Electron. Commun. Probab. **29** (2024), article no. 77, 1–10.
https://doi.org/10.1214/24-ECP645
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

The harmonic descent chain

David J. Aldous* Svante Janson† Xiaodan Li‡

Abstract

The decreasing Markov chain on {1,2,3, . . . } with transition probabilities $p(j, j - i) \propto 1/i$ arises as a key component of the analysis of the beta-splitting random tree model. We give a direct and almost self-contained "probability" treatment of its occupation probabilities, as a counterpart to a more sophisticated but perhaps opaque derivation using a limit continuum tree structure and Mellin transforms.

Keywords: Markov chain; occupation measure; coupling.
MSC2020 subject classifications: 60J10.
Submitted to ECP on May 6, 2024, final version accepted on November 22, 2024.

1 Introduction

Write $h_n := \sum_{i=1}^{n} \frac{1}{i}$ for the harmonic series. We will study the discrete-time Markov chain $(X_t, t = 0, 1, 2, \ldots)$ on states $\{1, 2, 3 \ldots\}$ with transition probabilities

$$p(j,i) = \frac{1}{(j-i)h_{j-1}},\ 1 \le i < j,\ j \ge 2 \tag{1.1}$$

and $p(1, 1) = 1$. So sample paths are strictly decreasing until absorption in state 1. The simple form (1.1, 3.1) of the transitions suggests that this chain might have been arisen previously in some different context, but we have not found any reference. Let us call this the *harmonic descent* (HD) chain.

As discussed at length elsewhere [2], the HD chain arises in a certain model of random n-leaf trees: the chain describes the number of descendant leaves of a vertex, as one moves along the path from the root to a uniform random leaf. In this article we study the "occupation probability", that is

$$a(n, i) := \text{probability that the chain started at state } n \text{ is ever in state } i. \tag{1.2}$$

So $a(n, n) = a(n, 1) = 1$. The motivation for studying $a(n, i)$ is that the mean number of i-leaf subtrees of a random n-leaf tree equals $na(n, i)/i$. There is a general notion of the *fringe distribution* [1, 7] of a tree as viewed from a random leaf. Knowing the

*Department of Statistics, 367 Evans Hall # 3860, U.C. Berkeley CA 94720; E-mail: aldousdj@berkeley.edu https://www.stat.berkeley.edu/users/aldous.

†Department of Mathematics, Uppsala University, P.O.Box 480, SE-751 06 Uppsala Sweden; E-mail: svante.janson@math.uu.se

‡Department of Mathematics, Shanghai University of Finance and Economics, China. E-mail: lixiaodan@mail.shufe.edu.cn

explicit value of the limits $\lim_{n\to\infty} a(n,i)$ enables one in [3] to describe explicitly the fringe distribution in our tree model.

We should also mention that $\sum_{i=2}^n a(n,i)$ is just the mean time $\mathbb{E}_n T_1$ for the chain started at n to be absorbed in state 1. A detailed analysis of the distribution of T_1 and related quantities, by very different methods, is given in [5].

It seems very intuitive (but not obvious at a rigorous level) that the limits $\lim_{n\to\infty} a(n,i)$ exist, however there seems no intuitive reason to think there should be some simple formula for the limits. But the purpose of this paper is to give an almost[1] self-contained proof of the following result.

Theorem 1.1. *For each $i = 2, 3, \ldots$,*

$$\lim_{n\to\infty} a(n,i) \quad = \quad \frac{6h_{i-1}}{\pi^2(i-1)} := a(i). \tag{1.3}$$

As the reader might guess, we will encounter *Euler's formula* $\sum_{i\geq 1} 1/i^2 = \pi^2/6$.

We prove Theorem 1.1 in two stages. In section 3 we will prove by coupling that the limits $a(i)$ exist. This is straightforward in outline, though somewhat tedious in detail. More interesting, and therefore presented first in section 2, is the explicit formula for the limits $a(i)$. The limits satisfy an infinite set of equations (2.2), and a solution was found by inspired guesswork. Then we need only to check that the solution is unique.

