

Longest Path Distance in Random Circuits

NICOLAS BROUTIN¹ and OMAR FAWZI²

¹INRIA Rocquencourt, 78153 Le Chesnay, France
(e-mail: nicolas.broutin@inria.fr)

²School of Computer Science, McGill University, H3A 2K6, Montreal, Canada
(e-mail: ofawzi@cs.mcgill.ca)

Received 27 January 2011; revised 27 May 2012; first published online 3 July 2012

We study distance properties of a general class of random directed acyclic graphs (DAGs). In a DAG, many natural notions of distance are possible, for there exist multiple paths between pairs of nodes. The distance of interest for circuits is the maximum length of a path between two nodes. We give laws of large numbers for the typical depth (distance to the root) and the minimum depth in a random DAG. This completes the study of natural distances in random DAGs initiated (in the uniform case) by Devroye and Janson. We also obtain large deviation bounds for the minimum of a branching random walk with constant branching, which can be seen as a simplified version of our main result.

AMS 2010 *Mathematics subject classification*: Primary 60C05; 05C05; 94D99

1. Introduction

Motivated by the circuit value problem, and the delay to evaluate the output, Diaz, Serna, Spirakis, Toran and Tsukiji [21] initiated the study of depths in random circuits. The model is that of uniform random circuits [19]: a random circuit is built by iterative addition of gates, each gate randomly choosing k inputs among the outputs of the gates already present. The model has been further studied by Arya, Golin and Mehlhorn [4] and Tsukiji and Xhafa [36]. Writing D_x for the depth of the gate reached at step x , Tsukiji and Xhafa [36] proved that the depth of a random circuit of n gates, $\max\{D_x : x \leq n\}$, is asymptotic to $ke \log n$ in probability (\log denotes the natural logarithm).

In fact, in a random circuit, many distinct directed paths may link two gates, and one can define different notions of distances. Devroye and Janson [18] started the systematic analysis of these distances. Among those, the distances defined by paths that are only allowed to look one step ahead (‘greedy’ distances) were studied by Mahmoud [29] and Devroye, Fawzi and Fraiman [20]. These greedy distances have permitted us to quantify the effect of the ‘power of choice’ for depths in random trees: like a gate, each node has k potential contacts, but only attaches to the most desirable according to a measure of

optimality. The models in which the choice of each node is made only according to the labels of its potential contacts has been studied by Mahmoud [29] and Devroye, Fawzi and Fraiman [20]. D’Souza, Krapivsky and Moore [23] studied more general rules of growing trees, where the choice of each node might depend on the degree or the distance to the root of the potential ancestors.

Unfortunately, the distance that has the most meaning in terms of performance of circuits cannot be defined in a greedy way: the number of layers of a circuit depends on the *maximum* length of a path between an output and an input. In this paper, we study this distance, and hence the number of layers required to evaluate the random circuit. Aside from the depth of the entire circuit studied by Arya, Golin and Mehlhorn [4] and Tsukiji and Xhafa [36], the typical depth D_n and the minimum depth of a gate are also of interest. For the latter quantity, since $\min\{D_x : 0 \leq x \leq n\} = 0$, we study $\min\{D_x : x \geq n/2\}$ to estimate the concentration of the depths in the circuits. Our main results are laws of large numbers for D_n (Theorem 4.1) and $\min\{D_x : x \geq n/2\}$ (Theorem 5.1). In particular, for the model of uniform random circuits of [36] and $k = 2$, we show that

$$\frac{D_n}{\log n} \rightarrow \lambda \quad \text{and} \quad \frac{\min_{n/2 \leq x \leq n} D_x}{\log n} \rightarrow \lambda_{\min} = \frac{\lambda}{2} \tag{1.1}$$

in probability, where $\lambda = 4.31107\dots$ is the only solution to the equation $\lambda \log(2e/\lambda) = 1$ that is greater than one. In doing so, we prove a conjecture of Devroye and Janson [18] about the value of λ , and we also identify λ_{\min} for which they did not make a guess. This completes the study of some natural distances in *uniform* random circuits started in [36] and [18]. In fact, our results apply to a more general class of random DAGs where the parent nodes of x are not necessarily chosen uniformly from $\{0, \dots, x - 1\}$ [20] (we will be more precise below). In general, the limit constants are characterized uniquely as the root of some (often implicit) equation that depends on the precise model of attachment.

The problem of distances in random DAGs is related to minima in branching random walks. The relation between distances in *random trees* such as random recursive trees, or binary search trees and minima in branching random walks, has been shown by Pittel [32] and Devroye [16]. Distances in *random circuits* can also be studied using the simpler setting of branching random walks, but the relation is much more intricate because the circuits do not have a real tree structure. Although we do not use the results for branching random walks directly, we think that the reader may find it useful to warm up with this simpler model. Moreover, the ideas leading to the tail bounds for minima in branching random walks presented in Theorem 3.2 also underlie the main argument behind our analysis of the behaviour of $\min_{n/2 \leq x \leq n} D_x$.

Further bibliographical remarks. For a slightly different random circuit model, Mahmoud and Tsukiji have investigated the asymptotic behaviour of the number of outputs, that is, gates that do not feed into any other gate [35, 30]. The profile of the related model of k -trees has been studied by Darrasse, Hwang, Bodini and Soria [14]. Depths in random circuits are also used by Codenotti, Gemmell and Simon [13] in relation to parallel computation time.

Outline of the paper. The model of a random DAG is formally introduced in Section 2. In Section 3, we study tail bounds for minima of branching random walks. The results presented there help us understand why the values for the limiting constants in the law of large numbers in (1.1) are what they are. To the best of our knowledge, the exponential rates in the tail estimates we derive for the branching random walk were not previously known. Section 4 is then devoted to the study of the typical number of edges D_n on the longest path between n and the root of a random DAG. Finally, the minimal distance between a node (with sufficiently large label) and the root is analysed in Section 5. Although we do not think that the results for branching random walks are obvious, the main difficulty consists, in the DAG model, in dealing with the intricate dependence between the different paths up the root that originate not only from a single node, but also from different nodes.

2. Definitions and notation

We consider the more general model of scaled attachment random recursive DAGs (k -SARRD) introduced in Devroye, Fawzi and Fraiman [20]. We are given a random variable X , with support in $[0, 1)$. In a k -SARRD with attachment X (or (X, k) -SARRD), every node x chooses k parents: $[xX_{x,1}]$, $[xX_{x,2}]$, \dots , $[xX_{x,k}]$, where $X_{0,1}, \dots, X_{0,k}, \dots, X_{n,1}, \dots, X_{n,k}$ are independent copies of X . In other words, the random variable $X_{x,p}$ determines the p th parent of node x . A random DAG of size n is then composed of the root 0 and the nodes $1, \dots, n$ and the edge set binding each node to its k parents. When X is uniform, one obtains the uniform recursive circuit (k -URRD) that is the subject of Tsukiji and Xhafa [36], Devroye and Janson [18] and Arya, Golin and Mehlhorn [4].¹

In the following, we always reserve k for the number of parents in the DAG. We let D_n represent the length of the longest path from node n to node 0.

We now introduce some notation to describe the DAG. The set of finite words on the alphabet $\mathcal{A} = [k] = \{1, \dots, k\}$ is denoted by

$$\mathcal{U} := \bigcup_{m \geq 0} \mathcal{A}^m.$$

The set \mathcal{U} is naturally endowed with a partial order: we write $v \leq u$ if v is a prefix of u . We will also think of \mathcal{U} as a k -ary tree, where $v \leq u$ if v is an ancestor of u .

For a node x and a string $s \in \mathcal{U}$, $L(x, s)$ is the label of the ancestor of x obtained by following the path labelled by s . For example, $L(x, 1)$ is the first parent of node x . Concatenation of strings (corresponding to paths) is denoted by \cdot . Note that in our model we have $L(x, s \cdot p) = [L(x, s)X_{L(x, s), p}]$. For a string s and $i \geq 0$, let \underline{s}_i be the string composed of the first i letters of s . In the special case where $i = |s| - 1$, we write $s^- := \underline{s}_{|s|-1}$ for the string where the last letter is dropped. The last letter of the word s is denoted by \bar{s} so that $s = s^- \cdot \bar{s}$.

Of course, the paths up the root corresponding to two strings s and s' might intersect, so that even when s and s' have no non-trivial common prefix, the random variables $L(n, s)$

¹ One can also consider close variants of the model where the parents are chosen without replacement and the DAG has k roots [19].

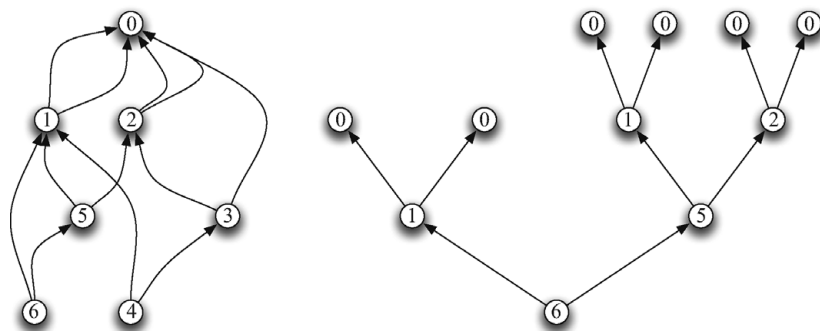


Figure 1. A binary dag on $\{0, 1, \dots, 6\}$ and the corresponding exploration tree from node 6.

and $L(n, s')$ are not independent in general (see Figure 1). However, the dependence of random variables $L(n, s)$, $s \in \mathcal{U}$ is ‘essentially caused’ by common prefixes of the strings s . We will justify this informal fact below, when we prove our main results in Sections 4 and 5. Understanding the lengths of the paths originating from a node n hence somewhat reduces to the analysis of the evolutions of the labels $L(n, s)$ in the k -ary tree \mathcal{U} . Furthermore, the reader should be intuitively convinced that the labels should satisfy the approximation

$$L(n, s) \approx n \prod_{u \leq s} X_u,$$

where $X_u, u \in \mathcal{U}$ are i.i.d. copies of X , which makes the connection between paths in X -DAGS and a branching random walk with step distribution $Y = -\log X$. Since we are interested in *maximum* lengths of paths, we should naturally study paths along which the $L(x, s)$ stays large; along these paths, the branching random walk should be small. This leads us to the study of asymptotics for *minima* in a branching random walk.

3. Large deviations for extremes of a branching random walk

In this section, we consider branching random walks with constant branching factor k and step distribution Y . Let $Y_u, u \in \mathcal{U} \setminus \emptyset$ be independent and identically distributed random variables distributed as Y . It is convenient to assume that $Y_\emptyset = 0$. Then, define the *position* of a word $u \in \mathcal{U}$ by $S_u = \sum_{v \leq u} Y_v$. We are interested in the minimum label over the k^m nodes at distance m from the root, $M_m = \min\{S_u : u \in \mathcal{A}^m\}$.

Asymptotics for M_m depend on rate functions associated with the step distribution Y . Let Λ be the cumulant generating function for the step distribution Y and define its convex dual Λ^* [15, 34] by

$$\Lambda(\lambda) = \log \mathbb{E}[e^{\lambda Y}] \quad \text{and} \quad \Lambda^*(z) = \sup_{\lambda \in \mathbb{R}} \{\lambda z - \Lambda(\lambda)\},$$

where \log denotes the natural logarithm. The following classical theorem describes the first-order asymptotics of M_m . In the entire document, we will always have $Y \geq 0$, so that the moment condition $\Lambda(\lambda) < \infty$ for some $\lambda < 0$ will always be satisfied.

