ with $b_0 = 0$. Up to a normalizing constant, the probability mass function of $\mathbf{h}_{m,n}$ is given by the function $p(h)$ defined below. Analyzing the asymptotics of $p(h)$ is not hard, and allows one show that $n^{-1/2}\mathbf{h}_{m,n} \to \mathbf{h}_m$ in distribution as $n \to \infty$ for a continuous random variable $\mathbf{h}_m$ with density proportional to $x^m e^{-x^2/2}\mathbf{1}_{x \geq 0}$. Furthermore, $\mathbf{h}_m - \sqrt{m}$ converges in distribution as $m \to \infty$ to a normal random variable with mean 0 and variance 1/2, suggesting that $(n/2)^{-1/2}(\mathbf{h}_{m,n} - \sqrt{mn})$ is approximately Gaussian when $1 \ll m \ll n$.

For the remainder of this section, fix positive integers m and n with n even. We will prove Theorem 1.6 by first analyzing $\mathbf{h}_{m,n}$ and then applying Theorem 1.5. To this end, define a function $p : [n/2] \to \mathbb{R}$ by

$$p(h) = (h + m - 1)_{(m-1)} \frac{h}{n - h} \prod_{i=1}^{h-1} \left(1 - \frac{i}{n - i}\right). \tag{4.1}$$

Then by Lemma 3.4, $\mathbb{P}\{\mathbf{h}_{m,n} = h\} \propto p(h)$, so for $h \in [n/2]$,

$$\mathbb{P}\{\mathbf{h}_{m,n} \geq h\} = \frac{\sum_{k \in [h, n/2]} p(k)}{\sum_{k \in [n/2]} p(k)}. \tag{4.2}$$

We will upper-bound the numerator and lower-bound the denominator in (4.2) by controlling the ratios $p(h+1)/p(h)$. In our application, we will have $m \ll h \ll n$, so the following lemma should be read as asserting that $p(h+1)/p(h) \approx \exp(m/h - h/n)$.

Lemma 4.1. *For all $h \in [1, n/2)$,*

$$\exp\left(\frac{m}{h+m} - \frac{h}{n-2h}\right) \le \frac{p(h+1)}{p(h)} \le \exp\left(\frac{m}{h} - \frac{h-2}{n}\right).$$

Proof. For the upper bound,

$$\begin{aligned}
\frac{p(h+1)}{p(h)} &= \left(1+\frac{m-1}{h+1}\right)\left(1+\frac{1}{h}\right)\left(1+\frac{1}{n-h-1}\right)\left(1-\frac{h}{n-h}\right)\\
&\le \exp\left(\frac{m-1}{h+1}+\frac{1}{h}+\frac{1}{n-h-1}-\frac{h}{n-h}\right)\\
&\le \exp\left(\frac{m-1}{h}+\frac{1}{h}+\frac{2}{n-h}-\frac{h}{n-h}\right)\\
&= \exp\left(\frac{m}{h}-\frac{h-2}{n-h}\right)\\
&\le \exp\left(\frac{m}{h}-\frac{h-2}{n}\right).
\end{aligned}$$

To see the second inequality, note that $n-h > 2h-h \ge 1$ since $h \in [1, n/2)$. Therefore $n-h \ge 2$, which implies $1/(n-h-1) \le 2/(n-h)$.

For the lower bound, we have

$$\begin{aligned}
\frac{p(h)}{p(h+1)} &= \left(1-\frac{m-1}{h+m}\right)\left(1-\frac{1}{h+1}\right)\left(\frac{n-h-1}{n-h}\right)\left(1+\frac{h}{n-2h}\right)\\
&\le \exp\left(-\frac{m-1}{h+m}-\frac{1}{h+1}+\frac{h}{n-2h}\right)\\
&\le \exp\left(-\frac{m-1}{h+m}-\frac{1}{h+m}+\frac{h}{n-2h}\right)\\
&= \exp\left(-\frac{m}{h+m}+\frac{h}{n-2h}\right).
\end{aligned}$$

□

Lemma 4.2. *For $h \ge 5$ and $t \ge 1$ such that $t+h \le n/2$,*

$$p(h+t) \le p(h)\exp\left(\frac{tm}{h} - \frac{th}{2n} - \frac{t^2}{2n}\right).$$

Proof. By the upper-bound in Lemma 4.1,

$$\begin{aligned}
p(h+t) &= p(h)\prod_{i=h}^{h+t-1}\frac{p(i+1)}{p(i)}\\
&\le p(h)\exp\left(\sum_{i=h}^{h+t-1}\left(\frac{m}{i}-\frac{i-2}{n}\right)\right)\\
&\le p(h)\exp\left(\frac{tm}{h}-\frac{1}{n}\sum_{i=h}^{h+t-1}(i-2)\right)\\
&= p(h)\exp\left(\frac{tm}{h}-\frac{1}{n}\left(t\left(h-\frac{5}{2}\right)+\frac{t^2}{2}\right)\right)\\
&\le p(h)\exp\left(\frac{tm}{h}-\frac{th}{2n}-\frac{t^2}{2n}\right),
\end{aligned}$$

where the final inequality holds because $h \ge 5$. □

Lemma 4.3. *For* $1 \le t \le h \le n/8$,

$$p(h+t) \ge p(h)\exp\left(-\frac{4th}{n}\right).$$

Proof. Note that $t+h-1 < n/2$ since $t \le h \le n/8$. Hence, by the lower-bound in Lemma 4.1,

$$\begin{aligned}
p(h+t) &= p(h)\prod_{i=h}^{h+t-1}\frac{p(i+1)}{p(i)} \\
&\ge p(h)\exp\left(\sum_{i=h}^{h+t-1}\left(\frac{m}{i+m}-\frac{i}{n-2i}\right)\right) \\
&\ge p(h)\exp\left(-\frac{t(h+t-1)}{n-2(h+t-1)}\right) \\
&\ge p(h)\exp\left(-\frac{2th}{n-4h}\right) \\
&= p(h)\exp\left(-\frac{4th}{n}\right),
\end{aligned}$$

where the last inequality holds because $h \le n/8$. □

Proof of Theorem 1.6. For now suppose that c is binary, so $\mathbb{P}\{\mathbf{h}=h\} \propto p(h)$. Note that $1.9\sqrt{mn} \ge \lceil 1.8\sqrt{mn}\rceil \ge 5$ since $m \ge 1$ and $n \ge 64m$, so we can apply Lemma 4.2 with $h = \lceil 1.8\sqrt{mn}\rceil$ to obtain

$$\begin{aligned}
&\sum_{1.9\sqrt{mn}+x\sqrt{n}\le k\le n/2} p(k) \\
&\le \sum_{x\sqrt{n}\le t\le n/2-1.9\sqrt{mn}} p(\lceil 1.8\sqrt{mn}\rceil + t) \\
&\le \sum_{x\sqrt{n}\le t\le n/2-1.9\sqrt{mn}} p(\lceil 1.8\sqrt{mn}\rceil)\exp\left(\frac{tm}{\lceil 1.8\sqrt{mn}\rceil}-\frac{t\lceil 1.8\sqrt{mn}\rceil}{2n}-\frac{t^2}{2n}\right) \\
&\le p(\lceil 1.8\sqrt{mn}\rceil)\sum_{t\ge x\sqrt{n}}\exp\left(\frac{tm}{1.8\sqrt{mn}}-\frac{1.8t\sqrt{mn}}{2n}-\frac{t^2}{2n}\right)
\end{aligned}$$

Using that for $t \ge x\sqrt{n}$,

$$\exp\left(\frac{tm}{1.8\sqrt{mn}}-\frac{1.8t\sqrt{mn}}{2n}\right) \le \exp\left(-\frac{t\sqrt{m}}{3\sqrt{n}}\right) \le \exp\left(-\frac{x\sqrt{m}}{3}\right),$$

it follows that

$$\sum_{1.9\sqrt{mn}+x\sqrt{n}\le k\le n/2} p(k) \le \exp\left(-\frac{x\sqrt{m}}{3}\right)p(\lceil 1.8\sqrt{mn}\rceil)\sum_{t\ge x\sqrt{n}}\exp\left(-\frac{t^2}{2n}\right) \tag{4.3}$$

Next, since $n \ge 64m$, we have $1 \le \lfloor\sqrt{n/m}\rfloor \le \lfloor\sqrt{mn}\rfloor \le n/8$ and hence $1 \le \lfloor\sqrt{n/m}\rfloor + \lfloor\sqrt{mn}\rfloor \le n/2$, so we can apply Lemma 4.3 with $h = \lfloor\sqrt{mn}\rfloor$ and $1 \le t \le \lfloor\sqrt{n/m}\rfloor$ to obtain

$$\sum_{k\in[n/2]} p(k) \ge \sum_{1\le t\le\sqrt{n/m}} p(\lfloor\sqrt{mn}\rfloor + t) \ge \sum_{1\le t\le\sqrt{n/m}} p(\lfloor\sqrt{mn}\rfloor)\exp\left(-\frac{4t\lfloor\sqrt{mn}\rfloor}{n}\right)$$

which since

$$\sum_{1\le t\le\sqrt{n/m}} \exp\left(-\frac{4t\lfloor\sqrt{mn}\rfloor}{n}\right) \ge \lfloor\sqrt{n/m}\rfloor \exp\left(-\frac{4\lfloor\sqrt{n/m}\rfloor\lfloor\sqrt{mn}\rfloor}{n}\right) \ge e^{-4-3/4}\sqrt{n/m}$$

yields that $\sum_{k\in[n/2]} p(k) \ge e^{-3/4}\sqrt{n/m}p(\lfloor\sqrt{mn}\rfloor)e^{-4}$. Also, applying Lemma 4.2 with $h = \lfloor\sqrt{mn}\rfloor$ and $t = \lceil 1.8\sqrt{mn}\rceil - h$, we find that $p(\lceil 1.8\sqrt{mn}\rceil) \le p(\lfloor\sqrt{mn}\rfloor)e^{1/4}$. Combining these two bounds with (4.3), we thus have

$$\begin{aligned}
\mathbb{P}\left\{\mathbf{h} \ge 1.9\sqrt{mn} + x\sqrt{n}\right\} &= \frac{\sum_{1.9\sqrt{mn}+x\sqrt{n/m}\le h\le n/2} p(h)}{\sum_{h\in[n/2]} p(h)} \\
&\le \frac{\exp\left(-\frac{x\sqrt{m}}{3}\right) p(\lceil 1.8\sqrt{mn}\rceil) \sum_{t\ge x\sqrt{n}} \exp\left(-\frac{t^2}{2n}\right)}{e^{-3/4}\sqrt{n/m}p(\lfloor\sqrt{mn}\rfloor)e^{-4}} \\
&\le \frac{\exp\left(-\frac{x\sqrt{m}}{3}\right) e^{1/4} \sum_{t\ge x\sqrt{n}} \exp\left(-\frac{t^2}{2n}\right)}{e^{-3/4}\sqrt{n/m}e^{-4}}
\end{aligned}$$

Now,

$$\sum_{t\ge x\sqrt{n}} e^{-\frac{t^2}{2n}}\frac{1}{\sqrt{n}} \le \int_{x-\frac{1}{\sqrt{n}}}^{\infty} e^{-t^2/2}\,\mathrm{d}t \le \int_{x-\frac{1}{8}}^{\infty} e^{-t^2/2}\,\mathrm{d}t\,,$$

the final inequality holding since $n \ge 64m \ge 64$, which together with the preceding bound yields that

$$\mathbb{P}\left\{\mathbf{h} \ge 1.9\sqrt{mn} + x\sqrt{n}\right\} \le \exp\left(5 - \frac{x\sqrt{m}}{3} + \frac{\log m}{2}\right)\int_{x-\frac{1}{8}}^{\infty} e^{-t^2/2}\,\mathrm{d}t.$$

It is easily checked that $-\frac{2.5\sqrt{m}}{3} + \frac{\log m}{2} \le -0.8$ for all $m \ge 1$. So, fixing $x \ge 2.5$, we have

$$\exp\left(5 - \frac{x\sqrt{m}}{3} + \frac{\log m}{2}\right) \le \exp(4.2).$$

Furthermore $x - 1/8 \ge e^{1/4}$, and so

$$\int_{x-\frac{1}{8}}^{\infty} e^{-t^2/2}\,\mathrm{d}t \le \frac{\exp\left(-\frac{1}{2}\left(x-\frac{1}{8}\right)\right)^2}{x-1/8} \le \exp\left(-\frac{x^2}{2} + \frac{x}{8} - \frac{1}{4}\right).$$

Thus,

$$\mathbb{P}\left\{\mathbf{h} \ge 1.9\sqrt{mn} + x\sqrt{n}\right\} \le \exp\left(-\frac{x^2}{2} + \frac{x}{8} + 3.95\right). \tag{4.4}$$

When $x < 2.5$, this bound is greater than 1. Therefore (4.4) holds for all $x \ge 0$. Now (4.4) is slightly stronger than the bound claimed in the theorem, so this completes the proof in the case where c is binary.

If c is not binary, then consider a binary child sequence $\mathrm{b} = (b_0, \ldots, b_{n+\varepsilon})$ with $b_0 = 0$, where ε is the parity of n. Take $(\mathbf{S}, \mathbf{Q}) \in_u \{(S, Q) : S \in \mathcal{T}_{\mathrm{b}}, Q \in \mathcal{P}_{m,\mathrm{ht}_S(0)}\}$, so $\mathbf{h} \preceq_{\mathrm{st}} \mathrm{ht}_{\mathbf{S}}(0)$ by Theorem 1.5. Note that $2\sqrt{mn} \ge 1.9\sqrt{m(n+1)}$ and $x\sqrt{n} \ge x\sqrt{n+1}/1.1$ since $n \ge$

$64m \geq 64$. Thus by (4.4),

$$\begin{aligned}\mathbb{P}\left\{\mathbf{h} \geq 2\sqrt{mn} + x\sqrt{n}\right\} &\leq \mathbb{P}\left\{\mathbf{h} \geq 1.9\sqrt{m(n+\varepsilon)} + \frac{x}{1.1}\sqrt{n+\varepsilon}\right\} \\ &\leq \mathbb{P}\left\{\mathrm{ht}_{\mathbf{S}}(0) \geq 1.9\sqrt{m(n+\varepsilon)} + \frac{x}{1.1}\sqrt{n+\varepsilon}\right\} \\ &\leq \exp\left(-\frac{(x/1.1)^2}{2} + \frac{x/1.1}{8} + 3.95\right) \\ &\leq \exp\left(-\frac{x^2}{3} + 4\right),\end{aligned}$$

as required. □

5. The diameter of the kernel of a core with no cycle components

The goal of this section is to prove Theorem 1.7, so throughout the section, we fix positive integers $n \leq \ell$ and a 1-free degree sequence d of the form $\mathrm{d} = (d_1, \ldots, d_\ell) = (d_1, \ldots, d_n, 2, \ldots, 2)$ with $d_v \geq 3$ for all $v \in [n]$. We take $\mathbf{C} \in_u \mathcal{G}_{\mathrm{d}}^-$ and let $\mathbf{K} = K(\mathbf{C})$, and aim to prove that, writing $2m = \sum_{v \in [n]} d_v$, then

$$\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{K}) \geq 500 \log m\right\} = O\left(\frac{\log^8 m}{m}\right).$$

As mentioned in the proof overview, we will prove the theorem by studying a breadth-first exploration of the kernel. For the description of this kernel exploration, it is useful to augment the core with the additional data of half-edge labels at each kernel vertex. The set of kernel *half-edges* is $\bigcup_{v \in [n]}\{v1, \ldots, vd_v\}$, where for $v \in [n]$ and $i \in [d_v]$ the notation vi is shorthand for the ordered pair (v, i). We say that a half-edge of the form vi is a half-edge *at* v, or that vi is *incident* to v. An *augmented core* A with degree sequence d is a set of pairs $(\{ui, vj\}, (p_1, \ldots, p_a))$ where ui and vj are distinct half-edges and $(p_1, \ldots, p_a)$ is a sequence of distinct elements of $[\ell] \setminus [n]$ such that the following hold. First, the *matching* of A, which we denote $M(A)$ and consists of the pairs $\{ui, vj\}$ such that $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ for some sequence $(p_1, \ldots, p_a)$, forms a partition of the set of half-edges. This implies $|A| = |M(A)| = m$. We allow $a = 0$ in which case $(p_1, \ldots, p_a) = \emptyset$. Second, $[\ell] \setminus [n]$ is partitioned by the sequences $(p_1, \ldots, p_a)$ such that $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ for some $\{ui, vj\} \in M(A)$. Given such a set of pairs A, one can construct a multigraph $C(A)$ with degree sequence d by introducing, for each $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ where $ui \prec_{\mathrm{lex}} vj$, a path from u to v whose sequence of internal vertices is $(p_1, \ldots, p_a)$. For example, for the graph G in Figure 1, one augmented core A with $C(A) = C(G)$ is given by

$$\begin{aligned}A = \{&(\{11, 21\}, (9, 5)), (\{12, 31\}, (8, 10, 7)), (\{13, 41\}, \varnothing), \\ &(\{22, 32\}, (6, 11)), (\{23, 42\}, \varnothing), (\{33, 43\}, \varnothing)\}.\end{aligned}$$

Note that $C(A)$ will have no cycle components because $d_v \geq 3$ for all $v \in [n]$. We say that an edge e in the kernel of $C(A)$ is *subdivided* if the corresponding element $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ has $a \geq 1$. If $a = 0$, we say that e is *non-subdivided.* For A to be an augmented core, we finally require that $C(A)$ be a simple graph. In other words, we insist that $a \geq 2$ for all $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ such that $u = v$, and that $a \geq 1$ for all but at most one of the elements $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$ whose pair of half-edges are at a fixed pair of distinct vertices u and v.

We write $\mathcal{A}_{\mathrm{d}}$ for the set of all augmented cores with degree sequence d. Let $A \in \mathcal{A}_{\mathrm{d}}$ and write $C = C(A)$ and $M = M(A)$ and observe that the matching M is sufficient for describing the kernel of C. Indeed if $K(M)$ is constructed from M by introducing an edge between u and v for each $\{ui, vj\} \in M$, then $K(M) = K(C)$. We use the shorthand

$K(A) = K(M(A)) = K(C(A))$. The following observation allows us to freely augment a random core with this additional half-edge information without changing its distribution.

Observation 5.1. *If $\mathbf{A} \in_u \mathcal{A}_{\mathrm{d}}$, then $C(\mathbf{A}) \in_u \mathcal{G}_{\mathrm{d}}^-$.*

Proof. For $C \in \mathcal{G}_{\mathrm{d}}^-$, let $\mathcal{A}_{\mathrm{d}}(C) = \{A \in \mathcal{A}_{\mathrm{d}} : C(A) = C\}$, and let $\Sigma_{\mathrm{d}}(C)$ be the set of all lists $(\sigma_u, u \in [n])$ of bijections $\sigma_u : [d_u] \to N(u)$. Here $N(u)$ denotes the set of neighbours of u in C. We will provide a bijection $\mathcal{A}_{\mathrm{d}}(C) \to \Sigma_{\mathrm{d}}(C)$, whose existence will imply that

$$\mathbb{P}\{C(\mathbf{A}) = C\} \propto |\mathcal{A}_{\mathrm{d}}(C)| = |\Sigma_{\mathrm{d}}(C)| = \prod_{u \in [n]} d_u!.$$

It will then follow that $\mathbb{P}\{C(\mathbf{A}) = C\}$ does not depend on C, and hence that $C(\mathbf{A}) \in_u \mathcal{G}_{\mathrm{d}}^-$.

Given $A \in \mathcal{A}_{\mathrm{d}}(C)$, we need to construct a list $(\sigma_u, u \in [n]) \in \Sigma_{\mathrm{d}}(C)$. For $u \in [n]$ and for $i \in [d_u]$, define $\sigma_u(i)$ as follows. Let vj be the unique half-edge such that $\{ui, vj\} \in M(A)$, and let $(p_1, \ldots, p_a)$ be the unique sequence such that $(\{ui, vj\}, (p_1, \ldots, p_a)) \in A$. Then set

$$\sigma_u(i) = \begin{cases} v & a = 0 \\ p_1 & a \geq 1, ui \prec_{\mathrm{lex}} vj \\ p_a & a \geq 1, vj \prec_{\mathrm{lex}} ui \end{cases}$$

Because $C(A) = C$ is a simple graph, each $\sigma_u : [d_u] \to N_C(u)$ is a well-defined bijection.

To argue that this map $\mathcal{A}_{\mathrm{d}}(C) \to \Sigma_{\mathrm{d}}(C)$ is bijective, we now describe its inverse $\Sigma_{\mathrm{d}}(C) \to \mathcal{A}_{\mathrm{d}}(C)$. Given $(\sigma_u, u \in [n]) \in \Sigma_{\mathrm{d}}(C)$, we need to construct an augmented core $A \in \mathcal{A}_{\mathrm{d}}(C)$. Let $K = K(C)$ and for each edge $e \in E(K)$, introduce $(\{ui, vj\}, (p_1, \ldots, p_a))$ to A, where $u = e^-$, $v = e^+$, and i, j, and $(p_1, \ldots, p_a)$ are defined as follows. If $u \neq v$, let i and j be such that $\sigma_u(i)$ and $\sigma_u(j)$ are the neighbours of u and v respectively on $C(e)$, and let $(p_1, \ldots, p_a)$ be the sequence of internal vertices on $C(e)$, oriented from u to v. If $u = v$, let $i < j$ be such that $\sigma_u(i)$ and $\sigma_v(j)$ are the two neighbours of $u = v$ on $C(e)$, and let $(p_1, \ldots, p_a)$ be the sequence of internal vertices on $C(e)$, oriented so that $\sigma_u(i) = p_1$ and $\sigma_v(j) = p_a$. One can check that this construction inverts the preceding map $\mathcal{A}_{\mathrm{d}}(C) \to \Sigma_{\mathrm{d}}(C)$. ☐

We are now prepared to describe the *breadth-first kernel exploration* of a given augmented core $A \in \mathcal{A}_{\mathrm{d}}$ started at a given vertex $v \in [n]$. Writing $M = M(A)$ and $K = K(A)$, this exploration consists of a sequence of pairs $((K_t, Q_t), t \geq 0)$ where for each $t \geq 0$, $K_t = K_t(A, v)$ is a subgraph of K called the *explored subgraph* at time t, and $Q_t = Q_t(A, v)$ is a set of half-edges called the *queue* at time t. Edges in K_t are called *explored* by time t, and if an explored edge corresponds to a pair of half-edges $\{ui, wj\} \in M$, then we also call ui and wj explored by time t. To start, K_0 is the edgeless graph whose only vertex is v, and $Q_0 = \{v1, \ldots, vd_v\}$. For $t \geq 0$, (K_{t+1}, Q_{t+1}) is constructed from (K_t, Q_t) as follows. If $Q_t = \emptyset$, then set $K_{t+1} = K_t$ and $Q_{t+1} = Q_t$. Otherwise, let ui be the lexicographically least half-edge in Q_t whose vertex u is at minimal distance to v. Let wj be the unique half-edge such that $\{ui, wj\} \in M$. Obtain K_{t+1} from K_t by adding the vertex w and the edge of K corresponding to $\{ui, wj\}$. Obtain Q_{t+1} from Q_t by adding all half-edges at w that are unexplored by time t, then removing ui and wj.

This process explores the component of K containing v one edge at a time in breadth-first order. In particular, there exists $s \leq m$ such that $e(K_t) = t$ for all $t \leq s$ and $e(K_t) = s$ for all $t \geq s$. Specifically, $s = \min(t : Q_t = \emptyset)$, which is also equal to the number of edges in the component of K containing v. For $t < s$, we denote the unique edge in $E(K_{t+1}) \setminus E(K_t)$ by $e_{t+1} = e_{t+1}(A, v)$. We say that e_{t+1} is a *back-edge* if both of its endpoints are in K_t, and that it is a *loop* if its endpoints are equal. If e_{t+1} is a back-edge, then $|Q_{t+1}| = |Q_t| - 2$. Crucially though, if e_{t+1} is not a back-edge, then writing w for the endpoint of e_{t+1} not in K_t, we have $|Q_{t+1}| = |Q_t| + d_w - 2 \geq |Q_t| + 1$. We also write $\mathrm{rad}(K_t) = \mathrm{rad}(K_t)$ for the maximum distance between v and any vertex in K_t.

