Annales de l'Institut Henri Poincaré - Probabilités et Statistiques
2020, Vol. 56, No. 1, 502–523
https://doi.org/10.1214/19-AIHP970

ANNALES
DE L'INSTITUT
HENRI
POINCARÉ
PROBABILITÉS
ET STATISTIQUES
www.imstat.org/aihp

Scaling limits of discrete snakes with stable branching

Cyril Marzouk[1]

Laboratoire de Mathématiques d'Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France. E-mail: cyril.marzouk@u-psud.fr

Received 28 February 2018; revised 10 January 2019; accepted 13 February 2019

Abstract. We consider so-called discrete snakes obtained from size-conditioned critical Bienaymé–Galton–Watson trees by assigning to each node a random spatial position in such a way that the increments along each edge are i.i.d. When the offspring distribution belongs to the domain of attraction of a stable law with index $\alpha \in (1, 2]$, we give a necessary and sufficient condition on the tail distribution of the spatial increments for this spatial tree to converge, in a functional sense, towards the Brownian snake driven by the α-stable Lévy tree. We also study the case of heavier tails, and apply our result to study the number of inversions of a uniformly random permutation indexed by the tree.

Résumé. Nous considérons des « serpents stables » obtenus à partir d'arbres de Bienaymé–Galton–Watson critiques conditionnés par la taille en assignant à chaque nœud une position spatiale de sorte que les accroissements le long des arêtes sont i.i.d. Lorsque la loi de reproduction appartient au bassin d'attraction d'une loi stable d'indice $\alpha \in]1, 2]$, nous donnons une condition nécessaire et suffisante sur la queue de distribution des accroissements spatiaux pour que cet arbre converge – en un sens fonctionnel – vers le serpent brownien sur l'arbre de Lévy α-stable. Nous étudions également le cas de queues plus lourdes et nous appliquons nos résultats au nombre d'inversions d'une permutation aléatoire indexée par l'arbre.

MSC: 60J80; 60F17; 60G50

Keywords: Brownian snake; Discrete snakes; Invariance principles; Branching random walk

1. Introduction and main results

We investigate scaling limits of large size-conditioned random Bienaymé–Galton–Watson trees equipped with spatial positions, when the offspring distribution belongs to the domain of attraction of a stable law. Our results extend previous ones established by Janson and Marckert [15] when the offspring distribution admits finite exponential moments. Relaxing this strong assumption to even a finite variance hypothesis is often challenging, and our key result is a tight control on the geometry of the trees, which is of independent interest. Let us present precisely our main result, assuming some familiarities with Bienaymé–Galton–Watson trees and their coding by paths. The basic definitions are recalled in Section 2.1 below. See Figure 1 for an example of a large random spatial tree.

1.1. *Large Bienaymé–Galton–Watson trees*

Throughout this work, we fix a probability measure μ on $\mathbf{Z}_+ = \{0, 1, \dots\}$ such that $\mu(0) > 0$ and $\sum_{k \geq 0} k\mu(k) = 1$. To simplify the exposition, we also assume that μ is *aperiodic*, in the sense that its support generates the whole group $\mathbf{Z}$, not just a strict subgroup; the general case only requires mild modifications. For every $n \geq 1$, we denote by T_n a random plane tree distributed as a *Bienaymé–Galton–Watson* tree with offspring distribution μ and conditioned to have $n + 1$ vertices,[2] which is well defined for every n large enough from the aperiodicity of μ. Finally, we assume that there exists $\alpha \in (1, 2]$ such that μ *belongs to the domain of attraction of an α-stable law*, which means that there exists an increasing sequence $(B_n)_{n \geq 1}$ such that if $(\xi_n)_{n \geq 1}$ is a sequence of i.i.d. random variables sampled from μ, then

---

[1] This work was supported by a public grant as part of the Fondation Mathématique Jacques Hadamard.

[2] We may more generally condition the trees to have n vertices with out-degree in a fixed set $A \subset \mathbf{Z}_+$, appealing to Kortchemski [18].

