CrossMark

Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees

Igor Kortchemski[1]

Abstract We establish uniform sub-exponential tail bounds for the width, height and maximal outdegree of critical Bienaymé–Galton–Watson trees conditioned on having a large fixed size, whose offspring distribution belongs to the domain of attraction of a stable law. This extends results obtained for the height and width by Addario-Berry, Devroye and Janson in the finite variance case.

Keywords Random trees · Bienaymé–Galton–Watson trees · Spectrally positive stable Lévy processes · Non-crossing trees

Mathematics Subject Classification Primary 60J80 · 05C05 · 05C07; Secondary 60F05 · 60G52

1 Introduction

We are interested in the geometric structure of large Bienaymé–Galton–Watson trees, which are an important well-studied class of random trees in probability theory. They arise as building blocks of many different models of random graphs, such as Erdős–Rényi random graphs or random maps, and appear in combinatorics under the term of simply-generated trees. Addario-Berry et al. [2] established sub-Gaussian tail bounds for the width and height of critical finite-variance Bienaymé–Galton–Watson trees

✉ Igor Kortchemski
igor.kortchemski@normalesup.org

[1] CNRS and CMAP, École polytechnique, Palaiseau, France

🍋 Springer

conditioned on having a fixed size, these bounds being uniform in the size. Such uniform bounds are often challenging to prove, and are useful tools to establish scaling limits for various families of random graphs, see for example [9,12,39,42,45]. Our goal is to obtain similar sub-exponential bounds for the width, height and also maximal outdegree of critical Bienaymé–Galton–Watson trees conditioned on having a fixed size, but whose offspring distribution belongs to the domain of attraction of a stable law and may have infinite variance. We believe that our results should find applications in the study of scaling limits of random graphs exhibiting heavy tail phenomena (see e.g. [6]).

Bienaymé–Galton–Watson trees belonging to domains of attraction of stable laws have recently appeared in a number of two-dimensional statistical physics models, in connection with scaling limits of random maps with large faces [34], scaling limits of critical site-percolation clusters on infinite random triangulations [15] via looptrees [14] and Liouville quantum gravity [17].

1.1 Stable Bienaymé–Galton–Watson trees

We consider critical offspring distributions belonging to domains of attraction of stable laws. Specifically, we fix a parameter $\alpha \in (1, 2]$ and let $\mu = (\mu(j); j \geqslant 0)$ be a probability distribution on the nonnegative integers satisfying the following two conditions:

(i) μ is critical, meaning that $\sum_{j=0}^{\infty} j\mu(j) = 1$.
(ii) μ belongs to the domain of attraction of a stable law of index $\alpha \in (1, 2]$.

By [20, Theorem XVII.5.2], assertion (ii) means that if X is a random variable with distribution μ, then $\mathrm{Var}(X \cdot \mathbb{1}_{X \leqslant n}) = n^{2-\alpha}L(n)$, where $L : \mathbb{R}_+ \to \mathbb{R}_+$ is a function such that $\lim_{x\to\infty} L(tx)/L(x) = 1$ for all $t > 0$ (such a function is called slowly varying). Equivalently, either the variance of μ is finite, or $n^{\alpha}\mu([n, \infty))$ is another slowly varying function (see [20, Eq. (5.16) and Theorem 2 in Sect. XVII.5]).

In addition, we always implicitly suppose that $\mu(0)+\mu(1) < 1$ to avoid degenerate cases, and always assume that μ is aperiodic, in the sense that the additive subgroup of the integers $\mathbb{Z}$ spanned by $\{j; \mu(j) \neq 0\}$ is $\mathbb{Z}$. We let $\mathbb{P}_{\mu}$ denote the law of a (plane, rooted) Bienaymé–Galton–Watson tree with offspring distribution μ, and $|\tau|$ be total number of vertices, or size, of a tree τ. For every $n \geqslant 1$ such that $\mathbb{P}_{\mu}(|\tau| = n) > 0$, $\mathfrak{t}_n$ will denote a BGW_{μ} tree conditioned on having n vertices. The aperiodicity of μ guarantees that $\mathbb{P}_{\mu}(|\tau| = n) > 0$ for every n sufficiently large (our results carry out to the periodic case with mild modifications, and we only focus on the aperiodic case for simplicity).

1.2 Asymptotic behavior of stable Bienaymé–Galton–Watson trees

The asymptotic behavior of $\mathfrak{t}_n$ is well understood, in particular through scaling limits of different functions coding $\mathfrak{t}_n$. Specifically, if $u(0), u(1), \ldots, u(n-1)$ denote the vertices of $\mathfrak{t}_n$ listed in lexicographical order (see Sect. 3 for precise definitions), define the height function $(\mathcal{H}_i(\mathfrak{t}_n); 0 \leqslant i \leqslant n)$ by letting $\mathcal{H}_i(\mathfrak{t}_n)$ be the generation of $u(i)$ in $\mathfrak{t}_n$

for $0 \leqslant i \leqslant n-1$ and setting $\mathcal{H}_n(\mathfrak{t}_n) = 0$ by convention. Define also the Łukasiewicz path $(\mathcal{W}_i(\mathfrak{t}_n); \, 0 \leqslant i \leqslant n)$ of $\mathfrak{t}_n$ by setting $\mathcal{W}_0(\mathfrak{t}_n) = 0$ and, for $0 \leqslant i \leqslant n-1$, $\mathcal{W}_{i+1}(\mathfrak{t}_n) - \mathcal{W}_i(\mathfrak{t}_n) + 1$ to be the outdegree (i.e. number of children) of $u(i)$. Let $(B_n)_{n \geqslant 1}$ be an increasing sequence such that if $(X_i)_{i \geqslant 1}$ is a sequence of i.i.d. random variables with distribution μ, $(X_1 + X_2 + \cdots + X_n - n)/B_n$ converges in distribution as $n \to \infty$ to a random variable Y_α with Laplace exponent given by $\mathbb{E}\left[\exp(-\lambda Y_\alpha)\right] = \exp(\lambda^\alpha)$ for every $\lambda > 0$ ([20, Sect. XVII.5] garanties its existence). Duquesne [18] (see also [30]) showed that the convergence

$$\left(\frac{1}{B_n} \cdot \mathcal{W}_{\lfloor nt \rfloor}(\mathfrak{t}_n), \, \frac{n}{B_n} \cdot \mathcal{H}_{\lfloor nt \rfloor}(\mathfrak{t}_n)\right)_{0 \leqslant t \leqslant 1} \quad \xrightarrow[n \to \infty]{(d)} \quad \left(X_t^{\mathrm{exc}}, H_t^{\mathrm{exc}}\right)_{0 \leqslant t \leqslant 1} \qquad (1)$$

holds in distribution in $\mathbb{D}([0, 1], \mathbb{R})^2$, where $\mathbb{D}([0, 1], \mathbb{R})$ is the space of real-valued càdlàg functions on $[0, 1]$ equipped with the Skorokhod J_1 topology, where X^{exc} is the normalized excursion of a spectrally positive strictly stable Lévy process of index α and H^{exc} its associated continuous height function (we refer to [18] for their construction as we will not use them) which codes the α-stable Lévy tree introduced by Le Gall and Le Jan [33]. In the particular case $\alpha = 2$, we have $(X^{\mathrm{exc}}, H^{\mathrm{exc}}) = (\sqrt{2} \cdot \mathbb{e}, \sqrt{2} \cdot \mathbb{e})$, where $\mathbb{e}$ is the normalized Brownian excursion. The scaling factor B_n is of order $n^{1/\alpha}$ (more precisely, $B_n/n^{1/\alpha}$ is slowly varying), and one may take $B_n = \sigma\sqrt{n/2}$ when μ has finite variance σ^2.

1.3 Uniform bounds on the width and height of stable Bienaymé–Galton–Watson trees

If τ is a (plane, rooted) tree and $k \geqslant 0$, we denote by $Z_k(\tau)$ the number of vertices of τ at generation k, and let

$$W(\tau) = \sup\{Z_k(\tau); \, k \geqslant 0\}, \qquad H(\tau) = \sup\{k; \, Z_k(\tau) > 0\}$$

be respectively the width and height of τ. We also denote by $\Delta(\tau)$ the maximum outdegree of τ. To simplify notation, we will sometimes write Z_k, W, H instead of respectively $Z_k(\tau), W(\tau)$ and $H(\tau)$. Let $\Delta^*(X^{\mathrm{exc}}) = \sup_{0 < t \leqslant 1}(X^{\mathrm{exc}}_t - X^{\mathrm{exc}}_{t-})$ be the maximum jump of X^{exc}. Since the maximum jump of $\mathcal{W}(\mathfrak{t}_n)$ is equal to $\Delta(\mathfrak{t}_n) - 1$ and the largest jump is a continuous functional on $\mathbb{D}([0, 1], \mathbb{R})$ (see e.g. [23, Proposition 2.4 in Chapter VI]; we emphasize that we are always working with the J_1 topology), (1) immediately implies that

$$\left(\frac{1}{B_n} \cdot \Delta(\mathfrak{t}_n), \, \frac{n}{B_n} \cdot H(\mathfrak{t}_n)\right) \quad \xrightarrow[n \to \infty]{(d)} \quad (\Delta^*(X^{\mathrm{exc}}), \sup H^{\mathrm{exc}}).$$

It is also plausible that this convergence holds jointly with that of $W(\mathfrak{t}_n)/B_n$ to a positive random variable, see [26] (unfortunately, we have not managed to found a published reference of this fact). As a consequence, for every $u \geqslant 0$, the quantities $\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right), \mathbb{P}\left(W(\mathfrak{t}_n) \geqslant uB_n\right)$ and $\mathbb{P}\left(\Delta(\mathfrak{t}_n) \geqslant uB_n\right)$ should converge as $n \to$

∞ to functions of u that tend to 0 as $u \to \infty$. It is therefore natural to ask if it is possible to bound $\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right)$, $\mathbb{P}\left(W(\mathfrak{t}_n) \geqslant uB_n\right)$ and $\mathbb{P}\left(\Delta(\mathfrak{t}_n) \geqslant uB_n\right)$ by functions of u which do *not* depend on n.

In the case where μ is critical and has finite positive variance, such bounds have been established by Addario-Berry et al. [2, Theorems 1.1 and 1.2], who show the existence of constants $C_1, c_1 > 0$ (depending only on μ) such that the inequalities

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u\sqrt{n}\right) \leqslant C_1 e^{-c_1 u^2}, \qquad \mathbb{P}\left(W(\mathfrak{t}_n) \geqslant u\sqrt{n}\right) \leqslant C_1 e^{-c_1 u^2}$$

hold for every $n \geqslant 1$ and $u \geqslant 0$. Addario-Berry [1] establishes similar bounds for uniform random trees with a given outdegree sequence satisfying a "finite variance" type condition. When μ is critical and belongs to the domain of attraction of a stable law, Haas and Miermont [21, Lemma 33] show that for every $p > 0$, there exists a constant $C_p > 0$ such that

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant \frac{un}{B_n}\right) \leqslant \frac{C_p}{u^p}$$

for every $n \geqslant 1$ and $u \geqslant 1$ (this reference actually treats the more general case of so-called Markov-branching trees).

We are now in position to state our main results. Recall that $\mathfrak{t}_n$ denotes a Bienaymé–Galton–Watson tree with a critical offspring distribution in the domain of attraction of a stable law of index $\alpha \in (1, 2]$, conditioned on having n vertices.

Theorem 1 (Bounds for the width) *For every $\gamma \in (0, \alpha/(\alpha - 1))$, there exist positive constants $C_1, C_2 > 0$ such that for every $u \geqslant 0$ and every $n \geqslant 1$:*

$$\mathbb{P}\left(W(\mathfrak{t}_n) \geqslant uB_n\right) \leqslant C_1 \exp(-C_2 u^\gamma).$$

The exponent $\alpha/(\alpha - 1)$ is optimal. We will see this by explicitly calculating the tail of the supremum of the stable bridge (Theorem 12) and evaluating its asymptotic behavior (Corollary 13), which are results of independent interest. See also [46] for a study of the width of (non-conditioned) stable Bienaymé–Galton–Watson trees.

Theorem 2 (Bounds for the height) *For every $\delta \in (0, \alpha)$, there exist positive constants $C_1, C_2 > 0$ such that for every $u \geqslant 0$ and every $n \geqslant 1$:*

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n}\right) \leqslant C_1 \exp(-C_2 u^\delta).$$

Here, the exponent α is optimal. Indeed, for every $u \geqslant 0$, we have $\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right) \to \mathbb{P}\left(H^{\mathrm{exc}} \geqslant u\right)$ as $n \to \infty$ by (1), and it is shown in [19, Theorem 1.5] that

$$\mathbb{P}\left(\sup H^{\mathrm{exc}} > u\right) \underset{u \to \infty}{\sim} \beta \cdot u^{1 + \frac{\alpha}{2}} e^{-(\alpha - 1)^{\frac{1}{\alpha-1}} u^\alpha}, \tag{2}$$

where $\beta > 0$ is a positive constant depending only on α.

As noted in [2], since $H(\mathsf{t}_n)W(\mathsf{t}_n) \geqslant n - 1$, the previous results also yield, for every $\gamma \in (0, \alpha/(\alpha - 1))$ and $\delta \in (0, \alpha)$, the existence of constants $C_1, C_2 > 0$ such that

$$\mathbb{P}\left(W(\mathsf{t}_n) \leqslant \frac{B_n}{u}\right) \leqslant C_1 \exp\left(-C_2 u^\delta\right) \quad \text{and} \quad \mathbb{P}\left(H(\mathsf{t}_n) \leqslant \frac{1}{u} \cdot \frac{n}{B_n}\right) \leqslant C_1 \exp\left(-C_2 u^\gamma\right)$$

for every $u > 0$ and $n \geqslant 1$. We believe that the exponent α is optimal for the first inequality. Also, the exponent $\alpha/(\alpha - 1)$ is optimal for the second inequality, since, by [19, Theorem 1.8],

$$\mathbb{P}\left(\sup H^{\text{exc}} < \frac{1}{u}\right) \underset{u \to \infty}{\sim} \lambda \cdot \frac{1}{u^{\alpha + 2 + \frac{1}{\alpha - 1}}} \exp\left(-\left(\frac{\pi/\alpha}{\sin(\pi/\alpha)}\right)^{\frac{\alpha}{\alpha - 1}} \cdot u^{\frac{\alpha}{\alpha - 1}}\right), \quad (3)$$

where $\lambda > 0$ is a positive constant depending only on α.

1.4 Application to the maximal outdegree of stable Bienaymé–Galton–Watson forests

By using Theorem 1, we establish the following result.

Theorem 3 (Bounds for the maximal outdegree) *For every $\delta \in (0, \alpha/(\alpha - 1))$, there exist positive constants $C_1, C_2 > 0$ such that for every $u \geqslant 0$ and every $n \geqslant 1$:*

$$\mathbb{P}\left(\Delta(\mathsf{t}_n) \geqslant u B_n\right) \leqslant C_1 \exp(-C_2 u^\delta).$$

Here we believe that the exponent $\alpha/(\alpha - 1)$ is also optimal. See [4,5,38,41] for results concerning the maximal degree of stable Bienaymé–Galton–Watson trees conditioned on non-extinction at high generation or for the maximal degree of forests.

In addition, we establish the following bounds, which are sharper when μ has finite variance and which also apply to forests of Bienaymé–Galton–Watson trees. For $j \geqslant 1$, denote by $\mathbb{P}_{\mu,j}$ the law of j independent BGW_μ trees.

Theorem 4 (Bounds for the maximal outdegree of a forest) *For every $M > 0$, there exist constants $C_1, C_2 > 0$ such that, for every $n, k \geqslant 1$,*

$$\sup_{1 \leqslant j \leqslant M B_n} \mathbb{P}_{\mu,j}\left(\Delta(F) \leqslant k \mid |F| = n\right) \leqslant C_1 \exp\left(-C_2 n \mu([k+1, \infty))\right) \quad (4)$$

and

$$\sup_{1 \leqslant j \leqslant M B_n} \mathbb{P}_{\mu,j}\left(\Delta(F) \geqslant k \mid |F| = n\right) \leqslant C_1 n \mu([k, \infty)) \quad (5)$$

As an application of these bounds, we obtain concentration inequalities for the maximum outdegree of a large uniform non-crossing tree (Theorem 20 below), improving a result by Deutsch and Noy [16].

1.5 Sizes of generations in stable Bienaymé–Galton–Watson trees

Theorem 2 estimates the probability that t_n has a large height, namely at least un/B_n. One may then wonder: what is the size of the generation at level un/B_n, on the event that t_n has height at least un/B_n? In this direction, we establish the following bounds.

Theorem 5 *For every* $\gamma \in (0, \alpha - 1)$ *and* $\eta > 0$, *there exists a positive constant* $C_1 > 0$ *such that for every* $u \geqslant \eta$, $v \geqslant 0$ *and every* $n \geqslant 1$:

$$\mathbb{P}\left(0 < Z_{u\frac{n}{B_n}}(t_n) < vB_n\right) \leqslant C_1 v^{\gamma}.$$

When u takes values in a compact subset of $(0, \infty)$, the exponent $\alpha - 1$ is optimal (see Remark 25). In proving this, we crucially rely on the following uniform estimate, concerning the size Z_n^* at generation n of a BGW_μ tree conditioned to survive (see Sect. 4.1 for a definition), and which is of independent interest.