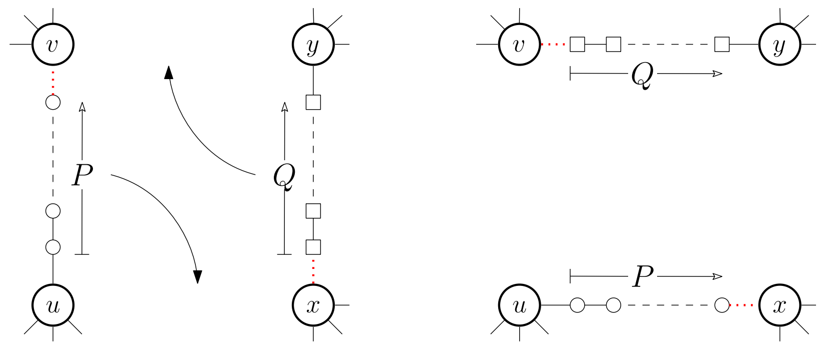

FIGURE 2. Left, a portion of $C(A)$ for an augmented core A, corresponding to pairs $(\{ui, vj\}, (p_1, \dots, p_a))$ and $(\{xk, y\ell\}, (q_1, \dots, q_b))$. If $ui \prec_{\mathrm{lex}} vj$ then P is $(p_1, \dots, p_a)$, and otherwise P is $(p_a, \dots, p_1)$, and likewise Q is either $(q_1, \dots, q_b)$ or $(q_b, \dots, q_1)$ depending on whether or not $xk \prec_{\mathrm{lex}} y\ell$. Right, the same portion of $C(A)$ after switching on $\vec{e} = (ui, vj)$ and $\vec{f} = (xk, y\ell)$. The curved arrows in the left figure indicate how the paths P and Q "pivot" as a result of the switching operation. The red dotted edge attached to v corresponds to half-edge vj, which is "detached from the head of P and reattached to the tail of Q"; likewise, the edge corresponding to half-edge xk is detached from the tail of Q and reattached to the head of P.

The principal tool we use for analyzing the exploration of a random augmented core is the notion of *switching on paths*, mentioned in the proof overview, and introduced with a slightly different formalism in [24]. Consider an augmented core $A \in \mathcal{A}_{\mathrm{d}}$ and write $C = C(A)$, $M = M(A)$, and $K = K(A)$. By an *oriented edge* in K, we mean an ordered pair of half-edges (ui, vj) such that $\{ui, vj\} \in M$. The *reversal* of an oriented edge $\vec{e} = (ui, vj)$ is the oriented edge (vj, ui), and is denoted $\overleftarrow{e}$. Given a pair of oriented edges $(\vec{e}, \vec{f})$ in K, *switching on* $(\vec{e}, \vec{f})$ *in* A is the following operation, depicted in Figure 2. Write $\vec{e} = (ui, vj)$ and $\vec{f} = (xk, y\ell)$, and let $(p_1, \dots, p_a)$ and $(q_1, \dots, q_b)$ be the sequences such that $(\{ui, vj\}, (p_1, \dots, p_a)) \in A$ and $(\{xk, y\ell\}, (q_1, \dots, q_b)) \in A$. Delete the pairs $(\{ui, vj\}, (p_1, \dots, p_a))$ and $(\{xk, y\ell\}, (q_1, \dots, q_b))$ from A, then introduce either the pair $(\{ui, xk\}, (p_1, \dots, p_a))$ or the pair $(\{ui, xk\}, (p_a, \dots, p_1))$ and either $(\{vj, y\ell\}, (q_1, \dots, q_b))$ or $(\{vj, y\ell\}, (q_b, \dots, q_1))$, to form A' according to the following rule. Writing $C' = C(A')$, we insist that the the sequences of internal vertices on the paths from u to v and from x to y in C become the sequences of internal vertices on the paths from u to x and from v to y respectively in C'. Now C' may be a multigraph, but we say that $(\vec{e}, \vec{f})$ is a *valid pair* in A, or that $(\vec{e}, \vec{f})$ is a *switching*, if $A' \in \mathcal{A}_{\mathrm{d}}$ (i.e., if C' is simple). Switching on $(\vec{e}, \vec{f})$ in A can be understood as a simpler operation taking place on the matching M. Namely, if $M' = M(A')$, then M' can be obtained from M by deleting $\{ui, vj\}$ and $\{xk, y\ell\}$, then introducing $\{ui, xk\}$ and $\{vj, y\ell\}$. However, the data in A is relevant in switching because without it, we cannot tell which pairs are valid.

An important observation is that switching is reversible. By this we mean that if $\vec{e}\,' = (ui, xk)$ and $\vec{f}\,' = (vj, y\ell)$, then $(\vec{e}\,', \vec{f}\,')$ is a valid pair in A' and switching on this pair yields the original augmented core A. We call $(\vec{e}\,', \vec{f}\,')$ the *reversal* of the switching $(\vec{e}, \vec{f})$. Switching on $(\vec{e}, \vec{f})$ and $(\overleftarrow{f}, \overleftarrow{e})$ results in the same augmented core A', and we therefore consider them to be *equivalent* switchings. Thus, there are at most four *distinct* switchings involving the underlying edges of $\vec{e}$ and $\vec{f}$; all are equivalent to one of $(\vec{e}, \vec{f})$, $(\vec{e}, \overleftarrow{f})$, $(\overleftarrow{e}, \vec{f})$, $(\overleftarrow{e}, \overleftarrow{f})$. We note that the only way to create a cycle component by switching is to switch on $(\vec{e}, \overleftarrow{e})$; any other switching preserves membership in $\mathcal{G}_{\mathrm{d}}^{-}$. We do not switch on a pair $(\vec{e}, \overleftarrow{e})$ at any point in the paper.

In the remainder of the section we will repeatedly apply the following special case of Lemma 1.8.

Lemma 5.2. *Let $\mathcal{A}$ and $\mathcal{B}$ be disjoint subsets of a set $\mathcal{C} \subseteq \mathcal{A}_{\mathrm{d}}$. Suppose that for each $A \in \mathcal{A}$, there are at least a switchings from A to an element of $\mathcal{B}$, and for each $B \in \mathcal{B}$, there are at most b switchings from B to an element of $\mathcal{A}$. Then with $\mathbf{A} \in_u \mathcal{A}_{\mathrm{d}}$, we have*

$$a\mathbb{P}\{\mathbf{A} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\} \leq b\mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\}.$$

In order to lower-bound the number of switchings from a given $A \in \mathcal{A}$ to an augmented core in $\mathcal{B}$, it will be useful for us to have sufficient conditions for a pair of oriented edges $(\vec{e}, \vec{f})$ in $K(A)$ to be valid. The most straightforward of these conditions is that if neither endpoint of $\vec{e}$ is adjacent to an endpoint of $\vec{f}$ in K, then $(\vec{e}, \vec{f})$ is a valid pair. One way to ensure that $\vec{e}$ and $\vec{f}$ do not have any adjacent endpoints is as follows. For a subgraph L of K and for $r \geq 0$, write $B_K(L, r)$ for the set of edges in K both of whose endpoints have distance at most r from some vertex in L. If $\vec{e}$ is in $B_K(L, r)$ but $\vec{f}$ is not in $B_K(L, r+2)$, then $\vec{e}$ and $\vec{f}$ cannot have any edges between their endpoints, and so $(\vec{e}, \vec{f})$ is a valid pair. More subtly, we have the following sufficient condition for a given pair of edges to have orientations that form a valid pair.

Lemma 5.3. *Fix $A \in \mathcal{A}_{\mathrm{d}}$, and let $\vec{e}$ and $\vec{f}$ be oriented edges in $K(A)$ with underlying edges e and f respectively. Suppose that neither $(\vec{e}, \vec{f})$ nor $(\vec{e}, \overleftarrow{f})$ is a valid pair in A. Then for some $(e', f') \in \{(e, f), (f, e)\}$, e' is non-subdivided and has an endpoint adjacent to both endpoints of f' via two non-subdivided edges in $K(A)$. In particular, if e is either subdivided or has an endpoint adjacent to at most one endpoint of f, and f is either subdivided or has an endpoint adjacent to at most one endpoint of e, then either $(\vec{e}, \vec{f})$ or $(\vec{e}, \overleftarrow{f})$ is a valid pair in A.*

Recall that $\mathrm{d} = (d_1, \ldots, d_n, 2, \ldots, 2)$ is a 1-free degree sequence such that $d_v \geq 3$ for all $v \in [n]$, and that $2m = \sum_{v \in [n]} d_v$. By adjusting the implicit constant in Theorem 1.7, we can and will assume that m is sufficiently large to make various inequalities hold which are scattered throughout this section. We also fix the following objects and notation for the remainder of this section. Choose a starting vertex $v \in [n]$, let $A^* \in \mathcal{A}_{\mathrm{d}}$ and take $\mathbf{A} \in_u \mathcal{A}_{\mathrm{d}}$. Let $\mathbf{K} = K(\mathbf{A})$, and let

$$((\mathbf{K}_t, \mathbf{Q}_t), t \geq 0) = ((K_t(\mathbf{A}, v), Q_t(\mathbf{A}, v)), t \geq 0) \text{ and}$$
$$((K_t, Q_t), t \geq 0) = ((K_t(A^*, v), Q_t(A^*, v)), t \geq 0)$$

be the breadth-first kernel explorations of $\mathbf{A}$ and A^* respectively started from v. For $t \geq 1$, let $\mathbf{e}_t = e_t(\mathbf{A}, v)$ if it exists.

Our proof of Theorem 1.7 is rather technical, so we take a moment to give a heuristic argument before getting into the details. In the process described above, at each time $t \geq 0$ we explore an edge $\mathbf{e}_{t+1}$. If $\mathbf{e}_{t+1}$ is a back-edge, then $|\mathbf{Q}_{t+1}| = |\mathbf{Q}_t| - 2$, and if not, then $|\mathbf{Q}_{t+1}| \geq |\mathbf{Q}_t| + 1$. Thus, assuming that the other half-edge of $\mathbf{e}_{t+1}$ is roughly uniform over all unexplored half-edges, then at the early stages of the exploration when most half-edges are unexplored and outside the queue, one should expect

$$\begin{aligned}
&\mathbb{E}\left[|\mathbf{Q}_{t+1}| - |\mathbf{Q}_t| \mid \mathbf{K}_t, \mathbf{Q}_t\right] \\
&\geq \mathbb{P}\{\mathbf{e}_{t+1} \text{ not back-edge} \mid \mathbf{K}_t, \mathbf{Q}_t\} - 2\mathbb{P}\{\mathbf{e}_{t+1} \text{ back-edge} \mid \mathbf{K}_t, \mathbf{Q}_t\} \\
&= 1 - 3\mathbb{P}\{\mathbf{e}_{t+1} \text{ back-edge} \mid \mathbf{K}_t, \mathbf{Q}_t\} \\
&\approx 1 - \frac{3|\mathbf{Q}_t|}{2m}.
\end{aligned}$$

Until a significant fraction of the $2m$ half-edges in $\mathbf{K}$ are either explored or in $\mathbf{Q}_t$, we should therefore expect an average net gain of at least some $\varepsilon > 0$ from each step of the exploration. Then, letting s be the number of time steps until all half-edges in $\mathbf{Q}_t$ are explored, we should expect $|\mathbf{Q}_{t+s}| \geq (1+\varepsilon)|\mathbf{Q}_t|$. Because our exploration is *breadth-first*, for all $s \geq 0$ such that $\mathbf{e}_{t+s}$ corresponds to an element of $M(\mathbf{A})$ with at least one half-edge in $\mathbf{Q}_t$, the radius cannot have increased by more than 1; that is, $\mathrm{rad}(\mathbf{K}_{t+s}) \leq \mathrm{rad}(\mathbf{K}_t)+1$. We thus expect that with high probability the queue size grows exponentially as a function of the radius until the queue has order m half-edges. In other words, after only exploring to distance $O(\log m)$ from v, we should have encountered a non-vanishing fraction of the edges of $\mathbf{K}$. Again assuming that the match of a given half-edge is roughly uniformly distributed over the other half-edges, if two such explorations (started from different vertices but run on the same kernel) have a linear number of half-edges in their queues, then with high probability some edge of $\mathbf{K}$ should have one half-edge in one of the queues and the other half-edge in the other queue, joining the explorations. If all starting vertices have a linear queue size after exploring to logarithmic radius, and any two linear-size queues meet, then $\mathrm{diam}^+(\mathbf{K}) = O(\log m)$.

The only part of this argument which resists formalization is the assumption of approximate uniformity of the match of a given half-edge. This is essentially equivalent to the distribution of $\mathbf{K}$ being well approximated by a sample from the configuration model, and this heuristic is false when ℓ is not much larger than n and d contains sufficiently large degrees. We can however prove the following two lemmas using the switching method, both of which are weak forms of the assertion that $\mathbb{P}\{\mathbf{e}_{t+1} \text{ back-edge} \mid \mathbf{K}_t, \mathbf{Q}_t\} \approx |\mathbf{Q}_t|/2m$.

Lemma 5.4. *For all t,*

$$\mathbb{P}\{\mathbf{e}_{t+1} \textit{ is a back-edge}, |B_{\mathbf{K}}(K_t,2)| \leq m/2 \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\} \leq \frac{2|Q_t|}{m}.$$

Lemma 5.5. *For all t such that $t + |Q_t| + \binom{|Q_t|}{2} \leq m/2$,*

$$\mathbb{P}\{\mathbf{e}_{t+1} \textit{ is a loop} \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\} \leq \frac{4|Q_t|}{m}.$$

Unfortunately, not all back-edges are loops, and the event that $|B_{\mathbf{K}}(K_t,2)| \leq m/2$ is not measurable with respect to the σ-algebra generated by the process up to time $t+1$. This means that Lemma 5.4 is insufficient as an input to Azumas inequality or any of its relatives. We avoid this problem by using analogues of Lemma 5.4 (Lemma 5.10 and Lemma 5.13) where we explore many edges at a time. This way we can obtain failure probabilities that are small enough to be integrated with a union bound.

Proofs that invoke switching are often rather technical, so for the sake of brevity we have relegated some of the switching proofs to the appendix. We will however give three switching arguments in the main text, each of which illustrate a different aspect of the technique. For now we will prove Lemma 5.4 and Lemma 5.5. The proof of Lemma 5.4 shows how intersecting with the event that the ball of some small radius around the explored subgraph can be taken advantage of. The proof of Lemma 5.5 shows why edges that are known to be subdivided (such as loops) are easier to work with in switching arguments. Later, we will prove Lemma 5.15, which serves as an example of a situation where we work with sequences of switches, and use the more general Lemma 1.8, rather than Lemma 5.2.

Proof of Lemma 5.4. Let

$$\begin{aligned}
\mathcal{C} &= \{A \in \mathcal{A}_{\mathrm{d}} : K_t(A,v) = K_t, Q_t(A,v) = Q_t\},\\
\mathcal{A} &= \{A \in \mathcal{C} : e_{t+1}(A,v) \text{ is a back-edge}, |B_{K(A)}(K_t,2)| \leq m/2\}, \text{ and}\\
\mathcal{B} &= \{A \in \mathcal{C} : e_{t+1}(A,v) \text{ is not a back-edge}\}.
\end{aligned}$$

For $A \in \mathcal{A}$ and $B \in \mathcal{B}$, we will lower-bound the number of switchings from A to an element of $\mathcal{B}$ and upper-bound the number of switchings from B to an element of $\mathcal{A}$.

Choose an orientation $\vec{e}$ of $e_{t+1}(A,v)$. There are 2 choices for $\vec{e}$. Choose an oriented edge $\vec{f}$ in $K(A)$ outside $B_{K(A)}(K_t,2)$. Since $A\in\mathcal{A}$, there are at least $m/2$ edges in $K(A)$ outside $B_{K(A)}(K_t,2)$, and thus at least $2(m/2)=m$ choices for $\vec{f}$. Since $A\in\mathcal{A}$, $e_{t+1}(A,v)$ is a back-edge, and hence $\vec{e}$ is in $B_{K(A)}(K_t,0)$. It follows that $(\vec{e},\vec{f})$ is a valid pair. Writing A' for the augmented core obtained from A by switching on $(\vec{e},\vec{f})$, the edge $e_{t+1}(A',v)$ shares one half-edge with $e_{t+1}(A,v)$ which is in Q_t, and the other half-edge of $e_{t+1}(A',v)$ is incident to an endpoint of $\vec{f}$. Since both endpoints of $\vec{f}$ are outside K_t, it follows that $A'\in\mathcal{B}$. This proves that there are at least $2m$ switchings from A to an element of $\mathcal{B}$.

Any switching from B to an element of $\mathcal{A}$ is equivalent to a pair $(\vec{e},\vec{f})$ where $\vec{e}$ is an orientation of $e_{t+1}(B,v)$ and $\vec{f}$ is an orientation of an edge in $K(B)$ with a half-edge in Q_t. There are at most $4|Q_t|$ such pairs. This proves that there are at most $4|Q_t|$ switchings from B to an element of $\mathcal{A}$.

Applying Lemma 5.2, we obtain

$$\begin{aligned}
&\mathbb{P}\{\mathbf{e}_{t+1}\text{ back-edge}, |B_{\mathbf{K}}(K_t,2)|\le m/2 \mid \mathbf{K}_t=K_t,\mathbf{Q}_t=Q_t\}\\
&=\mathbb{P}\{\mathbf{A}\in\mathcal{A}\mid \mathbf{A}\in\mathcal{C}\}\\
&\le \frac{4|Q_t|}{2m}\mathbb{P}\{\mathbf{A}\in\mathcal{B}\mid \mathbf{A}\in\mathcal{C}\}\\
&\le \frac{2|Q_t|}{m},
\end{aligned}$$

as required. □

Proof of Lemma 5.5. Let

$$\begin{aligned}
\mathcal{C}&=\{A\in\mathcal{A}_{\mathrm{d}}: K_t(A,v)=K_t, Q_t(A,v)=Q_t\},\\
\mathcal{A}&=\{A\in\mathcal{C}: e_{t+1}(A,v)\text{ is a loop}\},\text{ and}\\
\mathcal{B}&=\{A\in\mathcal{C}: e_{t+1}(A,v)\text{ is not a loop}\}
\end{aligned}$$

For $A\in\mathcal{A}$ and $B\in\mathcal{B}$, we lower-bound the number of switchings from A to an element of $\mathcal{B}$, and upper-bound the number of switchings from B to an element of $\mathcal{A}$.

Consider the set E of all edges in $K(A)$ not incident to K_t that are either subdivided or have an endpoint not adjacent to K_t. The number of edges in $K(A)$ with an endpoint in K_t is at most $t+|Q_t|$. Furthermore, there are at most $|Q_t|$ vertices outside K_t that are adjacent to K_t. Because the non-subdivided edges span a simple subgraph of $K(A)$, it follows that there are at most $\binom{|Q_t|}{2}$ non-subdivided edges both of whose endpoints are outside K_t but adjacent to K_t. This implies

$$|E|\ge m-|Q_t|-t-\binom{|Q_t|}{2}\ge m/2.$$

Choose an orientation $\vec{e}$ of $e_{t+1}(A,v)$. There are 2 choices for $\vec{e}$. Let $\vec{f}$ be an orientation of an edge in E. There are at least $2(m/2)=m$ choices for $\vec{f}$. Since $A\in\mathcal{A}$, $\vec{e}$ is a loop, so is subdivided. Since the underlying edge of $\vec{f}$ is in E, it is either subdivided or has an endpoint not adjacent to either endpoint of $\vec{e}$. Therefore by Lemma 5.3, either $(\vec{e},\vec{f})$ or $(\vec{e},\overleftarrow{f})$ is a valid pair. Moreover switching on whichever of these pairs is valid yields an element of $\mathcal{B}$. It follows that there are at least $2m/2=m$ switchings from A to an element of $\mathcal{B}$.

Any switching from B to an element of $\mathcal{A}$ is equivalent to a pair $(\vec{e},\vec{f})$ such that $\vec{e}$ is an orientation of $e_{t+1}(A,v)$ and $\vec{f}$ is an orientation of some edge with a half-edge in Q_t (in fact one of the endpoints of $\vec{f}$ must be an endpoint of $e_{t+1}(A,v)$). There are at most $4|Q_t|$ such pairs. This proves that there are at most $4|Q_t|$ switchings from B to an element of $\mathcal{A}$.

Applying Lemma 5.2, we obtain

$$\begin{aligned}
&\mathbb{P}\left\{\mathbf{e}_{t+1} \text{ is a loop} \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\right\} \\
&= \mathbb{P}\left\{\mathbf{A} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\right\} \\
&\le \frac{4|Q_t|}{m}\mathbb{P}\left\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\right\} \\
&\le \frac{4|Q_t|}{m},
\end{aligned}$$

as required. □

5.1. **Getting the queue going.** The goal of this subsection is to show that with probability $1 - O\left(m^{-2}\right)$, the queue reaches size 11 at a time when the radius is bounded. This result will be important later, when we establish that linear queue growth in time implies exponential queue growth in radius. In that argument, we will require the initial queue size to exceed a certain threshold (which we take to be 11).

Lemma 5.6.

$$\mathbb{P}\left\{|\mathbf{Q}_t| \ge 11,\ \mathrm{rad}(\mathbf{K}_t) \le 20 \textit{ for some } t \le 75\right\} = 1 - O\left(m^{-2}\right).$$

As we saw in the proof of Lemma 5.4, in switching arguments, when bounding the probability of some "bad" queue growth event, it can be useful to intersect it with the event that a ball of small radius around the explored subgraph is not too large. The following observation shall allow us to make this move safely.

Observation 5.7. *If* $\mathrm{rad}(K_t) \le r$ *and* $|B_K(K_t, 3)| \ge m/3$*, then there exists* s *such that* $\mathrm{rad}(K_s) \le r + 3$ *and* $|Q_s| \ge m/(3r + 12)$*.*

Proof. For $r' \ge 0$, write $t(r') = \max(t' : \mathrm{rad}(K_{t'}) \le r')$. If $e \in B_K(K_t, 3)$, then since $\mathrm{rad}(K_t) \le r$, both endpoints of e are within distance $r + 3$ of v, so e has a half-edge in $Q_{t(r')}$ for some $r' \le r + 3$. Hence,

$$\frac{m}{3} \le |B_K(K_t, 3)| \le \sum_{r'=0}^{r+3} |Q_{t(r')}| \le (r + 4)\max(|Q_{t(r')}| : r' \in [0, r + 3]),$$

so $|Q_{t(r')}| \ge m/(3(r + 4))$ for some $r' \le r + 3$, and we take $s = t(r')$. □

Our first step toward proving Lemma 5.6 is to establish that with high probability the queue reaches size 5 in the first few time steps.

Lemma 5.8.

$$\mathbb{P}\left\{|\mathbf{Q}_t| < 5,\ |B_{\mathbf{K}}(\mathbf{K}_t, 2)| \le m/2 \textit{ for all } t \le 5\right\} = O\left(m^{-2}\right).$$

Lemma 5.8 itself relies on Lemma 5.4, Lemma 5.5 and the following additional preliminary, whose proof we postpone. Collectively, these three results allow us to understand the queue at early times.

Lemma 5.9. *For all* $q \ge 0$ *such that* $1 + q + \binom{q}{2} \le m/2$*,*

$$\mathbb{P}\left\{\mathbf{e}_2 \textit{ is a back-edge}, |\mathbf{Q}_1| \le q\right\} \le \frac{16q + 32}{m}.$$

Given these bounds, we can deduce that the queue quickly reaches size 5.