Theorem 3.1 ([26, 28, 7]). *Suppose Y is such that $\Lambda(\lambda) < \infty$ for some $\lambda < 0$. Let $\gamma = \inf\{z \leq \mathbb{E}[Y] : \Lambda^*(z) < \log k\}$. Then, γ is finite and with probability 1,*

$$\lim_{m \rightarrow \infty} \frac{M_m}{m} = \gamma.$$

Moreover, if $\mathbb{E}[Y] < \infty$, by the dominated convergence theorem, we have

$$\lim_{m \rightarrow \infty} \frac{\mathbb{E}[M_m]}{m} = \gamma.$$

Results on the minima in branching random walks have many applications in the study of random trees; see Devroye [17] for a survey. Here, we are interested in tail bounds for the distribution of M_m/m . McDiarmid [31] and Addario-Berry and Reed [1] have proved general exponential tail bounds for the deviations of M_m/m ; however, the exponential rates in the bounds there are not optimal, and our setting requires us to identify them. In the following, we assume that Y is asymptotically exponential in the following sense: there exists a constant $\alpha \in (0, \infty)$ such that

$$\frac{\log \mathbb{P}\{Y \geq x\}}{x} \rightarrow -\alpha, \tag{3.1}$$

as $x \rightarrow \infty$.

The two sides of the distribution of M_m/m have in general very different behaviour. Quite intuitively, if one wants to make the minimum value at level m smaller, it suffices to modify the random variables Y_u on a single path of length m , so one expects that the tail should have exponential tails with a scale of m on the left. On the other hand, to make M_m larger, one needs to modify *all* k^m paths of length m , and it is not immediately clear how one should do this in order to optimize the probability. We will show that when the random variable Y has exponential tails in the sense of (3.1), it is essentially best to modify the random variables on the first level of the tree by a huge amount. We now turn to formalizing this intuition.

Theorem 3.2 (Right tail). *Suppose that $Y \geq 0$, $\log \mathbb{P}\{Y \geq x\} = -\alpha x + o(x)$, as $x \rightarrow \infty$, for $\alpha \in (0, \infty)$. Let $\varepsilon, \delta > 0$. Then there exist constants $c, c' \in (0, \infty)$ such that, for all m large enough,*

$$c \cdot e^{-kx\varepsilon(1+\delta)m} \leq \mathbb{P}\{M_m \geq (\gamma + \varepsilon)m\} \leq c' \cdot e^{-kx\varepsilon(1-\delta)m}.$$

Proof. We start with the lower bound. Consider the k nodes connected to the root. Each one of these nodes is the root of a tree of depth $m - 1$. In order to have $M_m \geq (\gamma + \varepsilon)m$, it suffices that all k (independent) trees have a minimum $M_{m-1} \geq (\gamma - \frac{\varepsilon\delta}{2})m$ and all the steps between the root and its children are such that $Y_1, \dots, Y_k \geq (\varepsilon + \frac{\varepsilon\delta}{2})m$. Therefore,

$$\begin{aligned} \mathbb{P}\{M_m \geq (\gamma + \varepsilon)m\} &\geq \mathbb{P}\{M_{m-1} \geq (\gamma - \varepsilon\delta/2)m\}^k \cdot \mathbb{P}\{Y \geq (\varepsilon + \varepsilon\delta/2)m\}^k \\ &\geq c \cdot e^{-kx(\varepsilon + \frac{\varepsilon\delta}{2})m + o(m)} \\ &\geq c \cdot e^{-kx\varepsilon(1+\delta)m}, \end{aligned}$$

for m large enough, by Theorem 3.1 and our assumption on the tail of Y .

We now prove the upper bound. Let $h = \lfloor 10 \log_k m \rfloor$. We start by proving an exponential tail bound for S_u where $u \in \mathcal{A}^h$. Notice first that

$$\mathbb{E}[e^{\alpha(1-\delta)Y}] = \int_0^\infty \mathbb{P}\{e^{\alpha(1-\delta)Y} \geq t\} dt = \int_0^\infty t^{-1/(1-\delta)+o(1)} dt < \infty.$$

Thus, by Markov's inequality, we obtain

$$\mathbb{P}\{S_u \geq (1 - \delta)\epsilon m\} \leq \mathbb{E}[e^{\alpha(1-\delta)Y}]^h e^{-\alpha(1-\delta)(1-\delta)\epsilon m} \leq e^{-(1-\delta)^3 \alpha \epsilon m},$$

for m large enough. We can now prove exponential tails for M_m . Using a decomposition according to the values of some nodes at level $h + 1$, we obtain

$$\begin{aligned} \mathbb{P}\{M_m \geq (\gamma + \epsilon)m\} &\leq \mathbb{P}\{\forall p \in [k], \exists u \in \mathcal{A}^h : S_{p \cdot u} > (1 - \delta)\epsilon m\} \\ &\quad + \mathbb{P}\{\exists p \in [k], \forall u \in \mathcal{A}^h : S_{p \cdot u} \leq (1 - \delta)\epsilon m \text{ and } M_m \geq (\gamma + \epsilon)m\}. \end{aligned} \tag{3.2}$$

For the first event in (3.2) to hold, there must be k paths u^1, \dots, u^k such that, for the k disjoint paths $p \cdot u^k$ at level $h + 1$, the values are rather large:

$$\begin{aligned} \mathbb{P}\{\forall p \in [k], \exists u \in \mathcal{A}^h : S_{p \cdot u} > (1 - \delta)\epsilon m\} &\leq \mathbb{P}\{\forall p \in [k], \exists u \in \mathcal{A}^h : S_{p \cdot u} > (1 - \delta)\epsilon m\} \\ &\leq k \cdot k^{kh} e^{-k(1-\delta)^3 \alpha \epsilon m}. \end{aligned} \tag{3.3}$$

On the other hand,

$$\begin{aligned} &\mathbb{P}\{\exists p \in [k], \forall u \in \mathcal{A}^h : S_{p \cdot u} \leq (1 - \delta)\epsilon m \text{ and } M_m \geq (\gamma + \epsilon)m\} \\ &= \mathbb{P}\{\exists p \in [k], \forall u \in \mathcal{A}^h : [S_{p \cdot u} \leq (1 - \delta)\epsilon m \text{ and } \forall v \in \mathcal{A}^{m-h-1} : S_{p \cdot u \cdot v} \geq (\gamma + \epsilon)m]\} \\ &\leq \mathbb{P}\{M_{m-h-1} \geq (\gamma + \delta\epsilon)m\}^{k^h} \\ &\leq c_0^{k^h}, \end{aligned}$$

for m large enough and some constant $c_0 < 1$ (using Theorem 3.1). As a consequence, for fixed $\delta, \epsilon > 0$, we have for all m large enough

$$\mathbb{P}\{M_m \geq (\gamma + \epsilon)m\} \leq k^{kh+1} e^{-k(1-\delta)^3 \alpha \epsilon m} + c_0^{k^h}.$$

By our choice of $h = \lfloor 10 \log_k m \rfloor$ we have $k^h \geq m^{10}/k$. Finally, since $\delta > 0$ was arbitrary, the desired upper bound follows. □

At this point, we should comment on tail bounds for minima in branching random walks. Although the estimates for the upper tail of M_m in Theorem 3.2 are tight, we must mention that the strength of the result does not really compare to recent results on the precise location of minima in branching random walks. Indeed, Theorem 3.2 only provides decent estimates for $\mathbb{P}\{M_m \geq \gamma m + t\}$ for positive t , while $\mathbb{E}[M_m] = \gamma m - \beta \log m + O(1)$ [1]; for the study of branching random walks the tail bounds of interest are in fact those for $\mathbb{P}\{M_m \geq \mathbb{E}[M_m] + t\}$ or $\mathbb{P}\{M_m \geq \gamma m - \beta \log m + t\}$. See also the related results about tightness and weak convergence for M_m [8, 27, 6, 2, 9]. Equivalent results for the height of random trees (binary search trees and m -ary search trees) were proved in Reed [33], Drmota [22] and Chauvin and Drmota [10]. Similar comments apply to our estimates on the left tail which follow.

The left tail for the minimum at level m in a branching random walk is essentially governed by the level m : to change the minimum M_m , it suffices to change any single one of the k^m paths of length m , and this way of proceeding is essentially optimal.

Theorem 3.3 (Left tail). *Suppose that there exists $\lambda < 0$ such that $\mathbb{E}[e^{\lambda Y}] < \infty$. Then, for any $\delta > 0$, there exist constants c, c' such that*

$$ck^m e^{-m(\Lambda^*(\gamma-\varepsilon)+\delta)} \leq \mathbb{P}\{M_m \leq (\gamma - \varepsilon)m\} \leq c'k^m e^{-m\Lambda^*(\gamma-\varepsilon)}.$$

Proof. The upper bound follows easily from the union and Chernoff’s bound [11, 15]:

$$\begin{aligned} \mathbb{P}\{M_m \leq (\gamma - \varepsilon)m\} &\leq k^m \mathbb{P}\left\{\sum_{i=1}^m Y_i \leq (\gamma - \varepsilon)m\right\} \\ &\leq k^m e^{-n\Lambda^*(\gamma-\varepsilon)}. \end{aligned}$$

The lower bound is proved using a branching process argument. Let $L \geq 1$ be an arbitrary integer to be chosen later. The potential individuals of our branching process are the nodes of \mathcal{U} at levels $iL, i \geq 0$. A node u is called *good* if it is either the root (i.e., it lies at level 0) or it lies at level $(i + 1)L$ for some $i \geq 0$, its ancestor v at level iL is good and $S_u - S_v \leq (\gamma - \varepsilon)L$. Let Z_i denote the number of good nodes at level iL in the tree ; $\{Z_i, i \geq 0\}$ is a Galton–Watson process. Clearly, if there is a good node at level iL , i.e., $Z_i > 0$ then $M_{iL} \leq (\gamma - \varepsilon)iL$. As a consequence,

$$\mathbb{P}\{M_{iL} \leq (\gamma - \varepsilon)iL\} \geq \mathbb{P}\{Z_i > 0\}. \tag{3.4}$$

By the second moment method [3], more precisely the Chung–Erdős inequality [12], we have

$$\mathbb{P}\{Z_i > 0\} \geq \frac{\mathbb{E}[Z_i]^2}{\mathbb{E}[Z_i^2]} = \frac{\mathbb{E}[Z_i]^2}{\text{Var}(Z_i) + \mathbb{E}[Z_i]^2}. \tag{3.5}$$