Proposition 6 *Set* $p_n = \mathbb{P}_\mu (H \geqslant n)$. *For every* $\beta \in (0, \alpha)$, *there exists a constant* $C > 0$ *such that*

$$\text{for every } n \geqslant 1, x \geqslant 0, \quad \mathbb{P}\left(p_n Z_n^* \leqslant x\right) \leqslant C \cdot x^{\beta}. \tag{6}$$

Let us mention that using a different approach, Croydon and Kumagai [11, Proposition 2.6] show the weaker result that for every $\beta \in (0, \alpha - 1)$, there exists a constant $C > 0$ such that (6) holds. However, in our case, it is important to be able to choose $\beta > \alpha - 1$. The exponent α is optimal, since by [37, Theorem 4], $p_n Z_n^*$ converges in distribution to a random variable Z^* with Laplace transform given by

$$\mathbb{E}\left[e^{-\lambda Z^*}\right] = \frac{1}{\left(1 + \lambda^{\alpha-1}\right)^{\frac{\alpha}{\alpha-1}}}, \quad \lambda \geqslant 0,$$

which shows that for every $\epsilon > 0$, there exists $C > 0$ such that $\mathbb{P}(Z^* \leqslant x) \geqslant Cx^{\alpha+\epsilon}$ for every $0 \leqslant x \leqslant 1$. It would be interesting to know whether (6) holds for $\beta = \alpha$.

Corollary 7 (i) *For every* $\gamma \in (0, \alpha/(\alpha - 1))$ *and* $\delta \in (0, \alpha)$, *there exist positive constants* $C_1, C_2 > 0$ *such that for every* $u, v \geqslant 0$ *and every* $n \geqslant 1$:

$$\mathbb{P}\left(Z_{u\frac{n}{B_n}}(t_n) > vB_n\right) \leqslant C_1 \exp(-C_2(u^{\delta} + v^{\gamma})).$$

(ii) *For every* $\gamma \in (0, (\alpha - 1)/2)$ *and* $\delta \in (0, \alpha)$, *there exist positive constants* $C_1, C_2 > 0$ *such that for every* $u \geqslant \eta$, $v \geqslant 0$ *and every* $n \geqslant 1$:

$$\mathbb{P}\left(0 < Z_{u\frac{n}{B_n}}(t_n) < vB_n\right) \leqslant C_1 v^{\gamma} \exp(-C_2 u^{\delta})$$

The second assertion gives a better bound than Theorem 5 when u is large, but we believe that the exponent $(\alpha - 1)/2$ is not optimal here.

1.6 Techniques

We now comment on the main tools involved in the proof of the bounds for the width and height, and in particular on their connections with [2]. The main tool, as in [2], is the coding of conditioned Bienaymé–Galton–Watson trees by their Łukasiewicz paths, which are, roughly speaking, nonnegative spectrally positive random walks conditioned on a late return to 0. In order to establish the bounds on the width and maximal outdegrees, we establish bounds on the supremum of such walks, following the idea of [2] that since these walks are spectrally positive, reaching high values and then returning to 0 has a sub-exponential cost. However, the implementation of this idea is different, since [2] uses a bound that is only known to hold in the finite variance case (see the discussion after the statement of Theorem 9). The starting observation for proving the bounds for the height is the same as in [2]: if a conditioned Bienaymé–Galton–Watson tree has a large height, then vertices at high generations will have a lot of children branching off their ancestral line to the root, resulting in a large value for the supremum of the Łukasiewicz path, which we already know to have an sub-exponential cost. However, a major difference is that the proof of [2] crucially uses the fact that the width and height of t_n are of the same order $\sqrt{n}$, which breaks down when μ has infinite variance, and the proof thus requires new ideas.

2 Estimates for random walks

Recall that μ is a critical offspring distribution belonging to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. Let $(W_n)_{n \geqslant 0}$ be a random walk with starting point $W_0 = 0$ and jump distribution given by $\mathbb{P}(W_1 = i) = \mu(i + 1)$ for $i \geqslant -1$. Observe that $\mathbb{E}[W_1] = 0$ since μ is critical.

In this section, we study statistics of the random walk $(W_n)_{n \geqslant 0}$ under different types of conditioning. They will play an important role since we will later see that Bienaymé–Galton–Watson related are coded by such walks. It may be useful to refer to Table 1 to keep track of the main notation of this section.

2.1 Large deviations for left-continuous random walks

Recall from the Introduction that $(B_n)_{n \geqslant 1}$ is an increasing sequence such that W_n/B_n converges in distribution as $n \to \infty$ to the random variable Y_α with Laplace exponent given by $\mathbb{E}\left[\exp(-\lambda Y_\alpha)\right] = \exp(\lambda^\alpha)$ for every $\lambda > 0$. We let $d_\alpha(x)$ denote the density of Y_α at $x \in \mathbb{R}$. Note that $d_2(x) = e^{-x^2/4}/\sqrt{4\pi}$ is the density of a centered Gaussian distribution with variance 2, and that $d_\alpha(0) = |\Gamma(-1/\alpha)|^{-1}$ (see [20, Lemma XVII.6.1]).

In addition,

$$\frac{nL(B_n)}{B_n^\alpha} \xrightarrow[n \to \infty]{} \frac{1}{(2 - \alpha)\Gamma(-\alpha)}, \tag{7}$$

where we recall that L is the slowly varying function such that $\mathrm{Var}(X \cdot \mathbb{1}_{X \leqslant n}) = n^{2-\alpha}L(n)$ with X a random variable distributed according to μ (by continuity, the

Table 1 Table of the main notation and symbols appearing in Sect. 2

μ	Critical offspring distribution on $\mathbb{Z}_+$ in the domain of attraction of a stable law of index $\alpha \in (1, 2]$
$L(n)$	Slowly varying function such that $\mathrm{Var}(X \cdot \mathbb{1}_{X \leqslant n}) = n^{2-\alpha}L(n)$, with X distributed as μ
$(W_n; n \geqslant 0)$	Random walk with $W_0 = 0$ and jump distribution $\mathbb{P}(W_1 = i) = \mu(i+1)$ for $i \geqslant -1$
ζ_j	is $\inf\{n \geqslant 1 : W_n = -j\}$
Y_α	Stable random variable with Laplace exponent given by $\mathbb{E}\left[\exp(-\lambda Y_\alpha)\right] = \exp(\lambda^\alpha)$ for $\lambda > 0$
$d_\alpha(x)$	Density of Y_α at $x \in \mathbb{R}$
$(B_n)_{n \geqslant 1}$	Increasing sequence such that $W_n/B_n \to Y_\alpha$ in distribution
X^{exc}	Normalized excursion of a spectrally positive strictly stable Lévy process of index α
X^{br}	Bridge of a spectrally positive strictly stable Lévy process of index α

quantity $((2-\alpha)\Gamma(-\alpha))^{-1}$ is interpreted as equal to 2 for $\alpha = 2$). Indeed, in the notation of [47, Sect. 4.5.1], Y_α is the stable random variable $S_\alpha(|\cos(\pi\alpha/2)|^{1/\alpha}, 1, 0)$ by [47, Eq. (5.17)]. First assume that $1 < \alpha < 2$. It follows from [20, Eq. (5.16)] that $\mathbb{P}(|W_1| \geqslant n) \sim \mathbb{P}(W_1 \geqslant n) \sim \frac{2-\alpha}{\alpha}L(n)n^{-\alpha}$ as $n \to \infty$. Thus, by [47, Theorem 4.5.1],

$$\frac{nL(B_n)}{B_n^\alpha} \xrightarrow[n\to\infty]{} \frac{\alpha}{2-\alpha} \cdot \frac{1-\alpha}{\Gamma(2-\alpha)} = \frac{1}{(2-\alpha)\Gamma(-\alpha)}.$$

If $\alpha = 2$, we have $\mathbb{E}\left[W_1^2 \mathbb{1}_{|W_1|\leqslant n}\right] \sim L(n)n^{2-\alpha}$ as $n \to \infty$, since $\sum_{i=0}^n i^2\mu(i) - 1 \sim L(n)n^{2-\alpha}$ by the definition of L. It then follows from [20, Theorem XVII.5.3] that $nL(B_n)/B_n^2 \to 2$.

The local limit theorem [22, Theorem 4.2.1] shows that

$$\sup_{k\in\mathbb{Z}} \left| B_n\mathbb{P}(W_n = k) - d_\alpha\left(\frac{k}{B_n}\right) \right| \xrightarrow[n\to\infty]{} 0. \tag{8}$$

By the so-called representation theorem (see e.g. [7, Theorem 1.3.1]), we can write

$$L(x) = c(x)\exp\left(\int_1^x \frac{\eta(u)}{u}du\right), \quad x \geqslant 0,$$

where c is a non-negative measurable function having a finite positive limit at infinity and η is a measurable function tending to 0 at infinity. It easily follows that if

$(x_n, y_n)_{n \geqslant 1}$ are two sequences tending to infinity, for every $\epsilon > 0$, there exists a constant $C > 1$ such that for every integer n sufficiently large:

$$C^{-1} \left(\frac{\max(x_n, y_n)}{\min(x_n, y_n)} \right)^{-\epsilon} \leqslant \frac{L(y_n)}{L(x_n)} \leqslant C \left(\frac{\max(x_n, y_n)}{\min(x_n, y_n)} \right)^{\epsilon}, \qquad (9)$$

Similar bounds hold with $L(n)$ replaced by $B_n / n^{1/\alpha}$, since the latter quantity is slowly varying. In the literature, these inequalities are known as the Potter bounds.

We shall establish the following estimate.

Proposition 8 *For every $\delta \in (0, \alpha/(\alpha - 1))$ there exists $C_1, C_2 > 0$ such that for every $u \geqslant 0$ and every $n \geqslant 1$:*

$$\mathbb{P} \left(\min_{0 \leqslant i \leqslant n} W_i \leqslant -u B_n \right) \leqslant C_1 \exp(-C_2 u^\delta).$$

Proof Here C will stand for a positive constant which may vary from expression to expression (but that is independent of u and n). Note that $\mathbb{P} \left(\min_{0 \leqslant i \leqslant n} W_i \leqslant -u B_n \right) = 0$ if $u B_n > n$, so that we can suppose without loss of generality that $1 \leqslant u \leqslant n/B_n$. Write, for $h > 0$:

$$\mathbb{P} \left(\min_{0 \leqslant i \leqslant n} W_i \leqslant -u B_n \right) = \mathbb{P} \left(\max_{0 \leqslant i \leqslant n} e^{h(-W_i)} \geqslant e^{hu B_n} \right)$$
$$\leqslant e^{-hu B_n} \mathbb{E} \left[e^{-h W_n} \right] = e^{-hu B_n} \mathbb{E} \left[e^{-h W_1} \right]^n \qquad (10)$$

where we have used Doob's maximal inequality with the submartingale $\left(e^{h(-W_n)}; n \geqslant 0 \right)$ for the inequality. Fix $\eta \in (0, 1/(\alpha - 1))$ and note that $\eta\alpha < 1 + \eta$. We shall apply the inequality (10) with $h = h_n(u) = u^\eta / B_n$. Observe that by the Potter bounds, $u^\eta / B_n \to 0$ as $n \to \infty$, uniformly in $1 \leqslant u \leqslant n/B_n$. Therefore, by the estimate (42) of the Appendix, for every $n \geqslant 1$ and $1 \leqslant u \leqslant n/B_n$,

$$\mathbb{E} \left[e^{-\frac{u^\eta}{B_n} W_1} \right]^n \leqslant \exp \left(C n L \left(\frac{B_n}{u^\eta} \right) \frac{u^{\eta\alpha}}{B_n^\alpha} \right)$$

Now choose $\epsilon > 0$ such that $1 + \eta > \eta(\alpha + \epsilon)$. For every n sufficiently large and $u \geqslant 1$, we have $L(B_n/u^\eta) \leqslant C u^{\epsilon\eta} L(B_n)$ by the Potter bounds, so that

$$\exp \left(C n L \left(\frac{B_n}{u^\eta} \right) \frac{u^{\eta\alpha}}{B_n^\alpha} \right) \leqslant \exp \left(C \frac{n L (B_n)}{B_n^\alpha} u^{\eta(\alpha + \epsilon')} \right).$$

By (7), $n L (B_n)/B_n^\alpha$ is bounded as n varies. It follows that for n sufficiently large and $1 \leqslant u \leqslant n/B_n$,

$$\mathbb{E} \left[e^{-\frac{u^\eta}{B_n} W_1} \right]^n \leqslant \exp \left(C u^{\eta(\alpha + \epsilon)} \right). \qquad (11)$$

Putting together (10) and (11), we get

$$\mathbb{P}\left(\min_{0 \leqslant i \leqslant n} W_i \leqslant -u B_n\right) \leqslant \exp\left(-u^{1+\eta} + C u^{\eta(\alpha+\epsilon)}\right).$$

Setting $\delta = 1 + \eta$, the conclusion readily follows from the choice of ϵ.   □

2.2 Conditioned random walks and cyclic shifts

In view of applications for Bienaymé–Galton–Watson trees, we will need estimates on conditioned random walks. We establish in particular a conditioned version of Proposition 8. For every $j \geqslant 1$, set $\zeta_j = \inf\{n \geqslant 1 : W_n = -j\}$.

Theorem 9 *For every $\delta \in (0, \alpha/(\alpha-1))$, there exist $C_1, C_2 > 0$ such that for every $u \geqslant 0$ and every $n \geqslant 1$:*

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant n} W_i \geqslant u B_n \mid \zeta_1 = n\right) \leqslant C_1 \exp(-C_2 u^\delta).$$

When $\sigma^2 < \infty$, this result is established in [2, Eq. (32)] by using a sub-exponential upper bound due to Janson [24] on $\mathbb{P}(W_n = -m)$ valid for every $n \geqslant 1$ and $m \geqslant 0$. In the infinite variance case, a similar bound has been established in [15, Lemma 6.6] when $\mu(n) \sim C \cdot n^{-(1+\alpha)}$ as $n \to \infty$, but is not known to hold in general. For this reason, we combine Proposition 8 with results of [1] for the proof of Theorem 9.

A useful tool for the proof of Theorem 9 is the Vervaat transform, which we now introduce. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$ and $i \in \mathbb{Z}/n\mathbb{Z}$, denote by $\mathbf{x}^{(i)}$ the ith cyclic shift of $\mathbf{x}$ defined by $x_k^{(i)} = x_{i+k \mod n}$ for $1 \leqslant k \leqslant n$. Let $n \geqslant 1$ be an integer and let $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$. Set $w_j = x_1 + \cdots + x_j$ for $1 \leqslant j \leqslant n$ and let the integer $i_*(\mathbf{x})$ be defined by $i_*(\mathbf{x}) = \inf\{j \geqslant 1; w_j = \min_{1 \leqslant i \leqslant n} w_i\}$. The Vervaat transform of $\mathbf{x}$, denoted by $\mathbf{V}(\mathbf{x})$, is defined to be $\mathbf{x}^{(i_*(\mathbf{x}))}$. The following fact is well known (see e.g. [40]):

Proposition 10 *Under the conditional probability distribution $\mathbb{P}(\cdot \mid W_n = -1)$, the vector $\mathbf{V}(W_1 - W_0, \ldots, W_n - W_{n-1})$ has the same distribution as $(W_1 - W_0, \ldots, W_n - W_{n-1})$ under the conditional probability distribution $\mathbb{P}(\cdot \mid \zeta_1 = n)$.*

Proof of Theorem 9 To simplify notation, set $(\overline{X}_1, \ldots, \overline{X}_n) = \mathbf{V}(W_1 - W_0, \ldots, W_n - W_{n-1})$. Noting that

$$\max_{1 \leqslant i \leqslant n} \left(\overline{X}_1 + \cdots + \overline{X}_i\right) \leqslant \max_{1 \leqslant i \leqslant n} W_i - \min_{1 \leqslant i \leqslant n} W_i \quad \text{on the event } W_n = -1,$$

Proposition 10 gives that

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant n} W_i \geqslant i + 3 \mid \zeta_1 = n\right) \leqslant \mathbb{P}\left(\max_{1 \leqslant i \leqslant n} W_i - \min_{1 \leqslant i \leqslant n} W_i \geqslant i + 3 \mid W_n = -1\right)$$

(12)

for every $i \geqslant 1$ and $n \geqslant 1$. Fix $m \geqslant i$. The proof of Eq. (3) in [1] shows that on the event $\max_{1 \leqslant i \leqslant n} W_i - \min_{1 \leqslant i \leqslant n} W_i = m + 3$, at least one of the following three events hold:

$$\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} (W_{\lfloor n/2 \rfloor} - W_{\lfloor n/2 \rfloor - i}) \leqslant -(m + 3)/3,$$
$$\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} (W_n - W_{n-i}) \leqslant -(m + 3)/3$$

or $\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor}(W_i - W_{\lfloor n/2 \rfloor}) \leqslant -(m + 3)/3$ (a close inspection indicates that condition d. in [1] should actually be $\max_{\lfloor n/2 \rfloor < i \leqslant n} S_i > (m + 1)/3$ instead of $\max_{\lfloor n/2 \rfloor < i \leqslant n} S_i > 2(m + 3)/3$. As a consequence, by monotonicity,

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant n} W_i - \min_{1 \leqslant i \leqslant n} W_i \geqslant m + 3 \,\middle|\, W_n = -1 \right)$$
$$\leqslant 3\mathbb{P}\left(\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} W_i \leqslant -(m/3 + 1) \mid W_n = -1 \right).$$

Then, setting $\varphi_n(j) = \mathbb{P}(W_n = -j)$ to simplify notation, the Markov property for the random walk W applied at time $\lfloor n/2 \rfloor$ entails that

$$\mathbb{P}\left(\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} W_i \leqslant -\left(\frac{m}{3} + 1 \right) \mid W_n = -1 \right)$$
$$= \mathbb{E}\left[1_{\{\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} W_i \leqslant -(\frac{m}{3}+1)\}} \frac{\varphi_{n-\lfloor n/2 \rfloor}(W_{\lfloor n/2 \rfloor} + 1)}{\varphi_n(1)} \right].$$

But the local limit theorem (8) yields the existence of a constant $C > 0$ such that $\varphi_{n-\lfloor n/2 \rfloor}(k)/\varphi_n(1) \leqslant C$ for every $n \geqslant 1$ and $k \in \mathbb{Z}$. Hence, by the previous estimates,

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant n} W_i \geqslant 3u B_n \,\middle|\, \zeta_1 = n \right) \leqslant 3C\mathbb{P}\left(\min_{0 \leqslant i \leqslant \lfloor n/2 \rfloor} W_i \leqslant -u B_n \right)$$

for every $u \geqslant 0$ and $n \geqslant 1$. The conclusion then follows by an application of Proposition 8. □

Recall from the Introduction that X^{exc} denotes the normalized excursion of a spectrally positive strictly stable Lévy process of index α. It is well known that the random walk W, conditionally on $\zeta_1 = n$ and appropriately rescaled, converges in distribution to X^{exc} for the Skorokhod topology on $\mathbb{D}([0, 1], \mathbb{R})$ (see e.g. [18, Proof of Theorem 3.1]):

$$\left(\frac{1}{B_n} W_{\lfloor nt \rfloor}; 0 \leqslant t \leqslant 1 \right) \quad \text{under} \quad \mathbb{P}(\cdot \mid \zeta_1 = n) \xrightarrow[n \to \infty]{(d)} (X_t^{\text{exc}}; 0 \leqslant t \leqslant 1).$$

Since the supremum is a continuous function on $\mathbb{D}([0, 1], \mathbb{R})$, we then get from Theorem 9 that for every $\delta \in (0, \alpha/(\alpha - 1))$ there exists $C_1, C_2 > 0$ such that for

every $u \geqslant 0$ and every $n \geqslant 1$:

$$\mathbb{P}\left(\sup X^{\mathrm{exc}} \geqslant u\right) \leqslant C_1 \exp(-C_2 u^{\delta}).$$

It would be interesting to obtain an asymptotic expansion of $\mathbb{P}\left(\sup X^{\mathrm{exc}} \geqslant u\right)$ as $u \to \infty$, similar to the one known for $\alpha = 2$ (i.e. the Brownian excursion) involving a Theta function.