Regarding the "inspired guesswork", numerical approximation of $a(i)$ (by computing $a(n,i)$ for large n) suggested that $a(i)$ grows as order $(\log i)/i$. Because we are studying $i \geq 2$, it is in retrospect natural to try $h_{i-1}/(i-1)$.

Could one prove Theorem 1.1 without guesswork? In the random tree model, there is a limit continuum tree structure within which there is a continuous analog of the "hypothetical" chain described below. By analysis of that process and the exchangeability properties of the continuum tree, forthcoming work [3, 4] relates the $(a(i))$ to the $x \downarrow 0$ behavior of a certain function $f(x)$ determined by its Mellin transform. Then by technically intricate analysis one can re-prove Theorem 1.1 via a "proof by calculation".

In very recent unpublished work, Iksanov [8] observes that one can exploit an exact relationship with regenerative composition structures [6], enabling a shorter derivation of Theorem 1.1 from known results in that theory.

2 The explicit limit

In section 3 we will prove by coupling the following result. Here $a(i)$ is defined by (2.1).

Proposition 2.1. *For each $i = 1, 2, 3, \ldots$,*

$$\lim_{n\to\infty} a(n,i) \quad := \quad a(i) \text{ exists} \tag{2.1}$$

$$a(i) \quad = \quad \sum_{j>i} a(j)p(j,i) \tag{2.2}$$

and

$$a(1) = 1. \tag{2.3}$$

Such a coupling proof does not give any useful quantitative information about the limits $a(i)$. In this section we show how to derive the value of the limits, granted Proposition 2.1.

[1] We quote one sharp estimate from [5] as our Theorem 3.1.

For fixed $n \geq 2$ it is clear that the values of $(a(n,i), 1 \leq i \leq n)$ are determined by the natural system of equations

$$\begin{aligned} a(n,n) &= 1, && (2.4)\\ a(n,i) &= \sum_{n \geq j > i} a(n,j)p(j,i), \quad n-1 \geq i \geq 1, && (2.5)\\ a(n,1) &= 1. && (2.6) \end{aligned}$$

So the equations (2.2)–(2.3) are what one expects as the $n \to \infty$ analog of (2.5)–(2.6). It is not obvious that these equations have a unique solution, but the following solution was found by inspired guesswork, rather than calculation. Define

$$b(1) = 1; \quad b(i) = \frac{6h_{i-1}}{\pi^2(i-1)}, \quad i \geq 2. \tag{2.7}$$

Lemma 2.2. *Equations* (2.2)–(2.3) *are satisfied by* $b(i)$ *in* (2.7)*.*

Proof. Equation (2.2) for $b(i)$ is explicitly

$$b(i) = \sum_{j>i} \frac{b(j)}{h_{j-1}(j-i)}, \quad i \geq 1. \tag{2.8}$$

To verify (2.8), consider $i > 1$. We have

$$\begin{aligned} \sum_{j>i} \frac{b(j)}{h_{j-1}(j-i)} &= \frac{6}{\pi^2} \sum_{j>i} \frac{1}{(j-i)(j-1)} \\ &= \frac{6}{\pi^2(i-1)} \sum_{j>i} \left(\frac{1}{j-i} - \frac{1}{j-1} \right) \\ &= \frac{6}{\pi^2(i-1)} \lim_{k \to \infty} \left(\sum_{j=1}^{k-i} \frac{1}{j} - \sum_{j=i}^{k-1} \frac{1}{j} \right) \\ &= \frac{6}{\pi^2(i-1)} \; h_{i-1} = b(i). \end{aligned} \tag{2.9}$$

For $i = 1$:

$$\sum_{j>1} \frac{b(j)}{h_{j-1}(j-1)} = \frac{6}{\pi^2} \sum_{j>1} \frac{1}{(j-1)^2} = 1 = b(1). \qquad \square$$

As noted before, at first sight we do not know that these equations (2.2)–(2.3) have a *unique* solution. We need a further careful argument to prove that $a(i) \equiv b(i)$, which then (granted Proposition 2.1) completes a proof of Theorem 1.1.