For the Galton–Watson process $\{Z_i, i \geq 0\}$, we have

$$\mathbb{E}[Z_i] = \mu^i \quad \text{and} \quad \text{Var}(Z_i) = \sigma^2 \frac{1 - \mu^i}{1 - \mu} \mu^{i-1},$$

where $\mu := \mathbb{E}[Z_1]$ and $\sigma^2 := \text{Var}(Z_1)$ [5]. We now move on to choosing L to obtain a good lower bound on the right-hand side in (3.5). The mean number of children of an individual is

$$\mu = \mathbb{E}[Z_1] = k^L \mathbb{P}\left\{\sum_{j=1}^L Y_j \leq (\gamma - \varepsilon)L\right\} = k^L e^{-L\Lambda^*(\gamma-\varepsilon)+o(L)},$$

as $L \rightarrow \infty$. Note that since $\varepsilon > 0$, we have $\mu < 1$ for any $L \geq 1$. So, for any $\delta > 0$ there exist L large enough that

$$\mu \geq k^L \cdot e^{-L(\Lambda^*(\gamma-\varepsilon)+\delta)}. \tag{3.6}$$

Fixing this value for L , and since $\mu < 1$, (3.5) yields $\mathbb{P}\{Z_i > 0\} \geq C\mu^i$, for some fixed constant $C > 0$ independent of i . More precisely, together with (3.4), we obtain

$$\mathbb{P}\{M_{iL} \leq (\gamma - \varepsilon)iL\} \geq Ck^{iL}e^{-iL(\Lambda^*(\gamma - \varepsilon) + \delta)},$$

which proves the lower bound for all $m = iL$, for some $i \geq 0$. Finally, when $m = iL + r$ for $r \in \{1, \dots, L - 1\}$ it suffices that $M_{iL} \leq (\gamma - \varepsilon)iL$ and $\sum_{j=1}^r Y_j \leq (\gamma - \varepsilon)r$, where $Y_j, j \geq 1$ are i.i.d. copies of Y . Since

$$\inf_{1 \leq r < L} \mathbb{P}\left\{\sum_{i=1}^r Y_i \leq (\gamma - \varepsilon)r\right\} > 0,$$

the result follows easily for general integers m . □

On the limiting constants for the DAG model. Before proceeding to the analysis of the circuit model, we use the tail bounds we have just devised to provide rough arguments that explain the values of the constants in our laws of large numbers. (The symbol \approx is used in a very informal way.) The first main idea is that the dependence is small enough that the upper bound given by the union bound essentially yields the correct constant. By Theorem 3.3 we (should) have

$$\begin{aligned} \mathbb{P}\{D_n \geq c \log n\} &\approx \mathbb{P}\{M_{\lfloor c \log n \rfloor} \leq \log n\} \\ &\approx k^{c \log n} \cdot e^{-c \log n \Lambda^*(1/c)} \\ &\approx n^{c(\log k - \Lambda^*(1/x))}. \end{aligned}$$

In particular, we have $\lim_{n \rightarrow \infty} \mathbb{P}\{D_n \geq c \log n\} = 0$ for every c such that $\Lambda^*(1/c) > \log k$. This suggests that we might have $D_n \sim \lambda_k \log n$ in probability for $\lambda_k = \sup\{x : \Lambda^*(1/c) < \log k\}$. The proof of this fact is the topic of Section 4. Similarly, using Theorem 3.2 we obtain

$$\begin{aligned} \mathbb{P}\left\{\min_{x \geq n/2} D_x \leq c \log n\right\} &\leq n\mathbb{P}\{D_{n/2} \leq c \log n\} \\ &\approx \mathbb{P}\{M_{\lfloor c \log n \rfloor} \geq \log(n/2)\} \\ &\approx ne^{-kxc \log n(1/c - \gamma)} \\ &\approx n^{1 - k\alpha(1 - c\gamma)}, \end{aligned}$$

so it should be the case that

$$\frac{\min_{n/2 \leq x \leq n} D_x}{\log n} \rightarrow \left(1 - \frac{1}{k\alpha}\right)\lambda_k \quad \text{in probability.}$$

We prove this formally (when the limit constant above is non-negative, $\min_x D_x$ is non-negative) in Section 5. Similarly, we can see that Theorem 3.3 leads to a correct guess that $\max_{x \leq n} \frac{D_x}{\log n} \rightarrow ke$ in probability when the attachment distribution is uniform [36].

4. Longest paths in random k -DAGs: typical distance D_n

We start by studying the length D_n of the longest path from node n to the root of the tree. The typical distance D_n is studied using methods similar to the study of the typical

Table 1. Some numerical values for the constant λ_k for the case of uniform recursive circuits.

k	2	3	4	5	10
λ_k	4.311070407...	7.080786915...	9.820440021...	12.55049054...	26.16346184...

shortest path distance in [18]. Let Λ^* denote the rate function associated with the random variable $Y = -\log X$. Define

$$\lambda_k = \sup\{z \geq 1/\mathbb{E}[-\log X] : \Lambda^*(1/z) \leq \log k\}. \tag{4.1}$$

Note that we will consider attachment distributions with bounded density f and thus

$$\mathbb{E}[-\log X] = - \int_0^1 (\log x)f(x)dx < \infty.$$

Moreover, as $X < 1$, $\mathbb{E}[-\log X] > 0$.

Theorem 4.1. *Let $k \geq 1$. Suppose $X \in (0, 1)$ has a bounded density. Then, the longest path from node n to the root in a k -SARRD with attachment X satisfies, as $n \rightarrow \infty$,*

$$\frac{D_n}{\log n} \rightarrow \lambda_k$$

in probability.

The theorem follows from Lemma 4.2 and Lemma 4.3 below. From Theorem 4.1, we easily obtain the asymptotics for the typical distance in the uniform circuit model discussed by Arya, Golin and Mehlhorn [4] and Tsukiji and Xhafa [36]. Indeed, if X is uniform $(0, 1)$, then the parents of a node i are i.i.d. uniform in $\{0, 1, \dots, i - 1\}$. In this case, $-\log X \sim \text{Exponential}(1)$ and $\Lambda^*(z) = z - 1 - \log z$. Then λ_k is the only solution of the equation in z

$$\left(\frac{ke}{z}\right)^z = e$$

that is at least 1. Numerical values are presented in Table 1.

4.1. Upper bound on D_n

Lemma 4.2. *For any $c > \lambda_k$, there exist C and $\eta > 0$ such that, for all n large enough, we have*

$$\mathbb{P}\{D_n \geq c \log n\} \leq Cn^{-\eta}.$$

Proof. To simplify the notation, we prove $\mathbb{P}\{D_n \geq c \log n + 2\} \leq Cn^{-\eta}$ for all $c > \lambda_k$, which is an equivalent statement. Let R_n denote the number of hops to the root for a random path from n , i.e., a path that chooses a uniformly random edge at every step. By the union bound, we get

$$\mathbb{P}\{D_n \geq c \log n + 2\} \leq k^{c \log n + 2} \cdot \mathbb{P}\{R_n \geq c \log n + 2\}. \tag{4.2}$$

Let X_1, X_2, \dots be i.i.d. copies of X . Then, to bound the right-hand side, we let $t = \lceil c \log n \rceil$ and apply Markov's inequality:

$$\begin{aligned}
 \mathbb{P}\{R_n \geq c \log n + 2\} &\leq \mathbb{P}\{R_n > t\} \\
 &\leq \mathbb{P}\{nX_1 \dots X_t \geq 1\} \\
 &\leq \inf_{\lambda \geq 0} n^\lambda \mathbb{E}[X_1^\lambda \dots X_t^\lambda] \\
 &= \inf_{\lambda \geq 0} n^\lambda \mathbb{E}[X^\lambda]^t \\
 &= \inf_{\lambda \geq 0} \exp(\lambda \log n + \Lambda(-\lambda)t) \\
 &\leq \inf_{\lambda \geq 0} \exp(\lambda \log n + \Lambda(-\lambda)c \log n) \quad (\text{as } \Lambda(-\lambda) \leq 0) \\
 &\leq \exp\left(-\sup_{\lambda \geq 0} \left\{-\frac{\lambda}{c} - \Lambda(-\lambda)\right\} c \log n\right) \\
 &= \exp(-c\Lambda^*(1/c) \log n). \tag{4.3}
 \end{aligned}$$

Combining (4.2) and (4.3), we get

$$\begin{aligned}
 \mathbb{P}\{D_n \geq c \log n + 2\} &\leq k^2 \exp(c \log n(\log k - \Lambda^*(1/c))) \\
 &= k^2 n^{c(\log k - \Lambda^*(1/c))}.
 \end{aligned}$$

As Λ^* is a decreasing function on $(-\infty, \mathbb{E}[-\log X])$ and $c > \lambda_k$, we have $\Lambda^*(1/c) > \log k$, which completes the proof. □

4.2. Lower bound on D_n

The objective of this section is to prove the following lemma.

Lemma 4.3. *For any fixed $\varepsilon > 0$, we have*

$$\mathbb{P}\{D_n < (1 - \varepsilon)\lambda_k \log n\} \rightarrow 0,$$

as $n \rightarrow \infty$. Furthermore, if there exists $\lambda > 0$ such that $\mathbb{E}[X^{-\lambda}] < \infty$, then there exist constants C and $\eta > 0$ such that, for all n large enough,

$$\mathbb{P}\{D_n < (1 - \varepsilon)\lambda_k \log n\} \leq Cn^{-\eta}. \tag{4.4}$$

We start by describing the proof strategy and each subsection covers part of the proof. Consider the set of ancestors of node n . These nodes form a directed acyclic graph where each node has out-degree k except for the root. We start by considering the following model for the ancestor DAG. Define a family of independent random variables X'_s for all strings s on $\mathcal{A} = \{1, \dots, k\}$ that have the same distribution as X . For a string s of length ℓ , write $s_0, s_1, s_2, \dots, s_\ell$ for the prefixes of s of length $0, 1, \dots, \ell$ respectively. The ancestor of node n obtained by following the path s is labelled

$$L'(n, s) = \lfloor \dots \lfloor \lfloor nX'_{s_1} \rfloor X'_{s_2} \rfloor \dots X'_s \rfloor.$$

This defines a tree (indexed by the strings s) where distinct nodes can share the same label. In this ideal model, the length of the longest path that reaches a node labelled 0 can be obtained with little effort from Theorem 3.1.

In our SARRD, the labels actually correspond to nodes, and there is a unique node labelled i , for $i = 0, \dots, n$. The tree of ancestors of a node is then a DAG, as the in-degree of a node can be more than one. In particular, this creates dependences between the random variables X along different paths even if they have no common prefix.

In order to avoid having to deal with these dependences, we use the following strategy to find a long path. Note that in our ideal setting, a path to the root corresponds to a path to a node labelled 0. Starting at node $V_0 = n$, we look at all the ancestors of n of order ℓ (i.e., ℓ jumps away from n) for some well-chosen large constant ℓ . From all these possible paths, we pick the path (of length ℓ) that reaches the node V_1 with the largest label. Then the same process is repeated for this node until a node with label 0 is reached. This strategy defines a path, and we show that the length of this path can be made as large as $(\lambda_k - \varepsilon) \log n$ with high probability. The advantage of using this method is that only a small portion of the tree is visited, so it is easier to bound the probability of ‘collision’. The path constructed for the ideal model can then be shown to be exactly the same in a real SARRD with high probability.