Finally, we will need a well-known result in the folklore of exchangeability (see [28, Sect. 1]), for which we give a proof for completeness. Fix $n \geqslant 1$. A function $F : \mathbb{R}^n \to \mathbb{R}$ is said to be invariant under cyclic shifts if $F(\mathbf{x}) = F(\mathbf{x}^{(i)})$ for every $\mathbf{x} \in \mathbb{R}^n$ and $i \in \mathbb{Z}/n\mathbb{Z}$. For $n \geqslant 1$, introduce $X_n = W_n - W_{n-1}$. Finally, recall that $\zeta_j = \inf\{n \geqslant 1 : W_n = -j\}$.

Lemma 11 *Let $F : \mathbb{R}^n \to \mathbb{R}$ be a function invariant under cyclic shifts. Then*

$$\mathbb{E}\left[F(X_1, \ldots, X_n)\mathbb{1}_{\zeta_j=n}\right] = \frac{j}{n}\mathbb{E}\left[F(X_1, \ldots, X_n)\mathbb{1}_{W_n=-j}\right].$$

In particular,

$$F(X_1, \ldots, X_n) \;\; under \;\; \mathbb{P}\left(\cdot \mid \zeta_j = n\right) \;\; \overset{(d)}{=} \;\; F(X_1, \ldots, X_n) \;\; under \;\; \mathbb{P}\left(\cdot \mid W_n = -j\right). \tag{13}$$

This result will be later used to study the maximal outdegree of (a forest of) Bienaymé–Galton–Watson trees. Its proof uses the so-called Cyclic Lemma. Before stating it, we need to introduce some notation. For $j \geqslant 1$, define:

$$\mathcal{S}_n^{(j)} = \left\{(x_1, \ldots, x_n) \in \{-1, 0, 1, 2, \ldots\}^n; \; \sum_{i=1}^{n} x_i = -j\right\}$$

and

$$\overline{\mathcal{S}}_n^{(j)} = \left\{(x_1, \ldots, x_n) \in \mathcal{S}_n^{(j)}; \; \sum_{i=1}^{m} x_i > -j \text{ for all } m \in \{0, 1, \ldots, n-1\}\right\}.$$

For $\mathbf{x} \in \mathcal{S}_n^{(j)}$, finally set $\mathcal{I}_{\mathbf{x}} = \{i \in \mathbb{Z}/n\mathbb{Z}; \; \mathbf{x}^{(i)} \in \overline{\mathcal{S}}_n^{(j)}\}$. The so-called Cyclic Lemma states that we have $\mathrm{Card}(\mathcal{I}_{\mathbf{x}}) = j$ for every $\mathbf{x} \in \mathcal{S}_n^{(j)}$ (see [40, Lemma 6.1] for a proof).

Proof of Lemma 11 Set $\mathbf{X}_n = (X_1, \ldots, X_n)$, and note that $W_n = -j$ if and only if $\mathbf{X}_n^{(i)} \in \mathcal{S}_n^{(j)}$ for a (or, equivalently, every) $i \in \mathbb{Z}/n\mathbb{Z}$, that $\zeta_j = n$ if and only if $\mathbf{X}_n \in \overline{\mathcal{S}}_n^{(j)}$ and finally that $\mathbf{X}_n^{(i)}$ has the same distribution as $\mathbf{X}_n$ for every $i \in \mathbb{Z}/n\mathbb{Z}$. Then write

$$\mathbb{E}\left[F\left(X_1,\ldots,X_n\right)\mathbb{1}_{\{\zeta_j=n\}}\right] = \mathbb{E}\left[F\left(\mathbf{X}_n\right)\mathbb{1}_{\left\{\mathbf{X}_n\in\overline{\mathcal{S}}_n^{(j)}\right\}}\right]$$

$$= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[F\left(\mathbf{X}_n^{(i)}\right)\mathbb{1}_{\left\{\mathbf{X}_n^{(i)}\in\overline{\mathcal{S}}_n^{(j)}\right\}}\right]$$

$$= \frac{1}{n}\sum_{i=0}^{n-1}\mathbb{E}\left[F\left(\mathbf{X}_n\right)\mathbb{1}_{\left\{\mathbf{X}_n^{(i)}\in\overline{\mathcal{S}}_n^{(j)}\right\}}\right]$$

$$= \frac{1}{n}\mathbb{E}\left[F\left(\mathbf{X}_n\right)\left(\sum_{i=0}^{n-1}\mathbb{1}_{\left\{\mathbf{X}_n^{(i)}\in\overline{\mathcal{S}}_n^{(j)}\right\}}\right)\mathbb{1}_{\left\{\mathbf{X}_n\in\mathcal{S}_n^{(j)}\right\}}\right]$$

$$= \frac{1}{n}\mathbb{E}\left[F\left(\mathbf{X}_n\right)\mathcal{I}_{\mathbf{X}_n}\mathbb{1}_{\left\{\mathbf{X}_n\in\mathcal{S}_n^{(j)}\right\}}\right]$$

$$= \frac{j}{n}\mathbb{E}\left[F\left(\mathbf{X}_n\right)\mathbb{1}_{\left\{\mathbf{X}_n\in\mathcal{S}_n^{(j)}\right\}}\right].$$

For the third equality, we have used the fact that F is invariant under cyclic shifts, and for the last equality we have used the Cyclic Lemma, which tells us that $\mathrm{Card}(\mathcal{I}_{\mathbf{X}_n}) = j$ on the event $\mathbf{X}_n \in \mathcal{S}_n^{(j)}$. This completes the proof of the first assertion.

The second one readily follows after noting that

$$\mathbb{P}\left(\zeta_j = n\right) = \frac{j}{n}\cdot\mathbb{P}\left(W_n = -j\right) \tag{14}$$

by taking F to be the constant function equal to 1. The identity (14) is often referred to as Kemperman's formula in the literature. $\qquad\square$

2.3 Bridge estimates

We now establish a tail estimate for the supremum of a stable Lévy bridge, which will allow us to see that the exponent $\alpha/(\alpha-1)$ is optimal in Theorem 9 and which is also of independent interest. We will in addition see that this gives a tail bound for the value of the stable excursion X^{exc} evaluated at a uniform point.

Denote by $(X_s^{\mathrm{br}}; 0 \leqslant s \leqslant 1)$ the stable Lévy bridge of index α, which is roughly speaking the α-stable Lévy process normalized such that its law at time 1 is Y_α, and conditioned to return to 0 at time 1 (see [10] or [3, Chapter VIII] for a rigorous construction). Recall that d_α denotes the density of Y_α.

Theorem 12 *For every $u > 0$, we have*

$$\mathbb{P}\left(\sup_{0\leqslant s\leqslant 1}X_s^{\mathrm{br}} \geqslant u\right) = |\Gamma(-1/\alpha)|\cdot u\int_0^1 ds\frac{1}{s^{1/\alpha}(1-s)^{1+1/\alpha}}d_\alpha\left(\frac{u}{s^{1/\alpha}}\right)$$

$$\times d_\alpha\left(-\frac{u}{(1-s)^{1/\alpha}}\right).$$

For $\alpha = 2$, we will see in Remark 14 below that this quantity is actually equal to e^{-u^2}. However, for $\alpha \in (1, 2)$ we have the following interesting asymptotic behavior:

Corollary 13 *For $\alpha \in (1, 2)$, we have*

$$\mathbb{P}\left(\sup_{0 \leqslant s \leqslant 1} X_s^{\mathrm{br}} \geqslant u\right) \underset{u \to \infty}{\sim} \frac{|\Gamma(-1/\alpha)|}{\Gamma(-\alpha)} \cdot \frac{\alpha^{(4\alpha-1)/(2\alpha-2)}}{\sqrt{2\pi(\alpha-1)}} \cdot u^{-\frac{(2+\alpha)(2\alpha-1)}{2(\alpha-1)}}$$
$$\times e^{-(\alpha-1)\alpha^{-\alpha/(\alpha-1)}u^{\alpha/(\alpha-1)}}.$$

Proof For $\alpha \in (1, 2)$, it is known that

$$d_\alpha(x) \underset{x \to \infty}{\sim} \frac{1}{\Gamma(-\alpha)} \cdot \frac{1}{x^{1+\alpha}}, \quad d_\alpha(-x) \underset{x \to \infty}{\sim} \frac{\alpha^{-1/(2\alpha-2)}}{\sqrt{2\pi(\alpha-1)}} x^{-1+\frac{\alpha}{2(\alpha-1)}}$$
$$\times e^{-(\alpha-1)\alpha^{-\alpha/(\alpha-1)}x^{\alpha/(\alpha-1)}}.$$

This follows from the first two terms of the asymptotic series [48, Eqs. (2.5.4) and (2.5.17)], since in the notation of the latter reference $d_\alpha(x) = g(x, \alpha, 1)$ and $d_\alpha(-x) = g(x, \alpha, -1)$ (see Sect. I.4. in [48] for the definition of g).

Since, for $s \in (0, 1)$ we have $u/s^{1/\alpha} > u$ and $u/(1 - s)^{1/\alpha} > u$, we can replace $d_\alpha(u/s^{1/\alpha})$ and $d_\alpha(-u/(1 - s)^{1/\alpha})$ by their asymptotic equivalents and get that

$$\mathbb{P}\left(\sup_{0 \leqslant s \leqslant 1} X_s^{\mathrm{br}} \geqslant u\right) \underset{u \to \infty}{\sim} |\Gamma(-1/\alpha)| \cdot \frac{1}{\Gamma(-\alpha)} \cdot \frac{\alpha^{-1/(2\alpha-2)}}{\sqrt{2\pi(\alpha-1)}} \cdot u^{\frac{\alpha}{2(\alpha-1)}-1-\alpha}$$
$$\times \int_0^1 ds \, \frac{s}{(1-s)^{\frac{1}{2(\alpha-2)}+1}} e^{-c_\alpha \frac{u^{\alpha/(\alpha-1)}}{(1-s)^{1/(\alpha-1)}}}, \tag{15}$$

where $c_\alpha = (\alpha - 1)\alpha^{-\alpha/(\alpha-1)}$. By making the change of variable $t = (1 - s)^{-1/(\alpha-1)}$, we see that

$$\int_0^1 ds \, \frac{s}{(1-s)^{\frac{1}{2(\alpha-1)}+1}} e^{-c_\alpha \frac{u^{\alpha/(\alpha-1)}}{(1-s)^{1/(\alpha-1)}}}$$
$$= (\alpha - 1) \int_1^\infty dt \, t^{-1/2} \left(1 - \frac{1}{t^{\alpha-1}}\right) e^{-c_\alpha u^{\alpha/(\alpha-1)}t}.$$

It is a simple matter to check that

$$\int_1^\infty dt \, t^{-1/2} \left(1 - \frac{1}{t^{\alpha-1}}\right) e^{-xt} \underset{x \to \infty}{\sim} (\alpha - 1)\frac{e^{-x}}{x^2}.$$

Hence, applying this with $x = c_\alpha u^{\alpha/(\alpha-1)}$,

$$\int_0^1 ds \; \frac{s}{(1-s)^{\frac{1}{2(\alpha-1)}+1}} e^{-c_\alpha \frac{u^{\alpha/(\alpha-1)}}{(1-s)^{1/(\alpha-1)}}} \underset{u\to\infty}{\sim} \alpha^{2\alpha/(\alpha-1)} \cdot u^{-2\alpha/(\alpha-1)} \cdot e^{-c_\alpha u^{\alpha/(\alpha-1)}}.$$

The desired estimate then follows from (15). □

Remark 14 Assume that $\alpha \in (1, 2]$. For $0 \leqslant s \leqslant 1$, set

$$X_t^{\mathrm{br},(s)} = \begin{cases} X_{s+t}^{\mathrm{br}} - X_s^{\mathrm{br}} & \text{if } 0 \leqslant t \leqslant 1-s \\ X_{s+t-1}^{\mathrm{br}} - X_s^{\mathrm{br}} & \text{if } 1-s \leqslant t \leqslant 1. \end{cases}$$

Set also $T^* = \inf\{t \in [0,1] : X_t^{\mathrm{br}} = \inf X^{\mathrm{br}}\}$. It is well known that for every fixed $s \in [0,1]$, $X^{\mathrm{br},(s)}$ has the same law as X^{br} and that $X^{\mathrm{br},(T^*)}$ has the same law as X^{exc}. In addition, if U is a uniform random variable on $[0,1]$, independent of X^{br}, then the random variable defined by

$$[T^* + U] \quad = \quad (T^* + U)\mathbb{1}_{\{T^*+U\leqslant 1\}} + (T^* + U - 1)\mathbb{1}_{T^*+U>1}$$

is uniform on $[0,1]$ and independent of X^{br}. Hence

$$X_U^{\mathrm{exc}} \overset{(d)}{=} X_U^{\mathrm{br},(T^*)} = X_{[T^*+U]}^{\mathrm{br}} - \inf X^{\mathrm{br}} \overset{(d)}{=} X_U^{\mathrm{br}} - \inf X^{\mathrm{br}}$$
$$= -\inf X^{\mathrm{br},(U)} \overset{(d)}{=} -\inf X^{\mathrm{br}} \overset{(d)}{=} \sup X^{\mathrm{br}}.$$

As a consequence, Theorem 12 and Corollary 13 hold with $\sup_{0\leqslant s\leqslant 1} X_s^{\mathrm{br}}$ replaced by X_U^{exc}.

Also, in the case $\alpha = 2$, we have $X^{\mathrm{exc}} = \sqrt{2}\mathrm{e}$, where e is the normalized Brownian excursion. It is well known that $2\mathrm{e}_U$ is distributed as a Rayleigh random variable. Specifically, $\mathbb{P}(2\mathrm{e}_U \geqslant x) = e^{-x^2/2}$ for $x \geqslant 0$. In particular, for $\alpha = 2$,

$$\mathbb{P}\left(X_U^{\mathrm{exc}} \geqslant u\right) = \mathbb{P}\left(\sup_{0\leqslant s\leqslant 1} X_s^{\mathrm{br}} \geqslant u\right) = e^{-u^2}.$$

It is possible to check that Theorem 12 gives indeed this expression for $\alpha = 2$.

Finally, since $X^{\mathrm{br},(T^*)}$ has the same law as X^{exc} and clearly $\sup X^{\mathrm{br},(T^*)} \geqslant \sup X^{\mathrm{br}}$, we have the inequality $\mathbb{P}\left(\sup X^{\mathrm{br}} \geqslant u\right) \leqslant \mathbb{P}\left(\sup X^{\mathrm{exc}} \geqslant u\right)$ for every $u \geqslant 0$. Corollary 13 thus implies that the exponent $\alpha/(\alpha-1)$ is optimal in Theorem 9.