Proof of Lemma 5.8. We write "$\mathbf{e}_t$ back" and "$\mathbf{e}_t$ loop" respectively as shorthand for "$\mathbf{e}_t$ is a back-edge" and "$\mathbf{e}_t$ is a loop". If neither $\mathbf{e}_1$ nor $\mathbf{e}_2$ is a back-edge, then $|\mathbf{Q}_2| \ge 5$. If exactly one of $\mathbf{e}_1$ or $\mathbf{e}_2$ is a back-edge but none of $\mathbf{e}_3$, $\mathbf{e}_4$, $\mathbf{e}_5$ are back-edges, then $|\mathbf{Q}_5| \geq 5$. Therefore

$$\begin{aligned}
&\mathbb{P}\left\{|\mathbf{Q}_t| < 5,\ |B_{\mathbf{K}}(K_t,2)| \leq m/2 \text{ for all } t \leq 5\right\} \\
&\leq \mathbb{P}\left\{\mathbf{e}_1 \text{ back}, |\mathbf{Q}_1| \leq 4, \mathbf{e}_2 \text{ back}\right\} \\
&+ \mathbb{P}\left\{\mathbf{e}_1 \text{ back}, \mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4\right\} \\
&+ \mathbb{P}\left\{\mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4, \mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4\right\}
\end{aligned}$$

We can assume that $d_v \leq 4$ since $|\mathbf{Q}_0| = d_v$ deterministically. If $\mathbf{e}_1$ is a back-edge, then it is a loop. Furthermore, if both $\mathbf{e}_1$ and $\mathbf{e}_2$ are back-edges, then they are both loops. By Lemma 5.5,

$$\begin{aligned}
&\mathbb{P}\left\{\mathbf{e}_1 \text{ back}, |\mathbf{Q}_1| \leq 4, \mathbf{e}_2 \text{ back}\right\} \\
&= \mathbb{P}\left\{\mathbf{e}_1 \text{ loop}, |\mathbf{Q}_1| \leq 4, \mathbf{e}_2 \text{ loop}\right\} \\
&\leq \mathbb{P}\left\{\mathbf{e}_1 \text{ loop}\right\} \mathbb{P}\left\{\mathbf{e}_2 \text{ loop} \mid \mathbf{e}_1 \text{ loop}, |\mathbf{Q}_1| \leq 4\right\} \\
&\leq \frac{4d_v}{m} \cdot \frac{4 \cdot 4}{m} \\
&\leq \frac{256}{m^2}.
\end{aligned}$$

By Lemma 5.5 and Lemma 5.4,

$$\begin{aligned}
&\mathbb{P}\left\{\mathbf{e}_1 \text{ back}, \mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4\right\} \\
&= \mathbb{P}\left\{\mathbf{e}_1 \text{ loop}\right\} \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4 \mid \mathbf{e}_1 \text{ loop}\right\} \\
&\leq \frac{4d_v}{m} \sum_{t=2}^{4} \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back}, |B_{\mathbf{K}}(K_t,2)| \leq m/2, |\mathbf{Q}_t| \leq 4 \mid \mathbf{e}_1 \text{ loop}\right\} \\
&\leq \frac{4d_v}{m} \sum_{t=2}^{4} \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back}, |B_{\mathbf{K}}(K_t,2)| \leq m/2 \mid \mathbf{e}_1 \text{ loop}, |\mathbf{Q}_t| \leq 4\right\} \\
&\leq \frac{4d_v}{m} \cdot 3 \cdot \frac{2 \cdot 4}{m} \\
&\leq \frac{384}{m^2}.
\end{aligned}$$

By Lemma 5.9 and Lemma 5.4,

$$\begin{aligned}
&\mathbb{P}\left\{\mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4, \mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4\right\} \\
&= \mathbb{P}\left\{\mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4\right\} \\
&\cdot \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back},\ |B_{\mathbf{K}}(K_t,2)| \leq m/2,\ |\mathbf{Q}_t| \leq 4 \text{ for some } 2 \leq t \leq 4 \mid \mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4\right\} \\
&\leq \frac{16 \cdot 4 + 32}{m} \sum_{t=2}^{4} \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back}, |B_{\mathbf{K}}(K_t,2)| \leq m/2, |\mathbf{Q}_t| \leq 4 \mid \mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4\right\} \\
&\leq \frac{96}{m} \sum_{t=2}^{4} \mathbb{P}\left\{\mathbf{e}_{t+1} \text{ back}, |B_{\mathbf{K}}(K_t,2)| \leq m/2 \mid \mathbf{e}_2 \text{ back}, |\mathbf{Q}_1| \leq 4, |\mathbf{Q}_t| \leq 4\right\} \\
&\leq \frac{96}{m} \cdot 3 \cdot \frac{2 \cdot 4}{m} \\
&= \frac{2304}{m^2}.
\end{aligned}$$

Combining these three bounds we get the required inequality

$$\mathbb{P}\left\{|\mathbf{Q}_t| < 5,\ |B_{\mathbf{K}}(K_t,2)| \leq m/2 \text{ for all } t \leq 5\right\} \leq \frac{256}{m^2} + \frac{384}{m^2} + \frac{2304}{m^2} = O\left(m^{-2}\right).$$

□

The other fact we need in the proof of Lemma 5.6 is that when the queue has size at least 5, it is likely to grow in the next 5 time steps. This is the content of the following lemma, whose proof we also postpone.

Lemma 5.10. *For all t such that $|Q_t| \geq 5$,*

$$\mathbb{P}\left\{|B_{\mathbf{K}}(K_t,3)| \leq m/2,\ |\mathbf{Q}_{t+s}| \leq |Q_t| \text{ for all } s \in [5] \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\right\} \leq \left(\frac{32|Q_t| + 1760}{m}\right)^2 .$$

Proof of Lemma 5.6. We start by establishing that if the queue does not reach size 11 quickly with a small radius, then one of the low probability events described in Lemma 5.8 or Lemma 5.10 must occur.

Claim 5.11. *Suppose that either $|Q_{t_0}| \geq 5$ or $|B_K(K_{t_0}, 3)| \geq m/3$ for some $t_0 \leq 5$. Suppose also that for all $t \leq 35$ such that $5 \leq |Q_t| < 11$, either $|B_K(K_t, 3)| \geq m/3$ or $|Q_{t+s}| \geq |Q_t| + 1$ for some $s \in [5]$. Then there exists $t^* \leq 75$ such that $\mathrm{rad}(K_{t^*}) \leq 20$ and $|Q_{t^*}| \geq 11$.*

Before proving this claim, we deduce the conclusion of Lemma 5.6 from it. If there is no $t \leq 75$ such that $\mathrm{rad}(\mathbf{K}_t) \leq 20$ and $|\mathbf{Q}_t| \geq 11$, then by Claim 5.11, either $|\mathbf{Q}_t| < 5$ and $|B_{\mathbf{K}}(\mathbf{K}_t, 3)| < m/3$ for all $t \leq 5$, or, for some $t \leq 35$ such that $5 \leq |\mathbf{Q}_t| < 11$, $|B_{\mathbf{K}}(\mathbf{K}_t, 3)| < m/3$ and $|\mathbf{Q}_{t+s}| \leq |\mathbf{Q}_t|$ for all $s \in [5]$. Then by Lemma 5.8 and Lemma 5.10,

$$\begin{aligned}
&\mathbb{P}\left\{|\mathbf{Q}_t| < 11 \text{ or } \mathrm{rad}(\mathbf{K}_t) > 20 \text{ for all } t \leq 75\right\} \\
&\leq \mathbb{P}\left\{|\mathbf{Q}_t| < 5,\ |B_{\mathbf{K}}(\mathbf{K}_t, 2)| < m/3 \text{ for all } t \leq 5\right\} \\
&+ \sum_{t=0}^{35} \mathbb{P}\left\{|B_{\mathbf{K}}(\mathbf{K}_t, 3)| < m/3,\ |\mathbf{Q}_{t+s}| \leq |\mathbf{Q}_t| \text{ for all } s \in [5] \mid 5 \leq |\mathbf{Q}_t| < 11\right\} \\
&\leq O\left(m^{-2}\right) + 36\left(\frac{32 \cdot 10 + 1760}{m}\right)^2 \\
&= O\left(m^{-2}\right).
\end{aligned}$$

It therefore suffices to establish Claim 5.11.

Proof of Claim 5.11. First we show that if the conclusion fails, then there exists $t_0^+ \leq 35$ such that $\mathrm{rad}(K_{t_0^+}) \leq 17$ and $|B_K(K_{t_0^+})| \geq m/3$.

If $|Q_{t_0}| < 5$, then taking $t_0^+ = t_0$, we have $t_0^+ \leq 5 \leq 35$, $\mathrm{rad}(K_{t_0^+}) \leq 5 \leq 17$ and $|B_K(K_{t_0^+}, 3)| \geq m/3$. Otherwise, we define the sequence of time steps $(s_i, i \in [k])$ as follows. Set $s_0 = 0$. For $i \geq 0$, if $t_0 + s_i \leq 35$, $5 \leq |Q_{t_0+s_i}| < 11$, and $|B_K(K_{t_0+s_i}, 3)| < m/3$, then by the assumption of the claim, there exists $r_{t+1} \in [5]$ such that $|Q_{t_0+s_i+r_{i+1}}| \geq |Q_{t_0+s_i}| + 1$, and we take $s_{i+1} = s_i + r_{i+1}$. If any of these three conditions fail, then set $k = i$ and terminate. Then it follows by induction that $t_0 + s_i \leq t_0 + 5i \leq 5(i+1)$ and $|Q_{t_0+s_i}| \geq |Q_{t_0}| + i \geq 5 + i$ for all $i \in [k]$. For all $i \in [0, k)$, we have $|Q_{t_0+s_i}| \geq 5$, and since $r_{i+1} \in [5]$, this implies that $\mathrm{rad}(K_{t_0+s_{i+1}}) \leq \mathrm{rad}(K_{t_0+s_i}) + 2$. It therefore also follows by induction that $\mathrm{rad}(K_{t_0+s_i}) \leq \mathrm{rad}(K_{t_0}) + 2i$ for all $i \in [0, k]$. Now if $k \geq 1$, then $5 + (k-1) \leq |Q_{t_0+s_{k-1}}| < 11$. Rearranging yields $k \leq 6$, so $t_0 + s_k \leq 5(k+1) \leq 35$. Hence $\mathrm{rad}(K_{t_0+s_k}) \leq \mathrm{rad}(K_{t_0}) + 2k \leq 5 + 2 \cdot 6 = 17$. Furthermore, since $t_0 + s_k \leq 35$, by definition of k, either $|Q_{t_0+s_k}| \geq 11$ or $|B_K(K_{t_0+s_k}, 3)| \geq m/3$. If $|Q_{t_0+s_k}| \geq 11$, then taking $t^* = t_0 + s_k$, we have $t^* \leq 35 \leq 75$, $\mathrm{rad}(K_{t^*}) \leq 17 \leq 20$, and $|Q_{t^*}| \geq 11$, so we are done. Otherwise, with $t_0^+ = t_0 + s_k$, we have $t_0^+ \leq 35$, $\mathrm{rad}(K_{t_0^+}) \leq 17$, and $|B_K(K_{t_0^+}, 3)| \geq m/3$.

Suppose toward a contradiction that for all $t \le 75$ such that $\mathrm{rad}(K_t) \le 20$, we have $|Q_t| < 11$. Starting from any time t with $|Q_t| < 11$, it takes at most 10 time steps to start exploring half-edges at a radius greater than $\mathrm{rad}(K_t)$. Writing $r = \mathrm{rad}(K_{t_0^+})$ and using the notation from Observation 5.7, we thus have $t(r)+1 \le t_0^+ + 10 \le 45$, $t(r+1)+1 \le t(r)+1+10 \le 55$, $t(r+2)+1 \le t(r+1)+1+10 \le 65$, and $t(r+3)+1 \le t(r+2)+1+10 \le 75$. Now observe that that every edge in $B_K(K_{t_0^+}, 3)$ is either in $K_{t_0^+}$, has a half-edge in $Q_{t_0^+}$, or has a half edge in $Q_{t(r')+1}$ for some $r \le r' \le r+3$. Hence

$$\begin{aligned}\frac{m}{3} &\le |B_K(K_{t_0^+}, 3)| \\ &\le t_0^+ + |Q_{t_0^+}| + |Q_{t(r)+1}| + |Q_{t(r+1)+1}| + |Q_{t(r+2)+1}| + |Q_{t(r+3)+1}| \\ &\le 35 + 5 \cdot 10\end{aligned}$$

This contradicts our running assumption that m is large. ■

□

5.2. **Growing the queue.** In this subsection we show that with high probability, the exploration reaches queue length of order $m/\log^2 m$ within logarithmic radius.

Lemma 5.12.

$$\mathbb{P}\left\{|\mathbf{Q}_t| \ge m/8000\log^2 m,\ \mathrm{rad}(\mathbf{K}_t) \le 111\log m \textit{ for some } t\right\} = 1 - O\left(m^{-2}\right).$$

As previously mentioned, we will analyze the queue growth not one edge at a time but instead in batches of size b. Unsurprisingly, good choices for b depend on the current size of the queue. Once the queue has size 11, we will start with batches of size 5. Unfortunately the failure probability of $1 - O\left(m^{-2}\right)$ coming from Lemma 5.10 is not good enough for a union bound over all time and all starting vertices. This means that after using batches of size 5 for a short period, we must increase the batch size. This is the purpose of Lemma 5.13, whose proof we leave to the appendix.

Lemma 5.13. *If $b \ge 4$, then for all t such that $|Q_t| \ge b$ and $72b^3 + 24b|Q_t| \le m/2$,*

$$\begin{aligned}&\mathbb{P}\left\{|B_{\mathbf{K}}(K_t, 3)| \le m/3, |\mathbf{Q}_{t+s}| \le |Q_t| + b/4 \textit{ for all } s \in [b] \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\right\} \\ &\le 3\left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4}.\end{aligned}$$

To streamline the proof of Lemma 5.12, we isolate the following deterministic fact. The statement is technical but, roughly speaking, it says that if the queue is large enough compared to the batch size, then linear queue growth from each batch (which we should expect by Lemma 5.13) implies a queue size which is exponentially large as a function of radius until the process crosses a queue size threshold. Once this threshold is crossed, we will increase the batch size and apply Lemma 5.14 again.

Lemma 5.14. *Fix $\varepsilon \in (0,1)$, $\delta \in (0,1/2)$ and integers $t_0 \ge 0$, c, and $b,p,q \ge 1$ such that $q \ge 2p$, $b \le \delta p$, and $q(3(\log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 1) + 12) \le m$. Suppose that $\mathrm{rad}(K_{t_0}) \le \log(p)/\log(1+\varepsilon/2-\varepsilon\delta) - c$ and that $|Q_{t_0}| \ge p$. Suppose also that for all $t < (1/\varepsilon+1)q + t_0$ such that $p \le |Q_t| < q$, either $|B_K(K_t,3)| \ge m/3$ or $|Q_{t+s}| \ge |Q_t| + \varepsilon b$ for some $s \in [b]$. Then there exists t^* such that $\mathrm{rad}(K_{t^*}) \le \log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 4$ and $|Q_{t^*}| \ge q$.*

Proof. First we show that under the assumptions of the lemma, for all t such that $p \le |Q_t| < q$, $t + |Q_t|/2 \le (1/\varepsilon+1)q + t_0$, and $|B_K(K_t,3)| < m/3$, there exists $s \ge 1$ such that $s \le \min((|Q_{t+s}| - |Q_t|)/\varepsilon, |Q_t|/2)$, $\mathrm{rad}(K_{t+s}) \le \mathrm{rad}(K_t) + 1$, and either $|Q_{t+s}| \ge \min((1+\varepsilon/2-\varepsilon\delta)|Q_t|, q)$, or $|B_K(K_{t+s})| \ge m/3$.

Define an increasing sequence of integers $(s_i, i \in [0,k])$ recursively as follows. First, set $s_0 = 0$. For $i \geq 0$, if $t + s_i < (1/\varepsilon + 1)q + t_0$, $|Q_{t+s_i}| < \min((1+\varepsilon/2-\varepsilon\delta)|Q_t|, q)$, and $|B_K(K_{t+s_i})| < m/3$, then by the assumptions of the lemma, we can choose $r_{i+1} \in [b]$ such that $|Q_{t+s_i+r_{i+1}}| \geq |Q_{t+s_i}| + \varepsilon b$ and we take $s_{i+1} = s_i + r_{i+1}$. If any of these three conditions fail, set $k = i$ and terminate. We take $s = s_k$. By assumption, these conditions hold when $i = 0$, so $k \geq 1$, and hence $s \geq 1$. Furthermore, it follows by induction that $s_i \leq ib$ and $|Q_{t+s_i}| \geq |Q_t| + \varepsilon ib$ for all $i \in [0,k]$. We claim that $s < |Q_t|/2$. Indeed $|Q_t| + \varepsilon(k-1)b \leq |Q_{t+s_{k-1}}| < (1+\varepsilon/2-\varepsilon\delta)|Q_t|$, which combined with the fact that $|Q_t| \geq p \geq b/\delta$ yields $kb < |Q_t|/2$. Thus $s = s_k \leq kb < |Q_t|/2$ as claimed. This implies $t + s < t + |Q_t|/2 \leq (1/\varepsilon+1)q + t_0$, so by definition of k and since $s = s_k$, either $|Q_{t+s}| \geq \min((1+\varepsilon/2-\varepsilon\delta)|Q_t|, q)$ or $|B_K(K_{t+s})| \geq m/3$. That $s < |Q_t|/2$ also implies $\mathrm{rad}(K_{t+s}) \leq \mathrm{rad}(K_t) + 1$. Indeed at most two half-edges are explored in each time step, and in order to increase the radius of the exploration by two, we must explore every half-edge in the current queue. It only remains to check that $s \leq (|Q_{t+s}| - |Q_t|)/\varepsilon$, and this holds because $s = s_k \leq kb$, so $|Q_{t+s}| \geq |Q_t| + \varepsilon kb \geq |Q_t| + \varepsilon s$. Thus, s has the desired properties.

Starting with t_0, we recursively define times $(t_i, i \in [0,i^*])$ as follows. If $p \leq |Q_{t_i}| < q$, $t_i + |Q_{t_i}|/2 \leq (1/\varepsilon+1)q + t_0$, and $|B(K_{t_i}, 3)| < m/3$, then choose $t_{i+1} \geq t_i + 1$ such that $t_{i+1} - t_i \leq \min((|Q_{t_{i+1}}| - |Q_{t_i}|)/\varepsilon, |Q_{t_i}|/2)$, $\mathrm{rad}(K_{t_{i+1}}) \leq \mathrm{rad}(K_{t_i}) + 1$, and either $|Q_{t_{i+1}}| \geq \min((1+\varepsilon/2-\varepsilon\delta)|Q_{t_i}|, q)$ or $|B_K(K_{t_{i+1}}, 3)| \geq m/3$. Such an integer t_{i+1} exists by the preceding argument. Otherwise, set $i^* = i$ and terminate.

One of $|Q_{t_{i^*}}| \geq q$ or $t_{i^*} + |Q_{t_{i^*}}|/2 > (1/\varepsilon+1)q + t_0$ or $|B(K_{t_{i^*}}, 3)| \geq m/3$ must hold. Furthermore, it follows by induction that $|Q_{t_i}| \geq \min((1+\varepsilon/2-\varepsilon\delta)^i p, q)$ and $\mathrm{rad}(K_{t_i}) \leq \mathrm{rad}(K_{t_0}) + i$ for all $i \in [0,i^*]$. If $i^* \geq 1$ then $|Q_{t_{i^*-1}}| < q$, so $(1+\varepsilon/2-\varepsilon\delta)^{i^*-1}p < q$, and hence $i^* \leq \log(q/p)/\log(1+\varepsilon/2-\varepsilon\delta) + 1$. If $i^* = 0$ then this still holds. Therefore

$$\begin{aligned}
\mathrm{rad}(K_{t_{i^*}}) &\leq \mathrm{rad}(K_{t_0}) + i^* \\
&\leq \mathrm{rad}(K_{t_0}) + \log(q/p)/\log(1+\varepsilon/2-\varepsilon\delta) + 1 \\
&\leq \log(p)/\log(1+\varepsilon/2-\varepsilon\delta) - c + \log(q/p)/\log(1+\varepsilon/2-\varepsilon\delta) + 1 \\
&= \log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 1.
\end{aligned}$$

Now observe that $t_{i^*} \leq (1/\varepsilon + 1/2)q + t_0$. Indeed this holds if $i^* = 0$, and if $i^* \geq 1$, then

$$\begin{aligned}
t_{i^*} - t_0 &= (t_{i^*} - t_{i^*-1}) + \sum_{i=1}^{i^*-1}(t_i - t_{i-1}) \\
&\leq |Q_{t_{i^*-1}}|/2 + \sum_{i=1}^{i^*-1}(|Q_{t_i}| - |Q_{t_{i-1}}|)/\varepsilon \\
&= |Q_{t_{i^*-1}}|/2 + (|Q_{t_{i^*-1}}| - |Q_{t_0}|)/\varepsilon \\
&\leq (1/\varepsilon + 1/2)q.
\end{aligned}$$

Therefore if $|Q_{t_{i^*}}| < q$, then

$$t_{i^*} + |Q_{t_{i^*}}|/2 < (1/\varepsilon + 1/2)q + q/2 + t_0 = (1/\varepsilon + 1)q + t_0.$$

This implies $|B(K_{t_{i^*}})| \geq m/3$, but then by Observation 5.7, there exists t^* such that

$$\begin{aligned}
\mathrm{rad}(K_{t^*}) &\leq \mathrm{rad}(K_{t_{i^*}}) + 3 \leq \log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 1 + 3, \text{ and} \\
|Q_{t^*}| &\geq \frac{m}{3(\log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 1) + 12} \geq q,
\end{aligned}$$

the last bound holding by the assumptions of the lemma. Otherwise, $|Q_{t_{i^*}}| \geq q$ and with $t^* = t_{i^*}$, we have

$$\begin{aligned} \mathrm{rad}(K_{t^*}) &\leq \log(q)/\log(1+\varepsilon/2-\varepsilon\delta) - c + 1, \text{ and} \\ |Q_{t^*}| &\geq q. \end{aligned}$$

This completes the proof. □

We are now prepared to prove Lemma 5.12

Proof of Lemma 5.12. For $i \in \{1,2,3\}$, let $\varepsilon_i, \delta_i, t_i, c_i, b_i, p_i, q_i$ be given by the following table.

i	ε_i	δ_i	t_i	c_i	b_i	p_i	q_i
1	1/5	5/11	75	244	5	11	700
2	1/4	51/110	m	240	324	700	$\lfloor m^{1/3} \rfloor$
3	1/4	51/110	m	236	$\lceil 38 \log^2 m \rceil$	$\lfloor m^{1/3} \rfloor$	$\lfloor m/7500 \log^2 m \rfloor$

Let $\mathcal{A}^0$ be the set all $A \in \mathcal{A}_{\mathrm{d}}$ such that for some $t \leq 75$, we have $|Q_t(A,v)| \geq 11$ and $\mathrm{rad}(K_t(A,v)) \leq 20$. Then by Lemma 5.6,

$$\mathbb{P}\left\{\mathbf{A} \notin \mathcal{A}^0\right\} = O\left(m^{-2}\right).$$

For $i \in \{1,2,3\}$, let $\mathcal{A}^i$ be the set of all $A \in \mathcal{A}_{\mathrm{d}}$ such that for all $t \leq (1/\varepsilon_i+1)q_i + t_i$ with $p_i \leq |Q_t(A,v)| < q_i$, either $|B_{K(A)}(K_t(A,v),3)| > m/3$ or $|Q_{t+s}(A,v)| \geq |Q_t(A,v)| + \varepsilon_i b_i$ for some $s \in [b_i]$. Since $b_1 = 5$ and $\varepsilon_1 b_1 = 1$, by Lemma 5.10,

$$\begin{aligned} &\mathbb{P}\left\{\mathbf{A} \notin \mathcal{A}^1\right\} \\ &\leq \sum_{t=0}^{(1/\varepsilon_1+1)q_1+t_1-1} \mathbb{P}\left\{|B_{\mathbf{K}}(\mathbf{K}_t,3)| \leq m/3,\ |\mathbf{Q}_{t+s}| \leq |\mathbf{Q}_t| \text{ for all } s \in [5] \mid p_1 \leq |\mathbf{Q}_t| < q_1\right\} \\ &\leq \left(\left(\frac{1}{\varepsilon_1}+1\right)q_1 + t_1\right)\left(\frac{32q_1+1760}{m}\right)^2 \\ &= O\left(m^{-2}\right). \end{aligned}$$