4.2.1. The ideal tree. In this section, the objective is to obtain a good lower bound on the length of the longest path from node n to a node of label 0 in the ideal tree (Lemma 4.4). More precisely, let V_j denote the label obtained after j steps (each one composed of ℓ jumps); the string defining the path from n to the corresponding node is denoted $S_j \in \mathcal{A}^{\ell j}$. We have $V_0 = n, S_0 = \emptyset$, and for $j \geq 0$,

$$V_{j+1} = \max_{s \in \mathcal{A}^\ell} L'(n, S_j \cdot s) \quad \text{and} \quad S_{j+1} = S_j \cdot \arg \max_{s \in \mathcal{A}^\ell} L'(n, S_j \cdot s). \tag{4.5}$$

The objective is now to show that starting from n and after $q = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$ steps of ℓ jumps, no node of label zero has been reached with probability going to 1. If this happens, we clearly have a path of length at least $q\ell \geq (1 - \varepsilon)\lambda_k \log n$ between n and a node labelled zero.

Lemma 4.4. *Suppose $\mathbb{E}[-\log X] < \infty$. Let $\varepsilon \in (0, 1)$. For $q = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$, we have*

$$\mathbb{P}\{V_q \leq n^{\varepsilon/4}\} \rightarrow 0,$$

as $n \rightarrow \infty$. Furthermore, if there exists $\lambda > 0$ such that $\mathbb{E}[X^{-\lambda}] < \infty$, then there exists $\eta > 0$ such that, for all n large enough, we have

$$\mathbb{P}\{V_q \leq n^{\varepsilon/4}\} \leq n^{-\eta}. \tag{4.6}$$

Proof. Recall that for a string s , we write $s_0 = \emptyset, s_1, s_2, \dots$ for the prefixes of s of length $0, 1, 2, \dots$, respectively. Thus, for $j \geq 0$, we have by definition

$$\begin{aligned} V_{j+1} &= \max\{ \lfloor \lfloor V_j X'_{S_j \cdot s_1} \rfloor \cdots X'_{S_j \cdot s_\ell} \rfloor : s \in \mathcal{A}^\ell \} \\ &\geq V_j \cdot \max\{ X'_{S_j \cdot s_1} \times \cdots \times X'_{S_j \cdot s_\ell} : s \in \mathcal{A}^\ell \} - \ell. \end{aligned}$$

Letting $Z_j^{(\ell)} = \max\{X'_{S_j \mathfrak{S}_1} \times \cdots \times X'_{S_j \mathfrak{S}_\ell} : s \in \mathcal{A}^\ell\}$, we have

$$V_q \geq ((nZ_0^{(\ell)} - \ell)Z_1^{(\ell)} - \ell) \cdots Z_{q-1}^{(\ell)} - \ell \geq n \prod_{j=0}^{q-1} Z_j^{(\ell)} - q\ell.$$

It follows immediately that

$$\begin{aligned} \mathbb{P}\{V_q \leq n^{\varepsilon/4}\} &\leq \mathbb{P}\left\{n \prod_{j=0}^{q-1} Z_j^{(\ell)} - q\ell \leq n^{\varepsilon/4}\right\} \\ &\leq \mathbb{P}\left\{\sum_{j=0}^{q-1} -\log Z_j^{(\ell)} \geq \left(1 - \frac{\varepsilon}{3}\right) \log n\right\}, \end{aligned}$$

for all n large enough. Observe that $-\log Z_0^{(\ell)}$ is simply the minimum at generation ℓ of a branching random walk with increments distributed as $-\log X$. As $\mathbb{E}[-\log X] < \infty$, using Theorem 3.1, we get

$$\lim_{\ell \rightarrow \infty} \frac{\mathbb{E}[-\log Z_0^{(\ell)}]}{\ell} = \frac{1}{\lambda_k}. \tag{4.7}$$

We pick $\ell_0 = \ell_0(\varepsilon)$ so that for $\ell \geq \ell_0$, we have $(1 - \frac{\varepsilon}{8})\lambda_k \mathbb{E}[-\log Z_0^{(\ell)}] \leq \ell$. In the rest of the proof, $\ell = \ell_0$ is fixed and we let Z be a random variable having the same distribution as $Z_0^{(\ell)}$.

We can now use a limit theorem for renewal processes (see for example [24, Chapter 10.2, Theorem 1]; see also [25]). We have the convergence in probability

$$\frac{\max\{q : -\sum_{j=0}^{q-1} \log Z_j^{(\ell)} \leq (1 - \frac{\varepsilon}{3}) \log n\}}{(1 - \frac{\varepsilon}{3}) \log n} \xrightarrow{\mathcal{P}} \frac{1}{\mathbb{E}[-\log Z_0^{(\ell)}]} \geq \frac{(1 - \frac{\varepsilon}{8})\lambda_k}{\ell}, \tag{4.8}$$

where the last inequality follows by our choice of ℓ . Thus, with probability going to one, and for $q = \lceil (1 - \varepsilon)\lambda_k \log n / \ell \rceil$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}\{V_q \leq n^{\varepsilon/4}\} = 0. \tag{4.9}$$

We now move on to the proof of the tail bound (4.6). Note that, in a stochastic sense, we have $Z_0^{(\ell)} \geq X_1 \cdot X_2 \cdots X_\ell$, where X_i are i.i.d. copies of X . Hence, since ℓ is fixed, there exists $\eta > 0$ such that $\mathbb{E}[Z^{-\eta}] < \infty$ so that, by Markov's inequality,

$$\mathbb{P}\left\{n \prod_{j=0}^{q-1} Z_j^{(\ell)} \leq n^{\varepsilon/3}\right\} \leq \left(\frac{\mathbb{E}[e^{-\eta \log Z}]}{e^{\eta \ell (1-\varepsilon/3)/(1-\varepsilon/2)\lambda_k}}\right)^q,$$

for all n large enough. To complete the proof, observe that, as $\eta \rightarrow 0$, we have

$$\begin{aligned} \frac{\mathbb{E}[e^{-\eta \log Z}]}{e^{\eta \ell (1-\varepsilon/3)/(1-\varepsilon/2)\lambda_k}} &= 1 + \eta \left(\mathbb{E}[-\log Z] - \frac{(1 - \varepsilon/3)\ell}{(1 - \varepsilon/2)\lambda_k}\right) + o(\eta) \\ &\leq 1 + \eta \frac{\ell}{\lambda_k} \left(\frac{1}{1 - \frac{\varepsilon}{8}} - \frac{1 - \frac{\varepsilon}{3}}{1 - \frac{\varepsilon}{2}}\right) + o(\eta) < 1, \end{aligned}$$

for η small enough by our choice of ℓ ; it is routine to verify that the second term is indeed negative for any $\varepsilon \in (0, 1)$. Since $q = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$, the tail bound (4.6) follows readily. \square

We call the event studied in this section $A_n(q(n)) := \{V_{q(n)} > n^{\varepsilon/4}\}$. To make the notation lighter, the dependence in n will be omitted.

4.2.2. The real DAG: handling collisions. Recall that the statement in Lemma 4.4 only deals with the ideal tree model. In order to prove that a long enough path also exists in the real DAG, we couple the SARRD and the ideal tree in such a way that with high probability, the path S_q that we exhibited in the ideal tree can also be found inside the SARRD. What we mean here is the following: the path S_q defines a sequence of distinct labels in the ideal tree; in the SARRD the nodes corresponding to these labels are also along a path with high probability.

We use the random variables X'_s that were used in the ideal tree, and new independent random variables $X''_{x,p}$ for $x \in \{0, \dots, n\}$ and $p \in \mathcal{A}$. We start by defining an ordering of strings:

$$s \leq s' \quad \text{if and only if} \quad |s| < |s'| \quad \text{or} \quad |s| = |s'| \quad \text{and} \quad s \leq_{\text{lex}} s' \tag{4.10}$$

where \leq_{lex} is the lexicographic order on the strings on the alphabet \mathcal{A} . This is the *breadth-first* order in the k -ary tree. Then, if one of the paths from n has label x in the ideal tree, let T_x denote the first (or breadth-first) such path,

$$T_x = \min\{s : L'(n, s) = x, |s| \leq n\},$$

where the min is taken with respect to the order in (4.10) and we agree that $\min\{\emptyset\} = \infty$. The labels in the ideal tree are not distinct, and the corresponding vertices cannot directly represent nodes in the DAG. We now ensure that a node x in the SARRD corresponds to the *first* vertex in the ideal tree with label x (if such a vertex exists): for all $p \in \mathcal{A}$ let

$$X_{x,p} = \begin{cases} X'_{T_x \cdot p} & \text{if } T_x \neq \infty, \\ X''_{x,p} & \text{if } T_x = \infty. \end{cases}$$

Clearly, the random variables $X_{x,p}$ are independent and distributed as X , so that they define a proper (X, k) -SARRD.

We now define the event that the sequence of labels along the path S_q is the same in the DAG generated by the variables $(X_{x,p} : 0 \leq x \leq n, 1 \leq p \leq k)$ and in the ideal tree generated by $(X'_{x,p} : 0 \leq x \leq n, 1 \leq p \leq k)$:

$$B(q) = \{L'(n, s) = L(n, s) : s \leq S_q\}.$$

In particular, if $B(q)$ and $A(q)$ both hold, then the path S_q does not reach the root since its minimum label is at least $n^{\varepsilon/4}$, and since S_q is also a path in the DAG we have $D_n \geq q\ell$. Write E^c to denote the complement of an event E . Then, since $q = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$, we have

$$\begin{aligned} \mathbb{P}\{D_n < (1 - \varepsilon)\lambda_k \log n\} &\leq \mathbb{P}\{A(q)^c \cup B(q)^c\} \\ &= \mathbb{P}\{A(q)^c\} + \mathbb{P}\{B(q)^c \cap A(q)\}. \end{aligned} \tag{4.11}$$

Lemma 4.4 bounds the probability that $A(q)$ does not hold, so that to prove Lemma 4.3, it suffices to bound the probability that $B(q)^c \cap A(q)$ occurs.