Proof of Theorem 12 Fix $u > 0$. We will prove the result by a discrete approximation of the stable Lévy bridge. It is well-known that (see e.g. [18, Proposition 4.3])

$$\left(\frac{1}{B_n} W_{\lfloor nt \rfloor}; 0 \leqslant t \leqslant 1\right) \quad \text{under} \quad \mathbb{P}\left(\cdot \mid W_n = -1\right) \xrightarrow[n\to\infty]{(d)} (X_t^{\mathrm{br}}; 0 \leqslant t \leqslant 1). \quad (16)$$

It is therefore enough to estimate $\mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n \mid W_n = -1\right)$ for $u > 0$ and $n \geqslant 1$. To this end, fix $\eta \in (0, 1)$, set $T^{(n)} = \max\{i \geqslant 0; W_i = \lfloor uB_n\rfloor\}$. If $1 \leqslant j \leqslant n - 1$, note that conditionally on the event $\{T^{(n)} = j, W_n = -1\}$, the sequence $(W_n - W_{n-i}; 1 \leqslant i \leqslant n-j)$ has the same distribution as $(W_i; 1 \leqslant i \leqslant n-j)$ conditionally on the event $\{\zeta_{\lfloor uB_n\rfloor+1} = n - j\}$. Hence, using also (14), we can write

$$\mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n, \eta n \leqslant T^{(n)} \leqslant (1-\eta)n \mid W_n = -1\right)$$
$$= \frac{1}{\mathbb{P}(W_n = -1)} \sum_{j=\eta n}^{(1-\eta)n} \mathbb{P}\left(W_j = \lfloor uB_n\rfloor\right) \frac{1 + \lfloor uB_n\rfloor}{n - j} \mathbb{P}\left(W_{n-j} = -1 - \lfloor uB_n\rfloor\right).$$

Leaving details to the reader (see e.g. the proof of Theorem 3.1 in [29] for similar arguments), the dominated convergence theorem combined with the local limit theorem (8) yields

$$\mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n, \eta n \leqslant T^{(n)} \leqslant (1-\eta)n \mid W_n = -1\right)$$
$$\xrightarrow[n\to\infty]{} \frac{u}{d_\alpha(0)} \int_\eta^{1-\eta} ds \frac{1}{s^{1/\alpha}(1-s)^{1+1/\alpha}} d_\alpha\left(\frac{u}{s^{1/\alpha}}\right) d_\alpha\left(-\frac{u}{(1-s)^{1/\alpha}}\right).$$

In addition, by (16),

$$\mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n, T^{(n)} \notin [\eta n, (1-\eta)n] \mid W_n = -1\right)$$
$$\xrightarrow[n\to\infty]{} \mathbb{P}\left(\sup X^{\mathrm{br}} \geqslant u, \ \sup\{s \in [0, 1]; X_s^{\mathrm{br}} = u\} \notin [\eta n, (1-\eta)n]\right).$$

By monotone convergence, the last probability tends to 0 as $\eta \to 1$. Therefore, for every fixed $\epsilon > 0$, one may find $\eta \in (0, 1)$ such that for every n sufficiently large,

$$\left| \mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n \mid W_n = -1\right) \right.$$
$$\left. -\mathbb{P}\left(\max_{0\leqslant i\leqslant n} W_i \geqslant uB_n, \eta n \leqslant T^{(n)} \leqslant (1-\eta)n \mid W_n = -1\right) \right| \leqslant \epsilon$$

and

$$\frac{u}{d_\alpha(0)} \int_{[0,1]\setminus[\eta,1-\eta]} ds \frac{1}{s^{1/\alpha}(1-s)^{1+1/\alpha}} d_\alpha\left(\frac{u}{s^{1/\alpha}}\right) d_\alpha\left(-\frac{u}{(1-s)^{1/\alpha}}\right) \leqslant \epsilon.$$

Recalling that $d_\alpha(0) = |\Gamma(-1/\alpha)|^{-1}$, the conclusion readily follows.   □

3 Random walks and Bienaymé–Galton–Watson trees

We now explain how trees can be coded by different functions that allow to establish tail bounds for different statistics of large Bienaymé–Galton–Watson trees. In addition to Table 1, it may be useful to refer to Table 2 to keep track of the main notation of this section.

3.1 Definitions

Denote by $\mathbb{N} = \{1, 2, \ldots\}$ the set of the positive integers, set $\mathbb{N}^0 = \{\varnothing\}$ and let

$$\mathcal{U} = \bigcup_{n \geq 0} \mathbb{N}^n.$$

For $u = (u_1, \ldots, u_n) \in \mathcal{U}$, we denote by $|u| = n$ the length of u; if $n \geq 1$, we define $pr(u) = (u_1, \ldots, u_{n-1})$ and for $i \geq 1$, we let $ui = (u_1, \ldots, u_n, i)$; more generally, for $v = (v_1, \ldots, v_m) \in \mathcal{U}$, we let $uv = (u_1, \ldots, u_n, v_1, \ldots, v_m) \in \mathcal{U}$ be the concatenation of u and v. A plane rooted tree is a nonempty, finite subset $\tau \subset \mathcal{U}$ such that: (i) $\varnothing \in \tau$, (ii) if $u \in \tau$ with $|u| \geq 1$, then $pr(u) \in \tau$, (iii) if $u \in \tau$, then there exists an integer $k_u \geq 0$ such that $ui \in \tau$ if and only if $1 \leq i \leq k_u$. In the sequel, by tree we always mean plane rooted tree.

We will view each vertex u of a tree τ as an individual of a population whose τ is the genealogical tree. The vertex $\varnothing$ is called the root of the tree and for every $u \in \tau$, $k_u = k_u(\tau)$ is the number of children (or outdegree) of u, $|u|$ is its generation, $pr(u)$ is its parent and more generally, the vertices $u, pr(u), pr \circ pr(u), \ldots, pr^{|u|}(u) = \varnothing$ are its ancestors. If u is an ancestor of v, we let $[\![u, v]\!]$ be the shortest path between u and v. The size $|\tau|$ of a tree is its total number of vertices. If τ is a tree and $u \in \tau$, we define the shift of τ at u by $\theta_u \tau = \{v \in U;\ uv \in \tau\}$, which is itself a tree.

Since μ is critical, the law of the Bienaymé–Galton–Watson tree with offspring distribution μ is the unique probability measure $\mathbb{P}_\mu$ on the set of all finite plane trees such that for every $j \geqslant 0$, $\mathbb{P}_\mu(k_\varnothing = j) = \mu(j)$, and for every $j \geqslant 1$ with $\mu(j) > 0$, the shifted trees $\theta_1 \tau, \ldots, \theta_j \tau$ are independent under the conditional probability $\mathbb{P}_\mu(\cdot \mid k_\varnothing = j)$ and their conditional distribution is $\mathbb{P}_\mu$. A random tree whose distribution is $\mathbb{P}_\mu$ will be called a Bienaymé–Galton–Watson tree with offspring distribution μ, or in short a BGW_μ tree.

Table 2 Table of the main notation and symbols appearing in Sect. 3

$\mathcal{W}^*(\tau)$ for $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$	The coding path of a tree τ obtained by using respectively the lexicographical, reverse-lexicographical and breadth-first search ordering of the vertices of τ
$W(\tau)$	Width of a tree τ
$\Delta(\tau)$	Maximum outdegree of a vertex of tree τ
$\Delta^*(X^{\mathrm{exc}})$	Maximum jump of X^{exc}

3.2 Coding trees by left-continuous paths

We associate with every ordering $u(0) \prec u(1) \prec \cdots \prec u(|\tau|-1)$ of the vertices of τ a path $\mathcal{W}(\tau) = (\mathcal{W}_n(\tau), 0 \leqslant n \leqslant |\tau|)$ defined by $\mathcal{W}_0(\tau) = 0$ and for $0 \leqslant n \leqslant |\tau| - 1$:

$$\mathcal{W}_{n+1}(\tau) = \mathcal{W}_n(\tau) + k_{u(n)}(\tau) - 1.$$

Note that necessarily $\mathcal{W}_{|\tau|}(\tau) = -1$. As in [2], we will use three different orderings of the vertices of a tree:

(i) the lexicographical ordering, where $v \prec w$ if there exists $z \in \mathcal{U}$ such that $v = z(v_1, \ldots, v_n)$, $w = z(w_1, \ldots, w_m)$ and $v_1 < w_1$;

(ii) the reverse-lexicographical ordering, where $v \prec w$ if there exists $z \in \mathcal{U}$ such that $v = z(v_1, \ldots, v_n)$, $w = z(w_1, \ldots, w_m)$ and $v_1 > w_1$;

(iii) the breadth-first search ordering, where $v \prec w$ if $|v| < |w|$ or $v = z(v_1, \ldots, v_m)$ and $w = z(w_1, \ldots, w_m)$ with $v_1 < w_1$.

Denote by $\mathcal{W}^{\mathsf{lex}}(\tau), \mathcal{W}^{\mathsf{rev}}(\tau), \mathcal{W}^{\mathsf{bfs}}(\tau)$ the paths constructed by using respectively the lexicographical, reverse-lexicographical and breadth-first search ordering of the vertices of τ. The path $\mathcal{W}^{\mathsf{lex}}(\tau)$ is commonly called the Łukasiewicz path of τ.

We mention several useful elementary properties of these codings, which are left as an exercise to the reader.

Lemma 15 *Let τ be a tree of size n.*

(i) *For every $0 \leqslant i \leqslant n - 1$, $\mathcal{W}_i^{\mathsf{lex}}(\tau)$ is equal to the number of children of vertices of $[\![\emptyset, u(i)[\![$ branching to the right of $[\![\emptyset, u(i)[\![$.*

(ii) *For every $0 \leqslant i \leqslant n - 1$, $\mathcal{W}_i^{\mathsf{rev}}(\tau)$ is equal to the number of children of vertices of $[\![\emptyset, u(i)[\![$ branching to the left of $[\![\emptyset, u(i)[\![$.*

(iii) *For every $1 \leqslant i \leqslant n$, setting $k = |u(i - 1)|$, $\mathcal{W}_i^{\mathsf{bfs}}(\tau) + 1$ is equal to the number of children of vertices at generation k less than or equal to $u(i - 1)$ plus the number of vertices at generation k greater than $u(i - 1)$ (for the breadth-first search order).*

In particular, for every $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$, $\mathcal{W}_k^*(\tau) \geqslant 0$ for every $0 \leqslant k \leqslant |\tau| - 1$. Also, recalling that $W(\tau)$ is the width of τ, an immediate consequence of (iii) is that

$$W(\tau) \leqslant \max \mathcal{W}^{\mathsf{bfs}}(\tau) + 1 \leqslant 2W(\tau). \tag{17}$$

Indeed, of $u(i)$ is the largest (for the breadth-first ordering) vertex at generation k of τ, then $\mathcal{W}_{i+1}^{\mathsf{bfs}}(\tau) + 1 = Z_{k+1}(\tau)$, and if $|u(i)| = k$, then $\mathcal{W}_{i+1}^{\mathsf{bfs}}(\tau) + 1 \leqslant Z_k(\tau) + Z_{k+1}(\tau)$.

It is straightforward to adapt these codings to forests. By definition, a forest F is a finite ordered collection of trees $(\tau_1, \ldots, \tau_k)$. One naturally extends the different orderings to F, by declaring that $u \prec v$ if $u \in \tau_i, v \in \tau_j$ and $1 \leqslant i < j \leqslant k$, and the coding paths $\mathcal{W}^*(F)$ are obtained by concatenating the jumps of $\mathcal{W}^*(\tau_1), \ldots, \mathcal{W}^*(\tau_k)$.

Recall the definition of the random walk $(W_n)_{n \geqslant 0}$, the notation $\zeta_j = \inf\{n \geqslant 0; W_n = -j\}$ for $j \geqslant 1$, and that $\mathbb{P}_{\mu,j}$ denotes the law of a forest of j independent BGW_μ trees. The following proposition explains the importance of the different codings.

Proposition 16 *Fix* $1 \leqslant j \leqslant n$. *For every* $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$, *under* $\mathbb{P}_{\mu,j}$, $\mathcal{W}^*$ *has the same distribution as* $(W_0, W_1, \ldots, W_{\zeta_j})$. *In particular, the total progeny of a* $\mathbb{P}_{\mu,j}$ *forest has the same law as* ζ_j.

We leave the proof to the reader (see [32, Proposition 1.5] for the case of the lexicographical ordering).

3.3 Bounds on the supremum of coding paths

Recall that $\mathfrak{t}_n$ is a BGW_μ tree conditioned on having n vertices, with μ being a critical offspring distribution belonging to the domain of attraction of a stable law of index $\alpha \in (1, 2]$.

Since for every $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$, $\mathcal{W}^*(\mathfrak{t}_n)$ has the same distribution as $(W_0, W_1, \ldots, W_n)$ under the conditional probability distribution $\mathbb{P}(\cdot \mid \zeta_1 = n)$, the following tail bound on the maximum of $\mathcal{W}^*(\mathfrak{t}_n)$ is an immediate consequence of Theorem 9.

Theorem 17 *For every* $\delta \in (0, \alpha/(\alpha - 1))$, *there exist* $C_1, C_2 > 0$ *such that for every* $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$, *for every* $u \geqslant 0$ *and every* $n \geqslant 1$:

$$\mathbb{P}\left(\max_{1 \leqslant i \leqslant n} \mathcal{W}_i^*(\mathfrak{t}_n) \geqslant u B_n\right) \leqslant C_1 \exp(-C_2 u^\delta).$$

In particular, Theorem 1 follows from this result, since $W(\mathfrak{t}_n) \leqslant \mathcal{W}^{\mathsf{bfs}}(\mathfrak{t}_n) + 1$ by (17). Also note that (1) entails that, for every $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$,

$$\left(\frac{\mathcal{W}_{\lfloor nt \rfloor}^*(\mathfrak{t}_n)}{B_n}; 0 \leqslant t \leqslant 1\right) \xrightarrow[n \to \infty]{(d)} X^{\mathrm{exc}}. \tag{18}$$

In addition, by (17), $\lim_{n \to \infty} \mathbb{P}(W(\mathfrak{t}_n) \geqslant u B_n) \geqslant \mathbb{P}(\sup X^{\mathrm{exc}} \geqslant 2u)$ for every $u \geqslant 0$, which, combined with the last part of Remark 14 shows that the exponent $\alpha/(\alpha - 1)$ is optimal in Theorem 1.

3.4 Bounds on the maximum outdegree of Bienaymé–Galton–Watson trees

We are now interested in the maximum outdegree $\Delta(\mathfrak{t}_n)$ of $\mathfrak{t}_n$, and establish Theorems 3 and 4. The first one is a corollary of the results we have just established. Indeed, it is clear that $\Delta(\mathfrak{t}_n) \leqslant \max_{1 \leqslant i \leqslant n} \mathcal{W}_i(\mathfrak{t}_n) + 1$, so Theorem 9 entails Theorem 3.

In the case $\alpha < 2$, since the maximum jump $\Delta^*(X^{\mathrm{exc}})$ of X^{exc} is almost surely positive, we in addition get:

Proposition 18 *Assume that* $\alpha \in (1, 2)$. *Then, for every* $p \geqslant 1$,

$$\mathbb{E}\left[\frac{\Delta(\mathfrak{t}_n)^p}{B_n^p}\right] \xrightarrow[n \to \infty]{} \mathbb{E}\left[\Delta^*(X^{\mathrm{exc}})^p\right].$$

Proof Observe that $\Delta(\mathfrak{t}_n) - 1$ is equal to the maximum jump of $\mathcal{W}(\mathfrak{t}_n)$. Since the largest jump of a càdlàg function is a continuous functional on $\mathbb{D}([0, 1], \mathbb{R})$, the convergence (18) implies

$$\frac{\Delta(\mathfrak{t}_n)}{B_n} \xrightarrow[n\to\infty]{(d)} \Delta^*(X^{\mathrm{exc}}). \tag{19}$$

The claim then follows from Theorem 3. $\qquad\square$

For $p = 1$, it is known [14, Proposition 3.10] that $\mathbb{E}\left[\Delta^*(X^{\mathrm{exc}})\right] = \Gamma\left(1 - \frac{1}{\alpha}\right)\beta$, where $\beta > 0$ is the unique solution to the equation

$$\sum_{n=0}^{\infty} \frac{(-1)^n \beta^n}{(n-\alpha)n!} = 0.$$

It would be interesting to calculate the value of $\mathbb{E}\left[\Delta^*(X^{\mathrm{exc}})^p\right]$ for $p > 1$ and also to obtain asymptotic expansions of $\mathbb{P}\left(\Delta^*(X^{\mathrm{exc}}) > u\right)$ and $\mathbb{P}\left(\Delta^*(X^{\mathrm{exc}}) < 1/u\right)$ as $u \to \infty$.

We now establish Theorem 4.

Proof of Theorem 4 In this proof, $C > 0$ will denote a constant which may change from line to line (but does not depend on n or k). Fix $1 \leqslant j \leqslant n$. For $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n$, set $\max(\mathbf{x}) = \max_{1 \leqslant i \leqslant n} x_i$ and note that $\max$ is invariant under cyclic shifts. Recall that $X_n = W_n - W_{n-1}$ for $n \geqslant 1$. Finally, to simplify notation, set $M_i^j = \max(X_i, X_{i+1}, \ldots, X_j)$ for every $1 \leqslant i \leqslant j \leqslant n$. If F is a forest, observe that $\Delta(F) - 1$ is equal to the maximum jump of the associated path $\mathcal{W}^*(F)$ with $* \in \{\mathsf{lex}, \mathsf{rev}, \mathsf{bfs}\}$. It follows from by (13) and Proposition 16 that

$$\mathbb{E}_{\mu,j}\left[G(\Delta(F)) \,\big|\, |F| = n\right] = \mathbb{E}\left[G\left(M_1^n + 1\right) \,\big|\, W_n = -j\right] \tag{20}$$

for every measurable function $G : \mathbb{R} \to \mathbb{R}_+$.

We start by proving (4). Note that $\mathbb{P}\left(M_1^n \leqslant k - 1 \mid W_n = -j\right) \leqslant \mathbb{P}\left(M_1^{\lfloor n/2 \rfloor} \leqslant k - 1 \mid W_n = -j\right)$. Recalling that $\varphi_n(j) = \mathbb{P}(W_n = -j)$, the Markov property for the random walk W applied at time $\lfloor n/2 \rfloor$ yields:

$$\mathbb{P}_{\mu,j}\left(\Delta(F) \leqslant k \mid |F| = n\right) \leqslant \mathbb{P}\left(M_1^{\lfloor n/2 \rfloor} \leqslant k - 1 \mid W_n = -j\right)$$
$$= \mathbb{E}\left[\mathbb{1}_{\{M_1^{\lfloor n/2 \rfloor} \leqslant k-1\}} \frac{\varphi_{n-\lfloor n/2 \rfloor}(W_{\lfloor n/2 \rfloor} + j)}{\varphi_n(j)}\right].$$

By the local limit theorem (8), there exists a constant $C > 0$ such that

$$\sup_{1 \leqslant j \leqslant MB_n} \frac{\varphi_{n-\lfloor n/2 \rfloor}(W_{\lfloor n/2 \rfloor} + j)}{\varphi_n(j)} \leqslant C. \tag{21}$$

Hence

$$\sup_{1 \leqslant j \leqslant MB_n} \mathbb{P}_{\mu,j}\left(\Delta(F) \leqslant k \mid |F| = n\right) \leqslant C\mathbb{P}\left(M_1^{\lfloor n/2 \rfloor} \leqslant k - 1\right)$$

$$= C\mathbb{P}\left(X_1 \leqslant k - 1\right)^{\lfloor n/2 \rfloor}$$
$$= C_1(1 - \mu([k+1, \infty)))^{\lfloor n/2 \rfloor}.$$

The inequality $(1 - x)^a \leqslant \exp(-ax)$ valid for every $a \geqslant 0$ and $x \in [0, 1]$ yields (4).