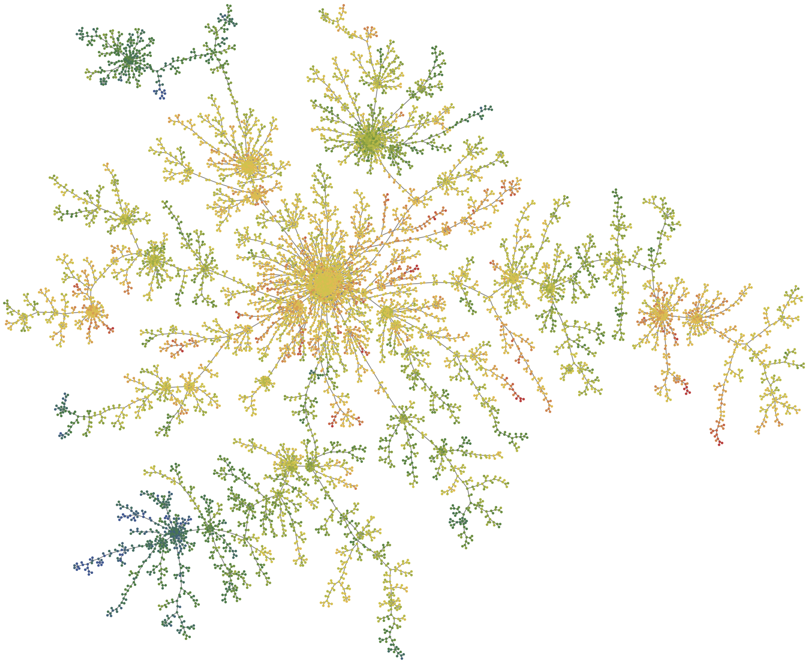

Fig. 1. A spatial stable Lévy tree with index $\alpha = 1,2$; colours indicate the spatial position of each vertex, from blue for the lowest (negative) ones to green then yellow and finally red for the highest ones. It corresponds to Theorem 1.1 with Y uniformly distributed on $\{-1, 0, 1\}$.

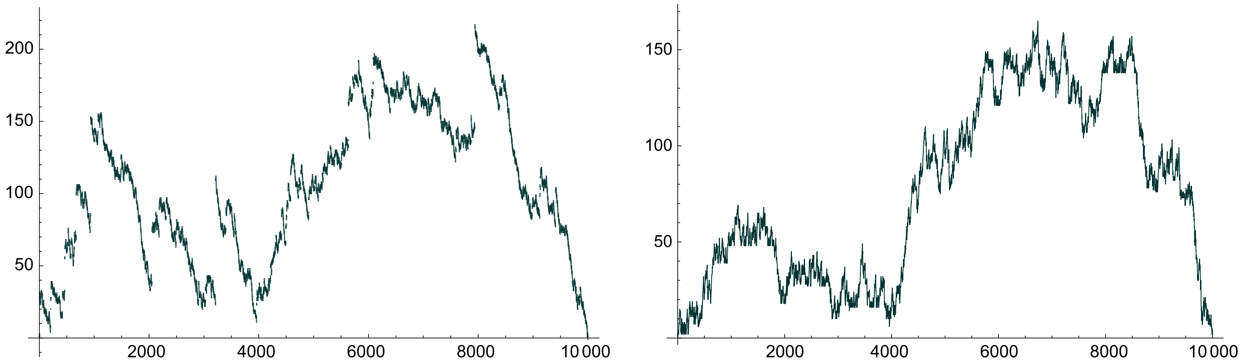

Fig. 2. The Łukasiewicz path and the height process of $T_{10,000}$ with $\alpha = 1,3$.

$B_n^{-1}(\xi_1 + \cdots + \xi_n - n)$ converges in distribution to a random variable $X^{(\alpha)}$ whose law is given by the Laplace exponent $\mathbf{E}[\exp(-\lambda X^{(\alpha)})] = \exp(\lambda^\alpha)$ for every $\lambda \geq 0$. The reader may want to keep in mind that B_n is of order $n^{1/\alpha}$ (up to a *slowly varying* sequence at infinity) and that if μ has finite variance σ_μ^2, then this falls in the case $\alpha = 2$ and we may take $B_n = (n\sigma_\mu^2/2)^{1/2}$.