Proposition 2.3. *For* $a(i)$ *defined by the limit* (2.1) *and* $b(i)$ *defined by* (2.7)*, we have* $a(i) = b(i), i \geq 1$*.*

Proof. Fix large k. By considering the chain started at n and decomposing at the jump over k:

$$a(n,i) = \sum_{n \geq m > k} \sum_{k \geq j \geq i} a(n,m)p(m,j)a(j,i), \; i \leq k < n.$$

Note we write $n \geq m > k$ in decreasing order, visualizing the chain as coming down from n. Letting $n \to \infty$ suggests

$$a(i) = \sum_{m>k} \sum_{k \geq j \geq i} a(m)p(m,j)a(j,i), \; i \leq k. \tag{2.10}$$

This does not have a direct Markov chain interpretation, because one cannot formalize the idea of starting a hypothetical version of the chain from $+\infty$ at time $-\infty$ and making its occupation probability be $(a(i))$. Nevertheless (2.10) is correct, and can be derived from the relation (2.2). Fix i and k with $1 \leq i \leq k$ and in (2.2) recursively substitute $a(j)$ by the same equation for every $j \leq k$. This yields, with summation over all paths of the form $\eta = (i_1 = m > i_2 = j > i_3 > \ldots > i_q > i)$ with $i_1 > k \geq i_2$:

$$a(i) = \sum_{\eta} a(m)p(m,j)p(j,i_3)p(i_3,i_4)\cdots p(i_q,i)$$

which collapses to (2.10). The key part of our proof is that the argument above for (2.10) depends only on $(a(i))$ satisfying (2.2), which by Lemma 2.2 also holds for $(b(i))$, and so the conclusion holds also for $(b(i))$:

$$b(i) = \sum_{m>k} \sum_{k \geq j \geq i} b(m)p(m,j)a(j,i), \ i \leq k. \tag{2.11}$$

Now consider

$$\hat{b}_k(j) := \sum_{m>k} b(m)p(m,j), \ 1 \leq j \leq k, \tag{2.12}$$

motivated as the overshoot (over k) distribution associated with our hypothetical version of a chain run from time $-\infty$. We will verify later the following technical lemma.

Lemma 2.4. *$(\hat{b}_k(j), 1 \leq j \leq k)$ is indeed a probability distribution, for each $k \geq 1$.*

Now let us re-write (2.11) as

$$b(i) = \sum_{k \geq j \geq i} \hat{b}_k(j)a(j,i) = \sum_{k \geq j \geq 1} \hat{b}_k(j)a(j,i), \quad i \leq k \tag{2.13}$$

and thus

$$b(i) - a(i) = \sum_{k \geq j \geq 1} \hat{b}_k(j)(a(j,i) - a(i)), \quad i \leq k. \tag{2.14}$$

It follows from (2.12) that $\hat{b}_k(j) \to 0$ as $k \to \infty$ for every fixed j, and thus the distributions $\hat{b}_k$ go off to infinity, that is $\lim_{k\to\infty} \hat{b}_k(L) = 0$ for each $L < \infty$. So letting $k \to \infty$ for fixed i we see that (2.14) and Proposition 2.1 imply that $b(i) = a(i)$, establishing Proposition 2.3. □

Proof of Lemma 2.4. From the definitions of $b(m)$ and $p(m,j)$ this reduces to proving that for each $k \geq 2$

$$\sum_{j=1}^{k} \sum_{m=k+1}^{\infty} \frac{1}{(m-j)(m-1)} = \frac{\pi^2}{6}. \tag{2.15}$$

For $k = 1$ this is Euler's formula. Furthermore, for $k \geq 1$, we have, writing s_k for the left side of (2.15) and $\phi(m,j) = \frac{1}{(m-j)(m-1)}$, and arguing similarly as in (2.9),

$$\begin{aligned} s_{k+1} - s_k &= \sum_{m=k+2}^{\infty} \phi(m,k+1) - \sum_{j=1}^{k} \phi(k+1,j) \\ &= \sum_{m=k+2}^{\infty} \frac{1}{(m-k-1)(m-1)} - \sum_{j=1}^{k} \frac{1}{(k+1-j)k} \\ &= \frac{1}{k} \sum_{m=k+2}^{\infty} \left(\frac{1}{m-k-1} - \frac{1}{m-1} \right) - \frac{1}{k} h_k \\ &= \frac{1}{k} h_k - \frac{1}{k} h_k = 0 \end{aligned}$$