Since $\varepsilon_2 b_2 = b_2/4$, $b_2 \geq 4$, $p_2 \geq b_2$, and $72b_2^3 + 24b_2q_2 \leq 8000m^{1/3} \leq m/2$, by Lemma 5.13,

$$\begin{aligned} &\mathbb{P}\left\{\mathbf{A} \notin \mathcal{A}^2\right\} \\ &\leq \sum_{t=0}^{m} \mathbb{P}\left\{|B_{\mathbf{K}}(\mathbf{K}_t,3)| \leq m/3, |\mathbf{Q}_{t+s}| \leq |\mathbf{Q}_t| + b_2/4 \text{ for all } s \in [b_2] \mid p_2 \leq |\mathbf{Q}_t| < q_2\right\} \\ &\leq 3(m+1)\left(\frac{72b_2^3+24b_2q_2}{m}\right)^{\sqrt{b_2}/4} \\ &\leq (m+1)3\left(\frac{8000m^{1/3}}{m}\right)^{9/2} \\ &= O\left(m^{-2}\right). \end{aligned}$$

Finally, since $\varepsilon_3 b_3 = b_3/4$, $b_3 \geq 4$, $p_3 \geq b_3$, $72b_3^3 + 24b_3q_3 \leq 0.13m \leq m/2$, and $\sqrt{b_3}/4 \geq 1.54\log m$, by Lemma 5.13

$$\begin{aligned}
&\mathbb{P}\left\{\mathbf{A} \notin \mathcal{A}^3\right\}\\
&\leq \sum_{t=0}^{m} \mathbb{P}\left\{|B_{\mathbf{K}}(\mathbf{K}_t, 3)| \leq m/3,\ |\mathbf{Q}_{t+s}| \leq |\mathbf{Q}_t| + b_3/4 \text{ for all } s \in [b_3] \mid p_3 \leq |\mathbf{Q}_t| < q_3\right\}\\
&\leq (m+1)3\left(\frac{72b_3^3 + 24b_3q_3}{m}\right)^{\sqrt{b_3}/4}\\
&\leq 3(m+1)\left(\frac{0.13m}{m}\right)^{1.54\log m}\\
&\leq 2(m+1)\left(\frac{1}{e^3}\right)^{\log m}\\
&= O\left(m^{-2}\right).
\end{aligned}$$

Combining these four bounds, we find that

$$\mathbb{P}\left\{\mathbf{A} \notin \bigcap_{i=0}^{3} \mathcal{A}^i\right\} = O\left(m^{-2}\right).$$

It therefore suffices to show that for all $A \in \bigcap_{i=0}^{3} \mathcal{A}^i$, there exists t such that $|Q_t(A, v)| \geq m/8000\log^2 m$ and $\mathrm{rad}(K_t(A, v)) \leq 111\log m$. For this, we apply Lemma 5.14. For $i \in \{1, 2, 3\}$, one can easily check that $\delta_i \leq 1/2$, $q_i \geq 2p_i$, $b_i \leq \delta_i p_i$, and finally $q_i(3(\log(q_i)/\log(1 + \varepsilon_i/2 - \varepsilon_i\delta_i) - c_i + 1) + 12) \leq m$. Furthermore, since $A \in \mathcal{A}^i$, for all $t \leq (1/\varepsilon_i + 1)q_i + t_i$ with $p_i \leq |Q_t(A, v)| < q_i$, either $|B_{K(A)}(K_t(A, v), 3)| > m/3$ or $|Q_{t+s}(A, v)| \geq |Q_t(A, v)| + \varepsilon_i b_i$ for some $s \in [b_i]$. For $i \in \{1, 2, 3\}$, the only remaining assumption required by Lemma 5.14 is that there exists $s_i \leq t_i$ such that $\mathrm{rad}(K_{s_i}(A, v)) \leq \log(p_i)/\log(1 + \varepsilon_i/2 - \varepsilon_i\delta_i) - c_i$ and $|Q_{s_i}(A, v)| \geq p_i$. If this assumption is established for some $i \in \{1, 2\}$, then applying Lemma 5.14 with $t_0 = s_i$ and $(\varepsilon, \delta, c, b, p, q) = (\varepsilon_i, \delta_i, c_i, b_i, p_i, q_i)$, there exists $s_{i+1} \leq m = t_{i+1}$ such that

$$\begin{aligned}
\mathrm{rad}(K_{s_{i+1}}(A, v)) &\leq \log(q_i)/\log(1 + \varepsilon_i/2 - \varepsilon_i\delta_i) - c_i + 4\\
&= \log(p_{i+1})/\log(1 + \varepsilon_{i+1}/2 - \varepsilon_{i+1}\delta_{i+1}) - c_{i+1}, \text{ and}\\
|Q_{s_{i+1}}(A, v)| &\geq q_i = p_{i+1},
\end{aligned}$$

establishing the assumption for $i + 1$. Now since $A \in \mathcal{A}^0$, there indeed exists $s_1 \leq 75 = t_1$ such that

$$\begin{aligned}
\mathrm{rad}(K_{s_1}(A, v)) &\leq 20 \leq \log(p_1)/\log(1 + \varepsilon_1/2 - \varepsilon_1\delta_1) - c_1, \text{ and}\\
|Q_{s_1}(A, v)| &\geq 11 = p_1.
\end{aligned}$$

Thus for all $A \in \bigcap_{i=0}^{3} \mathcal{A}^i$, there exists s_3 such that

$$\begin{aligned}
\mathrm{rad}(K_{s_3}(A, v)) &\leq \log(q_3)/\log(1 + \varepsilon_3/2 - \varepsilon_3\delta_3) - c_3 + 4 \leq 111\log m, \text{ and}\\
|Q_{s_3}(A, v)| &\geq q_3 \geq m/8000\log^2 m,
\end{aligned}$$

as required. □

5.3. **Connecting the explorations.** In this section we prove Theorem 1.7. The main work was already accomplished in the preceding subsection where we showed that the breadth-first kernel exploration started from a given vertex reaches queue size of order $m/\log^2 m$ within logarithmic diameter with high probability. The only missing ingredient is to establish that if we explore the same random kernel from two different starting vertices, then once the queues in the two explorations are large enough, with high probability the explored subgraphs will be connected by a short path. This is the content of the following lemma.

Lemma 5.15. *Let* $A^* \in \mathcal{A}_{\mathrm{d}}$ *and take* $\mathbf{A} \in_u \mathcal{A}_{\mathrm{d}}$. *For* $u \in [n]$, *let* $((\mathbf{K}_t^u, \mathbf{Q}_t^u), t \geq 0) = ((K_t(\mathbf{A}, u), Q_t(\mathbf{A}, u)), t \geq 0)$ *and* $((K_t^v, Q_t^u), t \geq 0) = ((K_t(A^*, u), Q_t(A^*, u)), t \geq 0)$ *be the breadth-first kernel explorations, started from* u, *of* $\mathbf{A}$ *and* A^* *respectively. Then for any vertices* $u, v \in [n]$ *and for any times* s *and* t *such that* $|Q_t^u|, |Q_t^v| \geq 16$,

$$\mathbb{P}\left\{\mathrm{dist}_{\mathbf{K}}(\mathbf{K}_t^u, \mathbf{K}_s^v) \geq 4 \mid \mathbf{A} \in \mathcal{C}\right\} \leq \left(\frac{45m}{|Q_t^u||Q_s^v|}\right)^2,$$

where $\mathcal{C}$ *is the set of all* $A \in \mathcal{A}_{\mathrm{d}}$ *such that* $(K_r(A,u), Q_r(A,u)) = (K_r^u, Q_r^u)$ *for all* $r \leq t$ *and* $(K_r(A,v), Q_r(A,v)) = (K_r^v, Q_r^v)$ *for all* $r \leq s$.

Proof. Let $\mathcal{A}$ be the set of all $A \in \mathcal{C}$ such that $\mathrm{dist}_{K(A)}(K_t^u, K_s^v) \geq 4$, and let $\mathcal{B}$ be the set of all $A \in \mathcal{C}$ such that there are between 2 and 4 edges joining K_t^u and K_s^v in $K(A)$.

Let $\mathcal{H}$ be the bipartite multigraph with parts $\mathcal{A}$ and $\mathcal{B}$ and an edge between $A \in \mathcal{A}$ and $B \in \mathcal{B}$ for each quadruple $(\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2)$ such that $\vec{e}_1$ and $\vec{e}_2$ have half-edges in Q_t^u, $\vec{f}_1$ and $\vec{f}_2$ have half-edges in Q_s^v, and switching on $(\vec{e}_1, \vec{f}_1)$ in A followed by switching on $(\vec{e}_2, \vec{f}_2)$ in the resulting augmented core A' yields B (in particular, we require that $(\vec{e}_1, \vec{f}_1)$ is valid in A and $(\vec{e}_2, \vec{f}_2)$ is valid in A').

Fix $A \in \mathcal{A}$ and choose oriented edges $\vec{e}_1$ and $\vec{f}_1$ in $K(A)$ with half-edges in Q_t^u and Q_s^v respectively. The number of edges in $K(A)$ with a half-edge in Q_t^u is at least $|Q_t^u|/2$, and similarly for Q_s^v. Thus, there are at least $2(|Q_t^u|/2) \cdot 2(|Q_s^v|/2) = |Q_t^u||Q_s^v|$ choices for the pair $(\vec{e}_1, \vec{f}_1)$. Moreover any such pair $(\vec{e}_1, \vec{f}_1)$ will be valid in A since $\mathrm{dist}_{K(A)}(K_t^u, K_s^v) \geq 4$. Let A' be the augmented core obtained by switching on $(\vec{e}_1, \vec{f}_1)$ in A. Next, choose oriented edges $\vec{e}_2$ and $\vec{f}_2$ in $K(A')$ with half-edges in Q_t^u and Q_s^v respectively not sharing a half-edge with either $\vec{e}_1$ or $\vec{f}_1$. There are at least $2(|Q_t^u|/2 - 1) \cdot 2(|Q_s^v|/2 - 1) = (|Q_t^u| - 2)(|Q_s^v| - 2)$ choices for the pair $(\vec{e}_2, \vec{f}_2)$. We claim that either $(\vec{e}_2, \vec{f}_2)$ or $(\vec{e}_2, \overleftarrow{f}_2)$ is a valid pair in A'. Write e_1, e_2, f_1, f_2 for the underlying edges of $\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2$. If neither $(\vec{e}_2, \vec{f}_2)$ nor $(\vec{e}_2, \overleftarrow{f}_2)$ is valid, then Lemma 5.3 yields an obstruction; by the symmetry between $\vec{e}_2$ and $\vec{f}_2$, we can assume that e_2 has an endpoint adjacent to both endpoints of f_2 in $K(A')$ via distinct non-subdivided edges e and f. If either of e or f was present in $K(A')$, we would have a path of length at most 3 in $K(A)$ between K_t^u and K_s^v. This is impossible because $A \in \mathcal{A}$. Thus, e and f are the two edges in $K(A')$ created by switching on $(\vec{e}_1, \vec{f}_1)$ in A. The only way to switch on a pair of edges and have the created pair of edges share an endpoint is if one of the switched edges was a loop in the original graph. Therefore either e_1 or f_1 is a loop in $K(A)$, contradicting the fact that neither e nor f is subdivided. This proves that there are at least $(|Q_t^u|-2)(|Q_s^v|-2)/2$ valid pairs $(\vec{e}_2, \vec{f}_2)$ of oriented edges in $K(A')$ with half-edges in Q_t^u and Q_s^v respectively not sharing a half-edge with either $\vec{e}_1$ or $\vec{f}_1$. Switching on any such pair in A' yields a graph in $\mathcal{B}$. This proves that A has degree at least $|Q_t^u||Q_s^v|(|Q_t^u| - 2)(|Q_s^v| - 2)/2$ in $\mathcal{H}$.

Now fix $B \in \mathcal{B}$ and consider an edge in $\mathcal{H}$ between B and some $A \in \mathcal{A}$ corresponding to a quadruple $(\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2)$. Let A' be the graph obtained from A by switching on $(\vec{e}_1, \vec{f}_1)$. Let $(\vec{e}'_1, \vec{f}'_1)$ be the switching from A' to A reversing $(\vec{e}_1, \vec{f}_1)$, and let $(\vec{e}'_2, \vec{f}'_2)$ be the switching from B to A' reversing $(\vec{e}_2, \vec{f}_2)$. Then the underlying edges of each of $\vec{e}'_1, \vec{e}'_2, \vec{f}'_1, \vec{f}'_2$ are present in $K(B)$, and at least two of these four edges have one endpoint in K_t^u and the other in K_s^v. Since $B \in \mathcal{B}$, there are at most 4 such edges. It follows that

B has degree at most $\binom{4}{2} \cdot 4 \cdot 3 \cdot m^2 \cdot 2^4$ in $\mathcal{H}$. We conclude that

$$\begin{aligned}\mathbb{P}\{\mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_t, \mathbf{K}^v_s) \geq 4 \mid \mathbf{A} \in \mathcal{C}\} &= \mathbb{P}\{\mathbf{A} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\} \\ &\leq \frac{\binom{4}{2} \cdot 4 \cdot 3 \cdot m^2 \cdot 2^4}{|Q^u_t||Q^v_s|(|Q^u_t|-2)(|Q^v_s|-2)/2}\mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\} \\ &\leq \left(\frac{45m}{|Q^u_t||Q^v_s|}\right)^2,\end{aligned}$$

where the final inequality holds because $|Q^u_t|, |Q^v_s| \geq 16$. □

We are now prepared to prove Theorem 1.7.

Proof of Theorem 1.7. By Observation 5.1, we can take $\mathbf{A} \in_u \mathcal{A}_{\mathrm{d}}$ and bound the diameter of $\mathbf{K} = K(\mathbf{A})$ instead of the diameter of $K(\mathbf{C})$ where $\mathbf{C} \in_u \mathcal{G}^-_{\mathrm{d}}$. For $u \in [n]$, let $((\mathbf{K}^u_t, \mathbf{Q}^u_t), t \geq 0) = ((K_t(\mathbf{A}, u), Q_t(\mathbf{A}, v)), t \geq 0)$ be the breadth-first kernel exploration of $\mathbf{A}$ started from u, and introduce the stopping time

$$\mathbf{t}_u = \inf\{t : |\mathbf{Q}^u_t| \geq m/8000\log^2 m, \mathrm{rad}(\mathbf{K}^u_t) \leq 111\log m\}.$$

By Lemma 5.12, for all $u \in [n]$,

$$\mathbb{P}\{\mathbf{t}_u = \infty\} = O\left(m^{-2}\right). \tag{5.1}$$

By Lemma 5.15, for all $u, v \in [n]$,

$$\begin{aligned}&\mathbb{P}\left\{\mathbf{t}_u < \infty, \mathbf{t}_v < \infty, \mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_{\mathbf{t}_u}, \mathbf{K}^v_{\mathbf{t}_v}) \geq 4\right\} \\ &= \sum_{s,t} \mathbb{P}\{\mathbf{t}_u = s, \mathbf{t}_v = t, \mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_s, \mathbf{K}^v_t) \geq 4\} \\ &= \sum_{s,t} \mathbb{P}\{\mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_s, \mathbf{K}^v_t) \geq 4 \mid \mathbf{t}_u = s, \mathbf{t}_v = t\}\mathbb{P}\{\mathbf{t}_u = s, \mathbf{t}_v = t\} \\ &\leq \sum_{s,t} \left(\frac{45m}{(m/8000\log^2 m)^2}\right)^2 \mathbb{P}\{\mathbf{t}_u = s, \mathbf{t}_v = t\} \\ &= O\left(\frac{\log^8 m}{m^2}\right).\end{aligned} \tag{5.2}$$

Now fix $v \in [n]$, and observe that if $\mathrm{diam}^+(\mathbf{K}) \geq 500\log m$, then either $\mathbf{t}_u = \infty$ for some $u \in [n]$, or, $\mathbf{t}_v < \infty$ and there exists $u \in [n]$ with $\mathbf{t}_u < \infty$ such that $\mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_{\mathbf{t}_u}, \mathbf{K}^v_{\mathbf{t}_v}) \geq 4$. Indeed if $\mathbf{t}_u < \infty$ and $\mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_{\mathbf{t}_u}, \mathbf{K}^v_{\mathbf{t}_v}) \leq 3$ for all $u \in [n]$, then for all $u, w \in [n]$,

$$\begin{aligned}\mathrm{dist}_{\mathbf{K}}(u, w) &\leq \mathrm{dist}_{\mathbf{K}}(u, v) + \mathrm{dist}_{\mathbf{K}}(w, v) \\ &\leq \mathrm{rad}(\mathbf{K}^u_{\mathbf{t}_u}) + \mathrm{rad}(\mathbf{K}^v_{\mathbf{t}_v}) + 3 + \mathrm{rad}(\mathbf{K}^w_{\mathbf{t}_w}) + \mathrm{rad}(\mathbf{K}^v_{\mathbf{t}_v}) + 3 \\ &\leq 4 \cdot 111\log m + 6 \\ &< 500\log m.\end{aligned}$$

Together with the bounds (5.1) and (5.2), this implies

$$\begin{aligned}&\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{K}) \geq 500\log m\right\} \\ &\leq \sum_{u\in[n]} \mathbb{P}\{\mathbf{t}_u = \infty\} + \sum_{u\in[n]} \mathbb{P}\left\{\mathbf{t}_u < \infty, \mathbf{t}_v < \infty, \mathrm{dist}_{\mathbf{K}}(\mathbf{K}^u_{\mathbf{t}_u}, \mathbf{K}^v_{\mathbf{t}_v}) \geq 4\right\} \\ &= O\left(\frac{n}{m^2}\right) + O\left(\frac{n\log^8 m}{m^2}\right) \\ &= O\left(\frac{\log^8 m}{m}\right),\end{aligned}$$

as required. □

6. Combining the bounds, and dealing with degree two vertices

Before reading this section, the reader may wish to reacquaint themselves with the definition of the simple kernel from Section 2, and to recall that $\mathcal{G}_{\mathrm{d}}^-$ is the set of all graphs with degree sequence d and no cycle components.

For a graph G, the *simple homeomorphic reduction* of G is the graph $H = H(G)$ obtained from G by deleting all cycle components and replacing each maximal path all of whose internal vertices lie outside $K^* = K^*(G)$ and have degree 2 by a single edge. Thus, the degree sequence of H restricted to $V(H) \setminus V(K^*)$ is 2-free. Note that every immutable edge $e \in E(K^*) \setminus E^*(K^*)$ in K^* is present in H, and we call e *immutable in* H as well. All other edges in H are *mutable*, and we write $E^*(H) = E(H) \setminus (E(K^*) \setminus E^*(K^*))$ for the set of mutable edges in H. Because $e^*(K^*) \geq e(K^*)/3$, we have $e(H)/3 \leq e^*(H) \leq e(H)$.

For each mutable edge $e^* \in E^*(H)$, let $G(e^*)$ denote the path in G that is replaced by e^* in constructing H, and let $|G(e^*)|$ be the number of internal vertices on $G(e^*)$. Because all leaves in G are present in H, we have $K^*(H) = K^*$. In particular, we can recover the set of mutable edges $E^*(H)$ without making reference to G. Observe that if G is connected and $s(G) \geq 2$, then writing $C = C(G)$, we have $H(C) = K^*$. For a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$, we write $\mathcal{G}_{\mathrm{d}}(H) = \{G \in \mathcal{G}_{\mathrm{d}} : H(G) = H\}$, $\mathcal{G}_{\mathrm{d}}^-(H) = \{G \in \mathcal{G}_{\mathrm{d}}^- : H(G) = H\}$, and $\mathcal{C}_{\mathrm{d}}(H) = \{G \in \mathcal{C}_{\mathrm{d}} : H(G) = H\}$. If H is connected, then $\mathcal{G}_{\mathrm{d}}^-(H) = \mathcal{C}_{\mathrm{d}}(H)$.

Fix a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$ and a graph H which is the simple homeomorphic reduction of some element of $\mathcal{G}_{\mathrm{d}}$. Given $G \in \mathcal{G}_{\mathrm{d}}(H)$, we can encode the data present in G but not H as follows. Write $h = n - v(H)$ and $m = e^*(H)$ and, for convenience, suppose that $V(H) = [n] \setminus [h]$. First, let C be the 2-regular graph given by the union of all cycle components of G. Next, let $P = (P_1, \ldots, P_m) \in \mathcal{P}_{m,h-v(C)}$ be given by $P_i = |G(e_i^*)|$ for $i \in [m]$, where $e_1^*, \ldots, e_m^*$ are the mutable edges of H listed in lexicographic order. Then, for $i \in [m]$, write $u_i^* = (e_i^*)^-$ and $v_i^* = (e_i^*)^+$, and let $\sigma^i = (\sigma_1^i, \ldots, \sigma_{P_i}^i)$ be the sequence of internal vertices on the path $G(e_i^*)$ from u_i^* to v_i^*. Finally, let $\sigma : [h - v(C)] \to [h] \setminus V(C)$ be the bijection where, for all $i \in [m]$, $\sigma(i)$ is equal to the ith entry in the concatenation $(\sigma^1, \ldots, \sigma^m)$.

Let $S_{[h]\setminus V}$ denote the set of all bijections $[h - |V|] \to [h] \setminus V$ and let $\mathcal{R}_V$ denote the set of all 2-regular graphs on V. Then setting

$$\mathcal{X}_{\mathrm{d}}(H) = \bigcup_{V \subseteq [h]} \mathcal{R}_V \times S_{[h]\setminus V} \times \mathcal{P}_{m,h-|V|},$$

we have $(C, \sigma, P) \in \mathcal{X}_{\mathrm{d}}(H)$. This encoding therefore defines a map $M : \mathcal{G}_{\mathrm{d}}(H) \to \mathcal{X}_{\mathrm{d}}(H)$ which we claim is a bijection. To justify this, we describe its inverse. Given $(G, \sigma, P) \in \mathcal{X}_n(H)$, we can recover $\sigma^i = (\sigma_1^i, \ldots, \sigma_{P_i}^i)$ for $i \in [m]$ as $(\sigma(j+1), \ldots, \sigma(j + P_i))$, where $j = |P_1| + \cdots + |P_{i-1}|$. Now replace each mutable edge e_i^* with a path from u_i^* to v_i^* whose sequence of internal vertices is σ^i, then take the union of the resulting graph with C. We obtain the unique graph $G \in \mathcal{G}_{\mathrm{d}}(H)$ which maps to (C, σ, P) under the map $\mathcal{G}_{\mathrm{d}}(H) \to \mathcal{X}_{\mathrm{d}}(H)$.

If G and (C, σ, P) correspond under M, then the number of vertices in cycle components of G is equal to $v(C)$ and $|G(e_i^*)| = P_i$ for all $i \in [m]$. In particular, G has no cycle components if and only if $v(C) = 0$. Defining $\mathcal{X}_{\mathrm{d}}^-(H) = S_{[h]} \times \mathcal{P}_{m,h}$, the map $M : \mathcal{G}_{\mathrm{d}}(H) \to \mathcal{X}_{\mathrm{d}}(H)$ therefore restricts to a bijection $M^- : \mathcal{G}_{\mathrm{d}}^-(H) \to \mathcal{X}_{\mathrm{d}}^-(H)$. Writing $\mathrm{cyc}(G)$ for the number of vertices in cycle components of G, this establishes the following distributional identities.