To check (5), write:

$$\mathbb{P}\left(M_1^n \geqslant k - 1 \mid W_n = -j\right) \leqslant \mathbb{P}\left(M_1^{\lfloor n/2 \rfloor} \geqslant k - 1 \mid W_n = -j\right)$$
$$+ \mathbb{P}\left(M_{\lfloor n/2 \rfloor + 1}^n \geqslant k - 1 \mid W_n = -j\right)$$
$$\leqslant 2\mathbb{P}\left(M_1^{\lfloor n/2 \rfloor + 1} \geqslant k - 1 \mid W_n = -j\right)$$

As before, using the Markov property for the random walk W applied at time $\lfloor n/2 \rfloor + 1$ combined with (21), we get that

$$\sup_{1 \leqslant j \leqslant MB_n} \mathbb{P}_{\mu, j}\left(\Delta(F) \geqslant k \mid |F| = n\right) \leqslant C\mathbb{P}\left(M_1^{\lfloor n/2 \rfloor + 1} \geqslant k - 1\right)$$
$$= C\left(1 - (1 - \mu([k, \infty)))^{\lfloor n/2 \rfloor + 1}\right).$$

The inequality $1 - (1 - x)^a \leqslant ax$ valid for every $x \in [0, 1]$ and $a \geqslant 1$ shows (5). This completes the proof. $\square$

Corollary 19 *Let $(p_n)_{n \geqslant 1}$ and $(q_n)_{n \geqslant 1}$ and be two sequence of real numbers such that $n\mu([p_n, \infty)) \to \infty$ and $n\mu([q_n, \infty)) \to 0$ as $n \to \infty$. Then:*

$$\mathbb{P}\left(p_n \leqslant \Delta(\mathfrak{t}_n) \leqslant q_n\right) \xrightarrow[n \to \infty]{} 1$$

This partially answers [25, Problem19.30]. Note also that (19) is stronger than Corollary 19. Indeed, using the Potter bounds, it is possible to verify that if $(p_n)_{n \geqslant 1}$ and $(q_n)_{n \geqslant 1}$ are two sequences of positive real numbers such that $n\mu([p_n, \infty)) \to \infty$ and $n\mu([q_n, \infty)) \to 0$ as $n \to \infty$, then $B_n/p_n \to \infty$ and $B_n/q_n \to 0$ as $n \to \infty$. By (19), this indeed implies that $\mathbb{P}\left(p_n \leqslant \Delta(\mathfrak{t}_n) \leqslant q_n\right) \to 1$.

3.5 Application to non-crossing trees

Let P_n be the convex polygon inscribed in the unit disk of the complex plane whose vertices are the nth roots of unity. By definition, a *non-crossing tree* of P_n is a tree drawn on the plane whose vertices are the of P_n and whose edges are non-crossing line segments. See Fig. 1 for an example.

Different combinatorial properties of non-crossing trees have been studied in the literature. Let $\mathcal{T}_n$ be uniformly distributed over the set of all non-crossing trees of P_n. Deutsch and Noy [16] have established that, for every $c > 0$, $\mathbb{P}\left(|\Delta(\mathcal{T}_n) - \log_3 n| \leqslant (1 + c) \log_3 \log_3 n\right)$ tends to 1 as $n \to \infty$. Using Theorem 3, we improve these bounds:
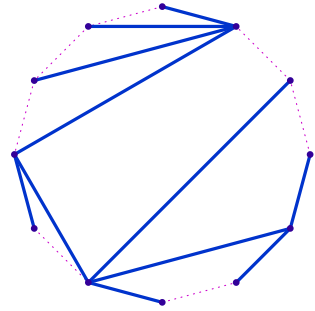
Theorem 20 *For every $c > 0$, there exists constants $C_1, C_2 > 0$ such that for every $n \geqslant 1$,*

$$\mathbb{P}\left(\Delta(\mathcal{T}_n) \geqslant \log_3 n + \log_3 \log_3 n + c \log_3 \log_3 n\right) \leqslant \frac{C_1}{(\log_3 n)^c}$$

and

$$\mathbb{P}\left(\Delta(\mathcal{T}_n) \leqslant \log_3 n + \log_3 \log_3 n - c \log_3 \log_3 n\right) \leqslant C_1 e^{-C_2 (\log_3 n)^c}.$$

In particular, for every $c > 0$, we have

$$\mathbb{P}\left(|\Delta(\mathcal{T}_n) - \log_3 n - \log_3 \log_3 n| \leqslant c \log_3 \log_3 n\right) \underset{n \to \infty}{\longrightarrow} 1.$$

Proof In this proof, $C > 0$ will denote a constant that may change from expression to expression (and which may only depend on c). By [36], $\mathcal{T}_n$, viewed as a plane tree rooted at vertex 1, is a modified Bienaymé–Galton–Watson tree $\widetilde{\mathcal{T}}$, conditioned on having n vertices, where the root has offspring distribution $\lambda(k) = 2/3^k$ for $k \geqslant 1$, and all the other vertices have offspring distribution $\mu(k) = 4(k+1)/3^{k+2}$ for $k \geqslant 0$. Note that $\mu([k, \infty)) = (3 + 2k)/3^{1+k}$ with the notation of Theorem 4. Also, by [36], $\mathbb{P}\left(|\widetilde{\mathcal{T}}| = n\right) \sim \sqrt{3/(4\pi)} \cdot n^{-3/2}$ as $n \to \infty$. Hence, for every $k \geqslant 1$,

$$\mathbb{P}\left(k_\emptyset(\widetilde{\mathcal{T}}) \geqslant k \mid |\widetilde{\mathcal{T}}| = n\right) = \frac{1}{\mathbb{P}\left(|\widetilde{\mathcal{T}}| = n\right)} \sum_{j \geqslant k} \lambda(j) \mathbb{P}_{\mu,j}\left(|F| = n - 1\right)$$

$$= \sum_{j \geqslant k} \lambda(j) \cdot \frac{j}{n-1} \frac{\mathbb{P}\left(W_{n-1} = -j\right)}{\mathbb{P}\left(|\widetilde{\mathcal{T}}| = n\right)}$$

$$\leqslant C \sum_{j=k}^{\infty} j\lambda(j) \leqslant C \frac{k}{3^k}. \tag{22}$$

For the second equality, we have used the well-known fact that $\mathbb{P}_{\mu,j}\left(|F| = n - 1\right) = j/(n-1) \cdot \mathbb{P}\left(W_{n-1} = -j\right)$, which is for instance a consequence of Proposition 16 and (14), and the first inequality uses the bound $\mathbb{P}\left(W_{n-1} = -j\right) \leqslant C/\sqrt{n}$

which is a consequence of the local limit theorem. Also, conditionally on the event $\{k_\emptyset(\widetilde{\mathcal{T}}) = j, |\widetilde{\mathcal{T}}| = n\}$ the forest of trees grafted on the root of $\widetilde{\mathcal{T}}$ is a $\mathbb{P}_{\mu,j}$ forest conditioned on having $n - 1$ vertices. Therefore, for every nonnegative measurable function f, we have

$$\mathbb{E}\left[f(\Delta(\widetilde{\mathcal{T}})), k_\emptyset(\widetilde{\mathcal{T}}) \leqslant k \mid |\widetilde{\mathcal{T}}| = n\right] \leqslant \sup_{1 \leqslant j \leqslant k} \mathbb{P}_{\mu,j}\left(f(\Delta(F)) \mid |F| = n - 1\right).$$

To prove the first inequality in Theorem 20, set $b_n = \log_3 n + \log_3 \log_3 n + c \log_3 \log_3 n$ to simplify notation, and take $k = \sqrt{n}$. Then, on one hand, by Theorem 4 (which we may apply because $B_n/\sqrt{n}$ converges to a finite positive value since μ has finite variance),

$$\sup_{1 \leqslant j \leqslant \sqrt{n}} \mathbb{P}_{\mu,j}\left(\Delta(F) \geqslant b_n \mid |F| = n - 1\right) \leqslant Cn\mu([a_n, \infty)) \leqslant \frac{C}{(\log_3 n)^c},$$

and, on the other hand, $\mathbb{P}\left(k_\emptyset(\widetilde{\mathcal{T}}) \geqslant \sqrt{n} \mid |\widetilde{\mathcal{T}}| = n\right) = o\left((\log_3 n)^{-c}\right)$ by (14).

Similarly, for the second inequality, set $a_n = \log_3 n + \log_3 \log_3 n - c \log_3 \log_3 n$ and also take $k = \sqrt{n}$. Then, on one hand, by Theorem 4

$$\sup_{1 \leqslant j \leqslant \sqrt{n}} \mathbb{P}_{\mu,j}\left(\Delta(F) \leqslant a_n \mid |F| = n - 1\right) \leqslant C_1 e^{-C_2 n\mu([a_n+1, \infty))} \leqslant C_1 e^{-C_2(\log_3 n)^c},$$

and, on the other hand, $\mathbb{P}\left(k_\emptyset(\widetilde{\mathcal{T}}) \geqslant \sqrt{n} \mid |\widetilde{\mathcal{T}}| = n\right) = o\left(e^{-C_2(\log_3 n)^c}\right)$ by (14). This completes the proof. □

See also [13,31] for a study of geometric properties of random non-crossing trees.

4 Sub-exponential bounds on the height of BGW trees

Our goal is now to establish Theorem 2. Recall that $\mathfrak{t}_n$ denotes a Bienaymé–Galton–Watson tree with a critical offspring distribution μ in the domain of attraction of a stable law of index $\alpha \in (1, 2]$, conditioned on having n vertices.

Fix $\delta \in (1, \alpha)$. Since $\mathbb{P}(H(\mathfrak{t}_n) \geqslant u \cdot n/B_n) = 0$ for $u > B_n$, we shall implicitly always restrict our attention to values of u such that $2 \leqslant u \leqslant B_n$. In addition, by monotonicity of $x \mapsto \mathbb{P}_\mu(H \geqslant x)$, we may and will always assume without loss of generality that un/B_n and $un/(2B_n)$ are both integers.

In addition to Tables 1 and 2, it may be useful to refer to Table 3 to keep track of the main notation used and introduced in this section.

4.1 Preliminary observations

We first introduce some notation. If τ is a tree and $z \in \tau$, as in [2], a key quantity will be $M(z)$, the number of vertices branching off $[\![\emptyset, z[\![$, that is the number of children of vertices of $[\![\emptyset, z[\![$ that do not belong to $[\![\emptyset, z[\![$ (the dependence in τ is implicit).

Table 3 Table of the main notation and symbols appearing in Sect. 4

$H(\tau)$	Height of a tree τ
$M(z)$	Number of vertices branching off $[\![\emptyset, z[\![$ for $z \in \tau$
$\mathrm{Cut}_z(\tau)$	Tree obtained by keeping only z and its descendance in τ
$Z_k = Z_k(\tau)$	Number of vertices of height k in τ
$\mathsf{U}_k = \mathsf{U}_k(\tau)$	Vertex chosen uniformly at random at height k in τ when $Z_k(\tau) > 0$
$\mathfrak{t}_n$	BGW_μ tree conditioned on having n vertices
$\mathcal{T}^*$	BGW_μ tree conditioned to survive
U_n^*	Vertex of the spine of $\mathcal{T}^*$ at generation n
Z_k^*	Number of vertices of height k in $\mathcal{T}^*$
M_k^*	Number of children branching off the spine of $\mathcal{T}^*$ up to generation $k - 1$
$(B_n')_{n \geqslant 1}$	An increasing scaling sequence such that $B'_{B_n} \sim B_{B'_n} \sim n$ as $n \to \infty$

Observe that by Lemma 15, for every $z \in \tau$, $M(z) \leqslant \max \mathcal{W}^{\mathrm{lex}}(\tau) + \max \mathcal{W}^{\mathrm{rev}}(\tau)$. In addition, by Theorem 17,

$$
\mathbb{P}\left(\max \mathcal{W}^{\mathrm{lex}}(\mathfrak{t}_n) + \max \mathcal{W}^{\mathrm{rev}}(\mathfrak{t}_n) \geqslant u^{\delta-1} B_n\right) \leqslant 2\mathbb{P}\left(\max \mathcal{W}^{\mathrm{lex}}(\mathfrak{t}_n) \geqslant u^{\delta-1} B_n/2\right)
$$
$$
\leqslant C_1 \exp(-C_2 u^\delta).
$$

To establish Theorem 2, it is therefore enough to show the existence of $C_1, C_2 > 0$ such that

$$
\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z \in \mathfrak{t}_n} M(z) \leqslant u^{\delta-1} B_n\right) \leqslant C_1 \exp(-C_2 u^\delta)
$$

for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$.

4.2 Size-biasing

A key ingredient will be the notion of *size-biasing*. We denote by $\mathcal{T}^*$ a random variable having the law of the BGW_μ tree conditioned to survive, which we will also call the size-biased tree. This tree was introduced by Kesten [27], and may informally be described as follows. Let X be a random variable with distribution μ, and let X^* be a random variable with the size-biased distribution $\mathbb{P}(X^* = i) = i\mu(i)$ for $i \geqslant 1$. In the tree $\mathcal{T}^*$, vertices are either normal, or mutant, and the root is a mutant vertex. Normal nodes have outdegree distributed according to independent copies of X, while mutant nodes have outdegree distributed according to independent copies of X^*. All children of a normal vertex are normal, while for each mutant node, all of its children are normal, except one, selected uniformly at random, which is mutant. In particular,

the tree $\mathcal{T}^*$ has a unique infinite simple path starting from the root, called the spine. Vertices on the spine have outdegree distributed according to X^*, while all the other vertices have outdegree distributed according to μ. For every $n \geqslant 0$, we let U_n^* be the vertex of the spine of $\mathcal{T}^*$ at generation n.

We now introduce some notation. Let τ be a tree and $z \in \tau$. Recall that $\theta_z \tau = \{w \in U; zw \in \tau\}$. Set $\mathrm{Cut}_z(\tau) = \tau \setminus \{zw; w \in \theta_z \tau \setminus \{\emptyset\}\}$, which is a tree such that $z \in \mathrm{Cut}_z(\tau)$. If $n \geqslant 0$ is such that $Z_n(\tau) > 0$, we let U_n be a vertex chosen uniformly at random among all those at generation n of τ. The term *size-biasing* comes from the following result, see [35]: for any nonnegative functions G_1, G_2:

$$\mathbb{E}_\mu \left[\mathbb{1}_{Z_n > 0} F(Z_n) G_1(\mathrm{Cut}_{\mathsf{U}_n}(\tau), \mathsf{U}_n) G_2(\theta_{\mathsf{U}_n} \tau) \right]$$
$$= \mathbb{E} \left[\frac{F(Z_n(\mathcal{T}^*))}{Z_n(\mathcal{T}^*)} G_1(\mathrm{Cut}_{\mathsf{U}_n^*}(\mathcal{T}^*), \mathsf{U}_n^*) \right] \cdot \mathbb{E}_\mu [G_2(T)]. \qquad (23)$$

This identity is usually written and used with F replaced by $x \mapsto x F(x)$, but we have written it in this form in view of future use.

4.3 Two technical estimates

We first establish an estimate describing the asymptotic behavior of B_n.

Lemma 21 *We have* $B_n \cdot \mathbb{P}_\mu \left(Z_{n/B_n} > 0 \right) \to (\alpha - 1)^{-1/(\alpha-1)}$ *as* $n \to \infty$.