establishing (2.15). □

3 Proof of Proposition 2.1

3.1 Outline of proof

As noted in other discussions of the random tree model [2, 5], it is often more convenient to work with the associated *continuous time* chain which holds in each state $i \geq 2$ for an Exponential(h_{i-1}) time, in other words has transition rates

$$\lambda(j,i) = \tfrac{1}{j-i}, \quad 1 \leq i < j < \infty. \tag{3.1}$$

We call this the *continuous HD chain* $(X_t, 0 \leq t < \infty)$. Switching to this continuous time chain does not affect our stated definition of $a(n,i)$, though of course the mean occupation time in state i changes from $a(n,i)$ to $a(n,i)/h_{i-1}$. And switching does more substantially change T_1, the absorption time to state 1. As mentioned before, a detailed study of absorption time distributions is given in [5], and we quote two results from there.

Theorem 3.1 ([5, Theorem 1.1])**.**

$$\textit{In continuous time, } \mathbb{E}_x[T_1] = \tfrac{6}{\pi^2}\log x + O(1) \textit{ as } x \to \infty. \tag{3.2}$$

Lemma 3.2 (From [5, Corollary 1.3])**.** *Let, for positive integers k,*

$$T_k := \min\{t : X_t \leq k\}. \tag{3.3}$$

Then, for any fixed $k \geq 1$ and $t \geq 0$, we have $\mathbb{P}_x(T_k \geq t) \to 1$ as $x \to \infty$; in other words, $T_k \xrightarrow{\mathrm{p}} \infty$.

Lemma 3.2 can in fact be obtained by a simple direct proof, given for completeness in section 4.

So now we consider the continuous time setting. We will use a *shift-coupling* [10]. For our purpose, a shift-coupling $((X_t, Y_t), 0 \leq t < \infty)$ started at (x_0, y_0) is a process such that, conditional on $X_t = x_t, Y_t = y_t$ and the past, either
(i) over $(t, t+dt)$ each component moves according to $\lambda(\cdot,\cdot)$, maybe dependently; or
(ii) one component moves as above while the other remains unchanged.

Such a process must reach state $(1,1)$ and stop, at some time $T_{(1,1)} < \infty$. So the *coupling time* is such that

$$T^{couple} := \min\{t : X_t = Y_t\} \leq T_{(1,1)}$$

and we can arrange that $X_t = Y_t$ for $t \geq T^{couple}$. Write $S^{couple} := X_{T^{couple}} = Y_{T^{couple}}$ for the coupling *state*. We will construct a shift-coupling in which, for each $i \geq 1$,

$$\mathbb{P}_{x_0,y_0}(S^{couple} < i) \to 0 \text{ as } x_0, y_0 \to \infty. \tag{3.4}$$

This is clearly sufficient to prove the main "limits exist" part (2.1) of Proposition 2.1, because $|a(x_0,i) - a(y_0,i)| \leq \mathbb{P}_{x_0,y_0}(S^{couple} < i)$.

In outline the construction is very simple. If the initial states x_0 and y_0 are not of comparable size, then we run the chain only from the larger state (as in (ii) above) until they are of comparable size; then at each time there is some non-vanishing probability that we can couple at the next transition (as in (i) above).

The details are given via two lemmas below. By symmetry, it suffices to consider the case $x_0 \leq y_0$, and by considering subsequences we may assume that $x_0/y_0 \to a$ for some $a \in [0,1]$. First we consider the "comparable size starts" case $a > 0$, and then the case $a = 0$. As in [2] and as suggested by (3.2) we analyze the processes on the log scale.