Proposition 6.1. *Fix a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$*, and let* H *be the simple homeomorphic reduction of a graph in* $\mathcal{G}_{\mathrm{d}}$*. Write* $m = e^*(H)$*, and list the mutable edges of* H *in lexicographic order as* $e_1^*, \ldots, e_m^*$*. Let* $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}(H)$*,* $\mathbf{G}^- \in_u \mathcal{G}_{\mathrm{d}}^-(H)$*,* $(\mathbf{C}, \boldsymbol{\sigma}, \mathbf{P}) \in_u \mathcal{X}_{\mathrm{d}}(H)$*,* *and* $(\boldsymbol{\sigma}^-, \mathbf{P}^-) \in_u \mathcal{X}^-_{\mathrm{d}}(H)$*. Then*

$$(\mathrm{cyc}(\mathbf{G}), |\mathbf{G}(e_1^*)|, \ldots, |\mathbf{G}(e_m^*)|) \stackrel{d}{=} (v(\mathbf{C}), \mathbf{P}_1, \ldots, \mathbf{P}_m), \text{ and}$$
$$(|\mathbf{G}^-(e_1^*)|, \ldots, |\mathbf{G}^-(e_m^*)|) \stackrel{d}{=} (\mathbf{P}_1^-, \ldots, \mathbf{P}_m^-).$$

This proposition will allow us to prove the following lemma controlling the tail of $\mathrm{cyc}(\mathbf{G})$.

Lemma 6.2. *Fix* $\varepsilon \in (0,1)$ *and a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$*. Let* H *be the homeomorphic reduction of a graph in* $\mathcal{G}_{\mathrm{d}}$*, let* $h = n - v(H)$*, and suppose that* $h < (1-\varepsilon)n$*. Let* $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}(H)$*. Then there is a constant* C *depending only on* ε *such that for all* $x \geq 0$*,*

$$\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) \geq x\} < \exp\left(-\frac{\varepsilon x}{10} + C\right).$$

Proof. Let c_k be the number of 2-regular graphs with vertex set $[k]$. It is shown in [18] that $c_k = (1+o(1))e^{-3/4}\pi^{-1/2}k!/\sqrt{k}$ as $k \to \infty$, and it follows immediately that

$$\frac{c_{k+1}}{c_k} = (1+o(1))(k+1). \tag{6.1}$$

Recalling that $m \geq e(H)/3$, we have

$$6m \geq 2e(H) = \sum_{v \in V(H)} d_v \geq v(H) = n - h \geq \frac{h}{1-\varepsilon} - h \geq \varepsilon h. \tag{6.2}$$

Now by Proposition 6.1,

$$\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k\} \propto \left| \bigcup_{V \subseteq [h], |V| = k} \mathcal{R}_V \times S_{[h]\setminus V} \times \mathcal{P}_{m,h-k} \right| = \binom{h}{k} c_k (h-k)! \binom{m-1+h-k}{m-1}.$$

Choose $k_0 = k_0(\varepsilon)$ such that for all $k \geq k_0$ it holds that $(k-3)/((1+\epsilon/6)k-4) < (1+\epsilon/50)/(1+\epsilon/6)$ and $c_{k+1}/c_k \leq (1+\epsilon/50)(k+1)$; the second is possible by (6.1). Then for all $h \geq k \geq k_0$, we have

$$\begin{aligned}
\frac{\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k+1\}}{\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k\}} &= \frac{h-k}{k+1} \cdot \frac{c_{k+1}}{c_k} \cdot \frac{1}{h-k} \cdot \frac{h-k}{m-1+h-k} \\
&= \frac{c_{k+1}/c_k}{k+1} \cdot \frac{h-k}{m-1+h-k} \\
&\leq \frac{c_{k+1}/c_k}{k+1} \cdot \frac{h-3}{m+h-4} \\
&\leq \frac{c_{k+1}/c_k}{k+1} \cdot \frac{h-3}{(1+\varepsilon/6)h-4} \\
&= \frac{(1+\epsilon/50)^2}{1+\varepsilon/6},
\end{aligned}$$

where the second inequality comes from (6.2). Since $\varepsilon \in (0,1)$, we have $(1+\epsilon/50)^2/(1+\varepsilon/6) < 1-\varepsilon/10 < \exp(-\varepsilon/10)$. Hence for all $k \geq k_0$,

$$\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k+1\} < \exp\left(-\frac{\varepsilon}{10}\right) \mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k\}.$$

Applying this inequality repeatedly, we obtain $\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k\} \leq \exp(-\varepsilon(k-k_0)/10)$ for all $k \geq k_0$. This bound also holds for $k < k_0$, so for all $x \geq 0$,

$$\begin{aligned}\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) \geq x\} &= \sum_{k\geq x} \mathbb{P}\{\mathrm{cyc}(\mathbf{G}) = k\} \\ &< \sum_{k\geq x} \exp(-\varepsilon(k-k_0)/10) \\ &< \frac{\exp(\varepsilon k_0/10)}{1-\exp(-\varepsilon/10)} \exp\left(-\frac{\varepsilon x}{10}\right).\end{aligned}$$

As the coefficient depends only on ε, this completes the proof. □

Our next lemma bounds the increase in diameter upon reintroducing suppressed degree 2 vertices. The "stretch factor" is roughly what one would expect if all edges were subdivided an equal number of times. For a random graph $\mathbf{G}$ as in Theorem 1.1 or as in Theorem 1.2, we will use this lemma to bound the diameter of $C(\mathbf{G})$ as a function of $K^*(\mathbf{G})$, and also to bound the diameter of $\mathbf{G}$ as a function of $H(\mathbf{G})$.

Lemma 6.3. *Fix a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$*. Let* H *be the simple homeomorphic reduction of a graph in* $\mathcal{G}_{\mathrm{d}}$ *and write* $h = n - v(H)$*. Suppose that* $m = e^*(H) \geq 2$*. Then with* $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}^-(H)$*, for all* $x \geq 16$*,*

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{G}) \geq 2\left(1+\frac{h}{m-1}\right)(\mathrm{diam}(H)+2+x\log m)\right\} \leq \exp\left(-\frac{x\log m}{8}+2\right).$$

Proof. For an edge $e \in E(H)$, let $\mathrm{len}_{\mathbf{G}}(e)$ denote the number of edges in the corresponding path in $\mathbf{G}$, so $\mathrm{len}_{\mathbf{G}}(e) = |\mathbf{G}(e)| + 1$ if $e \in E^*(H)$ and otherwise $\mathrm{len}(e) = 1$. For a set of edges $E \subseteq E(H)$, define $\mathrm{len}_{\mathbf{G}}(E) = \sum_{e\in E} \mathrm{len}_{\mathbf{G}}(e)$. Now fix any set of mutable edges $E^* \subseteq E^*(H)$, and let $k = |E^*|$. Let $\mathbf{P} \in_u \mathcal{P}_{m,h}$, so by Proposition 6.1, and since $\mathbf{P} = (\mathbf{P}_1, \ldots, \mathbf{P}_m)$ is an exchangeable sequence,

$$\mathrm{len}_{\mathbf{G}}(E^*) \stackrel{d}{=} \sum_{i=1}^{k}(\mathbf{P}_i + 1).$$

It therefore follows (for example, from the stars and bars encoding of $\mathcal{P}_{m,h}$) that $\mathrm{len}_{\mathbf{G}}(E^*)$ has the same distribution as that of the random variable $\mathbf{X}_k$ described by the following experiment. Consider an urn with $m-1$ white balls and h black balls. Repeatedly sample balls from the urn without replacement until none remain. Now, for $i \in [m-1+h]$, let $\mathbf{W}_i$ be the number of white balls observed after i balls in total have been drawn. Also set $\mathbf{W}_0 = 0$ and $\mathbf{W}_{m+s} = m$. Then let $\mathbf{X}_k = \min(i \in [0, m+h] : \mathbf{W}_i = k)$.

It follows from [4, Proposition 20.6] that, letting $\mathbf{W}_i'$ be a Binomial$(i, (m-1)/(m-1+h))$ random variable, then $\mathbb{E}[\phi(\mathbf{W}_i)] \leq \mathbb{E}[\phi(\mathbf{W}_i')]$ for any convex function $\phi : \mathbb{R} \to [0, \infty]$. This implies that tail bounds for $\mathbf{W}_i'$ that are proved by bounding exponential moments $\mathbb{E}[\exp(\lambda \mathbf{W}_i')]$ (in particular Chernoff bounds) can also be applied to $\mathbf{W}_i$. Now fix $y \geq 2$ and let $z = yk(1 + h/(m-1))$. Then $\mathbb{E}\left[\mathbf{W}_{\lceil z\rceil}\right] \geq yk \geq 2k$ and thus a Chernoff bound with $\lambda = 1/2$ gives

$$\begin{aligned}\mathbb{P}\left\{\mathbf{X}_k \geq yk\left(1+\frac{h}{m-1}\right)\right\} &\leq \mathbb{P}\left\{\mathbf{W}_{\lceil z\rceil} \leq k\right\} \\ &\leq \mathbb{P}\left\{\mathbf{W}_{\lceil z\rceil} \leq \frac{1}{2}\mathbb{E}\left[\mathbf{W}_{\lceil z\rceil}\right]\right\} \\ &\leq \exp\left(-\frac{yk}{8}\right). \qquad (6.3)\end{aligned}$$

It follows that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E^*) \geq 2\left(1+\frac{h}{m-1}\right)(|E^*|+x\log m)\right\} \leq \exp\left(-\frac{x\log m}{4}\right). \tag{6.4}$$

Indeed if $k = |E^*| \leq x\log m$, then taking $y = 2x\log m/k$ in (6.3), we obtain

$$\begin{aligned}\mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E^*) \geq 2x\log m\left(1+\frac{h}{m-1}\right)\right\} &= \mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E^*) \geq yk\left(1+\frac{h}{m-1}\right)\right\} \\ &\leq \exp\left(-\frac{yk}{8}\right) \\ &= \exp\left(-\frac{x\log m}{4}\right).\end{aligned}$$

If $|E^*| > x\log m$, then taking $y = 2$ in (6.3), we obtain

$$\mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E^*) \geq 2|E^*|\left(1+\frac{h}{m-1}\right)\right\} \leq \exp\left(-\frac{|E^*|}{4}\right) \leq \exp\left(-\frac{x\log m}{4}\right).$$

Now for every pair of edges $e, f \in E(H)$ in the same component of H, let E_{ef} be the union of e, f, and the edge set of a shortest path in H between e^- and f^-. Then $|E_{ef}| \leq \mathrm{diam}(H)+2$. Write $E^*_{ef} = E_{ef} \cap E^*(H)$ for the set of mutable edges in E_{ef}. Now if

$$\mathrm{len}_{\mathbf{G}}(E^*_{ef}) < 2\left(1+\frac{h}{m-1}\right)(|E^*_{ef}|+x\log m),$$

then

$$\begin{aligned}\mathrm{len}_{\mathbf{G}}(E_{ef}) &= |E_{ef}| - |E^*_{ef}| + \mathrm{len}_{\mathbf{G}}(E^*_{ef}) \\ &< |E_{ef}| - |E^*_{ef}| + 2\left(1+\frac{s}{m-1}\right)(|E^*_{ef}|+x\log m) \\ &\leq 2\left(1+\frac{h}{m-1}\right)(|E_{ef}|+x\log m) \\ &\leq 2\left(1+\frac{h}{m-1}\right)(\mathrm{diam}(H)+2+x\log m).\end{aligned}$$

Because $\mathbf{G}$ has no cycle components, every vertex $v \in V(\mathbf{G})$ lies on $\mathbf{G}(e)$ for some edge $e \in E(H)$. Thus every pair of vertices $u, v \in V(\mathbf{G})$ in the same component of $\mathbf{G}$ have a path between them of length at most $\mathrm{len}_{\mathbf{G}}(E_{ef})$ for some $e, f \in E(H)$. By a union bound and (6.4), for all $x \geq 16$,

$$\begin{aligned}
&\mathbb{P}\left\{\mathrm{diam}(\mathbf{G}) \geq 2\left(1+\frac{h}{m-1}\right)(\mathrm{diam}(H)+2+x\log m)\right\} \\
&\leq \sum_{e,f\in E(H)} \mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E_{ef}) \geq 2\left(1+\frac{h}{m-1}\right)(\mathrm{diam}(H)+2+x\log m)\right\} \\
&\leq \sum_{e,f\in E(H)} \mathbb{P}\left\{\mathrm{len}_{\mathbf{G}}(E^*_{ef}) \geq 2\left(1+\frac{h}{m-1}\right)(|E^*_{ef}|+x\log m)\right\} \\
&\leq \binom{e(H)}{2}\exp\left(-\frac{x\log m}{4}\right) \\
&\leq \frac{(3m)^2}{2}\exp\left(-\frac{x\log m}{4}\right) \\
&= \exp\left(-\frac{x\log m}{4}+2\log m+\log(4.5)\right) \\
&\leq \exp\left(-\frac{x\log m}{8}+2\right),
\end{aligned}$$

as required. □

Recall that $\mathcal{C}_{\mathrm{d}}$ is the set of connected graphs with degree sequence d. Lemma 6.4 is our first application of Lemma 6.3, where we bound the diameter of the core of a random graph as a function of its simple kernel. Note that if $G \in \mathcal{C}_{\mathrm{d}}(K^*)$ has surplus at least 2, then $H(C(G)) = K^*(G)$; this is what allows us to invoke Lemma 6.3.

Lemma 6.4. *Fix a degree sequence* $\mathrm{d} = (d_1, \ldots, d_n)$ *and let* K^* *be the (possibly empty) simple kernel of a graph in* $\mathcal{C}_{\mathrm{d}}$. *Suppose that* $\mathrm{d}|_{[n]\setminus V(K^*)}$ *is 2-free, and let* $m = e^*(K^*)$. *Let* $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}$ *and let* $\mathbf{C} = C(\mathbf{G})$. *Then for all* $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{C}) \geq 5\sqrt{mn} + x\sqrt{n}\right\} \leq \exp\left(-\frac{x^2}{3}+1\right).$$

Moreover, if $m \geq 2$, *then for all* $x \geq 24$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{C}) \geq \left(2+x\sqrt{n/m}\right)(\mathrm{diam}(K^*)+x^2\log m)\right\} \leq \exp\left(-\frac{x^2\log m}{2^9}+3\right),$$

Proof. First suppose that $m = 0$, so K^* is empty. Then $\mathbf{G}$ is either a cycle or a tree. But $\mathbf{G}$ cannot be a cycle because K^* being empty implies that $\mathrm{d} = \mathrm{d}|_{[n]\setminus V(K^*)}$ is 2-free. Thus $\mathbf{G}$ is a tree, so $\mathbf{C}$ is empty and both bounds are trivial.

Suppose, for the remainder of the proof, that $m \geq 1$. Letting $\mathbf{h} = v(\mathbf{C}) - v(K^*)$, we first claim that for all $y \geq 0$,

$$\mathbb{P}\left\{\mathbf{h} \geq 4\sqrt{mn} + y\sqrt{n}\right\} \leq \exp\left(-\frac{y^2}{3}+1\right). \tag{6.5}$$

To see this, let $(\mathbf{F},\mathbf{T},\mathbf{P}) \in_u \mathcal{X}_{\mathrm{d}}(K^*)$ so $\mathbf{h} \overset{\mathrm{d}}{=} \mathrm{ht}_{\mathbf{T}}(0)$ by Proposition 2.1. It suffices to show that for any $B \subseteq [0,n] \setminus V(K^*)$ with $\mathbb{P}\{V(\mathbf{T}) = B\} > 0$, the bound (6.5) holds of $\mathrm{ht}_{\mathbf{T}}(0)$ conditionally given that $V(\mathbf{T}) = B$. Fix any such set B, let $n' = |B| - 1$, and let $\mathrm{c} = (c_v, v \in B)$ be the restriction to B of the child sequence given by (2.1). Then $c_0 = 0$ and since $\mathrm{d}|_{[n]\setminus V(K^*)}$ is 2-free, c is 1-free. Moreover, we have $(\mathbf{T} \mid V(\mathbf{T}) = B) \in_u \{(T,P) : T \in \mathcal{T}_{\mathrm{c}}, P \in \mathcal{P}_{m,\mathrm{ht}_T(0)}\}$. Because c is 1-free, any tree $T \in \mathcal{T}_{\mathrm{c}}$ has $\mathrm{ht}_T(0) \leq n'/2$. Now if $n' < 64m$ then $4\sqrt{mn'} > n'/2$ and hence $\mathbb{P}\left\{\mathrm{ht}_{\mathbf{T}}(0) \geq 4\sqrt{mn'} + y\sqrt{n'} \mid V(\mathbf{T}) = B\right\} = 0$

for all $y \geq 0$. On the other hand if $n' \geq 64m$, then by Theorem 1.6, for all $y \geq 1.25$,

$$\begin{aligned}
&\mathbb{P}\left\{\mathrm{ht}_{\mathbf{T}}(0) \geq 4\sqrt{mn} + y\sqrt{n} \mid V(\mathbf{T}) = B\right\} \\
&\leq \mathbb{P}\left\{\mathrm{ht}_{\mathbf{T}}(0) \geq 4\sqrt{mn'} + y\sqrt{n'} \mid V(\mathbf{T}) = B\right\} \\
&= \mathbb{P}\left\{\mathrm{ht}_{\mathbf{T}}(0) \geq 2\sqrt{mn'} + \left(y + 2\sqrt{m}\right)\sqrt{n'} \mid V(\mathbf{T}) = B\right\} \\
&\leq \exp\left(-\frac{(y+2\sqrt{m})^2}{3} + 4\right) \\
&\leq \exp\left(-\frac{y^2}{3} + 1\right).
\end{aligned}$$

For $0 \leq y < 1.25$, this bound is greater than 1, and hence also holds. This proves (6.5).

For the first bound, we can assume that $m \leq n/9$, since otherwise $5\sqrt{mn} > n$ and $\mathbb{P}\{\mathrm{diam}(\mathbf{C}) \geq 5\sqrt{mn} + x\sqrt{n}\} = 0$ for all $x \geq 0$. Because K^* has minimum degree at least 2, we have $v(K^*) \leq e(K^*) \leq 3m$. Thus, for $x \geq 0$, if $\mathbf{h} < 4\sqrt{mn} + x\sqrt{n}$, then

$$\begin{aligned}
\mathrm{diam}(\mathbf{C}) &\leq \mathbf{h} + v(K^*) \\
&< 4\sqrt{mn} + x\sqrt{n} + 3m \\
&\leq 4\sqrt{mn} + x\sqrt{n} + 3\sqrt{mn/9} \\
&= 5\sqrt{mn} + x\sqrt{n}.
\end{aligned}$$

Applying (6.5), we therefore have

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{C}) \geq 5\sqrt{mn} + x\sqrt{n}\right\} \leq \mathbb{P}\left\{\mathbf{h} \geq 4\sqrt{mn} + x\sqrt{n}\right\} \leq \exp\left(-\frac{y^2}{3} + 1\right).$$

This completes the proof of the first bound.

For the proof of the second bound, suppose that $m \geq 2$. For any $V \subseteq [n]$ with $\mathbb{P}\{V(\mathbf{C}) = V\} > 0$, conditionally given that $V(\mathbf{C}) = V$, we have $\mathbf{C} \in_u \mathcal{C}_{\mathrm{d}'}(K^*)$, where $\mathrm{d}' = (d'_v, v \in V)$ is given by

$$d'_v = \begin{cases} \deg_{K^*}(v) & v \in V(K^*) \\ 2 & v \in V \setminus V(K^*) \end{cases}.$$

Now because $m \geq 2$, $s(\mathbf{C}) = s(K^*) \geq 2$, and hence $s(\mathrm{d}') \geq 2$. Since K^* is connected, this implies that

$$\mathcal{C}_{\mathrm{d}'}(K^*) = \mathcal{G}^-_{\mathrm{d}'}(K^*) = \{G \in \mathcal{G}^-_{\mathrm{d}'} : H(G) = K^*\}.$$

Thus, given that $V(\mathbf{C}) = V$, $\mathbf{C}$ is a uniformly random graph with degreee sequence d' and simple homeomorphic reduction K^*, which has $m \geq 2$ mutable edges. Because V was arbitrary, it then follows from Lemma 6.3 that for all $y \geq 4$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{C}) \geq 2\left(1 + \frac{\mathbf{h}}{m-1}\right)\left(\mathrm{diam}(K^*) + 2 + y^2 \log m\right)\right\} \leq \exp\left(-\frac{y^2 \log m}{8} + 2\right). \tag{6.6}$$

Now if $y \geq 4$ and $\mathbf{h} < 4\sqrt{mn} + y\sqrt{(3/8)n\log m}$, then

$$\begin{aligned}
2\left(1 + \frac{\mathbf{h}}{m-1}\right) &\leq 2 + \frac{4\mathbf{h}}{m} \\
&< 2 + 16\sqrt{n/m} + 4\sqrt{(3/8)\log m/m} \cdot y\sqrt{n/m} \\
&\leq 2 + 6y\sqrt{n/m}
\end{aligned}$$

Also, $2 + y^2 \log m \le (6y)^2 \log m$. Therefore by (6.5) and (6.6), for all $y \ge 4$,

$$\begin{aligned}
&\mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \ge \left(2 + 6y\sqrt{n/m}\right)\left(\operatorname{diam}(K^*) + (6y)^2 \log m\right)\right\} \\
&\le \mathbb{P}\left\{\mathbf{h} \ge 4\sqrt{mn} + y\sqrt{(3/8)n \log m}\right\} \\
&+ \mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \ge 2\left(1 + \frac{\mathbf{h}}{m-1}\right)\left(\operatorname{diam}(K^*) + 2 + y^2 \log m\right)\right\} \\
&\le \exp\left(-\frac{y^2 \log m}{8} + 1\right) + \exp\left(-\frac{y^2 \log m}{8} + 2\right) \\
&\le \exp\left(-\frac{y^2 \log m}{8} + 3\right)
\end{aligned}$$

Given $x \ge 24$, setting $y = x/6$, we find that

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \ge \left(2 + x\sqrt{n/m}\right)\left(\operatorname{diam}(K^*) + x^2 \log m\right)\right\} \le \exp\left(-\frac{x^2 \log m}{2^9} + 3\right)$$

This completes the proof of the second bound. $\square$

In Lemma 6.5, below, we prove tail bounds on the diameter of a random graph with a given simple kernel K^*; it is a consequence of Lemma 6.4, above, and Theorem 1.4. We obtain two bounds, the first of which holds with essentially no assumptions on K^*, and the second of which holds provided that the simple kernel has logarithmic diameter. The specific value $1000 \log m + 700$ arises for the following reason. Suppose $m = e^*(K^*) \ge 2$, and let $m' = e(K)$ where $K = K^*(K)$ so $m'/2 \le m \le m'$. Then by Theorem 1.7, typically $\operatorname{diam}(K) < 500 \log m'$. But if $\operatorname{diam}(K) < 500 \log m'$, then since $\operatorname{diam}(K^*) \le 2\operatorname{diam}(K)+2$, we have

$$\operatorname{diam}(K^*) < 2 \cdot 500 \log m' + 2 \le 1000 \log(2m) + 2 \le 1000 \log m + 700. \tag{6.7}$$

It is therefore useful to have improved bounds under the assumption that $\operatorname{diam}(K^*) \le 1000 \log m + 700$.