Proof By (41) and [44, Lemma 2],

$$\mathbb{P}_\mu (Z_n > 0)^{\alpha-1} L \left(\mathbb{P}_\mu (Z_n > 0)^{-1} \right) \underset{n \to \infty}{\sim} \frac{\alpha}{\Gamma(3 - \alpha)n}. \qquad (24)$$

To simplify notation, in this proof we set $R_n = \mathbb{P}_\mu \left(Z_{n/B_n} > 0 \right)^{-1}$, so that $R_n \to \infty$, and we aim at showing that $B_n / R_n \to (\alpha - 1)^{-1/(\alpha-1)}$ as $n \to \infty$. By replacing n with n/B_n in (24) and using (7), we get that

$$\left(\frac{B_n}{R_n} \right)^{\alpha-1} \underset{n \to \infty}{\sim} (\alpha - 1)^{-1} \cdot \frac{L(B_n)}{L(R_n)}. \qquad (25)$$

Since R_n and B_n tend to ∞ as $n \to \infty$, this implies that $(B_n/R_n)^{\alpha-1} \to (\alpha-1)^{-1}$. Indeed, if $(B_n/R_n)^{\alpha-1}$ converges to a positive limit along a subsequence, then $L(B_n)/L(R_n) \to 1$ along this subsequence by properties of slowly varying functions. If $B_n/R_n \to \infty$ along a subsequence, then it is a simple matter to see that a contradiction arises from (25) by the Potter bounds. One similarly treats the case $B_n/R_n \to 0$ along a subsequence, and this completes the proof. $\qquad \square$

Let $(S_i)_{i \geqslant 1}$ be a sequence of i.i.d. random variables having the distribution of the total size of a BGW_μ tree. Recall that d_α is the density of a random variable Y_α with

Laplace exponent given by $\mathbb{E}\left[\exp(-\lambda Y_\alpha)\right] = \exp(\lambda^\alpha)$ for every $\lambda > 0$. It is a standard fact that

$$\mathbb{P}_\mu\left(|\tau| = n\right) \underset{n \to \infty}{\sim} \frac{d_\alpha(0)}{n B_n}. \tag{26}$$

This can for instance be seen by combining the equality $\mathbb{P}_\mu\left(|\tau| = n\right) = \mathbb{P}\left(W_n = -1\right)/n$ with the local limit theorem. In particular, $\mathbb{P}_\mu\left(|\tau| \geqslant n\right) \sim \alpha d_\alpha(0)/B_n$ as $n \to \infty$. Hence under $\mathbb{P}_\mu$, $|\tau|$ belongs to the domain of attraction of a stable law of index $1/\alpha$. Therefore, setting

$$B'_n = \inf\left\{x \geqslant 0; \ \mathbb{P}_\mu\left(|\tau| \geqslant x\right) \leqslant \frac{\alpha d_\alpha(0)}{n}\right\}, \tag{27}$$

the proof of Theorem 2.6.1 in [22] (in particular, our B'_n plays the role of B_n that is defined in [22, Eq. (2.6.6)]) shows that $(S_1 + \cdots + S_n)/B'_n$ converges in distribution to a stable law of index $1/\alpha$ as $n \to \infty$ (since $1/\alpha < 1$, by [20, Theorem 3 (i) in Chapter XVII.5] no centering procedure is required). Hence, if we set $\Psi_n(k) = \mathbb{P}\left(S_1 + \cdots + S_n = k\right)$, by the local limit theorem [22, Theorem 4.2.1],

$$\lim_{n \to \infty} \sup_{k \in \mathbb{Z}} \left| B'_n \Psi_n(k) - g\left(\frac{k}{B'_n}\right) \right| = 0,$$

where g is the density of a stable random variable with index $1/\alpha$. By [48, Eq. (I.20)]), g is bounded. As a consequence, there exists a constant $C > 0$ such that

$$\text{for every } n \geqslant 1, \quad \text{for every } k \geqslant 1, \qquad \Psi_n(k) \leqslant \frac{C}{B'_n}. \tag{28}$$

We will also later use the following asymptotic estimate involving both B_n and B'_n:

$$B'_{B_n} \underset{n \to \infty}{\sim} n. \tag{29}$$

To see this, note that, by (27), we have $\mathbb{P}_\mu\left(|\tau| \geqslant B'_n\right) \sim \alpha d_\alpha(0)/n$ as $n \to \infty$. Since $\mathbb{P}_\mu\left(|\tau| \geqslant m\right) \sim \alpha d_\alpha(0)/B_m$ as $m \to \infty$, we also get that $\mathbb{P}_\mu\left(|\tau| \geqslant B'_n\right) \sim \alpha d_\alpha(0)/B_{B'_n}$ as $n \to \infty$. This implies that $B_{B'_n} \sim n$ as $n \to \infty$. The estimate (29) then immediately follows from [43, p. 21].

4.4 Proof of Theorem 2

We start by explaining the main steps for proving Theorem 2. Recall that we assume that un/B_n and $un/(2B_n)$ are both integers. First, let $Z_k^* = Z_k(\mathcal{T}^*)$ be the size of the k-th generation of the size biased tree, let also $M_k^* = M(\mathsf{U}_k^*)$ be the number of children branching off the spine of $\mathcal{T}^*$ up to generation $k - 1$. By using a cutting and size-biasing argument, we shall prove the following bound.

Lemma 22 *For every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$, we have*

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z \in \mathfrak{t}_n} M(z) \leqslant u^{\delta-1} B_n \right)$$

$$\leqslant \frac{1}{\mathbb{P}_\mu \left(|\tau| = n \right)} \cdot \mathbb{E}\left[\frac{1}{Z^*_{un/(2B_n)}} \right] \cdot \mathbb{E}\left[\frac{1}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{M^*_{un/(2B_n)} \leqslant u^{\delta-1} B_n} \right]. \tag{30}$$

The most technical part of the proof consists in bounding the second expectation appearing in (30):

Lemma 23 *There exist $C_1, C_2 > 0$ such that*

$$n \cdot \mathbb{E}\left[\frac{1}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{M^*_{un/(2B_n)} \leqslant u^{\delta-1} B_n} \right] \leqslant C_1 \exp(-C_2 u^\delta).$$

for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$.

We now explain how Theorem 2 follows from these two results.

Proof of Theorem 2 Recall that it is enough to show the existence of $C_1, C_2 > 0$ such that

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z \in \mathfrak{t}_n} M(z) \leqslant u^{\delta-1} B_n \right) \leqslant C_1 \exp(-C_2 u^\delta). \tag{31}$$

for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$. By (23), $\mathbb{E}\left[1/Z^*_{un/(2B_n)} \right] = \mathbb{P}_\mu \left(Z_{un/(2B_n)} > 0 \right) \leqslant \mathbb{P}_\mu \left(Z_{n/B_n} > 0 \right)$ since $u \geqslant 2$. Hence, by Lemmas 22 and 23,

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z \in \mathfrak{t}_n} M(z) \leqslant u^{\delta-1} B_n \right)$$

$$\leqslant \frac{\mathbb{P}_\mu \left(Z_{n/B_n} > 0 \right)}{n \cdot \mathbb{P}_\mu \left(|\tau| = n \right)} \cdot C_1 \exp(-C_2 u^\delta).$$

But by (26) and Lemma 21, the quantity $\mathbb{P}_\mu \left(Z_{n/B_n} > 0 \right) / (n \cdot \mathbb{P}_\mu \left(|\tau| = n \right))$ converges to a positive limit as $n \to \infty$. The conclusion follows. $\square$

We now prove Lemma 22.

Proof of Lemma 22 By using the size-biasing relation (23), write

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z \in \mathfrak{t}_n} M(z) \leqslant u^{\delta-1} B_n \right)$$

$$\leqslant \mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad M(\mathsf{U}_{un/B_n}) \leqslant u^{\delta-1}B_n\right)$$

$$= \frac{1}{\mathbb{P}_\mu\left(|\tau|=n\right)}\mathbb{E}\left[\frac{1}{Z^*_{un/B_n}}\mathbb{1}_{M^*_{un/B_n}\leqslant u^{\delta-1}B_n}\mathbb{1}_{|\mathrm{Cut}_{\mathsf{U}^*_{un/B_n}}(\mathcal{T}^*)|+|\mathcal{T}|=n+1}\right],$$

where $\mathcal{T}$ is an independent BGW_μ tree. Observe that the random variables Z^*_{un/B_n} and M^*_{un/B_n} are of course not independent, which is a major issue. Also, forgetting the other terms, note that one should find a good number of independent trees in order to bound $\mathbb{P}\left(|\mathrm{Cut}_{\mathsf{U}^*_{un/B_n}}(\mathcal{T}^*)|+|\mathcal{T}|=n+1\right)$ by using (28).

The main idea is to introduce independence: roughly speaking, we cut the spine up to generation un/B_n in half and, denoting by respectively $\mathcal{S}_\downarrow$ and $\mathcal{S}^\uparrow$ its lower and upper part, we bound from below Z^*_{un/B_n} by the number of vertices at generation un/B_n in $\mathcal{T}^*$ that have an ancestor belonging to $\mathcal{S}^\uparrow$, we bound from below M^*_{un/B_n} by the number of children branching off $\mathcal{S}_\downarrow$, and the collection of independent trees we use for applying the local limit theorem are those branching off $\mathcal{S}_\downarrow$.

Specifically, define $\mathcal{T}^*_\downarrow = \mathrm{Cut}_{\mathsf{U}^*_{un/(2B_n)}}(\mathcal{T}^*)$ and $\mathcal{T}^*_\uparrow = \theta_{\mathsf{U}^*_{un/(2B_n)}}\mathrm{Cut}_{\mathsf{U}^*_{un/B_n}}(\mathcal{T}^*)$ (to simplify notation, we keep the dependence in u and n implicit). Note that $\mathcal{T}^*_\downarrow$ and $\mathcal{T}^*_\uparrow$ are independent and have same distribution, and also that $M^*_{un/(2B_n)}$ is a measurable function of $\mathcal{T}^*_\downarrow$. Finally, note that

$$\left|\mathrm{Cut}_{\mathsf{U}^*_{un/B_n}}(\mathcal{T}^*)\right| = |\mathcal{T}^*_\downarrow| + |\mathcal{T}^*_\uparrow| - 1$$

and that $|\mathcal{T}^*_\downarrow|$ is equal to $1 + un/(2B_n)$ plus the total size of a forest of $M^*_{un/(2B_n)}$ independent BGW_μ trees. Observing that $Z^*_{un/B_n} \geqslant Z_{un/(2B_n)}(\mathcal{T}^*_\uparrow)$ and $M^*_{un/B_n} \geqslant M^*_{un/(2B_n)}$, we therefore have

$$\mathbb{E}\left[\frac{1}{Z^*_{un/B_n}}\mathbb{1}_{M^*_{un/B_n}\leqslant u^{\delta-1}B_n}\mathbb{1}_{|\mathrm{Cut}_{\mathsf{U}^*_{un/B_n}}(\mathcal{T}^*)|+|\mathcal{T}|=n+1}\right]$$

$$\leqslant \mathbb{E}\left[\frac{1}{Z_{un/(2B_n)}(\mathcal{T}^*_\uparrow)}\mathbb{1}_{M^*_{un/(2B_n)}\leqslant u^{\delta-1}B_n}\Psi_{M^*_{un/(2B_n)}}(n+1-|\mathcal{T}^*_\uparrow|-|\mathcal{T}|-un/(2B_n))\right],$$

where we recall that $\Psi_n(k)$ be the probability that a forest of n independent BGW_μ trees has total size k. The desired result then follows by using (28) and noting that $Z_{un/(2B_n)}(\mathcal{T}^*_\uparrow)$ is independent of $M^*_{un/(2B_n)}$ and has the same distribution as $Z^*_{un/(2B_n)}$. $\qquad\square$

At this point, we make several comments concerning Lemma 22. First, in its proof, we chose to use (28) with the forest of trees branching off the lower half of the spine. It would perhaps been more natural to choose the forest of trees above generation $un/(2B_n)$ in the tree $\mathcal{T}^*_\uparrow$, and in this way get that for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$,

$$\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant u \cdot \frac{n}{B_n} \quad \text{and} \quad \max_{z\in\mathfrak{t}_n} M(z) \leqslant u^{\delta-1}B_n\right)$$

$$\leqslant \frac{1}{\mathbb{P}_\mu\left(|\tau| = n\right)} \cdot \mathbb{E}\left[\frac{1}{Z^*_{un/(2B_n)} B'_{Z^*_{un/(2B_n)}}}\right] \cdot \mathbb{P}\left(M^*_{un/(2B_n)} \leqslant u^{\delta-1} B_n\right).$$

Unfortunately, this does not allow to conclude, since

$$\mathbb{E}\left[\frac{n B_n}{Z^*_{un/(2B_n)} B'_{Z^*_{un/(2B_n)}}}\right] \xrightarrow[n\to\infty]{} \infty.$$

The reason is that $Z^*_{n/B_n}/B_n$ converges in distribution to a positive random variable whose density has polynomial decay near 0 (see the discussion after Proposition 6), while $M^*_{n/B_n}/B_n$ converges in distribution to a positive random variable whose density has exponential decay near 0.

Also, it would have been slightly simpler to bound from below $M^*_{un/(2B_n)} + 1$ by the maximal outdegree of a vertex of the lower half of the spine. One may indeed follow this path when $\alpha \in (1, 2)$. However, in the case $\alpha = 2$, these two quantities are not of the same order.

It thus remains to establish Lemma 23. The main technical estimate is the following.

Lemma 24 *Fix $c_0, C_0 > 0$ and $\alpha_1 > \alpha$. There exist $\gamma > 0$ and constants $C_1, C_2 > 0$ such that*

$$\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant C_0 \frac{u^{\delta-1} B_n}{t^{1/\alpha_1}}\right) \leqslant C_1 e^{-C_2 u^\delta t^\gamma}$$

for every $n \geqslant 1$, $2 \leqslant u \leqslant B_n$ such that $c_0 un/B_n \leqslant u^{\delta-1} B_n$ and $1 \leqslant t \leqslant B'_{u^{\delta-1} B_n}/B'_{c_0 un/B_n}$.

We postpone its proof and explain how Lemma 23 follows from Lemma 24.

Proof of Lemma 23 We start with a preliminary result. We claim that there exists $c_0, C_1, C_2 > 0$ such that

$$n\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant c_0 un/B_n\right) \leqslant C_1 e^{-C_2 u^\delta}$$

for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$. To show this, we will use the following simple fact, which follows from Markov's exponential inequality. Let S_n be a sum of n i.i.d Bernoulli random variables of parameter $p \in (0, 1)$. Then, for every $p_0 \in (0, p/2)$, there exists a constant $C > 0$ such that $\mathbb{P}\left(S_n \leqslant 2p_0 n\right) \leqslant \exp(-Cn)$ for every $n \geqslant 1$.

Now write $M^*_{un/(2B_n)} = \sum_{i=1}^{un/(2B_n)}(X^*_i - 1)$, where $(X^*_i)_{i\geqslant 1}$ are i.i.d. random variables having the same distribution as the size-biased random variable X^*, so that $S^*_n = \#\{1 \leqslant i \leqslant n; X^*_i > 1\}$, which is a sum of n i.i.d. Bernoulli random variables of parameter $1 - \mu(1)$, and note that $M^*_{un/(2B_n)} \geqslant S^*_{un/(2B_n)}$. Therefore, choosing

$c_0 \in (0, (1 - \mu(1))/2)$, by the previous discussion we get

$$\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant c_0 un/B_n\right) \leqslant \mathbb{P}\left(S^*_{un/(2B_n)} \leqslant c_0 un/B_n\right) \leqslant \exp(-Cun/B_n)$$

Then observe that

$$n \exp(-Cun/B_n) \leqslant C_1 \exp(-u^\delta),$$

for a certain $C_1 > 0$, for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$. Indeed, set $f_n(u) = Cun/B_n - u^\delta - \ln(n)$. Since $\delta \in (1, \alpha)$, f_n is concave so that $\inf_{[2, B_n]} f_n \geqslant \min(f_n(2), f_n(B_n))$. An application of the Potter bounds show that $f_n(2)$ and $f_n(B_n)$ both tend to infinity as $n \to \infty$. This completes the proof of the claim.

As a consequence, Lemma 23 will follow if we manage to show the existence of $C_1, C_2 > 0$ such that

$$\mathbb{E}\left[\frac{n}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{c_0 un/B_n \leqslant M^*_{un/(2B_n)} \leqslant u^{\delta-1}B_n}\right] \leqslant C_1 \exp(-C_2 u^\delta).$$

for every $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$ such that $c_0 un/B_n \leqslant u^{\delta-1}B_n$. To this end, using the formula

$$\mathbb{E}\left[\frac{1}{X}\mathbb{1}_{x \leqslant X \leqslant y}\right] = \frac{1}{y}\int_1^{y/x} dt\, \mathbb{P}(x \leqslant X \leqslant y/t) + \frac{1}{y} \cdot \mathbb{P}(x \leqslant X \leqslant y)$$

valid for every $0 \leqslant x \leqslant y$ and every nonnegative real valued random variable X, write

$$\mathbb{E}\left[\frac{n}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{c_0 un/B_n \leqslant M^*_{un/(2B_n)} \leqslant u^{\delta-1}B_n}\right]$$

$$\leqslant \frac{n}{B'_{u^{\delta-1}B_n}} \cdot \int_1^{\frac{B'_{u^{\delta-1}B_n}}{B'_{c_0 un/B_n}}} dt\, \mathbb{P}\left(B'_{M^*_{un/(2B_n)}} \leqslant \frac{B'_{u^{\delta-1}B_n}}{t}\right)$$

$$+ \frac{n}{B'_{u^{\delta-1}B_n}}\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant u^{\delta-1}B_n\right)$$

If $\alpha_1 > \alpha$ is fixed, by the Potter bounds there exists a constant $C > 0$ such that $t^{-1} \cdot B'_{u^{\delta-1}B_n} \leqslant B'_{Cu^{\delta-1}B_n \cdot t^{-1/\alpha_1}}$ for every $t \geqslant 1$, $n \geqslant 1$ and $2 \leqslant u \leqslant B_n$ such that $c_0 un/B_n \leqslant u^{\delta-1}B_n$. Therefore, using the fact that B' is increasing, we get that

$$\mathbb{P}\left(B'_{M^*_{un/(2B_n)}} \leqslant t^{-1} \cdot B'_{u^{\delta-1}B_n}\right) \leqslant \mathbb{P}\left(M^*_{un/(2B_n)} \leqslant C\frac{u^{\delta-1}B_n}{t^{1/\alpha_1}}\right).$$

In addition, by (29) and the Potter bounds, there exist $C, \eta > 0$ such that $n \cdot (B'_{u^{\delta-1}B_n})^{-1} \leqslant Cu^\eta$ for every $n \geqslant 1$ and $u \geqslant 2$. Therefore, by Lemma 24,

$$\mathbb{E}\left[\frac{n}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{c_0 un/B_n \leqslant M^*_{un/(2B_n)} \leqslant u^{\delta-1}B_n}\right]$$
$$\leqslant Cu^\eta \left(\int_1^\infty dt\, e^{-C_2 u^\delta t^\gamma} + e^{-C_2 u^\delta}\right).$$

Finally note that, for every $A > 0$,

$$\int_1^\infty dt\, e^{-At^\gamma} \leqslant \int_1^\infty dt\, \frac{A\gamma t^{\gamma-1}}{\gamma A} e^{-At^\gamma} = \frac{1}{\gamma A}\left[-e^{-At^\gamma}\right]_1^\infty = \frac{e^{-A}}{\gamma A}.$$

Therefore

$$\mathbb{E}\left[\frac{n}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{c_0 un/B_n \leqslant M^*_{un/(2B_n)} \leqslant u^{\delta-1}B_n}\right] \leqslant Cu^{\eta-\delta} e^{-C_2 u^\delta},$$

and the proof is complete. $\square$

It remains to establish Lemma 24.