3.2 The maximal coupling regime

In the maximal coupling regime, we construct the joint process $((X_t, Y_t), 0 \le t < \infty)$ as follows. From $(X_t, Y_t) = (x, y)$ with $x \le y$, each component moves according to $\lambda(\cdot,)$ but with the joint distribution that maximizes the probability that they move to the same state. That joint distribution is such that, for infinitesimal dt,

$$\mathbb{P}_{x,y}(X_{dt} = Y_{dt} = i) = \mathbb{P}_{x,y}(Y_{dt} = i) = \frac{dt}{y-i}, \quad 1 \le i < x.$$

So

$$\mathbb{P}_{x,y}(X_{dt} = Y_{dt}) = \sum_{i=1}^{x-1} \frac{dt}{y-i} \ge \frac{x-1}{y}.$$

Hence, for any $c \in (0,1)$, if $c \le x/y \le 1$, then

$$\mathbb{P}_{x,y}(X_{dt} = Y_{dt}) \ge (c - y^{-1})dt. \tag{3.5}$$

Lemma 3.3. *For the maximal coupling process, if $y_0 \to \infty$ and $x_0/y_0 \to a \in (0,1]$ then for each k we have $\mathbb{P}_{x_0,y_0}(S^{couple} \le k) \to 0$.*

Proof. Write $T_k^Y := \min\{t : Y_t \le k\}$, and note that $\{S^{couple} \le k\} = \{T_k^Y \le T^{couple}\}$. Consider the process (X_t, Y_t) using the maximal coupling with $(X_0, Y_0) = (x_0, y_0)$, where $x_0 \le y_0$. The coupling is stochastically monotone, so $X_t \le Y_t$ and the absorption times into state 1 satisfy $T_1^X \le T_1^Y$. By (3.2) we have

$$\mathbb{E}_{x_0,y_0}[T_1^Y - T_1^X] = \frac{6}{\pi^2} \log \frac{y_0}{x_0} + O(1) \tag{3.6}$$

where the $O(1)$ bound is uniform in all $x_0, y_0 \ge 1$.

Fix $0 < c < a \le 1$. Consider the stopping time

$$U_c := \min\{t : X_t/Y_t \le c\}.$$

Fix also a large τ and let $\ell \ge k$. Then

$$\begin{aligned} \mathbb{P}_{x_0,y_0}(S^{couple} \le k) &= \mathbb{P}_{x_0,y_0}(T^{couple} \ge T_k^Y) \le \mathbb{P}_{x_0,y_0}(T^{couple} \ge T_\ell^Y) \\ &\le \mathbb{P}_{x_0,y_0}(U_c < T_\ell^Y) + \mathbb{P}_{x_0,y_0}(T_\ell^Y \le \tau) + \mathbb{P}_{x_0,y_0}(T^{couple} \wedge U_c \wedge T_\ell^Y > \tau). \end{aligned} \tag{3.7}$$

We consider the three terms in (3.7) separately.

On the event $\{U_c < T_\ell^Y\}$ the conditional expectation of $(T_1^Y - T_1^X)$ is at least $\frac{6}{\pi^2} \log 1/c - O(1)$ by (3.6) and conditioning on (X_{U_c}, Y_{U_c}), noting that $Y_{U_c}/X_{U_c} \ge 1/c$. So by Markov's inequality and (3.6) again, there exists a constant C (not depending on a, c) such that, provided $\log 1/c > C$ and y_0 is large enough,

$$\mathbb{P}_{x_0,y_0}(U_c \le T_\ell^Y) \le \frac{\log(y_0/x_0) + O(1)}{\log 1/c - O(1)} \le \frac{\log 1/a + C}{\log 1/c - C}. \tag{3.8}$$

This holds for any sufficiently small $c > 0$, and can be made arbitrarily small by choosing c small.

As long as $t < T^{couple} \wedge U_c \wedge T_\ell^Y$, the coupling event happens at rate $\ge c - \ell^{-1}$ by (3.5), so

$$\mathbb{P}_{x_0,y_0}(T^{couple} \wedge U_c \wedge T_\ell^Y > \tau) \le \exp\left((-c + \ell^{-1})\tau\right), \tag{3.9}$$

which for fixed c can be made arbitrarily small by choosing $\ell \ge 2/c$ and τ large.

Finally, for fixed ℓ and τ, $\mathbb{P}_{x_0,y_0}(T_\ell^Y \le \tau) \to 0$ as $y_0 \to \infty$ by Lemma 3.2.