Lemma 4.5. *Let $\varepsilon \in (0, 1)$ and let $q = q(n) = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$. Then, there exists $\eta > 0$ such that, for n large enough,*

$$\mathbb{P}\{A(q) \cap B(q)^c\} \leq n^{-\eta}.$$

Proof. In the following, we write $n_1 = n^{\varepsilon/4}$. Recall that s^- denotes the prefix of s of length $|s| - 1$. If the path S_q is a breadth-first path in the ideal tree (in the sense that each of its vertices is the first with its label) then our coupling ensures that $B(q)$ occurs. So for $B(q)$ to fail, there must be a node in S_q that is not the first with its label. We have

$$\begin{aligned} \mathbb{P}\{B(q)^c \cap A(q)\} &\leq \mathbb{P}\{\exists j < q, |s| \leq \ell : T_{L'(V_j, s)} \neq s, T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\} \\ &\leq \sum_{j=0}^{q-1} \mathbb{P}\{\exists s : |s| \leq \ell, T_{L'(V_j, s)} \neq s, T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\}. \end{aligned} \tag{4.12}$$

For any fixed $j < q$, we have by the union bound

$$\begin{aligned} &\mathbb{P}\{\exists s, |s| \leq \ell : T_{L'(V_j, s)} \neq s, T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\} \\ &\leq \sum_{|s| \leq \ell} \mathbb{P}\{\exists s_0 < s : L'(V_j, s_0) = L'(V_j, s), T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\} \\ &\leq \sum_{|s| \leq \ell} \sum_{s_0 < s} \mathbb{P}\{L'(V_j, s_0) = \lfloor X'_s L'(V_j, s^-) \rfloor, T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\}. \end{aligned} \tag{4.13}$$

We now condition on $L'(V_j, s^-)$ and on $L'(V_j, s_0)$ and use the independence between the random variable X_s and $(L'(V_j, s^-), L'(V_j, s_0))$:

$$\begin{aligned} &\mathbb{P}\{L'(V_j, s_0) = \lfloor X'_s L'(V_j, s^-) \rfloor, T_{L'(V_j, s^-)} = s^-, V_q \geq n_1\} \\ &\leq \sum_{u, v \leq n} \mathbb{P}\{v = \lfloor uX'_s \rfloor, L'(V_j, s^-) = u, L'(V_j, s_0) = v, L'(V_j, s^-) \geq n_1\} \\ &\leq \sup_{u \geq n_1, v \leq n} \mathbb{P}\{v = \lfloor uX'_s \rfloor\}. \end{aligned}$$

Since X'_s has a density bounded by b , we have, for any $v \leq n$ and $u \leq n_1$,

$$\mathbb{P}\{v = \lfloor uX'_s \rfloor\} = \mathbb{P}\{v \leq uX'_s < v + 1\} \leq \frac{b}{u} \leq \frac{b}{n_1}.$$

Thus, going back to equations (4.12) and (4.13), since there are at most $qk^{2(\ell+1)}$ terms, each one at most b/n_1 , we obtain

$$\mathbb{P}\{B(q)^c \cap A(q)\} \leq \frac{bqk^{2(\ell+1)}}{n_1},$$

which readily yields the result since $n_1 = n^{\varepsilon/4}$, $q = \lceil (1 - \varepsilon)\lambda_k \log(n)/\ell \rceil$ and b, k, ℓ are fixed constants. □

Thus, recalling Lemma 4.4 and (4.11), we have

$$\lim_{n \rightarrow \infty} \mathbb{P}\{D_n < (1 - \varepsilon)\lambda_k \log n\} = 0.$$

In the case when the attachment distribution X is such that $\mathbb{E}[X^{-\lambda}] < \infty$ for some $\lambda > 0$, there exists $\eta > 0$ such that, for all n large enough, the stronger bound

$$\mathbb{P}\{D_n < (1 - \varepsilon)\lambda_k \log n\} \leq n^{-\eta}$$

holds, which finishes the proof of Lemma 4.3.

5. Longest paths in random k -DAGs: minimum distance $\min_{n/2 \leq x \leq n} D_x$

In this section, we assume that the attachment distribution X has a bounded density and satisfies

$$\mathbb{P}\{X \leq t\} = t^{\alpha+o(1)}, \tag{5.1}$$

as $t \rightarrow 0$, for some $\alpha \in (0, \infty)$. Note that this implies that

$$\mathbb{E}[X^{-\alpha/2}] = \int_0^\infty \mathbb{P}\{X^{-\alpha/2} \geq t\} dt < \infty.$$

The lower tails of the step size should clearly influence the distances; the present setting with an underlying branching structure points towards dependence in (5.1). The value of α determines the value of the minimum distance for a random recursive (k, X) -DAG. Define the constant

$$\beta := \max\left(1 - \frac{1}{k\alpha}, 0\right). \tag{5.2}$$

Theorem 5.1. *Let $k \geq 2$ and X be as above. The minimum longest path distance in a random recursive (k, X) -DAG satisfies*

$$\frac{\min_{n/2 \leq x \leq n} D_x}{\log n} \rightarrow \beta \lambda_k,$$

in probability, where $\lambda_k = \sup\{z : \Lambda^(1/z) \leq \log k\}$ is the constant defined in (4.1).*

In the case where $\mathbb{P}\{X \leq t\} \geq t^{\alpha+o(1)}$ for any $\alpha > 0$, we obtain a limit of 0. In fact, the upper bound in Lemma 5.2 only uses a lower bound on $\mathbb{P}\{X \leq t\}$ for t in the neighbourhood of 0. Similarly, if $\mathbb{P}\{X \leq t\} \leq t^{\alpha+o(1)}$ for any $\alpha > 0$, the limit becomes λ_k as the lower bound of Lemma 5.3 only uses an upper bound on $\mathbb{P}\{X \leq t\}$ for t in the neighbourhood of 0.

For $X \sim \text{uniform}[0, 1)$, we have $\alpha = 1$ and $\beta = 1 - 1/k$. This yields $\min_{n/2 \leq x \leq n} D_x = (k - 1)/k \cdot \lambda_k \log n$ for uniform random recursive DAGs. In particular, for $k = 2$, we obtain

$$\frac{D_n}{\log n} \rightarrow \lambda_2 \quad \text{and} \quad \frac{\min_{n/2 \leq x \leq n} D_x}{\log n} \rightarrow \frac{\lambda_2}{2},$$

where $\lambda_2 = 4.31107\dots$. This completes the table of asymptotic properties of different natural distances in uniform random DAGs [18, Table 1].

The reader will easily be convinced when reading the proof that the result remains unchanged if one considers $\min\{D_i : \delta n \leq i \leq n\}$, for some $\delta > 0$. We keep the current statement for simplicity.

5.1. Upper bound on $\min_{n/2 \leq x \leq n} D_x$

Lemma 5.2. *For any $\varepsilon > 0$, there exists $\eta > 0$ such that, for all n large enough,*

$$\mathbb{P}\left\{ \min_{n/2 \leq x \leq n} D_x \geq (\beta + \varepsilon)\lambda_k \log n \right\} \leq n^{-\eta}.$$

Proof. The strategy is to show that there exists a node $x \in \{[n/2], \dots, n\}$ such that all its parents have label at most $n^{\beta+\varepsilon}$. We then conclude using Theorem 4.1. For $n/2 \leq x \leq n$, we look at the k parents $L(x, 1), \dots, L(x, k)$ of x in the DAG. Since $\beta \geq 1 - \frac{1}{k\alpha}$ and $x \leq n$, we have

$$\begin{aligned} \mathbb{P}\{\max\{L(x, 1), \dots, L(x, k)\} \leq n^{\beta+\varepsilon}\} &= \mathbb{P}\{xX \leq n^{\beta+\varepsilon}\}^k \\ &\geq \mathbb{P}\{X \leq n^{-1/(k\alpha)+\varepsilon}\}^k \\ &\geq n^{-1+k\alpha\varepsilon/2} \end{aligned}$$

for all n large enough, as X satisfies (5.1). Define

$$N = \left\{ x : [n/2] \leq x \leq n, \max_{1 \leq i \leq k} L(x, i) \leq n^{\beta+\varepsilon} \right\}.$$

We have

$$\mathbb{E}[N] \geq \left(n - \left\lfloor \frac{n}{2} \right\rfloor + 1 \right) \cdot n^{-1+k\alpha\varepsilon/2} \geq \frac{n^{k\alpha\varepsilon/2}}{2}. \tag{5.3}$$

Observe that the events $\{\max(L(x, 1), \dots, L(x, k)) \leq n^{1-1/(k\alpha)+\varepsilon}\}$ are independent for different nodes x . So we can also compute the variance of N :

$$\text{Var}(N) \leq \sum_{x=[n/2]}^n \mathbb{P}\{\max\{L(x, 1), \dots, L(x, k)\} \leq n^{\beta+\varepsilon}\} = \mathbb{E}[N].$$

Thus, using the second moment inequality, sometimes called Chung–Erdős inequality [12], and since $\mathbb{E}[N] \rightarrow \infty$ by (5.3), we obtain

$$\mathbb{P}\{N = 0\} \leq \frac{\text{Var}(N)}{\text{Var}(N) + \mathbb{E}[N]^2} \leq \frac{1}{1 + \mathbb{E}[N]} = O(n^{-k\alpha\varepsilon/2}). \tag{5.4}$$

Now, define the random node

$$V = \max\{\{0\} \cup \{x : [n/2] \leq x \leq n, \max(L(x, 1), \dots, L(x, k)) \leq n^{\beta+\varepsilon}\}\}.$$

It only remains to show that such a node V has a small longest path to the root. It is sufficient to bound the following probability, with $\varepsilon < 1$. We have

$$\begin{aligned} &\mathbb{P}\left\{ \min_{n/2 \leq x \leq n} D_x \geq (\beta + 3\varepsilon)\lambda_k \log n + 1 \right\} \\ &\leq \mathbb{P}\{D_V \geq (\beta + 3\varepsilon)\lambda_k \log n + 1, N > 0\} + \mathbb{P}\{N = 0\} \end{aligned}$$

$$\begin{aligned}
 &\leq \sum_{s=1}^k \mathbb{P}\{D_{L(V,s)} \geq (\beta + 3\varepsilon)\lambda_k \log n, N > 0\} + \mathbb{P}\{N = 0\} \\
 &\leq \sum_{s=1}^k \mathbb{P}\{D_{L(V,s)} \geq (\beta + 3\varepsilon)\lambda_k \log n, L(V,s) \leq n^{\beta+\varepsilon}\} + \mathbb{P}\{N = 0\} \\
 &\leq k\mathbb{P}\{D_{\lfloor n^{\beta+\varepsilon} \rfloor} \geq (1 + \varepsilon)(\beta + \varepsilon)\lambda_k \log n\} + \mathbb{P}\{N = 0\} \quad \text{for } \varepsilon < 1.
 \end{aligned}$$

The last inequality holds because we can condition on $L(V, s)$ and use the independence of the events $\{L(V, s) = v\}$ and $\{D_v \geq (1 + \varepsilon)(\beta + \varepsilon)\lambda_k \log n\}$ for any v . We can then conclude using Lemma 4.2 and (5.4). □

5.2. Lower bound on $\min_{n/2 \leq x \leq n} D_x$

The objective of this section is to prove the following lower bound. Note that when $\beta = 0$, no lower bound is needed to prove Theorem 5.1, so that we can safely assume here that $\beta > 0$. In particular, we have $\beta = 1 - \frac{1}{k\alpha}$.

Lemma 5.3. *Suppose that $\beta > 0$. For any $\varepsilon > 0$, there exists $\eta > 0$ such that*

$$\mathbb{P}\left\{ \min_{n/2 \leq x \leq n} D_x \leq (\beta - \varepsilon)\lambda_k \log n \right\} = O(n^{-\eta}).$$

Clearly, it is sufficient to consider $\varepsilon < \beta$; we do so until the end of the proof. By the union bound, we have

$$\begin{aligned}
 \mathbb{P}\left\{ \min_{n/2 \leq x \leq n} D_x \leq (\beta - \varepsilon)\lambda_k \log n \right\} &\leq \sum_{x=\lceil n/2 \rceil}^n \mathbb{P}\{D_x \leq (\beta - \varepsilon)\lambda_k \log n\} \\
 &\leq n \cdot \sup_{n/2 \leq x \leq n} \mathbb{P}\{D_x \leq (\beta - \varepsilon)\lambda_k \log n\}.
 \end{aligned}$$

It suffices to show $\mathbb{P}\{D_x \leq (\beta - \varepsilon)\lambda_k \log n\} = o(1/n)$ uniformly for all $x \in \{\lceil n/2 \rceil, \dots, n\}$. In the rest of the proof we fix $x \in \{\lceil n/2 \rceil, \dots, n\}$.