Proof of Lemma 24 In this proof, C denotes a positive constant that may change from line to line (and that does not depend on u and n). We mention that the condition $c_0 un/B_n \leqslant u^{\delta-1}B_n$ restricts u to take smaller values than possible and will play a crucial role.

First, fix $\alpha_+ \in (\alpha, \alpha_1)$ so that by the Potter bounds, if $a_n \leqslant b_n$ are positive sequences tending to infinity, $B'_{b_n}/B'_{a_n} \leqslant C(b_n/a_n)^{\alpha_+}$ for every $n \geqslant 1$. Then let $\gamma > 0$ be such that

$$0 < \gamma < \min\left(\frac{1}{\alpha_+(2-\delta)}, \frac{1}{\alpha_1(2-\alpha)}\right).$$

Since we may write $M^*_{un/(2B_n)} = \sum_{i=1}^{un/(2B_n)}(X_i^* - 1)$, where $(X_i^*)_{i \geqslant 1}$ are i.i.d. random variables having the same distribution as the size-biased random variable X^*, by Markov's exponential inequality, for $\lambda > 0$,

$$\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant C_0 \frac{u^{\delta-1}B_n}{t^{1/\alpha_1}}\right) \leqslant \exp\left(\lambda C_0 \frac{u^{\delta-1}B_n}{t^{1/\alpha_1}} + \frac{un}{2B_n} \ln \mathbb{E}\left[e^{-\lambda(X^*-1)}\right]\right).$$

We take $\lambda = ut^\gamma/B_n$ (the dependence in n and u of λ is implicit).

We check that $\lambda \to 0$. First, since $c_0 u n / B_n \leqslant u^{\delta-1} B_n$ and $t \leqslant B'_{u^{\delta-1} B_n} \cdot (B'_{c_0 u n / B_n})^{-1}$, we have

$$2 \leqslant u \leqslant C \frac{B_n^{2/(2-\delta)}}{n^{1/(2-\delta)}}, \qquad 1 \leqslant t \leqslant \frac{B_n^{2\alpha_+}}{u^{\alpha_+(2-\delta)} n^{\alpha_+}}. \tag{32}$$

By combining the previous two estimates, we get that

$$\frac{u t^\gamma}{B_n} \leqslant u^{1-\alpha_+(2-\delta)\gamma} \cdot \frac{B_n^{2\alpha_++\gamma-1}}{n^{\alpha_++\gamma}} \leqslant \left(C \frac{B_n^{2/(2-\delta)}}{n^{1/(2-\delta)}} \right)^{1-\alpha_+(2-\delta)\gamma}$$

$$\times \frac{B_n^{2\alpha_++\gamma-1}}{n^{\alpha_++\gamma}} = C \left(\frac{B_n^\delta}{n} \right)^{\frac{1}{2-\delta}},$$

where we have used the fact that $1 - \alpha_+(2-\delta)\gamma > 0$ for the second inequality. Since $\delta < \alpha$, $B_n^\delta / n \to 0$ as $n \to \infty$ by the Potter bounds. This shows that $\lambda \to 0$. Note that the convergence $\lambda \to 0$ does not hold without the restriction $c_0 u n / B_n \leqslant u^{\delta-1} B_n$ (take e.g. $u = B_n$ and $t = 1$).

Now let $\epsilon > 0$ be such that $\delta + \epsilon < \alpha$ and $\gamma(\alpha - 1) - \epsilon\gamma > \gamma - 1/\alpha_1$. By the Potter bounds, $L(B_n/(ut^\gamma)) \geqslant CL(B_n)(ut^\gamma)^{-\epsilon}$. Thus, using (46) and (7),

$$\frac{un}{2B_n} \ln \mathbb{E}\left[e^{-\lambda(X^*-1)} \right] \leqslant -Cu^\alpha t^{\gamma(\alpha-1)} \frac{n}{B_n^\alpha} \cdot L \left(\frac{B_n}{ut^\gamma} \right)$$

$$= -Cu^\alpha t^{\gamma(\alpha-1)} \frac{nL(B_n)}{B_n^\alpha} \cdot \frac{L\left(\frac{B_n}{ut^\gamma} \right)}{L(B_n)}$$

$$\leqslant -Cu^{\alpha-\epsilon} t^{\gamma(\alpha-1)-\epsilon\gamma}.$$

Hence

$$\mathbb{P}\left(M^*_{un/(2B_n)} \leqslant C \frac{u^{\delta-1} B_n}{t^{1/\alpha_1}} \right) \leqslant \exp\left(Cu^\delta t^{\gamma-1/\alpha_1} - Cu^{\alpha-\epsilon} t^{\gamma(\alpha-1)-\epsilon\gamma} \right).$$

Since $\alpha - \epsilon > \delta$ and $\gamma(\alpha - 1) - \epsilon\gamma > \gamma - 1/\alpha_1$, the proof is complete. $\qquad\square$

4.5 Bounds on generation sizes

Our goal is now to establish Theorem 5 by using Proposition 6, whose proof is postponed to the end of this section. We will use the following simple consequence of Proposition 6: for every $\beta \in (0, \alpha)$, there exists a constant $C > 0$ such that

$$\text{for every } n \geqslant 1, x \geqslant 0, \qquad \mathbb{E}\left[\frac{1}{p_n Z_n^*} \mathbb{1}_{p_n Z_n^* \leqslant x} \right] \leqslant C \cdot x^{\beta-1}. \tag{33}$$

Proof of Theorem 5 Fix $\gamma \in (0, \alpha - 1)$ and $\eta > 0$. The same size-biasing argument that led us to Lemma 22 shows that

$$\mathbb{P}\left(0 < Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) < vB_n\right) \leqslant \frac{1}{\mathbb{P}_\mu(|\tau| = n)} \mathbb{E}\left[\frac{\mathbb{1}_{Z^*_{un/(2B_n)} \leqslant vB_n}}{Z^*_{un/(2B_n)}}\right] \cdot \mathbb{E}\left[\frac{1}{B'_{M^*_{un/(2B_n)}}}\right].$$

Lemma 21 combined with (33) shows that the first expectation in the previous product is bounded above by Cv^γ/B_n (for every $v \geqslant 0$, $u \geqslant \eta$ and $n \geqslant 1$). Also, using (29), observe that

$$\mathbb{E}\left[\frac{1}{B'_{M^*_{un/(2B_n)}}} \mathbb{1}_{M^*_{un/(2B_n)} \geqslant u^{\delta-1}B_n}\right] \leqslant \frac{1}{B'_{\eta^{\delta-1}B_n}} \leqslant \frac{C}{n}$$

for every $v \geqslant 0$, $u \geqslant \eta$ and $n \geqslant 1$. Hence, by Lemma 23, $\mathbb{E}\left[1/B'_{M^*_{un/(2B_n)}}\right] \leqslant C/n$ for every $v \geqslant 0$, $u \geqslant \eta$ and $n \geqslant 1$ (actually Lemma 23 is stated for $u \geqslant 2$, but the arguments carry through). The conclusion the follows since $nB_n\mathbb{P}_\mu(|\tau| = n)$ converges to a positive limit as $n \to \infty$. □

By adapting the arguments of this proof in order to control the estimates for small values of u, it is possible to establish an upper bound for $\mathbb{P}\left(0 < Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) < vB_n\right)$ valid for every $u, v \geqslant 0$. For brevity, we shall not enter such considerations.

Proof of Corollary 7 For the first assertion, note that

$$\mathbb{P}\left(Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) > vB_n\right) \leqslant \mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right),$$
$$\mathbb{P}\left(Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) > vB_n\right) \leqslant \mathbb{P}\left(W(\mathfrak{t}_n) \geqslant vB_n\right).$$

Hence

$$\mathbb{P}\left(Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) > vB_n\right) \leqslant \sqrt{\mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right) \cdot \mathbb{P}\left(W(\mathfrak{t}_n) \geqslant vB_n\right)}.$$

The desired result then follows from Theorems 1 and 2.

The second assertion is established similarly, by combining Theorem 5 with the observation that

$$\mathbb{P}\left(0 < Z_{u\frac{n}{B_n}}(\mathfrak{t}_n) < vB_n\right) \leqslant \mathbb{P}\left(H(\mathfrak{t}_n) \geqslant un/B_n\right).$$

This completes the proof. □

Remark 25 There is an analog of Proposition 6 for Z_n instead of Z_n^*. In the proof of Proposition 2.6 in [11], Croydon and Kumagai show that for every $\beta \in (0, \alpha - 1)$,

there exists a constant $C > 0$ such that $\mathbb{E}_\mu \left[p_n Z_n \leqslant x \mid Z_n > 0 \right] \leqslant x^\beta$ for every $n \geqslant 1$, $x \geqslant 0$. In this case, the exponent $\alpha - 1$ is optimal. Indeed, by [44, Theorem 1],

$$\mathbb{E}_\mu \left[e^{-\lambda p_n Z_n} \mid Z_n > 0 \right] \xrightarrow[n \to \infty]{} 1 - \frac{\lambda}{\left(1 + \lambda^{\alpha-1} \right)^{\frac{1}{\alpha-1}}}, \qquad \lambda \geqslant 0. \qquad (34)$$

As before, if Z is a random variable having this Laplace transform, for every $\epsilon > 0$, there exists $C > 0$ such that $\mathbb{P}\left(Z \leqslant x \right) \geqslant C x^{\alpha - 1 + \epsilon}$ for every $0 \leqslant x \leqslant 1$. In particular, combined with Lemma 21, this shows that the exponent $\alpha - 1$ is optimal in Theorem 5 when u takes there values in a compact subset of $(0, \infty)$.

We finally establish Proposition 6.

Proof of Proposition 6 Note that $p_n = \mathbb{P}_\mu \left(Z_n > 0 \right)$. It is clear that we may assume that $x \geqslant p_n$. In turn, it is sufficient to check the existence of $\lambda_0 > 0$, $N = N(\lambda_0)$ and a constant $C = C(\lambda_0) > 0$ such that

$$\mathbb{E}\left[e^{-\lambda p_n Z_n^*} \right] \leqslant \frac{C}{\lambda^\beta} \qquad (35)$$

holds for every $\lambda_0 \leqslant \lambda \leqslant 1/p_n$ and $n \geqslant N$. Indeed, if (35) holds, then we have, for every $x \in [p_n, 1/\lambda_0]$ and $n \geqslant N$,

$$\mathbb{P}\left(p_n Z_n^* \leqslant x \right) = \mathbb{P}\left(e^{-p_n Z_n^*/x} \geqslant 1 \right) \leqslant \mathbb{E}\left[e^{-p_n Z_n^*/x} \right] \leqslant C x^\beta.$$

From now on, we assume that $0 \leqslant \lambda \leqslant 1/p_n$.

By definition of Z_n^*, if $f_m(s) = \mathbb{E}_\mu \left[s^{Z_m} \right]$, we have $\mathbb{E}\left[s^{Z_n^*} \right] = s f_n'(s) = s \prod_{i=0}^{n-1} f_1'(f_i(s))$. In particular,

$$\mathbb{E}\left[e^{-\lambda p_n Z_n^*} \right] = e^{-\lambda p_n} \prod_{i=0}^{n-1} \widehat{F}\left(\mathbb{E}_\mu \left[e^{-\lambda p_n Z_i} \right] \right)$$

$$\leqslant \exp\left(\int_{2/n}^1 du\, n \ln\left(\widehat{F}\left(\mathbb{E}_\mu \left[e^{-\lambda p_n Z_{\lfloor un \rfloor}} \right] \right) \right) \right), \qquad (36)$$

where we set $\widehat{F}(s) = \sum_{i=1}^\infty i \mu(i) s^{i-1}$.

Fix $\epsilon \in (0, 1)$ and choose $\lambda_0 > 0$ such that $1 - \lambda_0 (1 + \lambda_0^{\alpha-1})^{-1/(\alpha-1)} < \epsilon$. By (34) and since $\lambda \mapsto \mathbb{E}_\mu \left[e^{-\lambda p_n Z_n} \mid Z_n > 0 \right]$ is decreasing, there exists $N = N(\lambda_0)$ such that

$$\text{for every } n \geqslant N, \quad \lambda \geqslant \lambda_0, \quad \mathbb{E}_\mu \left[e^{-\lambda p_n Z_n} \mid Z_n > 0 \right] \leqslant \epsilon. \qquad (37)$$

From now on, we assume that $\lambda_0 \leqslant \lambda \leqslant 1/p_n \leqslant C n^{\frac{1}{\alpha-1}+\epsilon}$, where the last inequality follows from (24) and the Potter bounds. To simplify notation, we set

$\gamma = 1/\left(\frac{1}{\alpha-1} + 2\epsilon\right)$. In particular, notice that $n/\lambda^\gamma \to \infty$ as $n \to \infty$. Indeed, since $\lambda \leqslant Cn^{\frac{1}{\alpha-1}+\epsilon}$, we have

$$\frac{n}{\lambda^\gamma} \geqslant C^{-\gamma} \cdot n \cdot n^{-\gamma\left(\frac{1}{\alpha-1}+\epsilon\right)} \xrightarrow[n\to\infty]{} \infty \tag{38}$$

because $\gamma\left(1/(\alpha-1) + \epsilon\right) < 1$.

Again by (24) and the Potter bounds, there exist two constants a constant $C_1 = C_1(\epsilon)$, $C_2 = C_2(\epsilon) > 0$ such that

$$\lambda\frac{p_n}{p_{\lfloor un \rfloor}} \geqslant C_1\lambda\left(\frac{\lfloor un \rfloor}{n}\right)^{\frac{1}{\alpha-1}+2\epsilon} \geqslant C_2\lambda u^{\frac{1}{\alpha-1}+2\epsilon}.$$

for every $2/n \leqslant u \leqslant 1$ and $n \geqslant 2$. Set $C_3 = (\lambda_0/C_2)^\gamma$, so that in particular by (37), for every n sufficiently large,

$$u \geqslant \frac{C_3}{\lambda^\gamma} \implies \lambda\frac{p_n}{p_{\lfloor un \rfloor}} \geqslant \lambda_0 \implies \mathbb{E}_\mu\left[e^{-\lambda\frac{p_n}{p_{\lfloor un \rfloor}}\cdot p_{\lfloor un \rfloor}Z_{\lfloor un \rfloor}} \,\middle|\, Z_{\lfloor un \rfloor} > 0\right] \leqslant \epsilon.$$

Thus, by observing that $\mathbb{E}_\mu\left[e^{-cZ_i}\right] = 1 - \mathbb{P}_\mu(Z_i > 0) + \mathbb{P}_\mu(Z_i > 0) \cdot \mathbb{E}_\mu\left[e^{-cZ_i}|Z_i > 0\right]$ for every $c > 0$ and $i \geqslant 0$, for every n sufficiently large and for every $u \in [C_3\lambda^{-\gamma}, 1]$, we get that

$$\mathbb{E}_\mu\left[e^{-\lambda p_n Z_{\lfloor un \rfloor}}\right] = 1 - p_{\lfloor un \rfloor}\left(1 - \mathbb{E}_\mu\left[e^{-\lambda\frac{p_n}{p_{\lfloor un \rfloor}}\cdot p_{\lfloor un \rfloor}Z_{\lfloor un \rfloor}} \,\middle|\, Z_{\lfloor un \rfloor} > 0\right]\right)$$
$$\leqslant 1 - p_{\lfloor un \rfloor}(1 - \epsilon). \tag{39}$$

For every n sufficiently large, we have $C_3\lambda^{-\gamma} \geqslant 2/n$ by (38), so that (36) and (39) yield

$$\mathbb{E}\left[e^{-\lambda p_n Z_n^*}\right] \leqslant \exp\left(\int_{C_3\lambda^{-\gamma}}^1 du\, n \ln\left(\widehat{F}\left(1 - p_{\lfloor un \rfloor}(1-\epsilon)\right)\right)\right). \tag{40}$$

Now, if X^* is a random variable with distribution given by $\mathbb{P}(X^* = i) = i\mu(i)$ for $i \geqslant 0$, we have $\widehat{F}(s) = \mathbb{E}\left[s^{X^*-1}\right]$. Hence, by (46),

$$1 - \widehat{F}(1-s) \underset{s\downarrow 0}{\sim} \frac{\Gamma(3-\alpha)}{\alpha-1} \cdot s^{\alpha-1}L(1/s),$$

so that

$$\ln\left(\widehat{F}(1-s)\right) \underset{s\downarrow 0}{\sim} -\frac{\Gamma(3-\alpha)}{\alpha-1} \cdot s^{\alpha-1}L(1/s).$$

In particular, we may choose $\eta > 0$ sufficiently small in such a way that

$$s \in (0, \eta) \implies \ln\left(\widehat{F}(1-s)\right) \leqslant -(1-\epsilon)\frac{\Gamma(3-\alpha)}{\alpha-1} \cdot s^{\alpha-1}L(1/s).$$

For every n sufficiently large and for every $u \in [C_3 \lambda^{-\gamma}, 1]$, we have $p_{\lfloor un \rfloor} \leqslant \eta$ by (38), so (40) implies that

$$\mathbb{E}\left[e^{-\lambda p_n Z_n^*}\right]$$
$$\leqslant \exp\left(-(1-\epsilon)\frac{\Gamma(3-\alpha)}{\alpha-1} \cdot \int_{C_3\lambda^{-\gamma}}^1 du \, n \left(p_{\lfloor un \rfloor}(1-\epsilon)\right)^{\alpha-1} L\left(\left(p_{\lfloor un \rfloor}(1-\epsilon)\right)^{-1}\right)\right).$$