Consequently, (3.7) shows that $\mathbb{P}_{x_0,y_0}(S^{couple} \le k) \to 0$ as $y_0 \to \infty$ and $x_0/y_0 \to a$. □

3.3 The large discrepancy regime

Given Lemma 3.3, to establish (3.4) it remains only to consider the case $x_0/y_0 \to 0$. Here we first run the chain (Y_t) starting from $y_0 \gg x_0$ while holding $X_t = x_0$ fixed. The next lemma shows that the (Y_t) process does not overshoot x_0 by far, on the log scale.

For $x \ge 1$ write as above

$$T_x^Y := \min\{t : Y_t \le x\}.$$

Let also

$$V_x := \log x - \log Y_{T_x^Y} \quad \text{(the overshoot factor).}$$

Lemma 3.4. *There exists an absolute constant K such that*

$$\mathbb{E}_y[V_x] \le K, \quad 1 \le x < y < \infty.$$

Proof. Here we work in discrete time, which does not change $\mathbb{E}_y[V_x]$. Consider a single transition $y \to Y_1$. From the transition probabilities (1.1) we obtain the exact formula, for any $a > 0$, and writing $b = e^{-a} \in (0,1)$,

$$\mathbb{P}_y(\log y - \log Y_1 \ge a) = \mathbb{P}_y(Y_1 \le e^{-a}y) = \frac{\sum_{y-\lfloor by\rfloor}^{y-1} \frac{1}{i}}{h_{y-1}} = \frac{h_{y-1} - h_{y-\lfloor by\rfloor-1}}{h_{y-1}}. \tag{3.10}$$

If, say, $a \ge 1$ is fixed, then it follows by the formula

$$h_n = \log n + \gamma + o(1) \text{ as } n \to \infty \tag{3.11}$$

that as $y \to \infty$

$$\mathbb{P}_y(\log y - \log Y_1 \ge a) \to \frac{\theta[a,\infty)}{\log y}, \tag{3.12}$$

where θ is the measure on $(0,\infty)$ defined by

$$\theta[a,\infty) := -\log(1 - e^{-a}). \tag{3.13}$$

We will show that the approximation (3.12) holds within some constant factors uniformly for all $y \ge 2$ and $a \in [a_1(y), a_2(y)]$ where $a_1(y) := \log y - \log(y-1) = -\log(1-1/y)$ and $a_2(y) := \log y$. (Note that for $a \le a_1(y)$, trivially $\mathbb{P}_y(\log y - \log Y_1 \ge a) = 1$, and for $a > a_2(y)$, trivially $\mathbb{P}_y(\log y - \log Y_1 \ge a) = 0$.) That is, for some $C_1, C_2 > 0$,

$$C_1 \frac{\theta[a,\infty)}{\log y} \le \mathbb{P}_y(\log y - \log Y_1 \ge a) \le C_2 \frac{\theta[a,\infty)}{\log y}, \qquad a \in [a_1(y), a_2(y)]. \tag{3.14}$$

To verify (3.14), note first that by (3.11) we only have to estimate the numerator in (3.10). We have

$$\int_{y-\lfloor by\rfloor}^{y} \frac{dx}{x} \le \sum_{y-\lfloor by\rfloor}^{y-1} \frac{1}{i} \le \int_{y-\lfloor by\rfloor}^{y-1} \frac{dx}{x} + \frac{1}{y-\lfloor by\rfloor} \tag{3.15}$$

where $b = e^{-a} \in [e^{-a_2(y)}, e^{-a_1(y)}] = [\frac{1}{y}, \frac{y-1}{y}]$, and it is easily seen that both sides of (3.15) are within constant factors of $-\log(1-b) = \theta[a,\infty)$, thus showing (3.14).