If $D_x \leq (\beta - \varepsilon)\lambda_k \log n$, this means that all the ancestors a of x of order h (think of $h = \lfloor 10 \log \log n \rfloor$) have depth $D_a \leq (\beta - \varepsilon)\lambda_k \log n - h$. To bound the probability of such an event, we show that there are many distinct ancestors of order h that have labels at least $n^{\beta-\delta}$ (Proposition 5.8, which is based on Lemma 5.7 and 5.5). Then, using the explicit bound in Lemma 4.3, we prove that the probability that none of these ancestors have direct parents that have typical depth is small.

More precisely, we build a small DAG of ancestors of x up to $h = \lfloor 10 \log \log n \rfloor$ generations. We start by showing that many of the ancestors of order h have labels at least $n^{\beta-\delta}$. As a warm-up, we first bound the probability that a single ancestor h levels away from x has a low label.

Lemma 5.4. *Let $x \geq n/2$ be fixed. Let $s \in [k]^h$, where $h = \lceil 10 \log \log n \rceil$ and $u \geq h$. For any $\eta > 0$ and n large enough,*

$$\mathbb{P}\{L(x, s) \leq u\} \leq \left(\frac{u}{n}\right)^{\alpha-\eta}.$$

In particular, for $\delta \in (0, \beta)$ and $u = n^{\beta-\delta}$, we get

$$\mathbb{P}\{L(x, s) \leq n^{\beta-\delta}\} \leq n^{-1/k-2\delta/2}. \tag{5.5}$$

Proof. We have the following bound:

$$\begin{aligned} \mathbb{P}\{L(x, s) \leq u\} &\leq \mathbb{P}\{nX_1 \dots X_h - h \leq u\} \\ &\leq \mathbb{P}\left\{X_1 \dots X_h \leq \frac{2u}{n}\right\} \\ &= \mathbb{P}\{(X_1 \dots X_h)^\gamma > (2u/n)^\gamma\}, \end{aligned}$$

for all sufficiently large n and any $\gamma < 0$. By Markov’s inequality and the independence of X_1, \dots, X_h , it follows that

$$\begin{aligned} \mathbb{P}\{L(x, s) \leq n^{\beta-\delta}\} &\leq (2u/n)^{-\gamma} \cdot \mathbb{E}[X_1^\gamma \dots X_h^\gamma] \\ &= (2u/n)^{-\gamma} \cdot \mathbb{E}[X^\gamma]^h. \end{aligned}$$

Now observe that, for any $\varepsilon > 0$, we have

$$\mathbb{E}[X^{-\alpha+\varepsilon}] = \int_0^\infty \mathbb{P}\{X^{-\alpha+\varepsilon} > t\} dt < \infty,$$

by our assumption on the tail of X in (5.1). Choosing $\gamma = -\alpha + \varepsilon$ for small enough ε , we obtain

$$\mathbb{P}\{L(x, s) \leq u\} \leq (2u/n)^{\alpha-\eta/2} \leq (u/n)^{\alpha-\eta}$$

for n large enough. We obtain (5.5) by choosing η small enough. □

In order to handle the dependence between different paths of the DAG, we bound the number of path intersections far from the root. More precisely, order the strings s according to the order defined in (4.10) (breadth-first order). Define the set of nodes that are ancestors of x along paths indexed by the words $s' < s$: $\mathcal{V}(x, s) = \{L(x, s'), s' < s\}$. Then we say that a path labelled by s *collides* if

$$L(x, s) \in \mathcal{V}(x, s) \quad \text{and} \quad L(x, \underline{s}_t) \notin \mathcal{V}(x, \underline{s}_t), \quad 1 \leq t < |s|,$$

where \underline{s}_t is the prefix of s of length t . Of course, the chance of colliding is greater if the labels are small; for us it will be sufficient to consider the nodes with label at least $n^{\beta-\delta}$. Define the number of paths of length at most h colliding at nodes with label at least $n^{\beta-\delta}$:

$$N_c = |\{s : |s| \leq h, L(x, s) \geq n^{\beta-\delta} \text{ and } s \text{ collides}\}|.$$

Lemma 5.5. *Let $\delta > 0$. For all $n \geq 1$ and all $i \geq 0$, we have*

$$\mathbb{P}\{N_c \geq i\} \leq k^{2i(h+1)} b^i n^{-i(\beta-\delta)}.$$

Proof. Let $i \geq 0$. By definition, we have

$$\begin{aligned} \mathbb{P}\{N_c \geq i\} &= \mathbb{P}\{\exists s_1 < \dots < s_i : s_1, \dots, s_i \text{ collide}\} \\ &\leq k^{i(h+1)} \sup_{s_1 < \dots < s_i} \mathbb{P}\{s_1, \dots, s_i \text{ collide}\}. \end{aligned} \tag{5.6}$$

We prove by induction that for all $i \geq 0$

$$\sup_{s_1 < \dots < s_i} \mathbb{P}\{s_1, \dots, s_i \text{ collide}\} \leq (bk^{h+1}n^{\delta-\beta})^i. \tag{5.7}$$

Clearly, (5.6) and (5.7) together imply the result. The base case, $i = 0$, is clear. Suppose now that $i \geq 1$. Let $s_1 < s_2 < \dots < s_i$. Write $S_i = \{s_1, \dots, s_i\}$ and $S_{i-1} = S_i \setminus \{s_i\}$. Note that S_{i-1} is empty if $i = 1$. We say that a set S collides if all its elements collide. Recall that \bar{s} denotes the last symbol of the word s . Then

$$\begin{aligned} &\mathbb{P}\{s_1, \dots, s_i \text{ collide}\} \\ &= \mathbb{P}\{S_i \text{ collides}\} \\ &= \mathbb{P}\{S_{i-1} \text{ collides, } L(x, s_i) \geq n^{1-1/k-\delta}, L(x, s_i^-) \notin \mathcal{V}(x, s_i^-), L(x, s_i) \in \mathcal{V}(x, s_i)\} \\ &= \sum_{u, W} \mathbb{P}\{S_{i-1} \text{ collides, } L(x, s_i^-) = u, \mathcal{V}(x, s_i) = W, u \notin \mathcal{V}(x, s_i^-), [uX_{u, \bar{s}_i}] \in W\}, \end{aligned}$$

where the sum in the last line ranges on $u \geq n^{1-1/(k\alpha)-\delta}$ and $W \subseteq \{0, \dots, n\}$. In the rest of the proof, W will always be a subset of $\{0, \dots, n\}$ and, to avoid cumbersome notation, we do not always include it.

We claim that for any fixed $W \subseteq \{0, \dots, n\}$ and any u , the events

$$\{S_{i-1} \text{ collides, } L(x, s_i^-) = u, \mathcal{V}(x, s_i) = W, u \notin \mathcal{V}(x, s_i^-)\} \quad \text{and} \quad \{[uX_{u, \bar{s}_i}] \in W\}$$

are independent. The latter event is clearly determined by $X_{u,p}$, $1 \leq p \leq k$. Now consider the other event. In order to determine whether it occurs, it suffices to look at the ancestors $L(x, s')$ of x for $s' < s_i$. If the value of X_{u, \bar{s}_i} is needed to compute one of these ancestors, then the event does not hold because we would then have $u \in \mathcal{V}(x, s^-)$. Otherwise, we can determine whether the event holds without looking at X_{u, \bar{s}_i} .

For $|s| \leq h$, we have $|\mathcal{V}(x, s)| \leq k^{h+1}$ and we can then write

$$\begin{aligned} \mathbb{P}\{S_i \text{ collides}\} &\leq \mathbb{P}\{S_{i-1} \text{ collides}\} \sup_{u \geq n^{\beta-\delta}, |W| \leq k^{h+1}} \mathbb{P}\{[uX] \in W\} \\ &\leq \mathbb{P}\{S_{i-1} \text{ collides}\} \cdot k^{h+1} b n^{-\beta+\delta} \\ &\leq (k^{h+1} b n^{-\beta+\delta})^i, \end{aligned}$$

by the induction hypothesis. This completes the proof. □

Recall that our aim is to prove that, with probability at least $1 - o(1/n)$, x has many distinct ancestors of order h , all of which should have a large enough label. To make this

precise, define the event

$$E_x = \{x \text{ has at least } k^{h-3} \text{ distinct ancestors of order } h \text{ with labels at least } n^{\beta-\delta}\}.$$

We want to prove that $\mathbb{P}\{E_x\} = 1 - o(1/n)$. We first show the following decomposition for E_x . We have already proved that $\mathbb{P}\{N_c \geq 3\} = o(1/n)$, and it will then suffice to bound the probability that the first event in (5.8) below does not occur.

Lemma 5.6. *For every $x \geq n/2$ we have*

$$\{\exists p \in [k], \forall s_p \in [k]^h : s_p[1] = p \text{ and } L(x, s_p) > n^{\beta-\delta}\} \cap \{N_c \leq 2\} \subseteq E_x. \tag{5.8}$$

Proof. The first event in (5.8) ensures that all the ancestors of $L(x, p)$ of order $h - 1$ have labels greater than $n^{\beta-\delta}$ for some $p \in [k]$. These k^{h-1} nodes need not be distinct. Consider the paths in the order defined by (4.10). By definition, if a path does not collide and has no prefix that collides, then its label is distinct from all the previous ones. It follows that only the paths that do have a prefix counted by N_c might not have distinct labels. Here $N_c \leq 2$, and it is simple to see that a pair of paths whose collision maximizes the number of potential duplicates is $p \cdot 2$ and $p \cdot 3$ if $k \geq 3$ and $p \cdot 2$ and $p \cdot 12$ if $k = 2$. In any case, of the k^{h-1} , there are at least k^{h-3} nodes with distinct labels. \square

To complete the proof that $\mathbb{P}\{E_x\} = 1 - o(1/n)$, it only suffices to bound the probability of the first event in (5.8) not occurring.

Lemma 5.7. *Let $\delta \in (0, \beta/4)$. For n large enough,*

$$\mathbb{P}\{\exists s^1, \dots, s^k \in [k]^{h-1} : \forall p \in [k], L(x, p \cdot s^p) \leq n^{\beta-\delta}\} \leq n^{-1-\delta/4}.$$

Proof. Let A_i be the event that $L(x, s^j) \leq n^{\beta-\delta}$ for $j \leq i$. We prove by induction on $i \geq 1$, for n large enough,

$$\mathbb{P}\{A_i\} \leq n^{-i/k-\delta/8}. \tag{5.9}$$

The base case, for $i = 1$, follows from Lemma 5.4. Suppose now that $i \geq 2$. The difficulty in proving the induction step relies on the dependence between the events A_{i-1} and $L(x, s^i) \leq n^{\beta-\delta}$. We introduce the following notation for the paths from node x :

$$P(x, s) = \{L(x, \underline{s}_1), L(x, \underline{s}_2), \dots, L(x, s)\} \text{ and } P_{i-1} = \bigcup_{1 \leq j < i} P(x, s^j).$$

Note that x is not included in the paths. To upper-bound the left-hand side in (5.9), we condition on the first time that the path $P(x, s_i)$ reaches the set P_{i-1} :

$$\mathbb{P}\{A_i\} = \mathbb{P}\{A_i, P(x, s^i) \cap P_{i-1} = \emptyset\} + \sum_{1 \leq t \leq h} \mathbb{P}\{A_i, L(x, \underline{s}_{t-1}^i) \notin P_{i-1}, L(x, \underline{s}_t^i) \in P_{i-1}\}. \tag{5.10}$$

In the following, W and Q will always denote a subset of $\{0, \dots, n\}$; we do not always include it.