Lemma 6.5. *Fix a degree sequence* $\mathrm{d} = (d_1, \dots, d_n)$*, let* K^* *be the (possibly empty) simple kernel of a graph in* $\mathcal{C}_{\mathrm{d}}$*, and suppose that* $\mathrm{d}|_{[n]\setminus V(K^*)}$ *is 2-free. Let* $m = e^*(K^*)$ *and let* $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}(K^*)$*. Then for all* $x \ge 0$

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \ge 5\sqrt{mn} + x\sqrt{n}\right\} \le \exp\left(-\frac{x^2}{2^{11}} + 3\right).$$

Moreover, if $m \ge 2$ *and* $\operatorname{diam}(K^*) \le 1000 \log m + 700$*, then for all* $x \ge 0$*,*

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \ge x\sqrt{n}\right\} \le \exp\left(-\frac{x^{2/3}}{2^{12}} + 4\right).$$

Proof. Let $\mathbf{F} = F(\mathbf{G})$. Then for any set $V \subseteq [n]$ with $\mathbb{P}\{V(\mathbf{F}) = V\} > 0$, conditionally given that $V(\mathbf{G}) = V$, we have $\mathbf{F} \in_u \mathcal{F}_{\mathrm{c}}$, where $\mathrm{c} = (d_v - 1, v \in V)$. Now since $V(\mathbf{F}) \cap V(\mathbf{C}) = \emptyset$, we have $V \cap V(K^*) = \emptyset$. Then c is 1-free because $\mathrm{d}|_{[n]\setminus V(K^*)}$ is 2-free. By Theorem 1.4, for all $y \ge 0$,

$$\mathbb{P}\left\{\operatorname{ht}(\mathbf{F}) \ge y\sqrt{|V|} \mid V(\mathbf{F}) = V\right\} \le \exp\left(-\frac{y^2}{2^8} + 2\right).$$

As $|V| \le n$ and V was arbitrary, it follows that for all $y \ge 0$,

$$\mathbb{P}\left\{\operatorname{ht}(\mathbf{F}) \ge y\sqrt{n}\right\} \le \exp\left(-\frac{y^2}{2^8} + 2\right). \tag{6.8}$$

By Lemma 6.4, for all $y \geq 0$,

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \geq 5\sqrt{mn} + y\sqrt{n}\right\} \leq \exp\left(-\frac{y^2}{3} + 1\right). \tag{6.9}$$

Now for $y \geq 2$, if $\operatorname{ht}(\mathbf{F}) < 10y\sqrt{n}$ and $\operatorname{diam}(\mathbf{C}) < 5\sqrt{mn} + y\sqrt{n}$, then

$$\begin{aligned}
\operatorname{diam}(\mathbf{G}) &\leq 2\operatorname{ht}(\mathbf{F}) + 2 + \operatorname{diam}(\mathbf{C}) \\
&< 20y\sqrt{n} + 2 + 5\sqrt{mn} + y\sqrt{n} \\
&\leq 5\sqrt{mn} + 22y\sqrt{n}
\end{aligned}$$

Thus for all $x \geq 44$, setting $y = x/22$, by (6.8) and (6.9),

$$\begin{aligned}
\mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \geq 5\sqrt{mn} + x\sqrt{n}\right\} &\leq \mathbb{P}\left\{\operatorname{ht}(\mathbf{F}) \geq 9y\sqrt{n}\right\} + \mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \geq 5\sqrt{mn} + y\sqrt{n}\right\} \\
&\leq \exp\left(-\frac{(10y)^2}{2^8} + 2\right) + \exp\left(-\frac{y^2}{3} + 1\right) \\
&\leq \exp\left(-\frac{y^2}{3} + 3\right) \\
&\leq \exp\left(-\frac{x^2}{2^{11}} + 3\right),
\end{aligned}$$

When $0 \leq x < 44$, this bound is greater than 1, and hence also holds. This completes the proof of the first bound.

For the proof of the second bound, suppose that $m \geq 2$ and $\operatorname{diam}(K^*) \leq 1000\log m + 700$. Then by Lemma 6.4, for all $y \geq 24$,

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{C}) \geq \left(2 + y\sqrt{n/m}\right)\left(\operatorname{diam}(K^*) + y^2\log m\right)\right\} \leq \exp\left(-\frac{y^2\log m}{2^9} + 3\right). \tag{6.10}$$

Fix $z \geq 24^3$, and let $y = z^{1/3} \geq 24$. Since K^* is a simple graph on at most n vertices, $m \leq e(K^*) \leq n^2/2$. Therefore if $\operatorname{ht}(\mathbf{F}) < z\sqrt{n}$ and $\operatorname{diam}(\mathbf{C}) < \left(2 + y\sqrt{n/m}\right)(\operatorname{diam}(K^*) + y^2\log m)$, then

$$\begin{aligned}
\operatorname{diam}(\mathbf{G}) &\leq 2\operatorname{ht}(\mathbf{F}) + 2 + \operatorname{diam}(\mathbf{C}) \\
&< 2z\sqrt{n} + 2 + \left(2 + y\sqrt{n/m}\right)(1000\log m + 700 + y^2\log m) \\
&\leq 2z\sqrt{n} + 2 + 5y^2\log m\left(2 + y\sqrt{n/m}\right) \\
&\leq 2z\sqrt{n} + 2 + 20y^2\log n - 10y^2\log 2 + 5y^3\log m\sqrt{n/m} \\
&\leq 2z\sqrt{n} + y^3\log n + 4y^3\sqrt{n} \\
&\leq 2z\sqrt{n} + y^3\sqrt{n} + 4y^3\sqrt{n} \\
&\leq 7z\sqrt{n}.
\end{aligned}$$

Hence, for all $x \geq 7 \cdot 24^3$, letting $z = x/7$ and $y = z^{1/3}$, we find that

$$\begin{aligned}
\mathbb{P}\left\{\mathrm{diam}(\mathbf{G}) \geq x\sqrt{n}\right\} &= \mathbb{P}\left\{\mathrm{diam}(\mathbf{G}) \geq 7z\sqrt{n}\right\} \\
&\leq \mathbb{P}\left\{\mathrm{ht}(\mathbf{F}) \geq z\sqrt{n}\right\} \\
&+ \mathbb{P}\left\{\mathrm{diam}(\mathbf{C}) \geq \left(2 + y\sqrt{n/m}\right)\left(\mathrm{diam}(K^*) + y^2 \log m\right)\right\} \\
&\leq \exp\left(-\frac{z^2}{2^8} + 2\right) + \exp\left(-\frac{y^2 \log m}{2^9} + 3\right) \\
&\leq \exp\left(-\frac{z^2}{2^8} + 2\right) + \exp\left(-\frac{z^{2/3} \log m}{2^9} + 3\right) \\
&\leq \exp\left(-\frac{x^2}{2^{14}} + 2\right) + \exp\left(-\frac{x^{2/3}}{2^{12}} + 3\right) \\
&\leq \exp\left(-\frac{x^{2/3}}{2^{12}} + 4\right)
\end{aligned}$$

When $0 \leq x < 7 \cdot 24^3$, this bound is greater than 1 and therefore also holds. This completes the proof of the second bound. □

We are now prepared to prove Theorems 1.1 and 1.2

Proof of Theorem 1.1. Fix $\varepsilon \in (0,1)$ and fix a degree sequence $\mathrm{d} = (d_1, \ldots, d_n)$ such that $n_2(\mathrm{d}) < (1-\varepsilon)n$, and let $\mathbf{G} \in_u \mathcal{C}_{\mathrm{d}}$.

Let $\mathbf{H} = H(\mathbf{G})$ be the simple homeomorphic reduction of $\mathbf{G}$, and let $\mathbf{K}^* = K^*(\mathbf{G}) = K^*(\mathbf{H})$. We proceed in two steps, first showing that $\mathbb{E}[\mathrm{diam}(\mathbf{H})] = O(\sqrt{n})$, and then showing that $\mathbb{E}[\mathrm{diam}(\mathbf{G}) \mid \mathbf{H}] = O((\mathrm{diam}(\mathbf{H}) + \log n)/\varepsilon)$. To start, we bound the expected diameter of $\mathbf{H}$, given its vertex set and given $\mathbf{K}^*$, as a function of $\mathbf{K}^*$. So, fix a set $V \subseteq [n]$ and a simple kernel K^* with $\mathbb{P}\{V(\mathbf{H}) = V, \mathbf{K}^* = K^*\} > 0$. Then $\mathrm{d}|_{V \setminus V(K^*)}$ is 2-free because by definition of the simple homeomorphic reduction, the only vertices in $\mathbf{H}$ of degree 2 lie in $\mathbf{K}^*$. Given that $V(\mathbf{H}) = V$ and that $\mathbf{K}^* = K^*$, we have $\mathbf{H} \in_u \mathcal{C}_{\mathrm{d}|_V}(K^*)$. Let $m = e^*(K^*)$. Noting that $v(\mathbf{H}) \leq n$, it follows from Lemma 6.5 that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}) \geq 5\sqrt{mn} + x\sqrt{n} \mid \mathbf{K}^* = K^*\right\} \leq \exp\left(-\frac{x^2}{2^{11}} + 3\right). \tag{6.11}$$

Again because $v(\mathbf{H}) \leq n$ and by Lemma 6.5, if $m \geq 2$ and $\mathrm{diam}(K^*) \leq 1000 \log m + 700$, then for all $x \geq 0$

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}) \geq x\sqrt{n} \mid \mathbf{K}^* = K^*\right\} \leq \exp\left(-\frac{x^{2/3}}{2^{12}} + 4\right). \tag{6.12}$$

First suppose that $s(\mathrm{d}) \leq 1$; then $m \leq 1$, and hence by (6.11),

$$\begin{aligned}
\mathbb{E}[\mathrm{diam}(\mathbf{H}) \mid \mathbf{K}^* = K^*] &= \int_0^\infty \mathbb{P}\{\mathrm{diam}(\mathbf{H}) > x \mid \mathbf{K}^* = K^*\}\,\mathrm{d}x \\
&\leq \sqrt{n}\int_0^\infty \mathbb{P}\left\{\mathrm{diam}(\mathbf{H}) \geq 5\sqrt{mn} + (x-5)\sqrt{n} \mid \mathbf{K}^* = K^*\right\}\mathrm{d}x \\
&\leq \sqrt{n}\int_0^\infty \min\left(1, \exp\left(-\frac{(x-5)^2}{2^{11}} + 3\right)\right)\mathrm{d}x \\
&= O\left(\sqrt{n}\right).
\end{aligned}$$

Because this holds for all V and K^*, upon averaging over $\mathbf{K}^*$, we obtain $\mathbb{E}[\mathrm{diam}(\mathbf{H})] \leq O(\sqrt{n})$ provided that $s(\mathrm{d}) \leq 1$.

Now suppose that $s(\mathrm{d}) \geq 2$. Repeating almost the exact same calculation as above, again using (6.11), yields

$$\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})\geq 5\sqrt{mn}]} \mid \mathbf{K}^* = K^*\right] = O\left(\sqrt{n}\right)$$

Let $K = K(K^*)$ and write $m' = e(K)$. Then $m, m' \geq 2$ and $m'/2 \leq m \leq m'$. Also let $\mathbf{K} = K(\mathbf{K}^*)$. Then by the preceding bound,

$$\begin{aligned}
&\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})\geq 5\sqrt{m'n}]} \mid e(\mathbf{K}) = m'\right] \\
&\leq \mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})\geq 5\sqrt{e^*(\mathbf{K}^*)n}]} \mid e(\mathbf{K}) = m'\right] \\
&= O\left(\sqrt{n}\right)
\end{aligned} \tag{6.13}$$

Without any work, we also have

$$\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})< 5\sqrt{m'n}]} \mid \mathbf{K} = K\right] < 5\sqrt{m'n}, \tag{6.14}$$

Finally, if $\operatorname{diam}(K^*) \leq 1000\log m + 700$, then a similar calculation but instead applying (6.12) yields

$$\mathbb{E}\left[\operatorname{diam}(\mathbf{H}) \mid \mathbf{K}^* = K^*\right] = O\left(\sqrt{n}\right).$$

Now if $\operatorname{diam}(K) < 500\log m'$, then as in (6.7), $\operatorname{diam}(K^*) \leq 1000\log m + 700$. Therefore

$$\begin{aligned}
&\mathbb{E}\left[\operatorname{diam}(\mathbf{H}) \mid \operatorname{diam}(\mathbf{K}) < 500\log m', e(\mathbf{K}) = m'\right] \\
&\leq \sup(\mathbb{E}\left[\operatorname{diam}(\mathbf{H}) \mid \mathbf{K}^* = K^*\right] : \operatorname{diam}(K^*) \leq 1000\log m + 700) \\
&= O\left(\sqrt{n}\right)
\end{aligned} \tag{6.15}$$

Now let $\mathbf{C} = C(\mathbf{G})$ and note that $K(\mathbf{C}) = K(\mathbf{K}^*) = \mathbf{K}$ because $s(\mathrm{d}) \geq 2$. Write $\mathrm{d}(\mathbf{C})$ for the degree sequence of $\mathbf{C}$. Then for any degree sequence d' with $\mathbb{P}\{\mathrm{d}(\mathbf{C}) = \mathrm{d}', e(\mathbf{K}) = m'\} > 0$, conditionally given that $\mathrm{d}(\mathbf{C}) = \mathrm{d}'$, we have $\mathbf{C} \in_u \mathcal{C}_{\mathrm{d}'}$. Fix such a degree sequence d', and note that d' is 1-free and $e(K(C')) = m'$ for any $C' \in \mathcal{G}^-_{\mathrm{d}'}$, and let $\mathbf{C}' \in_u \mathcal{G}^-_{\mathrm{d}'}$. Then given that $\mathbf{C}'$ is connected, $\mathbf{C}' \in_u \mathcal{C}_{\mathrm{d}'}$. Thus, $(\mathbf{C} \mid \mathrm{d}(\mathbf{C}) = \mathrm{d}') \stackrel{\mathrm{d}}{=} (\mathbf{C}' \mid \mathbf{C}' \text{ is connected})$. Now writing $\mathbf{K}' = K(\mathbf{C}')$, by Theorem 1.7,

$$\mathbb{P}\left\{\operatorname{diam}^+(\mathbf{K}') \geq 500\log m'\right\} = O\left(\frac{\log^8 m'}{m'}\right).$$

In particular, $\mathbb{P}\{\mathbf{C}' \text{ is connected}\}$ is bounded away from 0 by a constant, independent of d'. Hence,

$$\begin{aligned}
\mathbb{P}\left\{\operatorname{diam}(\mathbf{K}) \geq 500\log m' \mid \mathrm{d}(\mathbf{C}) = \mathrm{d}'\right\} &= \mathbb{P}\left\{\operatorname{diam}(\mathbf{K}') \geq 500\log m' \mid \mathbf{C}' \text{ connected}\right\} \\
&= \mathbb{P}\left\{\operatorname{diam}^+(\mathbf{K}') \geq 500\log m' \mid \mathbf{C}' \text{ connected}\right\} \\
&\leq \frac{\mathbb{P}\left\{\operatorname{diam}^+(\mathbf{K}') \geq 500\log m'\right\}}{\mathbb{P}\{\mathbf{C}' \text{ connected}\}} \\
&= O\left(\frac{\log^8 m'}{m'}\right).
\end{aligned}$$

On the event that $\mathrm{d}(\mathbf{C}) = \mathrm{d}'$, we have $e(\mathbf{K}) = m'$. Because this holds for any d', we obtain

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{K}) \geq 500\log m' \mid e(\mathbf{K}) = m'\right\} = O\left(\frac{\log^8 m'}{m'}\right). \tag{6.16}$$

Combining (6.14) and (6.16) yields

$$\begin{aligned}
&\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})<5\sqrt{m'n}]}\mathbf{1}_{[\operatorname{diam}(\mathbf{K})\geq 500\log m']} \mid e(\mathbf{K})=m'\right] \\
&= \mathbb{P}\left\{\operatorname{diam}(\mathbf{K}) \geq 500\log m' \mid e(\mathbf{K})=m'\right\} \\
&\cdot \mathbb{E}\left[\operatorname{diam}(\mathbf{H})\mathbf{1}_{[\operatorname{diam}(\mathbf{H})<5\sqrt{m'n}]} \mid \operatorname{diam}(\mathbf{K}) \geq 500\log m', e(\mathbf{K})=m'\right] \\
&= O\left(\frac{\log^8 m'}{m'}\right)5\sqrt{m'n} \\
&= O\left(\sqrt{n}\right). \qquad (6.17)
\end{aligned}$$

Combining (6.13), (6.15), and (6.17), we have

$$\mathbb{E}\left[\operatorname{diam}(\mathbf{H}) \mid e(\mathbf{K})=m'\right] = O\left(\sqrt{n}\right).$$

Because this holds regardless of m', we conclude that $\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\right] = O(\sqrt{n})$.

We now show that $\mathbb{E}\left[\operatorname{diam}(\mathbf{G}) \mid \mathbf{H}\right]$ is not much larger than $\operatorname{diam}(\mathbf{H})$. So, fix a graph H with $\mathbb{P}\{\mathbf{H}=H\}>0$, let $m^*=e^*(H)$, and let $h=n-v(H)$. Then H is connected, so given that $\mathbf{H}=H$, we have $\mathbf{G}\in_u \mathcal{C}_{\mathrm{d}}(H)=\mathcal{G}^-_{\mathrm{d}}(H)$.

We claim that $m\geq 2$. Suppose toward a contradiction that $m^*\leq 1$. Then writing $K^*=K^*(H)$, we have $e^*(K^*)\leq e^*(H)\leq 1$, so K^* is either empty or is a triangle with a single mutable edge e^*. The second case is only possible if $\mathbf{G}$ has at least one vertex of degree 1, so H has at least one mutable pendant edge, which together with e^*, witnesses that $m^*\geq 2$. In the first case, where K^* is empty, H is either empty or is a tree. If H is empty, then some (and hence every) graph in $\mathcal{C}_{\mathrm{d}}$ is a cycle, contradicting our assumption that $n_2(\mathrm{d})<(1-\varepsilon)n$. If H is a tree, but does not have at least 2 pendant edges (which would suffice to imply $m^*\geq 2$), then H is a single edge. This implies some (and hence every) graph in $\mathcal{C}_{\mathrm{d}}$ is a path, which also contradicts our assumption that $n_2(\mathrm{d})<(1-\varepsilon)n$, provided that $\varepsilon n\geq 2$, which we may assume. This establishes that $m^*\geq 2$.

It now follows from Lemma 6.3 that for all $y\geq 16$,

$$\begin{aligned}
&\mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \geq 2\left(1+\frac{h}{m^*-1}\right)(\operatorname{diam}(H)+2+y\log m^*)\middle|\mathbf{H}=H\right\} \\
&\leq \exp\left(-\frac{y\log m^*}{8}+2\right).
\end{aligned}$$

Since all vertices of $\mathbf{G}$ not present in $\mathbf{H}$ have degree 2, $h\leq n_2(\mathrm{d})<(1-\varepsilon)n$. Therefore, exactly as in (6.2) from the proof of Lemma 6.2, we have $6m\geq\varepsilon h$. Since $m^*\geq 2$ and $\varepsilon<1$, for all $x\geq 20$, if we set $y=x/1.25\geq 16$, then $y/8=x/10$ and

$$\begin{aligned}
&2\left(1+\frac{h}{m^*-1}\right)(\operatorname{diam}(H)+2+y\log m^*) \\
&\leq \left(2+\frac{4h}{m^*}\right)(\operatorname{diam}(H)+1.25y\log m^*) \\
&\leq \frac{26}{\varepsilon}(\operatorname{diam}(H)+x\log m^*).
\end{aligned}$$

Therefore

$$\mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \geq \frac{26}{\varepsilon}(\operatorname{diam}(H)+x\log m^*) \mid \mathbf{H}=H\right\} \leq \exp\left(-\frac{x\log m^*}{10}+2\right). \qquad (6.18)$$

We can now bound the conditional expectation of $\mathbf{G}$, given that $\mathbf{H} = H$. By (6.18),

$$\begin{aligned}
&\mathbb{E}\left[\operatorname{diam}(\mathbf{G}) \mid \mathbf{H} = H\right] \\
&= \int_0^\infty \mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) > x \mid \mathbf{H} = H\right\} \mathrm{d}x \\
&\le \frac{26}{\varepsilon}(\operatorname{diam}(H) + 20 \log m^*) \\
&+ \frac{26}{\varepsilon} \log m \int_{20}^\infty \mathbb{P}\left\{\operatorname{diam}(\mathbf{G}) \ge \frac{26}{\varepsilon}(\operatorname{diam}(H) + x \log m^*) \mid \mathbf{H} = H\right\} \mathrm{d}x \\
&\le \frac{26}{\varepsilon}(\operatorname{diam}(H) + 20 \log m^*) + \frac{26}{\varepsilon} \log m^* \int_{20}^\infty \exp\left(-\frac{x \log m^*}{10} + 2\right) \mathrm{d}x \\
&= \frac{26}{\varepsilon}(\operatorname{diam}(H) + 20 \log m^*) + \frac{e^2 \cdot 10 \cdot 26}{\varepsilon (m^*)^2} \\
&\le \frac{1}{\varepsilon}(26 \operatorname{diam}(H) + 2^{12} \log n),
\end{aligned} \tag{6.19}$$

where in the final inequality we have used that $m^* \le n^2$. Since $\mathbb{E}\left[\operatorname{diam}(\mathbf{H})\right] = O\left(\sqrt{n}\right)$, and since (6.19) holds for all H, we conclude that

$$\begin{aligned}
\mathbb{E}\left[\operatorname{diam}(\mathbf{G})\right] &= \mathbb{E}\left[\mathbb{E}\left[\operatorname{diam}(\mathbf{G}) \mid \mathbf{H}\right]\right] \\
&\le \mathbb{E}\left[\frac{1}{\varepsilon}(26 \operatorname{diam}(\mathbf{H}) + 2^{12} \log n)\right] \\
&= O\left(\sqrt{n}/\varepsilon\right),
\end{aligned}$$

as required. □

The final ingredient we need for the proof of Theorem 1.2 is a tool for extending tail bounds on diameters of connected graphs to tail bounds on maxima over the diameters of multiple connected components. This is the purpose of the following easy lemma.

Lemma 6.6. *Fix a positive integer n and nonnegative integers $n_1, \dots, n_k$ with $n_1 + \cdots + n_k \le n$. Also fix real numbers α and β with $\beta \ge 1$. Let $(\mathbf{X}_1, \dots, \mathbf{X}_k)$ be random variables such that for all $i \in [k]$ and for all $x \ge 0$,*

$$\mathbb{P}\left\{\mathbf{X}_i \ge x\sqrt{n_i}\right\} \le \exp\bigl(-\alpha x^2 + \beta\bigr).$$

Then for all $x \ge 0$,

$$\mathbb{P}\left\{\max(\mathbf{X}_1, \dots, \mathbf{X}_k) \ge x\sqrt{n}\right\} \le \exp\bigl(-\alpha x^2 + \beta\bigr).$$

Proof. By a union bound,

$$\mathbb{P}\left\{\max(\mathbf{X}_1, \dots, \mathbf{X}_k) \ge x\sqrt{n}\right\} \le \sum_{i=1}^k \mathbb{P}\left\{\mathbf{X}_i \ge \left(x\sqrt{n/n_i}\right)\sqrt{n_i}\right\} \le e^\beta \sum_{i=1}^k \exp\left(-\frac{\alpha x^2 n}{n_i}\right).$$

It is a calculus exercise to show that for any real numbers $a > 0$ and $a_1, \dots, a_k \ge 0$ with $a_1 + \cdots + a_k \le a$,

$$\sum_{i=1}^k \exp\left(-\frac{a}{a_i}\right) \le \exp\left(-\frac{a}{a_1 + \cdots + a_k}\right). \tag{6.20}$$

We may assume that $\alpha x^2 \ge 1$ because otherwise the claimed bound is greater than 1. Now apply (6.20) with $a = \alpha x^2 n \ge n_1 + \cdots + n_k$ and $(a_1, \dots, a_k) = (n_1, \dots, n_k)$. □

Proof of Theorem 1.2. Fix $\varepsilon \in (0, 1)$ and fix a degree sequence $\mathrm{d} = (d_1, \dots, d_n)$ such that $n_2(\mathrm{d}) < (1 - \varepsilon)n$, and let $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}$.