It is well-known that a planar tree can be encoded by discrete paths; in Section 2.1, we recall the definition of the *Łukasiewicz path* W_n, the *height process* H_n and the *contour process* C_n associated with the tree T_n. Duquesne [8] (see also Kortchemski [18,19]) has proved that

$$\left(\frac{1}{B_n} W_n\big(\lfloor nt \rfloor\big), \frac{B_n}{n} H_n(nt), \frac{B_n}{n} C_n(2nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} (\mathcal{X}_t, \mathcal{H}_t, \mathcal{H}_t)_{t\in[0,1]} \tag{1}$$

in the Skorokhod space $\mathcal{D}([0,1], \mathbf{R}^3)$, where $\mathcal{X}$ is the normalised excursion of the α*-stable Lévy process* with no negative jump, whose value at time 1 has the law of $X^{(\alpha)}$, and $\mathcal{H}$ is the associated *height function*; see the references above for definitions and Figure 2 for an illustration. In the case $\alpha = 2$, the processes $\mathcal{X}$ and $\mathcal{H}$ are equal, both to $\sqrt{2}$ times the standard Brownian excursion. In any case, $\mathcal{H}$ is a non-negative, continuous function, which vanishes only at 0 and 1. As any such function, it encodes a 'continuum tree' called the α*-stable Lévy tree* $\mathcal{T}_\alpha$ of Duquesne, Le Gall and Le Jan [8,23], which generalises the celebrated Brownian tree of Aldous [2] in the case $\alpha = 2$. The convergence (1) implies that the tree

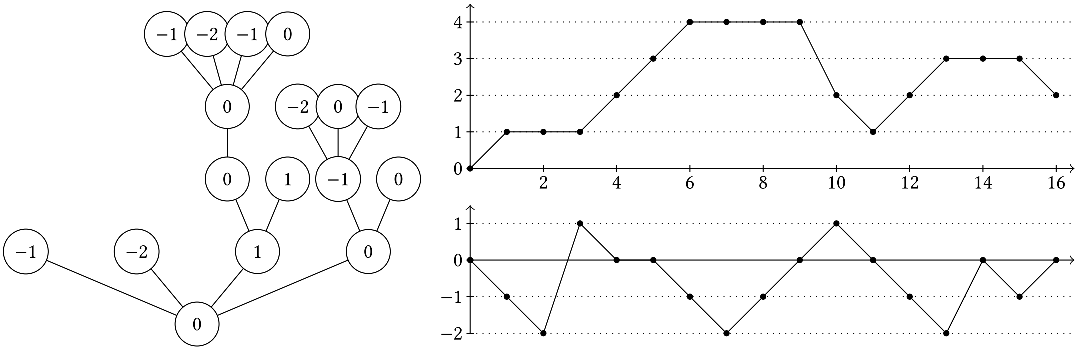

Fig. 3. A spatial tree, its height process H on top and its spatial height process H^{sp} below.