The measure θ has the property

$$\frac{\int_a^\infty (u-a)\theta(du)}{\theta[a,\infty)} \uparrow 1 \text{ as } a \uparrow \infty, \tag{3.16}$$

and thus the conditional mean excess over a is bounded above by 1. One can now use (3.14) to see that there exists $K < \infty$ such that the corresponding property holds for one transition $y \to Y_1$ of the discrete chain on the log scale:

$$\mathbb{E}_y[(\log y - \log Y_1 - a)^+] \leq K\mathbb{P}_y(\log y - \log Y_1 \geq a). \tag{3.17}$$

In fact, for $a_1(y) \leq a \leq a_2(y)$, this follows immediately from (3.14), (3.16), and

$$\mathbb{E}_y[(\log y - \log Y_1 - a)^+] = \int_a^{a_2(y)} \mathbb{P}_y(\log y - \log Y_1 > s)ds. \tag{3.18}$$

For $a < a_1(y)$, (3.17) follows from the case $a = a_1(y)$, with the right side equal to K, and for $a > a_2(y)$, both sides of (3.17) are 0.

Now fix x. For $y \geq x$ write

$$m(x,y) := \max_{x \leq z \leq y} \mathbb{E}_z[V_x]$$

and note that $m(x,x) = 0$. If $y > x$, by considering the first step $y \to Y_1$, which goes either into the interval $[1,x]$ (probability $q_{x,y}$ say) or into the interval $[x+1, y-1]$:

$$\mathbb{E}_y[V_x] \leq \mathbb{E}_y[(\log x - \log Y_1)^+] + \mathbb{P}_y(\log x - \log Y_1 < 0)m(x, y-1). \tag{3.19}$$

By (3.17) with $a = \log y - \log x$

$$\mathbb{E}_y[(\log x - \log Y_1)^+] \leq K\mathbb{P}_y(\log x - \log Y_1 \geq 0) = Kq_{x,y}. \tag{3.20}$$

Combining (3.19) and (3.20):

$$\mathbb{E}_y[V_x] \leq Kq_{x,y} + (1 - q_{x,y})m(x, y-1) \tag{3.21}$$

and so the bound $m(x,y) \leq K$ holds by induction on $y = x, x+1, x+2, \ldots$. □

3.4 Completing the proof of Proposition 2.1

As noted at the start of the previous section, given Lemma 3.3, to show (3.4), it remains only to consider the case $x_0/y_0 \to 0$. Use the shift regime dynamics (x_0, Y_t) in section 3.3 from $Y_0 = y_0$ until time

$$T^Y_{x_0} := \min\{t : Y_t \leq x_0\},$$

and then use the maximal coupling in section 3.2. Lemma 3.4 shows that the overshoot factor of the first phase

$$V_{x_0} := \log x_0 - \log Y_{T^Y_{x_0}}$$

has $\mathbb{E}_{y_0}[V_{x_0}] \leq K$. Hence the overshoot factors are tight, and by considering a subsequence we may assume that V_{x_0} converges in distribution to some random variable V. For convenience, we may also by the Skorohod coupling theorem [9, Theorem 4.30] assume that V_{x_0} converges to V almost surely, and thus

$$x_0/Y_{T^Y_{x_0}} \xrightarrow{\text{a.s.}} e^V > 0. \tag{3.22}$$

This allows us to condition on $Y_{T_{x_0}}$ and apply Lemma 3.3 (with X and Y interchanged) with starting state $(x_0, Y_{T_{x_0}})$. Hence, for every fixed k, we have

$$\mathbb{P}_{x_0,y_0}\left(S^{couple} \leq k \mid Y_{T_{x_0}}\right) \to 0 \qquad \text{a.s.} \tag{3.23}$$

and thus (3.4) follows by taking the expectation. As noted after (3.4), this completes the proof of the main “limits exist” part (2.1) of Proposition 2.1. To complete the proof of Proposition 2.1, and by section 2 thus also the proof of Theorem 1.1, we need to verify (2.2), that is

$$a(i) = \sum_{j>i} a(j)p(j,i), \quad i \geq 1. \tag{3.24}$$

From the pointwise convergence $a(n,j) \to a(j)$, Fatou's lemma gives

$$\begin{aligned} \sum_{j>i} a(j)p(j,i) &\leq \liminf_n \sum_{j>i} a(n,j)p(j,i) \\ &= \lim_n a(n,i) = a(i). \end{aligned}$$