(i) Path $P(x, s^i)$ does not collide. The first term in (5.10), on the event $P(x, s^i) \cap P_{i-1} \neq \emptyset$, is the easiest to deal with:

$$\mathbb{P}\{A_i, P(x, s^i) \cap P_{i-1} = \emptyset\} = \sum_{W, Q: W \cap Q = \emptyset} \mathbb{P}\{A_i, P(x, s^i) = Q, P_{i-1} = W\}.$$

As $Q \cap W = \emptyset$, the events

$$\{L(x, s^i) \leq n^{\beta-\delta}, P(x, s^i) = Q\} \quad \text{and} \quad \{L(x, s^j) \leq n^{\beta-\delta} \text{ for } j < i, P_{i-1} = W\}$$

are independent. In fact, the first event is in the sigma-algebra generated by $\{X_{v,p} : v \in Q, p \in [k]\}$ and the second in the one generated by $\{X_{v,p} : v \in W, p \in [k]\}$. Thus, we obtain

$$\begin{aligned} \mathbb{P}\{A_i, P(x, s^i) \cap P_{i-1} = \emptyset\} &\leq \mathbb{P}\{A_{i-1}\} \cdot \mathbb{P}\{L(x, s^i) \leq n^{\beta-\delta}\} \\ &\leq \mathbb{P}\{A_{i-1}\} \cdot n^{-1/k-\alpha\delta/2}, \end{aligned} \tag{5.11}$$

where the last inequality follows from Lemma 5.4.

(ii) Path $P(x, s^i)$ collides. We next look at the terms (5.10) that correspond to cases when there are some collisions, *i.e.*, $P(x, s^i) \cap P_{i-1} \neq \emptyset$. Recall that t is the location of the *first* collision on $P(x, s^i)$. In the following, we write a_t^i for the t th symbol of s^i (so $s^i = a_1^i a_2^i \dots a_{|s^i|}^i$). For $t \in \{1, \dots, h\}$, we have

$$\begin{aligned} &\mathbb{P}\{A_i, L(x, \underline{s}_{t-1}^i) \notin P_{i-1}, L(x, \underline{s}_t^i) \in P_{i-1}\} \\ &= \sum_{|W| \leq kh} \sum_{u \notin W} \mathbb{P}\{A_i, P_{i-1} = W, L(x, \underline{s}_{t-1}^i) = u, [uX_{u, a_t^i}] \in W\}. \end{aligned}$$

We separate this sum into two terms depending on whether $u \leq n^{\beta-\delta}$ or $u > n^{\beta-\delta}$. The sum over $u \leq n^{\beta-\delta}$ can be bounded as in (5.11): since $s_{i|t}$ is the first path that hits P_{i-1} , we have $P(x, \underline{s}_{t-1}^i) \cap P_{i-1} = \emptyset$. It follows that

$$\begin{aligned} &\sum_{|W| \leq kh} \sum_{u \leq n^{\beta-\delta}, u \notin W} \mathbb{P}\{A_i, P_{i-1} = W, L(x, \underline{s}_{t-1}^i) = u, [uX_{u, a_t^i}] \in W\} \\ &\leq \mathbb{P}\{A_{i-1}\} \cdot \mathbb{P}\{L(x, \underline{s}_t^i) \leq n^{\beta-\delta}\} \\ &\leq \mathbb{P}\{A_{i-1}\} \cdot n^{-1/k-\alpha\delta/2}, \end{aligned} \tag{5.12}$$

by Lemma 5.4. We now look at the sum over $u > n^{\beta-\delta}$:

$$\begin{aligned} &\sum_{|W| \leq kh} \sum_{u > n^{\beta-\delta}, u \notin W} \mathbb{P}\{A_i, P_{i-1} = W, L(x, \underline{s}_{t-1}^i) = u, [uX_{u, a_t^i}] \in W\} \\ &\leq \sum_{u > n^{\beta-\delta}} \sum_{|W| \leq kh, W \not\ni u} \sum_{Q \cap W = \emptyset} \mathbb{P}\{A_{i-1}, P_{i-1} = W, L(x, \underline{s}_{t-1}^i) = u, P(x, \underline{s}_{t-1}^i) = Q, [uX_{u, a_t^i}] \in W\} \end{aligned} \tag{5.13}$$

$$\begin{aligned} &= \sum_{u > n^{\beta-\delta}} \sum_{|W| \leq kh, W \not\ni u} \sum_{Q \cap W = \emptyset} \mathbb{P}\{A_{i-1}, P_{i-1} = W\} \cdot \mathbb{P}\{L(x, \underline{s}_{t-1}^i) = u, P(x, \underline{s}_{t-1}^i) = Q\} \cdot \mathbb{P}\{[uX] \in W\} \end{aligned} \tag{5.14}$$

$$\begin{aligned} &\leq \sum_{u > n^{\beta-\delta}} \left(\sup_{|W| \leq kh} \mathbb{P}\{[uX] \in W\} \right) \cdot \mathbb{P}\{L(x, \underline{s}_{t-1}^i) = u\} \cdot \mathbb{P}\{A_{i-1}\}. \end{aligned} \tag{5.15}$$

In the first equality, we used the independence of the events $\{A_{i-1}, P_{i-1} = W\}, \{L(x, s_{t-1}^i) = u, P(x, s_{t-1}^i) = Q\}$ and $\{X_{u, a_t^i} \in W\}$ if $u \notin W$ and $Q \cap W = \emptyset$. Using the fact that X has a density bounded by b , we have for any W of size at most kh

$$\mathbb{P}\{[uX] \in W\} \leq \frac{bkh}{u}.$$

The next step is to bound the sum in equation (5.15) by considering groups of nodes in the intervals $(2^\ell \lfloor n^{\beta-\delta} \rfloor, 2^{\ell+1} \lfloor n^{\beta-\delta} \rfloor]$ for non-negative integers ℓ . For any integer ℓ and any $\eta \in (0, \alpha)$, we have

$$\begin{aligned} & \sum_{u=2^\ell \lfloor n^{\beta-\delta} \rfloor + 1}^{2^{\ell+1} \lfloor n^{\beta-\delta} \rfloor} \left(\sup_{|W| \leq kh} \mathbb{P}\{[uX] \in W\} \right) \cdot \mathbb{P}\{L(x, s_{t-1}^i) = u\} \\ & \leq \mathbb{P}\{L(x, s_{t-1}^i) \leq 2^{\ell+1} n^{\beta-\delta}\} \cdot \frac{bkh}{2^\ell n^{\beta-\delta}} \\ & \leq \left(\frac{2^{\ell+1} n^{\beta-\delta}}{n} \right)^{\alpha-\eta} \frac{bkh}{2^\ell n^{\beta-\delta}} \\ & \leq 2^\alpha bkh \cdot (2^\ell n^{\beta-\delta})^{\alpha-\eta-1} n^{-\alpha+\eta}, \end{aligned}$$

for sufficiently large n . For the second inequality, we used Lemma 5.4. Now if $\alpha \leq 1$, recalling that $\beta = 1 - 1/(k\alpha)$, we can bound

$$\begin{aligned} n^{-\alpha+\eta} \cdot (2^\ell n^{\beta-\delta})^{\alpha-\eta-1} & \leq n^{-\alpha+\eta} (n^{\beta-\delta})^{\alpha-1} \\ & \leq n^{-\alpha+\eta} n^{\alpha-1/k-\alpha\delta-\beta+\delta} \\ & \leq n^{-1/k-\delta} \end{aligned}$$

for sufficiently small η and sufficiently large n . For the case $\alpha > 1$, we get

$$n^{-\alpha+\eta} \cdot (2^\ell n^{\beta-\delta})^{\alpha-\eta-1} \leq n^{-\alpha+\eta} \cdot (2^\ell n^{\beta-\delta})^{\alpha-1} \leq 2^{\alpha-1} n^{-1},$$

provided $2^\ell n^{\beta-\delta} \leq 2n$. This shows that total weight of an interval can be bounded by

$$\sup_{\ell} \sum_{u=2^\ell \lfloor n^{\beta-\delta} \rfloor + 1}^{2^{\ell+1} \lfloor n^{\beta-\delta} \rfloor} \left(\sup_{|W| \leq kh} \mathbb{P}\{[uX] \in W\} \right) \cdot \mathbb{P}\{L(x, s_{t-1}^i) = u\} \leq n^{-1/k-\delta/3}$$

for n large enough, where the supremum is taken over non-negative integers ℓ such that $2^\ell n^{\beta-\delta} \leq n$. This allows us to bound the expression in (5.15) by $\mathbb{P}\{A_{i-1}\} n^{-1/k-\delta/2}$, as there are at most $\log_2 n$ intervals. By summing this term with the term corresponding to the sum for $u \leq n^{\beta-\delta}$, we obtain

$$\mathbb{P}\{A_i, L(x, s_{t-1}^i) \notin P_{i-1}, L(x, s_t^i) \in P_{i-1}\} \leq \mathbb{P}\{A_{i-1}\} n^{-1/k-\min(\alpha, 1)\delta} \tag{5.16}$$

for any $t \in \{1, \dots, h\}$. Now it only remains to bound the sum (5.10) over the different values of t using (5.11) and (5.16). Using the induction hypothesis to bound $\mathbb{P}\{A_{i-1}\}$, we obtain the desired induction step (5.9). □

Putting these results together, we obtain the following.

Proposition 5.8. *Let $x \in \{\lfloor n/2 \rfloor, \dots, n\}$. Let $\delta > 0$. For n large enough,*

$$\mathbb{P}\{E_x\} \geq 1 - n^{-1-\delta/4} + n^{-3/2+4\delta}.$$

We are now in a position to prove the lower bound claimed in Lemma 5.3. Fix $\varepsilon, \delta \in (0, 1/10)$ and $h = \lfloor 10 \log \log n \rfloor$. First, using the lower bound on the typical distance, we can say that most of the nodes in the tree have $D_y \geq (1 - \varepsilon)\lambda_k \log n$. We want to show that, with high probability, for every y such that $n/2 \leq y \leq n$, we have $D_y \geq (\beta - \eta)\lambda_k \log n$ for arbitrarily small η . Define the set of ‘bad nodes’ that violate the property

$$\mathcal{B} = \{y : y \geq n/2, D_y < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n\},$$

which we decompose in a dyadic fashion: for any positive integer r , define

$$\mathcal{B}_r = \{y : 2^r \leq y < 2^{r+1}, D_y < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n\}.$$

Using Lemma 4.3 and Markov’s inequality, we have for every r such that $(\beta - 2\delta) \log_2 n \leq r \leq \log_2 n$

$$\mathbb{P}\left\{|\mathcal{B}_r| \geq \frac{2^r}{100 \cdot b}\right\} \leq 100b \cdot 2^{-r} \mathbb{E}[|\mathcal{B}_r|] \leq Cn^{-\eta},$$

for some C and $\eta > 0$ independent of r . Recall that b is a bound on the density of X . As a result the event

$$A = \{|\mathcal{B}_r| < 2^r/(100b) \text{ for } (\beta - 2\delta) \log_2 n \leq r \leq \log_2 n\} \tag{5.17}$$

is such that $\mathbb{P}\{A\} = 1 - O(n^{-\eta/2})$ for all n large enough.