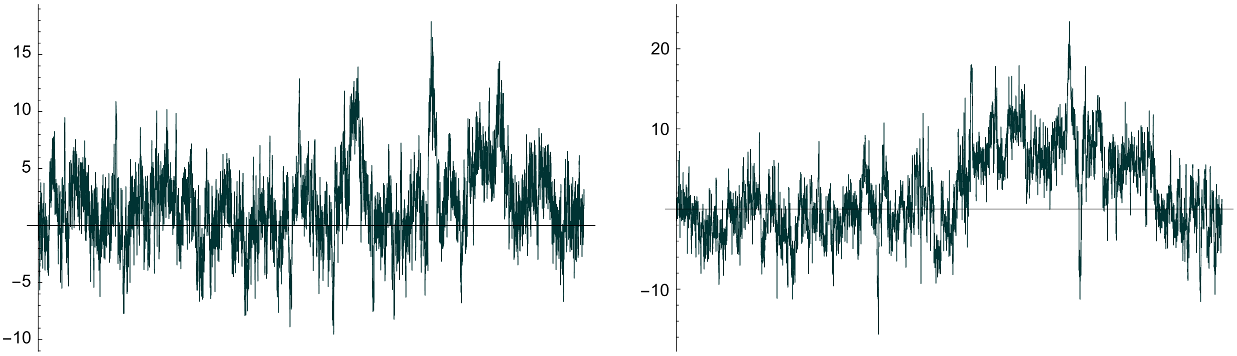

Fig. 4. Two instances of the spatial height process $\sigma_Y^{-1} H_n^{\mathrm{sp}}(n\cdot)$ associated with the height process of Figure 2: on the left, Y is uniformly distributed on $[-1, 1]$ and on the right, Y is symmetric and such that $\mathbf{P}(Y > y) = \frac{1}{2}(1 + y)^{-10}$ so both satisfy Theorem 1.1.

T_n, viewed as a metric space by endowing its vertex-set with the graph distance rescaled by a factor $\frac{B_n}{n}$, converges in distribution in the so-called Gromov–Hausdorff topology towards $\mathcal{T}_\alpha$, see e.g. Duquesne and Le Gall [11].

1.2. *Spatial trees and applications*

In this paper, we consider *spatial trees* (or *labelled trees*, or *discrete snakes*) which are plane trees in which each node u of the tree T carries a position S_u in $\mathbf{R}$. We shall always assume that the root $\varnothing$ of the tree has position $S_\varnothing = 0$ by convention so the spatial positions $(S_u)_{u\in T}$ are entirely characterised by the *displacements* $(Y_u)_{u\in T\setminus\{\varnothing\}}$ where Y_u is the difference between the position S_u of u and the position of its parent. Several models of such random spatial trees have been studied and the simplest one is the following: let Y be some random variable, then conditional on a random finite tree T, the spatial displacements $(Y_u)_{u\in T\setminus\{\varnothing\}}$ are i.i.d. copies of Y. Then the positions S_u form a branching random walk.

In the same way a tree T_n with $n + 1$ vertices is encoded by its height process H_n and its contour process C_n, the spatial positions are encoded by the *spatial height process* H_n^{sp} and the *spatial contour process* C_n^{sp}, as depicted in Figure 3. We consider scaling limits of these processes as $n \to \infty$. The most general such results are due to Janson and Marckert [15] who considered the case where the tree T_n is a size-conditioned Bienaymé–Galton–Watson tree whose offspring distribution has *finite exponential moments*. All their results extend to our setting. The main one is a necessary and sufficient condition for the convergence towards the so-called *Brownian snake* driven by the random excursion $\mathcal{H}$, which, similarly to the discrete setting, is interpreted as a Brownian motion indexed by the stable tree $\mathcal{T}_\alpha$; see Section 2.1 for a formal definition and Figure 4 for two simulations.

Theorem 1.1 (Convergence of discrete snakes). *Let $(\mathcal{H}, \mathcal{S})$ be the Brownian snake driven by the excursion $\mathcal{H}$. Suppose $\mathbf{E}[Y] = 0$ and $\sigma_Y^2 := \mathbf{E}[Y^2] \in (0, \infty)$, then the following convergence in distribution holds in the sense of finite-dimensional marginals*:

$$\left(\frac{B_n}{n} H_n(nt), \frac{B_n}{n} C_n(2nt), \left(\frac{B_n}{n\sigma_Y^2}\right)^{1/2} H_n^{\mathrm{sp}}(nt), \left(\frac{B_n}{n\sigma_Y^2}\right)^{1/2} C_n^{\mathrm{sp}}(2nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} (\mathcal{H}_t, \mathcal{H}_t, \mathcal{S}_t, \mathcal{S}_t)_{t\in[0,1]}.$$

It holds in $\mathcal{C}([0, 1], \mathbf{R}^4)$ if and only if $\mathbf{P}(|Y| \geq (n/B_n)^{1/2}) = o(n^{-1})$.