To prove equality we need to show that, for every fixed i,

$$\lim_{L\to\infty} \liminf_n \sum_{j>L} a(n,j)p(j,i) = 0. \tag{3.25}$$

This is an easy consequence of the overshoot bound, as follows. Assume, as we may, $L > i$. The sum above is the probability, starting at n, that the first entrance into $\{L, L-1, L-2, \dots\}$ is at i, which is bounded by the probability that first entrance is in $\{i, i-1, i-2, \dots\}$. The latter, in the notation of Lemma 3.4, is just $\mathbb{P}_n(V_L \geq \log L - \log i)$. By Markov's inequality and Lemma 3.4

$$\sum_{j>L} a(n,j)p(j,i) \leq \frac{K}{\log L - \log i}$$

implying (3.25).

4 Direct proof of Lemma 3.2

Let $Z_t := X_t^{-1/2}$. Then (3.1) implies that for an infinitesimal time dt, we have

$$\begin{aligned} \mathbb{E}_x[Z_{dt}] - x^{-1/2} &= dt \sum_{1\leq i<x} \lambda(x,i) \left(\frac{1}{\sqrt{i}} - \frac{1}{\sqrt{x}}\right) = dt \sum_{1\leq i<x} \frac{\sqrt{x}-\sqrt{i}}{(x-i)\sqrt{i}\sqrt{x}} \\ &= dt \sum_{1\leq i<x} \frac{1}{(\sqrt{x}+\sqrt{i})\sqrt{i}\sqrt{x}} \leq \frac{dt}{x} \sum_{1\leq i<x} \frac{1}{\sqrt{i}} \\ &\leq \frac{dt}{x} \int_0^x \frac{1}{\sqrt{y}}\, dy = 2\frac{dt}{\sqrt{x}} = 2Z_0 dt. \end{aligned} \tag{4.1}$$

Hence, by the Markov property,

$$\frac{d}{dt}\mathbb{E}_x[Z_t] \leq 2\mathbb{E}_x[Z_t], \qquad t \geq 0, \tag{4.2}$$

and consequently,

$$\mathbb{E}_x[Z_t] \leq e^{2t} Z_0 = e^{2t} x^{-1/2}, \qquad t \geq 0. \tag{4.3}$$

Finally, Markov's inequality yields

$$\mathbb{P}_x(T_k \leq t) = \mathbb{P}_x(X_t \leq k) = \mathbb{P}_x(Z_t \geq k^{-1/2}) \leq e^{2t}\sqrt{k/x}. \tag{4.4}$$

References

[1] David Aldous, *Asymptotic fringe distributions for general families of random trees*, Ann. Appl. Probab. **1** (1991), no. 2, 228–266. MR1102319

[2] David Aldous, *The critical beta-splitting random tree model II: Overview and open problems*, 2023, arXiv:2303.02529.

[3] David Aldous and Svante Janson, *The critical beta-splitting random tree model III*: the exchangeable partition representation and the fringe tree, 2024, In preparation.

[4] David Aldous and Svante Janson, *The critical beta-splitting random tree model IV*: Mellin analysis, 2024, In preparation.

[5] David Aldous and Boris Pittel, *The critical beta-splitting random tree I: Heights and related results*, arXiv:2302.05066, 2023. To appear in Ann. Appl. Probab.

[6] Alexander Gnedin and Jim Pitman, *Regenerative composition structures*, Ann. Probab. **33** (2005), 445–479. MR2122798

[7] Cecilia Holmgren and Svante Janson, *Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees*, Probab. Surv. **14** (2017), 53–154. MR3626585

[8] Alexander Iksanov, *A comment on the article 'The harmonic descent chain' by D. J. Aldous, S. Janson and X. Li.* Unpublished, 2024.

[9] Olav Kallenberg, *Foundations of modern probability*, second ed., Probability and its Applications (New York), Springer-Verlag, New York, 2002. MR1876169

[10] Hermann Thorisson, *Shift-coupling in continuous time*, Probab. Theory Related Fields **99** (1994), no. 4, 477–483. MR1288066

Acknowledgments. This article is part of a broad project with Boris Pittel. We thank Alexander Iksanov, Luca Pratelli and Pietro Rigo for discussion of alternate proofs, and we thank reviewers for careful reading and kind comments.