We proceed much like in the proof of Theorem 1.1, though there are some additional obstacles. Let $\mathbf{G}_1$ be obtained from $\mathbf{G}$ by removing all components of surplus at most 1, and let $\mathbf{H}_1 = H(\mathbf{G}_1)$ and $\mathbf{K}_1^* = K^*(\mathbf{G}_1) = K^*(\mathbf{H}_1)$. We first bound the diameter of $\mathbf{H}_1$ as a function of $\mathbf{K}_1^*$. So, fix a set $U \subseteq [n]$ and a simple kernel K_1^* with $\mathbb{P}\{V(\mathbf{H}_1) = U, \mathbf{K}_1^* = K_1^*\} > 0$. Then $\mathrm{d}|_{U \setminus V(K_1^*)}$ is 2-free. Now let $\mathbf{H}_1^1, \ldots, \mathbf{H}_1^{\mathbf{k}}$ and $(K_1^*)^1, \ldots, (K_1^*)^k$ be the components of $\mathbf{H}_1$ and K_1^* respectively, listed in (say) lexicographic order on vertex sets. Consider a partition $U^1, \ldots, U^k$ of U such that $V((K_1^*)^i) \subseteq U^i$ for all $i \in [k]$ and such that $\mathbb{P}\left\{(V(\mathbf{H}_1^1), \ldots, V(\mathbf{H}_1^{\mathbf{k}})) = (U^1, \ldots, U^k), \mathbf{K}_1^* = K_1^*\right\} > 0$. Then conditionally given that $(V(\mathbf{H}_1^1), \ldots, V(\mathbf{H}_1^{\mathbf{k}})) = (U^1, \ldots, U^k)$ and $\mathbf{K}_1^* = K_1^*$, we have $\mathbf{H}_1^i \in_u \mathcal{C}_{\mathrm{d}|_{U^i}}((K_1^*)^i)$ for all $i \in [k]$. Let $m_1 = e^*(K_1^*)$, and note that $m_1 \neq 1$. Then for all $i \in [k]$, because $|U^i| \leq n$ and $e^*((K_1^*)^i) \leq m_1$, it follows from Lemma 6.5 that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}_1^i) \geq 5\sqrt{m_1 n} + x\sqrt{|U^i|} \mid (V(\mathbf{H}_1^1), \ldots, V(\mathbf{H}_1^{\mathbf{k}})) = (U^1, \ldots, U^k), \mathbf{K}_1^* = K_1^*\right\}$$
$$\leq \exp\left(-\frac{x^2}{2^{11}} + 3\right).$$

Given that $(V(\mathbf{H}_1^1), \ldots, V(\mathbf{H}_1^{\mathbf{k}})) = (U^1, \ldots, U^k)$ and $\mathbf{K}_1^* = K_1^*$, we have $\mathrm{diam}(\mathbf{H}_1) = \max(\mathrm{diam}(\mathbf{H}_1^i) : i \in [k])$, so applying Lemma 6.6 to the conditioned random variables $(\mathrm{diam}(\mathbf{H}_1^i) - 5\sqrt{m_1 n}, i \in [k])$, then averaging over $(V(\mathbf{H}_1^1), \ldots, V(\mathbf{H}_1^{\mathbf{k}}))$, we find that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}_1) \geq 5\sqrt{m_1 n} + x\sqrt{n} \mid \mathbf{K}_1^* = K_1^*\right\} \leq \exp\left(-\frac{x^2}{2^{11}} + 3\right). \tag{6.21}$$

If $\mathrm{diam}^+(K_1^*) \leq 1000 \log m_1 + 700$, then in particular K_1^* is connected, so in this case, conditionally given that $V(\mathbf{H}_1) = U$ and that $\mathbf{K}_1^* = K_1^*$, we have $\mathbf{H}_1 \in_u \mathcal{C}_{\mathrm{d}|_U}(K_1^*)$. Therefore, if $m_1 \geq 2$ and $\mathrm{diam}^+(K_1^*) \leq 1000 \log m_1 + 700$, then it follows from Lemma 6.5 that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}_1) \geq x\sqrt{n} \mid \mathbf{K}_1^* = K_1^*\right\} \leq \exp\left(-\frac{x^{2/3}}{2^{12}} + 4\right). \tag{6.22}$$

But if $m_1 \leq 1$, then $m_1 = 0$; in this case, if $\mathbf{K}_1^* = K_1^*$ then $\mathbf{H}_1$ is empty. Therefore (6.22) holds provided that $\mathrm{diam}^+(K_1^*) \leq 1000 \log m_1 + 700$, even if $m_1 \leq 1$.

Now let $\mathbf{C}_1 = C(\mathbf{G}_1)$ and let $\mathbf{K}_1 = K(\mathbf{G}_1)$. Because all components of $\mathbf{G}_1$ have surplus at least 2, we have $\mathbf{K}_1 = K(\mathbf{C}_1)$. Fix a degree sequence d_1 such that $\mathbb{P}\{\mathrm{d}(\mathbf{C}_1) = \mathrm{d}_1\} > 0$, so d_1 is 1-free. Note that all kernels of graphs in $\mathcal{G}_{\mathrm{d}_1}^-$ have the same number of edges; that is, there exists $m_1' \in \mathbb{N}$ such that, for any $C_1 \in \mathcal{G}_{\mathrm{d}_1}^-$, it holds that $e(K(C_1)) = m_1'$. Given that $\mathrm{d}(\mathbf{C}_1) = \mathrm{d}_1$, we have $\mathbf{C}_1 \in_u \mathcal{G}_{\mathrm{d}_1}^-$, and hence by Theorem 1.7,

$$\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{K}_1) \geq 500 \log m_1' \mid \mathrm{d}(\mathbf{C}_1) = \mathrm{d}_1\right\} \leq O\left(\frac{\log^8 m_1'}{m_1'}\right).$$

This holds for any d_1 such that kernels of cores in $\mathcal{G}_{\mathrm{d}_1}^-$ have m_1 edges, so

$$\mathbb{P}\left\{\mathrm{diam}^+(\mathbf{K}_1) \geq 500 \log m_1' \mid e(\mathbf{K}_1) = m_1'\right\} \leq O\left(\frac{\log^8 m_1'}{m_1'}\right). \tag{6.23}$$

The bounds (6.21), (6.22), and (6.23), are sufficient to deduce, in precisely the same manner as in the proof of Theorem 1.1, that $\mathbb{E}[\mathrm{diam}(\mathbf{H}_1)] = O(\sqrt{n})$.

Next, let $\mathbf{G}_2$ be the union of the components of $\mathbf{G}$ with surplus at most 1, and let $\mathbf{H}_2 = H(\mathbf{G}_2)$ and $\mathbf{K}_2^* = K^*(\mathbf{H}_2)$. Fix a set $V \subseteq [n]$ and a simple kernel K_2^* with $\mathbb{P}\{V(\mathbf{H}_2) = V, \mathbf{K}_2^* = K_2^*\} > 0$, so the components of K_2^* are necessarily triangles with exactly one mutable edge. Then $\mathrm{d}|_{V \setminus V(K_2^*)}$ is 2-free. Let $\mathbf{H}_2^1, \ldots, \mathbf{H}_2^{\mathbf{j}}$ be the tree components of $\mathbf{H}_2$ listed in lexicographic order of vertex sets, and let $(K_2^*)^1, \ldots, (K_2^*)^j$ be empty graphs. Let $\mathbf{H}_2^{\mathbf{j}+1}, \ldots, \mathbf{H}_2^{\boldsymbol{\ell}}$ be the unicyclic components of $\mathbf{H}_2$, and let $(K_2^*)^{j+1}, \ldots, (K_2^*)^\ell$ be the triangle components of K_2^*, both listed in lexicographic order of vertex sets. Consider a partition $V^1, \ldots, V^\ell$ of V such that $V((K_2^*)^i) \subseteq V^i$ for all $i \in [\ell]$ and such that the event $E = \{(V(\mathbf{H}_2^1), \ldots, V(\mathbf{H}_2^{\boldsymbol{\ell}})) = (V^1, \ldots, V^\ell), (K^*(\mathbf{H}_2^1), \ldots, K^*(\mathbf{H}_2^{\boldsymbol{\ell}})) = ((K_2^*)^1, \ldots, (K_2^*)^\ell)\}$ has $\mathbb{P}\{E\} > 0$. Then conditionally given E, we have $\mathbf{H}_2^i \in_u \mathcal{C}_{\mathrm{d}|_{V^i}}((K_2^*)^i)$ for all $i \in [\ell]$. For each $i \in [\ell]$, we have $e^*((K_2^*)^i) \leq 1$, and it therefore follows from the first bound of Lemma 6.5 that for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}_2^i) \geq 5\sqrt{|V^i|} + x\sqrt{|V^i|} \mid E\right\} \leq \exp\left(-\frac{x^2}{2^{11}} + 3\right).$$

On the event E, $\mathrm{diam}(\mathbf{H}_2) = \max(\mathrm{diam}(\mathbf{H}_2^i) : i \in [\ell])$. Hence by Lemma 6.6, for all $x \geq 0$,

$$\mathbb{P}\left\{\mathrm{diam}(\mathbf{H}_2) \geq (5+x)\sqrt{n} \mid E\right\} \leq \exp\left(-\frac{x^2}{2^{11}} + 3\right),$$

As $(V^1, \ldots, V^\ell)$ and $((K_2^*)^1, \ldots, (K_2^*)^\ell)$ were arbitrary, it follows that $\mathbb{E}[\mathrm{diam}(\mathbf{H}_2)] = O(\sqrt{n})$. Defining $\mathbf{H} = H(\mathbf{G})$, we have $\mathrm{diam}(\mathbf{H}) = \max(\mathrm{diam}(\mathbf{H}_1), \mathrm{diam}(\mathbf{H}_2))$. Because $\mathbb{E}[\mathrm{diam}(\mathbf{H}_1)] = O(\sqrt{n})$, we conclude that $\mathbb{E}[\mathrm{diam}(\mathbf{H})] = O(\sqrt{n})$.

Next let $\mathbf{G}^-$ be the graph obtained from $\mathbf{G}$ by removing all cycle components. Then for any set $V^- \subseteq [n]$ and graph H with $\mathbb{P}\{V(\mathbf{G}^-) = V^-, \mathbf{H} = H\} > 0$, conditionally given that $V(\mathbf{G}^-) = V^-$ and that $\mathbf{H} = H$, we have $\mathbf{G}^- \in_u \mathcal{G}^-_{\mathrm{d}|_{V^-}}(H)$. It therefore follows, precisely as in the proof of Theorem 1.1, that $\mathbb{E}[\mathrm{diam}(\mathbf{G}^-) \mid \mathbf{H}] = O((\mathrm{diam}(\mathbf{H}) + \log n)/\varepsilon)$, and hence that $\mathbb{E}[\mathrm{diam}(\mathbf{G}^-)] = O(\sqrt{n}/\varepsilon)$.

Finally, we analyze $\mathrm{cyc}(\mathbf{G})$. Fix any graph H with $\mathbb{P}\{\mathbf{H} = H\} > 0$, and note that conditionally given that $\mathbf{H} = H$, we have $\mathbf{G} \in_u \mathcal{G}_{\mathrm{d}}(H)$. Letting $h = n - v(H)$, we also have $h \leq n_2(\mathrm{d}) < (1-\varepsilon)n$. Therefore, by Lemma 6.2, there is a constant $C = C(\varepsilon)$ such that for all $x \geq 0$,

$$\mathbb{P}\{\mathrm{cyc}(\mathbf{G}) \geq x \mid \mathbf{H} = H\} \leq \exp\left(-\frac{\varepsilon x}{10} + C\right).$$

Since this holds for all H, it follows immediately that $\mathbb{E}[\mathrm{cyc}(\mathbf{G})] = O(C/\varepsilon)$. Because $\mathrm{diam}(\mathbf{G}) \leq \max(\mathrm{diam}(\mathbf{G}^-), \mathrm{cyc}(\mathbf{G})/2)$, we conclude that

$$\mathbb{E}[\mathrm{diam}(\mathbf{G})] \leq O\left(\left(\sqrt{n} + C\right)/\varepsilon\right),$$

as required. □

Appendix A. Proofs of technical lemmas

Proof of Lemma 5.9. Let

$$\mathcal{A} = \{A \in \mathcal{A}_{\mathrm{d}} : e_2(A, v) \text{ is a back-edge}, |Q_1(A, v)| \leq q\}, \text{ and}$$
$$\mathcal{B} = \{A \in \mathcal{A}_{\mathrm{d}} : e_2(A, v) \text{ is not a back-edge}\}.$$

Fix $A \in \mathcal{A}$ and consider the set E of all edges in $K(A)$ not incident to $K_1(A, v)$ that are either subdivided or have an endpoint not adjacent to $K_1(A, v)$. As in the proof of Lemma 5.5,

$$|E| \geq m - |Q_1(A, v)| - 1 - \binom{|Q_1(A, v)|}{2} \geq m - q - 1 - \binom{q}{2} \geq m/2.$$

Note that either $e_1(A, v)$ and $e_2(A, v)$ are both loops at v, or for some vertex $w \neq v$, both $e_1(A, v)$ and $e_2(A, v)$ are edges between v and w. In the first case, let $\vec{e}$ be an orientation of $e_2(A, v)$. In the second case, either $e_1(A, v)$ or $e_2(A, v)$ is subdivided, and we let $\vec{e}$ be an orientation of whichever one is subdivided. In either case, there are at least 2 choices for $\vec{e}$. Next, let $\vec{f}$ be an orientation of an edge in E. There are at least $2(m/2) = m$ choices for $\vec{f}$. Because $A \in \mathcal{A}$, both endpoints of $\vec{e}$ are in $K_1(A, v)$, and therefore $\vec{f}$ is either subdivided or has an endpoint not adjacent to either endpoint of $\vec{e}$. Since $\vec{e}$ is subdivided, by Lemma 5.3, either $(\vec{e}, \vec{f})$ or $(\vec{e}, \overleftarrow{f})$ is a valid pair. Moreover switching on whichever of these pairs is valid yields an element of $\mathcal{B}$. It follows that there are at least $2m/2 = m$ switchings from A to an element of $\mathcal{B}$.

Now fix $B \in \mathcal{B}$ and note that $e_1(B, v)$ and $e_2(B, v)$ both have v as an endpoint since $d_v \geq 3$. Let w be the other endpoint of $e_1(B, v)$ and let u be the other endpoint of $e_2(B, v)$ (some of u, v, w may be equal). Any switching from B to an element $A \in \mathcal{A}$ is equivalent to a pair $(\vec{e}, \vec{f})$ where $\vec{e}$ is an orientation of either $e_1(B, v)$ or $e_2(B, v)$ and $\vec{f}$ is incident to some $x \in \{u, w\}$. There are at most 4 choices for $\vec{e}$. Moreover $x \in K_1(A, v)$ so $|Q_1(A, v)| \geq d_x - 2$. Since $A \in \mathcal{A}$, $|Q_1(A, v)| \leq q$, so it follows that $d_x \leq q + 2$. Hence, there are at most $2(q + 2)2$ choices for $\vec{f}$. This proves that there are at most $4 \cdot 2(q + 2)2 = 16q + 32$ switchings from B to an element of $\mathcal{A}$.

Applying Lemma 5.2 with $\mathcal{C} = \mathcal{A}_{\mathrm{d}}$, we obtain

$$\mathbb{P}\{\mathbf{e}_2 \text{ back-edge}, |\mathbf{Q}_1| \leq q\} = \mathbb{P}\{\mathbf{A} \in \mathcal{A}\} \leq \frac{16q + 32}{m} \mathbb{P}\{\mathbf{A} \in \mathcal{B}\} \leq \frac{16q + 32}{m},$$

as required. □

Proof of Lemma 5.10. Let H consist of the first 5 half-edges in Q_t. That is, no half-edge in $Q_t \setminus H$ has smaller distance to v than a half-edge in H, and subject to this, H contains the 5 lexicographically least half-edges. Let $\mathcal{C} = \{A \in \mathcal{A}_{\mathrm{d}} : K_t(A, v) = K_t, Q_t(A, v) = Q_t\}$ and for $A \in \mathcal{C}$, let $s(A)$ be the least nonnegative integer such that at time $t + s(A)$, all half-edges in H are explored. Then $3 \leq s(A) \leq 5$ because for all $s \in [s(A)]$, the edge $e_{t+s}(A, v)$ has either 1 or 2 half-edges in H. Now fix $\{i, j\} \in \binom{[5]}{2}$ and let $\mathcal{A}$ consist of all $A \in \mathcal{C}$ such that $i, j \leq s(A)$, both $e_{t+i}(A, v)$ and $e_{t+j}(A, v)$ are back-edges, $|B_{K(A)}(K_t, 3)| \leq m/2$, and none of $e_{t+1}(A, v), \ldots, e_{t+s(A)}(A, v)$ have an endpoint outside K_t with degree at least 11.

Consider the bipartite multigraph $\mathcal{H}$ with parts $\mathcal{A}$ and $\mathcal{C} \setminus \mathcal{A}$ and an edge between $A \in \mathcal{A}$ and $B \in \mathcal{C} \setminus \mathcal{A}$ for each quadruple $(\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2)$ such that $\vec{e}_1$ is an orientation of $e_{t+i}(A, v)$, $\vec{e}_2$ is an orientation of $e_{t+j}(A, v)$, $\vec{f}_1$ and $\vec{f}_2$ are orientations of edges in $K(A)$ outside $B_{K(A)}(K_t, 3)$, and by first switching on $(\vec{e}_1, \vec{f}_1)$ in A to obtain A', and then switching on $(\vec{e}_2, \vec{f}_2)$ in A', we obtain B. We can instead view $\mathcal{H}$ as having parts $\mathcal{A}$ and $\mathcal{B}$, where $\mathcal{B}$ is the subset of $\mathcal{C} \setminus \mathcal{A}$ consisting of all $B \in \mathcal{C} \setminus \mathcal{A}$ which are not isolated in $\mathcal{H}$.

Fix $A \in \mathcal{A}$. To specify an edge in $\mathcal{H}$ from A to some $B \in \mathcal{B}$, there are 2 choices for each of $\vec{e}_1$ and $\vec{e}_2$, and since $A \in \mathcal{A}$, there are at least $2(m/2) = m$ choices for $\vec{f}_1$, and at least $2(m/2 - 1) = m - 2$ choices for $\vec{f}_2$. Let A' be obtained from A by switching on $(\vec{e}_1, \vec{e}_2)$ in A. We claim that either $(\vec{e}_2, \vec{f}_2)$ or $(\vec{e}_2, \vec{f}_2)$ is a valid pair in A'. Write e_1, e_2, f_1, f_2 for the underlying edges of $\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2$. If neither $(\vec{e}_2, \vec{f}_2)$ nor $(\vec{e}_2, \vec{f}_2)$ is valid in A', then Lemma 5.3 yields an obstruction; by the symmetry between e_2 and f_2, we can assume that e_2 has an endpoint adjacent to both endpoints of f_2 in $K(A')$ via distinct non-subdivided edges e and f. If either e or f was present in $K(A)$, then this edge together with e_2 would form a path of length at most 2 from K_t to an endpoint of f_2, contradicting the fact that $f_2 \notin B_{K(A)}(K_t, 3)$. Therefore e and f were created by switching on $(\vec{e}_1, \vec{f}_1)$ in A. Since e and f share an endpoint, it follows that either e_1 or f_1 is a loop, and hence subdivided. This contradicts the fact that neither e nor f is subdivided. Hence, for any choice of $(\vec{e}_1, \vec{f}_1)$, there are at least $m(m-2)/2$ choices for $(\vec{e}_2, \vec{f}_2)$. This proves that the degree of A in $\mathcal{H}$ is at least $4m(m-2)/2 \geq m^2$.

Now fix $B \in \mathcal{B}$ and consider an edge between B and some $A \in \mathcal{A}$ corresponding to a quadruple $(\vec{e}_1, \vec{e}_2, \vec{f}_1, \vec{f}_2)$. Let A' be obtained from A by switching on $(\vec{e}_1, \vec{f}_1)$ in A. Write $(\vec{e}'_1, \vec{f}'_1)$ and $(\vec{e}'_2, \vec{f}'_2)$ for the reversals of the switchings $(\vec{e}_1, \vec{f}_1)$ and $(\vec{e}_2, \vec{f}_2)$ respectively. One of $\vec{e}'_2$ or $\vec{f}'_2$ is an orientation of $e_{t+j}(B, v)$, and the other is an orientation of an edge in $K(B)$ which either has a half-edge in $Q_t(B,v) = Q_t$, or has a half-edge incident to one of the at most 5 endpoints of $e_{t+1}(B,v), \dots, e_{t+s(B)}(B,v)$ outside $K_t(B,v) = K_t$, each of which have degree at most 11. Thus, there are at most $2 \cdot 2 \cdot 2(|Q_t| + 5 \cdot 11) = 8|Q_t| + 440$ choices for $(\vec{e}'_2, \vec{f}'_2)$. Similarly, there are at most $8|Q_t| + 440$ choices for $(\vec{e}'_1, \vec{f}'_1)$. Therefore the degree of B in $\mathcal{H}$ is at most $(8|Q_t| + 440)^2$.

By Lemma 1.8, we conclude that

$$\mathbb{P}\{\mathbf{A} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\} \le \frac{(8|Q_t| + 440)^2}{m^2} \mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\} \le \left(\frac{8|Q_t| + 440}{m}\right)^2.$$

To finish, just note that if $|\mathbf{Q}_{t+s(\mathbf{A})}| \le |\mathbf{Q}_t|$, then none of $\mathbf{e}_{t+1}, \dots, \mathbf{e}_{t+s(\mathbf{A})}$ have an endpoint outside K_t of degree at least 11, and for some $\{i,j\} \in \binom{[5]}{2}$ with $i, j \le s(\mathbf{A})$, both $\mathbf{e}_{t+i}$ and $\mathbf{e}_{t+j}$ are back-edges. Therefore

$$\begin{aligned}
&\mathbb{P}\{|B_{\mathbf{K}}(K_t,3)| \le m/2, |\mathbf{Q}_{t+s}| \le |\mathbf{Q}_t| \text{ for all } s \in [5] \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\} \\
&\le \binom{5}{2} \mathbb{P}\{\mathbf{A} \in \mathcal{A} \mid \mathbf{A} \in \mathcal{C}\} \\
&\le \binom{5}{2} \left(\frac{8|Q_t| + 440}{m}\right)^2 \\
&\le \left(\frac{32|Q_t| + 1760}{m}\right)^2,
\end{aligned}$$

as required. $\square$

Lemma A.1. *Any simple graph G contains a forest F with $e(F) \ge \sqrt{e(G)/2}$.*

Proof. First suppose that G is connected and let F be a spanning tree of G. Writing $n = v(G)$, we have $e(G) \le \binom{n}{2} \le n^2/2$ and $e(F) = n - 1$. Rearranging, we obtain $e(F) \ge \sqrt{2e(G)} - 1$, and hence $e(F) \ge \lceil \sqrt{2e(G)} - 1 \rceil \ge \sqrt{e(G)/2}$ as needed. We now proceed by induction on the number of components of G. We have already established the base case where G is connected, so suppose now that G has at least two components. Let G_1 be one of these components, and let G_2 be the union of the rest of the components. Then by the base case, G_1 has spanning tree F_1 with $e(F_1) \ge \sqrt{e(G_1)/2}$ and by induction, G_2 contains a forest F_2 with $e(F_2) \ge \sqrt{e(G_2)/2}$. Let F be the union of F_1 and F_2. Then F is a forest contained in G and

$$e(F) = e(F_1) + e(F_2) \ge \sqrt{e(G_1)/2} + \sqrt{e(G_2)/2} \ge \sqrt{e(G_1)/2 + e(G_2)/2} = \sqrt{e(G)/2}.$$

$\square$

Proof of Lemma 5.13. Let H consist of the first b half-edges in Q_t. That is, no half-edge in $Q_t \setminus H$ has smaller distance to v than a half-edge in H, and subject to this, H has the b lexicographically least half-edges. Set $\delta = 1/(12b)$, and let $\mathcal{C} = \{A \in \mathcal{A}_{\mathrm{d}} : K_t(A,v) = K_t, Q_t(A,v) = Q_t\}$. For $A \in \mathcal{C}$, let $s(A)$ be the least nonnegative integer such that all half-edges in H are explored by time $t + s(A)$. Then $b/2 \le s(A) \le b$ since for all $s \in [s(A)]$, the edge $e_{t+s}(A,v)$ has either 1 or 2 half-edges in H.