By (24), we have

$$L\left(\left(p_{\lfloor un \rfloor}(1-\epsilon)\right)^{-1}\right) \underset{n\to\infty}{\sim} \frac{\alpha}{\Gamma(3-\alpha)\lfloor un \rfloor \, p_{\lfloor un \rfloor}^{\alpha-1}},$$

and by (38) this estimate is uniform in $C_3\lambda^{-\gamma} \leqslant u \leqslant 1$. Hence, for every n sufficiently large,

$$\mathbb{E}\left[e^{-\lambda p_n Z_n^*}\right] \leqslant \exp\left(-(1-\epsilon)^{\alpha+1}\frac{\alpha}{\alpha-1} \cdot \int_{C_3\lambda^{-\gamma}}^1 du \frac{n}{\lfloor un \rfloor}\right)$$
$$\leqslant \exp\left(-(1-\epsilon)^{\alpha+1}\frac{\alpha}{\alpha-1} \cdot \int_{C_3\lambda^{-\gamma}}^1 du \frac{1}{u}\right)$$
$$= \exp\left((1-\epsilon)^{\alpha+1}\frac{\alpha}{\alpha-1} \cdot \ln\left(C_3\lambda^{-\gamma}\right)\right) = C_4 \cdot \lambda^{-(1-\epsilon)^{\alpha+1}\cdot\frac{\alpha}{\alpha-1}\cdot\gamma}$$

with $C_4 = \exp\left((1-\epsilon)^{\alpha+1}\frac{\alpha}{\alpha-1} \cdot \ln(C_3)\right)$. Finally, observe that

$$(1-\epsilon)^{\alpha+1} \cdot \frac{\alpha}{\alpha-1} \cdot \gamma = \alpha \cdot \frac{(1-\epsilon)^{\alpha+1}}{1+2\epsilon(\alpha-1)}.$$

This completes the proof of (35), since by choosing $\epsilon \in (0,1)$ small enough, the quantity $((1-\epsilon)^{\alpha+1})/(1+2\epsilon(\alpha-1))$ will be as close to 1 as desired.   □

5 Tail estimates for the stable excursion

Recall that the tail behavior of the supremum of the associated height process H^{exc} has been obtained by Duquesne and Wang [19], see (2) and (3). Recall also that we evaluated the asymptotic behavior of the supremum of the stable bridge (Corollary 13). Here we gather several open questions concerning tail estimates for statistics of the stable bridge X^{br} and stable excursion X^{exc} which have appeared throughout the text:

What is the asymptotic behavior of $\mathbb{P}\left(\sup X^{\mathrm{br}} \leqslant 1/u\right)$, $\mathbb{P}\left(\sup X^{\mathrm{exc}} \geqslant u\right)$, $\mathbb{P}\left(\sup X^{\mathrm{exc}} \leqslant 1/u\right)$, $\mathbb{P}\left(\Delta(X^*) \geqslant u\right)$, and $\mathbb{P}\left(\Delta(X^*) \leqslant 1/u\right)$ as $u \to \infty$? What are the values $\mathbb{E}\left[(\sup X^{\mathrm{br}})^p\right]$, $\mathbb{E}\left[(\sup X^{\mathrm{exc}})^p\right]$ and $\mathbb{E}\left[\Delta(X^*)^p\right]$ for $p \geqslant 1$? In the case $\alpha = 2$, X^{exc} and H^{exc} are multiples of the normalized Brownian excursion, and such estimates are well known (see e.g. Eq. (5) and Sect. 1.1 in [2]).

There also seems to be a duality between the behavior of sup X^{exc} at ∞ (resp. 0) and the behavior of sup H^{exc} at 0 (resp. ∞): indeed, the exponent governing the exponential decay should by $\alpha/(\alpha - 1)$ for $\mathbb{P}\left(\sup X^{\text{exc}} \geqslant u\right)$ and $\mathbb{P}\left(\sup H^{\text{exc}} \leqslant 1/u\right)$, and should by α for $\mathbb{P}\left(\sup X^{\text{exc}} \leqslant 1/u\right)$ and $\mathbb{P}\left(\sup H^{\text{exc}} \geqslant u\right)$. Can this be seen in a simple way directly in the continuous world?

Appendix

In this appendix, we prove several useful results concerning the asymptotic behavior of Laplace transforms of critical offspring distributions belonging to domains of attractions of stable laws and of their associated size-biased distributions. As before, assume that μ is a critical offspring distribution belonging to the domain of attraction of a stable law of index $\alpha \in (1, 2]$. Let $\sigma^2 \in (0, \infty]$ be the variance of μ and let X is a random variable with law μ. Recall from the Introduction that L is a slowly varying function such that $\text{Var}(X \cdot \mathbb{1}_{X \leqslant n}) = n^{2-\alpha}L(n)$. Note that $L(n) \to \sigma^2$ when $\sigma^2 < \infty$, and that $L(n) \to \infty$ when $\sigma^2 = \infty$ and $\alpha = 2$. Hence, if X is a random variable with law μ, we have

$$\mathbb{E}\left[X^2 \mathbb{1}_{X \leqslant n}\right] \underset{n \to \infty}{\sim} n^{2-\alpha}L(n) + 1.$$

since $\mathbb{E}\left[X\mathbb{1}_{X \leqslant n}\right] \to 1$ as $n \to \infty$. The term "$+1$" is not negligible only when $\sigma^2 < \infty$ (in which case $\alpha = 2$).

Offspring distributions. Set $G(s) = \sum_{i \geqslant 0} \mu(i)s^i$ for $0 \leqslant s \leqslant 1$. Then by e.g. [8, Lemma 4.7]

$$G(s) - s \underset{s \uparrow 1}{\sim} \frac{\Gamma(3 - \alpha)}{\alpha(\alpha - 1)} \cdot (1 - s)^\alpha L((1 - s)^{-1}). \tag{41}$$

We stress that this holds in the both cases $\sigma^2 < \infty$ and $\sigma^2 = \infty$.

Also, if W is a random variable with distribution $\mathbb{P}\left(W = i\right) = \mu(i+1)$ for $i \geqslant -1$, since $\mathbb{E}\left[e^{-\lambda W}\right] = e^\lambda G(e^{-\lambda})$ for $\lambda > 0$, we have

$$\mathbb{E}\left[e^{-\lambda W}\right] - 1 \underset{\lambda \downarrow 0}{\sim} \frac{\Gamma(3 - \alpha)}{\alpha(\alpha - 1)} \cdot \lambda^\alpha L(1/\lambda). \tag{42}$$

Again, this holds in the both cases $\sigma^2 < \infty$ and $\sigma^2 = \infty$.

Size-biased offspring distributions. Let μ^* be the so-called size-biased probability distribution on $\mathbb{Z}_+$ defined by $\mu^*(i) = i\mu(i)$ for $i \geqslant 0$. Note that μ^* is indeed a probability distribution since μ is critical. Let X^* be a random variable having law

μ^*. When μ has finite variance, we claim that

$$1 - \mathbb{E}\left[s^{X^*}\right] \underset{s \uparrow 1}{\sim} (1-s)(\sigma^2 + 1), \tag{43}$$

and when μ has infinite variance, we claim that

$$1 - \mathbb{E}\left[s^{X^*}\right] \underset{s \uparrow 1}{\sim} \frac{\Gamma(3-\alpha)}{\alpha - 1} \cdot (1-s)^{\alpha-1} L((1-s)^{-1}),$$

$$1 - \mathbb{E}\left[e^{-\lambda X^*}\right] \underset{\lambda \downarrow 0}{\sim} \frac{\Gamma(3-\alpha)}{\alpha - 1} \cdot \lambda^{\alpha-1} L(1/\lambda). \tag{44}$$

When μ has finite variance the claim (43) simply follows from the fact that $\mathbb{E}\left[X^*\right] = \sigma^2 + 1$.

Now assume that μ has infinite variance. Then there exists a slowly varying function L_1 such that $\mathbb{P}(X \geqslant n) = \mu([n, \infty)) = L_1(n)/n^\alpha$ (see [20, Corollary XVII.5.2 and (5.16)]) when $\alpha < 2$, we have $L_1(n) = \frac{2-\alpha}{\alpha} L(n)$, and $L_1(n)/L(n) \to 0$ as $n \to \infty$ when $\alpha = 2$. As a consequence, μ^* belongs to the domain of attraction of a stable law of index $\alpha - 1$, because

$$\mu^*([n, \infty)) \underset{n \to \infty}{\sim} \frac{\alpha}{\alpha - 1} \cdot \frac{L_1(n)}{n^{\alpha-1}} \tag{45}$$

since we can write $\mu^*([n, \infty)) = (n-1)\mu([n, \infty) + \sum_{j=n}^{\infty} \mu([j, \infty))$.

If $\alpha < 2$, (45) and [20, Corollary XVII.5.2 and (5.16)] give that

$$\mathbb{E}\left[(X^*)^2 \mathbb{1}_{X^* \leqslant n}\right] \underset{n \to \infty}{\sim} n^{3-\alpha} \cdot \frac{2-\alpha}{3-\alpha} L(n),$$

and (44) result follows e.g. by [8, Lemma 4.6].

Now assume that $\alpha = 2$ and set $q_i^* = \mathbb{P}(X^* > i)$ for $i \geqslant 0$. Then

$$\sum_{i=0}^{n} q_i^* = \mathbb{E}\left[X^2 \mathbb{1}_{X \leqslant n}\right] + (n+1)\mu^*([n+1, \infty)) \underset{n \to \infty}{\sim} L(n).$$

Indeed, we know that $L_1(n)/L(n) \to 0$ as $n \to \infty$. Hence, by [20, Theorem XIII.5.5], we have $\sum_{i=0}^{\infty} q_i^* s^i \sim L((1-s)^{-1})$ as $s \uparrow 1$. Then

$$1 - \mathbb{E}\left[s^{X^*}\right] = (1-s)\sum_{i=0}^{\infty} q_i^* s^i \underset{s \uparrow 1}{\sim} (1-s)L((1-s)^{-1}).$$

Our claim (44) then follows by taking $s = e^{-\lambda}$.

Finally, from (43) and (44) it is a simple matter to see that the estimates

$$1 - \mathbb{E}\left[s^{X^*-1}\right] \underset{s \uparrow 1}{\sim} \frac{\Gamma(3-\alpha)}{\alpha - 1} \cdot (1-s)^{\alpha-1} L((1-s)^{-1}),$$

$$1 - \mathbb{E}\left[e^{-\lambda(X^*-1)}\right] \underset{\lambda\downarrow 0}{\sim} \frac{\Gamma(3-\alpha)}{\alpha-1} \cdot \lambda^{\alpha-1} L(1/\lambda) \qquad (46)$$

hold in both the cases $\sigma^2 < \infty$ and $\sigma^2 = \infty$ (when μ has infinite variance and $\alpha = 2$ we use the fact that $L(n) \to \infty$ as $n \to \infty$).

References

1. Addario-Berry, L.: Tail bounds for the height and width of a random tree with a given degree sequence. Random Struct. Algorithms **41**, 253–261 (2012)
2. Addario-Berry, L., Devroye, L., Janson, S.: Sub-Gaussian tail bounds for the width and height of conditioned Galton-Watson trees. Ann. Probab. **41**, 1072–1087 (2013)
3. Bertoin, J.: Lévy Processes, vol. 121 of Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge (1996)
4. Bertoin, J.: On the maximal offspring in a critical branching process with infinite variance. J. Appl. Probab. **48**, 576–582 (2011)
5. Bertoin, J.: On largest offspring in a critical branching process with finite variance. J. Appl. Probab. **50**, 791–800 (2013)
6. Bhamidi, S., van der Hofstad, R., Sen, S.: The multiplicative coalescent, inhomogeneous continuum random trees, and new universality classes for critical random graphs. Preprint available on arxiv, arXiv:1508.04645
7. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular Variation, vol. 27 of Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge (1987)
8. Björnberg, J.E., Stefánsson, S.Ö.: Random walk on random infinite looptrees. J. Stat. Phys. **158**, 1234–1261 (2015)
9. Broutin, N., Marckert, J.-F.: Asymptotics of trees with a prescribed degree sequence and applications. Random Struct. Algorithms **44**, 290–316 (2014)
10. Chaumont, L.: Excursion normalisée, méandre et pont pour les processus de Lévy stables. Bull. Sci. Math. **121**, 377–403 (1997)
11. Croydon, D., Kumagai, T.: Random walks on Galton-Watson trees with infinite variance offspring distribution conditioned to survive. Electron. J. Probab. **13**(51), 1419–1441 (2008)
12. Curien, N., Haas, B., Kortchemski, I.: The CRT is the scaling limit of random dissections. Random Struct. Algorithms **47**, 304–327 (2015)
13. Curien, N., Kortchemski, I.: Random non-crossing plane configurations: a conditioned Galton-Watson tree approach. Random Struct. Algorithms **45**, 236–260 (2014)
14. Curien, N., Kortchemski, I.: Random stable looptrees. Electron. J. Probab. **19**(108), 35 (2014)
15. Curien, N., Kortchemski, I.: Percolation on random triangulations and stable looptrees. Probab. Theory Relat. Fields **163**, 303–337 (2015)
16. Deutsch, E., Noy, M.: Statistics on non-crossing trees. Discrete Math. **254**, 75–87 (2002)
17. Duplantier, B., Miller, J., Sheffield, S.: Liouville quantum gravity as a mating of trees. Preprint available on arxiv, arXiv:1409.7055
18. Duquesne, T.: A limit theorem for the contour process of conditioned Galton-Watson trees. Ann. Probab. **31**, 996–1027 (2003)
19. Duquesne, T., Wang, M.: Decomposition of Lévy trees along their diameter. To appear in Ann. Inst. H. Poincaré Probab. Statist
20. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. II, 2nd edn. John Wiley & Sons Inc., New York (1971)
21. Haas, B., Miermont, G.: Scaling limits of Markov branching trees, with applications to Galton-Watson and random unordered trees. Ann. Probab. **40**, 2589–2666 (2012)
22. Ibragimov, I.A., Linnik, Y.V.: Independent and stationary sequences of random variables. Wolters-Noordhoff Publishing, Groningen (1971). With a supplementary chapter by I. A. Ibragimov and V. V. Petrov, Translation from the Russian edited by J. F. C. Kingman
23. Jacod, J., Shiryaev, A.N.: Limit Theorems for Stochastic Processes, vol. 288 of Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 2nd edn. Springer-Verlag, Berlin (2003)

24. Janson, S.: Random cutting and records in deterministic and random trees. Random Struct. Algorithms **29**, 139–179 (2006)
25. Janson, S.: Simply generated trees, conditioned Galton-Watson trees, random allocations and condensation. Probab. Surv. **9**, 103–252 (2012)
26. Kersting, G.: On the Height Profile of a Conditioned Galton-Watson Tree. Unpublished manuscript
27. Kesten, H.: Subdiffusive behavior of random walk on a random cluster. Ann. Inst. H. Poincaré Probab. Stat. **22**, 425–487 (1986)
28. Kingman, J.F.C.: Uses of exchangeability. Ann. Probab. **6**, 183–197 (1978)
29. Kortchemski, I.: Invariance principles for Galton-Watson trees conditioned on the number of leaves. Stoch. Process. Appl. **122**, 3126–3172 (2012)
30. Kortchemski, I.: A simple proof of Duquesne's theorem on contour processes of conditioned Galton-Watson trees. In: Séminaire de Probabilités XLV, vol. 2078 of Lecture Notes in Math., pp. 537–558. Springer, Cham (2013)
31. Kortchemski, I., Marzouk, C.: Triangulating stable laminations. Electron. J. Probab. **21**(11), 1–31 (2016)
32. Le Gall, J.-F.: Random trees and applications. Probability Surveys (2005)
33. Le Gall, J.-F., Le Jan, Y.: Branching processes in Lévy processes: the exploration process. Ann. Probab. **26**, 213–252 (1998)
34. Le Gall, J.-F., Miermont, G.: Scaling limits of random planar maps with large faces. Ann. Probab. **39**, 1–69 (2011)
35. Lyons, R., Pemantle, R., Peres, Y.: Conceptual proofs of $L \log L$ criteria for mean behavior of branching processes. Ann. Probab. **23**, 1125–1138 (1995)
36. Marckert, J.-F., Panholzer, A.: Noncrossing trees are almost conditioned Galton-Watson trees. Random Struct. Algorithms **20**, 115–125 (2002)
37. Pakes, A.G.: Some new limit theorems for the critical branching process allowing immigration. Stoch. Process. Appl. **4**, 175–185 (1976)
38. Pakes, A.G.: Extreme order statistics on Galton-Watson trees. Metrika **47**, 95–117 (1998)
39. Panagiotou, K., Stufler, B., Weller, K.: Scaling Limits of Random Graphs from Subcritical Classes. To appear in Ann. Probab
40. Pitman, J.: Combinatorial Stochastic Processes, vol. 1875 of Lecture Notes in Mathematics. Springer-Verlag, Berlin (2006). Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7-24, 2002, With a foreword by Jean Picard
41. Rahimov, I., Yanev, G.P.: On maximum family size in branching processes. J. Appl. Probab. **36**, 632–643 (1999)
42. Ray, G.: Large unicellular maps in high genus. Ann. Inst. Henri Poincaré Probab. Stat. **51**, 1432–1456 (2015)
43. Seneta, E.: Regularly Varying Functions. Springer-Verlag, Berlin (1976)
44. Slack, R.S.: A branching process with mean one and possibly infinite variance. Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **9**, 139–145 (1968)
45. Stufler, B.: Random enriched trees with applications to random graphs. Preprint available on arxiv, arXiv:1504.02006
46. Vatutin, V.A., Wachtel, V., Vitalihtel, Fleishcmann, K.: Critical Galton-Watson branching processes: the maximum of the total number of particles within a large window. Teor. Veroyatn. Primen. **52**, 419–445 (2007)
47. Whitt, W.: Stochastic-process limits, Springer Series in Operations Research. Springer-Verlag, New York (2002). An introduction to stochastic-process limits and their application to queues
48. Zolotarev, V.M.: One-dimensional stable distributions, vol. 65 of Translations of Mathematical Monographs, vol. 65. American Mathematical Society, Providence, RI (1986). Translated from the Russian by H. H. McFaden, Translation edited by Ben Silver