Thus,

$$\begin{aligned} &\mathbb{P}\left\{\min_{n/2 \leq x \leq n} D_x < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + h + 1\right\} \\ &\leq \mathbb{P}\{A^c\} + \sum_{x=\lfloor n/2 \rfloor}^n \mathbb{P}\{A, D_x < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + h + 1\} \\ &\leq \mathbb{P}\{A^c\} + \sum_{x=\lfloor n/2 \rfloor}^n \mathbb{P}\{A, \forall s \in [k]^h : D_{L(x,s)} < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + 1\}. \end{aligned} \tag{5.18}$$

We now bound the term $\mathbb{P}\{A, \forall s \in [k]^h : D_{L(x,s)} < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + 1\}$ by conditioning on the event E_x . Let $S_1, \dots, S_{k^{h-3}}$ denote, when E_x holds, a set of paths that lead (when starting at x) to distinct nodes:

$$\begin{aligned} &\mathbb{P}\{A, \forall s \in \mathcal{A}^h : D_{L(x,s)} < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + 1\} \\ &\leq \mathbb{P}\{E_x^c\} + \mathbb{P}\{A, E_x, \forall s \in \mathcal{A}^h : D_{L(x,s)} < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + 1\} \\ &\leq \mathbb{P}\{E_x^c\} + \mathbb{P}\{A, E_x, \forall i \in \{1, \dots, k^{h-3}\} : D_{L(x,S_i)} < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + 1\} \\ &\leq \mathbb{P}\{E_x^c\} + \mathbb{P}\{A, E_x, \forall i \in \{1, \dots, k^{h-3}\}, p \in \mathcal{A} : [L(x, S_i)X_{L(x,S_i),p}] \in \mathcal{B}\}. \end{aligned}$$

By taking the worst possible set \mathcal{B} (compatible with the event A), we obtain the bound

$$\begin{aligned} &\mathbb{P}\{A, E_x, \forall i \in \{1, \dots, k^{h-3}\}, p \in \mathcal{A} : \lfloor L(x, S_i)X_{L(x, S_i), p} \rfloor \in \mathcal{B}\} \\ &\leq \left(\sup_{B, u \geq n^{\beta-\delta}} \mathbb{P}\{\lfloor uX \rfloor \in B\} \right)^{k \cdot k^{h-3}}, \end{aligned} \tag{5.19}$$

where the maximization is taken over all sets B such that $|B \cap [2^r, 2^{r+1})| < 2^r / (100b)$ for all r . Note that a key point here is that the nodes $L(x, S_i)$ for $i \in \{1, \dots, k^{h-3}\}$ are distinct and all have labels at least $n^{\beta-\delta}$. It now suffices to bound the right-hand side of (5.19).

Lemma 5.9. *Suppose $B \subseteq \{0, \dots, n\}$ such that, for all r satisfying $(\beta - 2\delta) \log_2 n \leq r \leq \log_2 n$, we have*

$$|B \cap [2^r, 2^{r+1})| < 2^r / (100b).$$

Then, for any $y \geq n^{\beta-\delta}$, we have for n large enough

$$\mathbb{P}\{\lfloor yX \rfloor \in B\} \leq 1/2.$$

Proof. Let r_{\max} be the largest integer such that $2^{r_{\max}} \leq y$ and let r_{\min} be the smallest integer at least as large as $(\beta - 2\delta) \log_2 n$. We then have

$$\begin{aligned} \mathbb{P}\{\lfloor yX \rfloor \in B\} &\leq \mathbb{P}\{\lfloor yX \rfloor \leq 2^{r_{\min}}\} + \sum_{r=r_{\min}}^{r_{\max}} \mathbb{P}\{\lfloor yX \rfloor \in B \cap [2^r, 2^{r+1})\} \\ &\leq \mathbb{P}\{yX \leq 2n^{\beta-2\delta} + 1\} + \frac{b}{y} \cdot \frac{1}{100b} (2^{r_{\max}+1} + 2^{r_{\max}} + \dots + 2^{r_{\min}}) \\ &\leq 3bn^{-\delta} + \frac{2^{r_{\max}+2}}{100y} \\ &\leq 1/2 \end{aligned}$$

for sufficiently large n . □

Getting back to equation (5.18), we obtain

$$\begin{aligned} &\mathbb{P}\left\{ \min_{n/2 \leq x \leq n} D_x < (1 - \varepsilon)(\beta - 2\delta)\lambda_k \log n + h + 1 \right\} \\ &\leq \mathbb{P}\{A^c\} + \sum_{x=\lceil n/2 \rceil}^n (\mathbb{P}\{E_x^c\} + (1/2)^{k^{h-2}}) \\ &= O(n^{-\eta}) \end{aligned}$$

for sufficiently small η , using Proposition 5.8. This concludes the proofs of Lemma 5.3 and Theorem 5.1.

6. Concluding remarks

We studied the longest path distance in a general class of random recursive DAGs. The parents of node x are chosen independently and distributed as $\lfloor xX \rfloor$ for some random

variable $X \in [0, 1)$. When X has a bounded density, we proved laws of large numbers for the typical and minimum distance. For both of these results, the upper bounds do not make any assumption on the attachment distribution X . We use the condition of bounded density for the lower bound when bounding the dependences between the different paths up to the root. It would be interesting to extend these results to more general distributions. More generally, under which conditions is it possible to translate a result in an ideal model like the branching random walk considered in Section 3 into a result about a real model that exhibits limited dependences?

Acknowledgements

We would like to thank Louigi Addario-Berry, Luc Devroye and Nicolas Fraiman for many helpful discussions as well as the anonymous referee for useful comments. Part of this work was completed while the authors were visiting McGill's Bellairs Research Institute, Barbados. NB's research is supported by the ANR-09-BLAN-0011 project Boole.

References

- [1] Addario-Berry, L. and Reed, B. (2009) Minima in branching random walks. *Ann. Probab.* **37** 1044–1079.
- [2] Aïdékon, É. (2011) Convergence in law of the minimum of a branching random walk. arXiv:1101.1810 [math.PR].
- [3] Alon, N. and Spencer, J. (2008) *The Probabilistic Method*, third edition, Wiley.
- [4] Arya, S., Golin, M. and Mehlhorn, K. (1999) On the expected depth of random circuits. *Combin. Probab. Comput.* **8** 209–228.
- [5] Athreya, K. B. and Ney, P. E. (1972) *Branching Processes*, Springer.
- [6] Bachmann, M. (2000) Limit theorems for the minimal position in a branching random walk with independent logconcave displacements. *Adv. Appl. Probab.* **32** 159–176.
- [7] Biggins, J. (1976) The first- and last-birth problems for a multitype age-dependent branching process. *Adv. Appl. Probab.* **8** 446–459.
- [8] Bramson, M. (1978) Maximum displacement of branching Brownian motion. *Comm. Pure Appl. Math.* **31** 531–581.
- [9] Bramson, M. and Zeitouni, O. (2009) Tightness for a family of recursive equations. *Ann. Probab.* **37** 615–653.
- [10] Chauvin, B. and Drmota, M. (2006) The random multisection problem, travelling waves, and the distribution of the height of m -ary search trees. *Algorithmica* **46** 299–327.
- [11] Chernoff, H. (1952) A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.
- [12] Chung, K. and Erdős, P. (1952) On the application of the Borel–Cantelli lemma. *Trans. Amer. Math. Soc.*, pp. 179–186.
- [13] Codenotti, B., Gemmell, P. and Simon, J. (1995) Average circuit depth and average communication complexity. In *Algorithms: ESA '95*, Vol. 979 of *Lecture Notes in Computer Science*, Springer, pp. 102–112.
- [14] Darrasse, A., Hwang, H.-K., Bodini, O. and Soria, M. (2010) The connectivity-profile of random increasing k -trees. In *Proc. ANALCO*, pp. 99–106.
- [15] Dembo, A. and Zeitouni, O. (1998) *Large Deviations Techniques and Applications*, Springer.
- [16] Devroye, L. (1986) A note on the height of binary search trees. *J. Assoc. Comput. Mach.* **33** 489–498.

- [17] Devroye, L. (1998) Branching processes and their applications in the analysis of tree structures and tree algorithms. In *Probabilistic Methods for Algorithmic Discrete Mathematics*, Vol. 16 of *Algorithms and Combinatorics*, Springer, pp. 249–314.
- [18] Devroye, L. and Janson, S. (2011) Long and short paths in uniform random recursive dags. *Arkiv för Matematik* **49** 61–77.
- [19] Devroye, L. and Lu, J. (1995) The strong convergence of maximal degrees in uniform random recursive trees and dags. *Random Struct. Alg.* **7** 1–14.
- [20] Devroye, L., Fawzi, O. and Fraiman, N. (2012) Depth properties of scaled attachment random recursive trees. *Random Struct. Alg.* **41** 66–98.
- [21] Diaz, J., Serna, M. J., Spirakis, P., Toran, J. and Tsukiji, T. (1994) On the expected depth of Boolean circuits. Technical Report LSI-94-7-R, Universitat Politècnica de Catalunya, Dep. LSI.
- [22] Drmota, M. (2003) An analytic approach to the height of binary search trees II. *J. Assoc. Comput. Mach.* **50** 333–374.
- [23] D’Souza, R. M., Krapivsky, P. L. and Moore, C. (2007) The power of choice in growing trees. *Eur. Phys. J. B* **59** 535–543.
- [24] Grimmett, G. and Stirzaker, D. (2001) *Probability and Random Processes*, Oxford University Press.
- [25] Gut, A. (2009) *Stopped Random Walks: Limit Theorems and Applications*, Springer.
- [26] Hammersley, J. M. (1974) Postulates for subadditive processes. *Ann. Probab.* **2** 652–680.
- [27] Hu, Y. and Shi, Z. (2009) Minimal position and critical martingale convergence in branching random walks, and directed polymers on disordered trees. *Ann. Probab.* **37** 742–789.
- [28] Kingman, J. (1975) The first birth problem for an age-dependent branching process. *Ann. Probab.* **3** 790–801.
- [29] Mahmoud, H. (2010) The power of choice in the construction of recursive trees. *Method. Comput. Appl. Probab.* **12** 763–773.
- [30] Mahmoud, H. and Tsukiji, T. (2004) Limit laws for terminal nodes in random circuits with restricted fan-out: a family of graphs generalizing binary search trees. *Acta Informatica* **41** 99–110.
- [31] McDiarmid, C. (1995) Minimal positions in a branching random walk. *Ann. Appl. Probab.* **5** 128–139.
- [32] Pittel, B. (1985) Asymptotical growth of a class of random trees. *Ann. Probab.* **13** 414–427.
- [33] Reed, B. (2003) The height of a random binary search tree. *J. Assoc. Comput. Mach.* **50** 306–332.
- [34] Rockafellar, R. (1970) *Convex Analysis*, Princeton University Press.
- [35] Tsukiji, T. and Mahmoud, H. (2001) A limit law for outputs in random recursive circuits. *Algorithmica* **31** 403–412.
- [36] Tsukiji, T. and Xhafa, F. (1996) On the depth of randomly generated circuits. In *Algorithms: ESA ’96*, Vol. 1136 of *Lecture Notes in Computer Science*, Springer, pp. 208–220.