Let $\mathcal{A}$ consist of all $A \in \mathcal{C}$ such that in $K(A)$, no vertex outside K_t of degree at least $3b$ is joined to H, and there are at least $m/3$ edges outside $B_{K(A)}(K_t,3)$ both of whose endpoints have degree at most δm. For $k \ge 0$, let $\mathcal{A}_k$ consist of all $A \in \mathcal{C}$ such that in $K(A)$, there are exactly k vertices outside K_t joined to H. Finally, let $\mathcal{B}$ be the set of all $A \in \mathcal{C}$ such that in $K(A)$, no vertex outside K_t of degree at least $3b$ is joined to H, there are at least $m/3$ edges outside $B_{K(A)}(K_t,3)$ with at least one endpoint of degree greater than δm, and there are at most $3b/4$ vertices outside K_t joined to H.

Claim A.2. *For all $A \in \mathcal{C}$, if $|B_{K(A)}(K_t,3)| \le m/3$ and $|Q_{t+s(A)}(A,v)| \le |Q_t| + b/4$, then either $A \in \mathcal{A} \cap \mathcal{A}_k$ for some $k \in [0, 3b/4]$, or $A \in \mathcal{B}$.*

Proof. Note that $A \in \bigcup_{k\in[0,3b/4]}(\mathcal{A}\cap\mathcal{A}_k)\cup\mathcal{B}$ if and only if in $K(A)$, (1) no vertex outside K_t of degree at least $3b$ is joined to H, (2) there are at most $3b/4$ vertices outside K_t joined to H, and (3) there are either at least $m/3$ edges outside $B_{K(A)}(K_t,3)$ with both endpoints of degree at most δm or there are at least $m/3$ edges outside $B_{K(A)}(K_t,3)$ with an endpoint of degree greater than δm.

If $|B_{K(A)}(K_t,3)| \le m/3$, then in $K(A)$ there must be at least $2m/3$ edges outside $B_{K(A)}(K_t,3)$, and either at least half of them have both endpoints of degree at most δm or at least half of them have an endpoint of degree greater than δm. Therefore if (3) fails, $|B_{K(A)}(K_t,3)| > m/3$. If (1) fails, then $|Q_{t+s(A)}(A,v)| \ge |Q_t| + b \ge |Q_t| + b/4$. Finally, if (2) fails, then in $K(A)$, there are distinct vertices $v_1,\ldots,v_{\lceil 3b/4\rceil}$ not in K_t joined to H. After exploring the half-edges in H which lead to each of the v_i, the queue size will have increased by at least $\sum_{i=1}^{\lceil 3b/4\rceil}(d_{v_i}-2) \ge 3b/4$. Exploring each of the at most $b/4$ remaining half-edges in H decreases the queue size by a total of at most $2(b/4) = b/2$. Therefore if (2) fails, then $|Q_{t+s(A)}(A,v)| \ge |Q_t| + 3b/4 - b/2 = |Q_t| + b/4$. ■

We now bound $\mathbb{P}\{\mathbf{A}\in\mathcal{A}\cap\mathcal{A}_k \mid \mathcal{A}\in\mathcal{C}\}$ for $k\in[0,3b/4]$ and $\mathbb{P}\{\mathbf{A}\in\mathcal{B} \mid \mathbf{A}\in\mathcal{C}\}$.

Claim A.3. *For $k \in [0,b)$,*

$$\mathbb{P}\{\mathbf{A}\in\mathcal{A}\cap\mathcal{A}_k \mid \mathbf{A}\in\mathcal{C}\} \le \left(\frac{36b^3+12b|Q_t|}{m}\right)^{\frac{b-k}{2}}.$$

Proof. Let $r = \lceil (b-k)/2\rceil$ and for each $A\in\mathcal{A}\cap\mathcal{A}_k$, fix a list of edges $(e^1(A),\ldots,e^r(A))$ in $K(A)$ such that for all $i\in[r]$, $e^i(A)$ has a half-edge in H, and the other half-edge of $e^i(A)$ is either in Q_t or is on a vertex outside K_t joined to H in $K(A)$ with an edge not among $e^1(A),\ldots,e^r(A)$. Such a list can be constructed as follows. Start by removing exactly one edge in $K(A)$ from H to each distinct vertex not in K_t. Since $A\in\mathcal{A}_k$, at least $(b-k)/2$ (and hence at least r), edges in $K(A)$ with a half-edge in H remain, and we can take $(e^1(A),\ldots,e^r(A))$ to be any ordering of any r of them.

We define a bipartite multigraph $\mathcal{H}$ with parts $\mathcal{X}=\mathcal{A}\cap\mathcal{A}_k$ and $\mathcal{Y}=\mathcal{C}\setminus\mathcal{X}$ and an edge between $A\in\mathcal{X}$ and $B\in\mathcal{Y}$ for each list of oriented edges $(\vec{e}_1,\ldots,\vec{e}_r,\vec{f}_1,\ldots,\vec{f}_r)$ meeting the following requirements. First, for all $i\in[r]$, $\vec{e}_i$ is an orientation of $e^i(A)$. Second, for all $i\in[r]$, in $K(A)$, both endpoints of $\vec{f}_i$ have degree at most δm and $\vec{f}_i$ is not in $B_{K(A)}(K_t,3)$. Third, for all $i,j\in[r]$, $\vec{f}_i$ and $\vec{f}_j$ are vertex disjoint. Finally, we require that $A_r = B$ where $A_0 = A$ and for $i\in[r]$, A_i is the result of switching on $(\vec{e}_i,\vec{f}_i)$ in A_{i-1} (in particular, we require $(\vec{e}_i,\vec{f}_i)$ to be a valid pair in A_{i-1} for all $i\in[r]$). We can instead consider $\mathcal{Y}$ to be the subset of $\mathcal{C}\setminus\mathcal{X}$ consisting of all $B\in\mathcal{C}\setminus\mathcal{X}$ which are not isolated in $\mathcal{H}$. For $A\in\mathcal{X}$ and $B\in\mathcal{Y}$, we lower-bound the degree of A in $\mathcal{H}$ and upper-bound the degree of B in $\mathcal{H}$.

There are 2^r choices for $(\vec{e}_1,\ldots,\vec{e}_r)$. Independently, there are at least $\prod_{i=1}^r 2(m/3 - 2\delta m(i-1))$ ways to choose a list $(\vec{f}_1,\ldots,\vec{f}_r)$ of vertex disjoint oriented edges in $K(A)$ outside $B_{K(A)}(K_t,3)$ both of whose endpoints have degree at most δm in $K(A)$. Now observe that for any such choice of $(\vec{e}_1,\ldots,\vec{e}_r,\vec{f}_1,\ldots,\vec{f}_r)$ and for all $i\in[r]$, $(\vec{e}_i,\vec{f}_i)$ is a valid pair in A_{i-1}. Indeed in $K(A_{i-1})$, $\vec{e}_i$ lies within $B_{K(A_{i-1})}(K_t,1)$, while $\vec{f}_i$ is outside $B_{K(A_{i-1})}(K_t,3)$, and $\vec{f}_1,\ldots,\vec{f}_r$ are vertex disjoint. Moreover, the resulting augmented core $B = A_r$ is in $\mathcal{C}\setminus\mathcal{X}$ and hence in $\mathcal{Y}$, because $r\ge 1$ and so in $K(B)$, there are more than k vertices outside K_t joined to H. Since $\delta = 1/(12b)$ and $r\le b$, we have $2(m/3 - 2\delta m(i-1)) \ge m/3$ for all $i\in[r]$. This proves that the degree of A in $\mathcal{H}$ is at least $2^r(m/3)^r = (2m/3)^r$.

To upper-bound the degree of B in $\mathcal{H}$, consider an edge from $A\in\mathcal{X}$ to B corresponding to a list of oriented edges $(\vec{e}_1,\ldots,\vec{e}_r,\vec{f}_1,\ldots,\vec{f}_r)$, or equivalently, corresponding to a sequence of switchings $(\vec{e}_i, \vec{f}_i)$ from A_{i-1} to A_i for $i \in [r]$, where $A_0 = A$ and $A_r = B$. For $i \in [r]$, let $(\vec{e}'_i, \vec{f}'_i)$ be the switching from A_i to A_{i-1} which is the reversal of $(\vec{e}_i, \vec{f}_i)$. Then in $K(A_i)$, one of $\vec{e}'_i$ or $\vec{f}'_i$ has a half-edge in H and the other has a half-edge either in Q_t or on a vertex outside K_t of degree at most $3b$ joined to H. There are at most 2 choices for which of $\vec{e}'_i$ or $\vec{f}'_i$ has a half-edge in H. Given this choices, there are at most $2b$ choices for which oriented edge with a half-edge in H it is. Finally, there are at most $2(3b^2 + |Q_t|) = 6b^2 + 2|Q_t|$ oriented edges with a half-edge either in Q_t or on an endpoint outside K_t of an edge joined to H. (This is because in $K(A_i)$, there are at most $b(3b) = 3b^2$ edges with an endpoint of degree at most $3b$ that is joined to H.) Hence, there are at most $2 \cdot 2b \cdot (6b^2 + 2|Q_t|) = 24b^3 + 8b|Q_t|$ choices for $(\vec{e}'_i, \vec{f}'_i)$. This holds for all $i \in [r]$, so the degree of B in $\mathcal{H}$ is at most $(24b^3 + 8b|Q_t|)^r$.

By Lemma 1.8, we conclude that

$$\begin{aligned}\mathbb{P}\{\mathbf{A} \in \mathcal{A} \cap \mathcal{A}_k \mid \mathcal{A} \in \mathcal{C}\} = \mathbb{P}\{\mathbf{A} \in \mathcal{X} \mid \mathcal{A} \in \mathcal{C}\} &\leq \frac{(24b^3 + 8b|Q_t|)^r}{(2m/3)^r}\mathbb{P}\{\mathbf{A} \in \mathcal{Y} \mid \mathcal{A} \in \mathcal{C}\} \\ &\leq \left(\frac{36b^3 + 12b|Q_t|}{m}\right)^r \\ &\leq \left(\frac{36b^3 + 12b|Q_t|}{m}\right)^{\frac{b-k}{2}},\end{aligned}$$

as claimed. ■

Claim A.4.

$$\mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\} \leq \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4}.$$

Proof. Let $r = \lceil \sqrt{b}/4 \rceil$, and for each $A \in \mathcal{B}$, fix a list of edges $(e^1(A), \ldots, e^r(A))$ in $K(A)$ such that for all $i \in [r]$, in $K(A)$, $e^i(A)$ has a half-edge in H and the other half-edge either in Q_t or on a vertex outside K_t joined to H both by $e^i(A)$ and by an edge not among $e^1(A), \ldots, e^r(A)$, and such that $e^i(A)$ is either subdivided or has an endpoint not shared by any of $e^1(A), \ldots, e^{i-1}(A)$. Such a list can be constructed in stages as follows. Let E_1 be the set of edges in $K(A)$ with a half-edge in H, and obtain E_2 from E_1 by removing exactly one edge to H from each vertex outside K_t joined to H. Then since $A \in \mathcal{B}$, $|E_2| \geq (b - 3b/4)/2 = b/8$. Let E_3 be the set of subdivided edges in E_2 and let $E_4 = E_2 \setminus E_3$. Then E_4 spans a simple subgraph of $K(A)$, so by Lemma A.1, there exists $F \subseteq E_4$ with $|F| \geq \sqrt{|E_4|/2}$ such that F spans a forest in $K(A)$. Let $E_5 = E_3 \cup F$, so

$$|E_5| \geq |E_3| + \sqrt{\frac{|E_4|}{2}} = |E_3| + \sqrt{\frac{|E_2| - |E_3|}{2}} \geq |E_3| + \sqrt{\frac{b}{16} - \frac{|E_3|}{2}} \geq \frac{\sqrt{b}}{4}.$$

Hence there exists $E \subseteq E_5$ with $|E| = r$. Finally, by exploring the edges of F in breadth-first order, followed by an arbitrary ordering of the edges in $E \setminus F$, we obtain a list of edges $(e^1(A), \ldots, e^r(A))$ with the desired properties. The fact that the half-edge of $e^i(A)$ not in H is either in Q_t or joined to H by and edge not among $e^1(A), \ldots, e^r(A)$ is assured by the removal of E_2 from E_1.

We define a bipartite multigraph $\mathcal{H}$ with parts $\mathcal{X} = \mathcal{B}$ and $\mathcal{Y} = \mathcal{C} \setminus \mathcal{X}$ and an edge between $A \in \mathcal{X}$ and $B \in \mathcal{Y}$ for each list of oriented edges $(\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_r)$ meeting the following requirements. First, for all $i \in [r]$, $\vec{e}_i$ is an orientation of $e^i(A)$. Second, for all $i \in [r]$, $\vec{f}_i$ is not in $B_{K(A)}(K_t, 3)$, has an endpoint of degree greater than δm, and is either subdivided or has an endpoint not shared by any of $\vec{f}_1, \ldots, \vec{f}_{i-1}$. Finally, we require that $A_r = B$, where $A_0 = A$ and for $i \in [r]$, A_i is the result of switching on $(\vec{e}_i, \vec{f}_i)$ in A_{i-1} (in particular, we require $(\vec{e}_i, \vec{f}_i)$ to be a valid pair in A_{i-1} for all $i \in [r]$). We can instead consider $\mathcal{Y}$ to be the subset of $\mathcal{C} \setminus \mathcal{X}$ consisting of all $B \in \mathcal{C} \setminus \mathcal{X}$ which are not isolated in $\mathcal{H}$. For $A \in \mathcal{X}$ and $B \in \mathcal{Y}$, we lower-bound the degree of A in $\mathcal{H}$ and upper-bound the degree of B in $\mathcal{H}$.

There are 2^r choices for $(\vec{e}_1, \ldots, \vec{e}_r)$. Independently, there are at least $\prod_{i=1}^{r} 2(m/3 - \binom{2(i-1)}{2})$ choices for a list $(\vec{f}_1, \ldots, \vec{f}_r)$ such that each $\vec{f}_i$ is outside $B_{K(A)}(K_t, 3)$, has an endpoint of degree at least δm, and is either subdivided or does not share an endpoint with any of $\vec{f}_1, \ldots, \vec{f}_{i-1}$. This is because for all $i \in [r]$, the number of vertices spanned by $\vec{f}_1, \ldots, \vec{f}_{i-1}$ is at most $2(i-1)$, so when choosing $\vec{f}_i$, there are at most $\binom{2(i-1)}{2}$ non-subdivided edges to be avoided among the at least $m/3$ edges outside $B_{K(A)}(K_t, 3)$ with an endpoint of degree at least δm. Now observe that for any such choice of $(\vec{e}_1, \ldots, \vec{e}_r, \vec{f}_1, \ldots, \vec{f}_r)$, for all $i \in [r]$, either $(\vec{e}_i, \vec{f}_i)$ or $(\vec{e}_i, \overleftarrow{f}_i)$ is a valid pair in A_{i-1}. Indeed because $\vec{e}_i$ is in $B_{K(A)}(K_t, 1)$ and $\vec{f}_i$ is outside $B_{K(A)}(K_t, 3)$, by construction $\vec{e}_i$ is either subdivided or has an endpoint not adjacent to either endpoint of $\vec{f}_i$ in $K(A_{i-1})$, and $\vec{f}_i$ is either subdivided or has an endpoint not adjacent to either endpoint of $\vec{e}_i$ in $K(A_{i-1})$. The validity of one of $(\vec{e}_i, \vec{f}_i)$ or $(\vec{e}_i, \overleftarrow{f}_i)$ therefore follows from Lemma 5.3. Since $r - 1 \leq \sqrt{b}/4$, we have $m/3 - \binom{2(i-1)}{2} \geq m/6$ for all $i \in [r]$. It follows that the degree of A in $\mathcal{H}$ is at least $2^r (m/6)^r = (m/3)^r$.

Following precisely the same reasoning as in the proof of Claim A.3, we find that the degree of B in $\mathcal{H}$ is at most $(24b^3 + 8b|Q_t|)^r$. By Lemma 1.8, we conclude that

$$\begin{aligned}
\mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\} = \mathbb{P}\{\mathbf{A} \in \mathcal{X} \mid \mathbf{A} \in \mathcal{C}\} &\leq \frac{(24b^3 + 8b|Q_t|)^r}{(m/3)^r} \mathbb{P}\{\mathbf{A} \in \mathcal{Y} \mid \mathbf{A} \in \mathcal{C}\} \\
&\leq \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^r \\
&\leq \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4}.
\end{aligned}$$

■

Recall that $b/2 \leq s(A) \leq b$. Thus, combining Claim A.2, Claim A.3, and Claim A.4 with a union bound, we obtain

$$\begin{aligned}
&\mathbb{P}\{|B_{\mathbf{K}}(K_t, 3)| \leq m/3,\ |\mathbf{Q}_{t+s}| \leq |Q_t| + b/4 \text{ for all } s \in [b] \mid \mathbf{K}_t = K_t, \mathbf{Q}_t = Q_t\} \\
&\leq \sum_{k \in [0, 3b/4]} \mathbb{P}\{\mathbf{A} \in \mathcal{A} \cap \mathcal{A}_k \mid \mathbf{A} \in \mathcal{C}\} + \mathbb{P}\{\mathbf{A} \in \mathcal{B} \mid \mathbf{A} \in \mathcal{C}\} \\
&\leq \sum_{k \in [0, 3b/4]} \left(\frac{36b^3 + 12b|Q_t|}{m}\right)^{\frac{b-k}{2}} + \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4} \\
&\leq \sum_{k \in [0, 3b/4]} \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\frac{b-k}{2}} + \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4} \\
&\leq 2\left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{b/8} + \left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4} \\
&\leq 3\left(\frac{72b^3 + 24b|Q_t|}{m}\right)^{\sqrt{b}/4},
\end{aligned}$$

where the second last inequality holds because $72b^3 + 24b|Q_t| \leq m/2$ and the final inequality holds because $b \geq 4$. □

References

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probab. Theory Related Fields*, 152(3-4):367–406, 2012. ↑4

[2] Louigi Addario-Berry and Serte Donderwinkel. Random trees have height $O(\sqrt{n})$. *Ann. Probab.*, 52(6):2238–2280, 2024. ↑3, 6, and 13

[3] Louigi Addario-Berry, Serte Donderwinkel, Mickaël Maazoun, and James Martin. The foata-fuchs proof of cayley's formula, and its probabilistic uses, 2022. ↑13

[4] David J. Aldous. Exchangeability and related topics. In *École d'été de probabilités de Saint-Flour, XIII—1983*, volume 1117 of *Lecture Notes in Math.*, pages 1–198. Springer, Berlin, 1985. ↑38

[5] Caio Alves, Rodrigo Ribeiro, and Rémy Sanchis. Diameter of P.A. random graphs with edge-step functions. *Random Structures Algorithms*, 57(3):612–636, 2020. ↑4

[6] B. Bollobás and W. Fernandez de la Vega. The diameter of random regular graphs. *Combinatorica*, 2(2):125–134, 1982. ↑4

[7] Béla Bollobás and Oliver Riordan. The diameter of a scale-free random graph. *Combinatorica*, 24(1):5–34, 2004. ↑4

[8] Ju.D̃. Burtin. Extremal metric characteristics of a random graph. I. Asymptotic estimates. *Teor. Verojatnost. i Primenen.*, 19:740–754, 1974. ↑3

[9] Xing Shi Cai and Guillem Perarnau. The diameter of the directed configuration model. *Ann. Inst. Henri Poincaré Probab. Stat.*, 59(1):244–270, 2023. ↑4

[10] Francesco Caravenna, Alessandro Garavaglia, and Remco van der Hofstad. Diameter in ultra-small scale-free random graphs. *Random Structures Algorithms*, 54(3):444–498, 2019. ↑4

[11] Fan Chung and Linyuan Lu. The diameter of sparse random graphs. *Adv. in Appl. Math.*, 26(4):257–279, 2001. ↑4

[12] Steffen Dereich, Christian Mönch, and Peter Mörters. Distances in scale free networks at criticality. *Electron. J. Probab.*, 22:Paper No. 77, 38, 2017. ↑4

[13] Sander Dommers, Remco van der Hofstad, and Gerard Hooghiemstra. Diameters in preferential attachment models. *J. Stat. Phys.*, 139(1):72–107, 2010. ↑4

[14] Serte Donderwinkel and Zheneng Xie. Universality for the directed configuration model: metric space convergence of the strongly connected components at criticality. *Electron. J. Probab.*, 29:Paper No. 87, 85, 2024. ↑4

[15] R. B. Eggleton. Graphic sequences and graphic polynomials: a report. In *Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vols. I, II, III*, volume Vol. 10 of *Colloq. Math. Soc. János Bolyai*, pages 385–392. North-Holland, Amsterdam-London, 1975. ↑8

[16] Daniel Fernholz and Vijaya Ramachandran. The diameter of sparse random graphs. *Random Structures Algorithms*, 31(4):482–516, 2007. ↑4

[17] Philippe Flajolet and Andrew Odlyzko. The average height of binary trees and other simple trees. *J. Comput. System Sci.*, 25(2):171–213, 1982. ↑2

[18] Philippe Flajolet and Robert Sedgewick. *Analytic combinatorics*. Cambridge University Press, Cambridge, 2009. ↑37

[19] Dominique Foata and Aimé Fuchs. Réarrangements de fonctions et dénombrement. *J. Combinatorial Theory*, 8:361–375, 1970. ↑13

[20] Nicolas Fraiman and Dieter Mitsche. The diameter of inhomogeneous random graphs. *Random Structures Algorithms*, 53(2):308–326, 2018. ↑4

[21] Tobias Friedrich and Anton Krohmer. On the diameter of hyperbolic random graphs. *SIAM J. Discrete Math.*, 32(2):1314–1334, 2018. ↑4

[22] Christina Goldschmidt and Robin Stephenson. The scaling limit of a critical random directed graph. *Ann. Appl. Probab.*, 33(3):2024–2065, 2023. ↑4

[23] Václav Havel. Poznámka o existenci konečných grafů. *Časopis pro pěstování matematiky*, 080(4):477–480, 1955. ↑8
[24] Felix Joos, Guillem Perarnau, Dieter Rautenbach, and Bruce Reed. How to determine if a random graph with a fixed degree sequence has a giant component. *Probability Theory and Related Fields*, 170(1–2):263–310, January 2017. ↑9 and 23
[25] Joost Jorritsma and Júlia Komjáthy. Weighted distances in scale-free preferential attachment models. *Random Structures Algorithms*, 57(3):823–859, 2020. ↑4
[26] Victor Klee and David Larman. Diameters of random graphs. *Canadian J. Math.*, 33(3):618–640, 1981. ↑4
[27] Brendan D. McKay. Subgraphs of random graphs with specified degrees. *Congr. Numer.*, 33:213–223, 1981. ↑8
[28] Brendan D. McKay. Subgraphs of random graphs with specified degrees. In *Proceedings of the International Congress of Mathematicians. Volume IV*, pages 2489–2501. Hindustan Book Agency, New Delhi, 2010. ↑8
[29] Asaf Nachmias and Yuval Peres. Critical random graphs: diameter and mixing time. *Ann. Probab.*, 36(4):1267–1286, 2008. ↑4
[30] Oliver Riordan and Nicholas Wormald. The diameter of sparse random graphs. *Combin. Probab. Comput.*, 19(5-6):835–926, 2010. ↑4
[31] James K. Senior. Partitions and their representative graphs. *Amer. J. Math.*, 73:663–689, 1951. ↑8
[32] Nobutaka Shimizu. The average distance and the diameter of dense random regular graphs. *Electron. J. Combin.*, 27(3):Paper No. 3.62, 20, 2020. ↑4
[33] Remco van der Hofstad. *Random graphs and complex networks. Vol. 1*, volume [43] of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 2017. ↑4
[34] Remco van der Hofstad and Gerard Hooghiemstra. Universality for distances in power-law random graphs. *J. Math. Phys.*, 49(12):125209, 14, 2008. ↑4
[35] Remco van der Hofstad and Júlia Komjáthy. When is a scale-free graph ultra-small? *J. Stat. Phys.*, 169(2):223–264, 2017. ↑4