In the finite-variance case $B_n = (n\sigma_\mu^2/2)^{1/2}$, the last assumption is equivalent to $\mathbf{P}(|Y| \geq y) = o(y^{-4})$, which is weaker than $\mathbf{E}[Y^4] < \infty$ but stronger than $\mathbf{E}[Y^{4-\varepsilon}] < \infty$ for any $\varepsilon > 0$; otherwise, when the tree is less regular (see Lemma 1.4 below), one needs more regularity from the spatial displacements.

Let us mention that general arguments show that H_n and C_n, once rescaled, are close, see e.g. Le Gall [22, Section 1.6]. The same arguments apply for their spatial counterparts H_n^{sp} and C_n^{sp} so we concentrate only on the joint convergence of H_n and H_n^{sp}.

Janson and Marckert [15] also discuss the case of heavier tails, in which case the spatial processes converge once suitably rescaled towards a 'hairy snake' with vertical peaks; statements are more involved and we defer them to Section 4. Let us only mention the next result, which extends Theorem 8 in [15].

Theorem 1.2 (Non centred snakes). *Suppose that* $m_Y := \mathbf{E}[Y] \neq 0$. *Then each process* $\frac{B_n}{n} H_n^{\mathrm{sp}}(n\cdot)$ *and* $\frac{B_n}{n} C_n^{\mathrm{sp}}(2n\cdot)$ *is tight in* $\mathcal{C}([0,1], \mathbf{R})$ if and only if $\mathbf{P}(|Y| \geq n/B_n) = o(n^{-1})$, *and in this case we have the convergence in distribution in* $\mathcal{C}([0,1], \mathbf{R}^4)$

$$\left(\frac{B_n}{n} H_n(nt), \frac{B_n}{n} C_n(2nt), \frac{B_n}{n} H_n^{\mathrm{sp}}(nt), \frac{B_n}{n} C_n^{\mathrm{sp}}(2nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} (\mathscr{H}_t, \mathscr{H}_t, m_Y \mathscr{H}_t, m_Y \mathscr{H}_t)_{t\in[0,1]}.$$

Again, in the finite-variance case, the assumption is equivalent to $\mathbf{P}(|Y| \geq y) = o(y^{-2})$, which is slightly weaker than $\mathbf{E}[Y^2] < \infty$. Let us comment on the result when $Y \geq 0$ almost surely and $m_Y > 0$. In this case, for every $u \in T_n$, the displacement Y_u can be interpreted as the length of the edge from u to its parent so H_n^{sp} and C_n^{sp} can be interpreted as the height and contour processes of the tree T_n with such random edge-lengths and Theorem 1.2 shows that this tree is close to the one obtained by assigning deterministic length m_Y to each edge of T_n, and it converges towards m_Y times the stable tree for the Gromov–Hausdorff topology, jointly with the original tree.

In another direction, the main result of [15] has been used very recently by Cai *et al.* [5] to study the asymptotic number of *inversions* in a random tree. Given the random tree T_n with $n+1$ vertices listed $u_0, u_1, \dots, u_n$ and an independent uniformly random permutation of $\{0, \dots, n\}$, say, σ, assign the label $\sigma(i)$ to the vertex u_i for every $i \in \{0, \dots, n\}$. The number of inversions of T_n is then defined by

$$I(T_n) = \sum_{0\leq i<j\leq n} \mathbf{1}_{\{u_i \text{ is a ancestor of } u_j\}} \mathbf{1}_{\{\sigma(i)>\sigma(j)\}}.$$

See Figure 5 for an example. This extends the classical definition of the number of inversions of a permutation, when the tree contains a single branch. We refer to [5] for a detailed review of the literature on this model. It is easy to see that $\mathbf{E}[I(T_n) \mid T_n]$ is half the so-called *total path length* of T_n, whose asymptotic behaviour is well-understood. Cai *et al.* [5, Theorem 5] study further the fluctuations of $I(T_n)$ when T_n is a size-conditioned Bienaymé–Galton–Watson tree whose offspring distribution admits exponential moments. Apparently, this model has nothing to do with spatial trees, but the key idea from [5] is a coupling which relates the number of inversions of a random permutation on a tree with spatial positions such that Y has the uniform distribution in $(-1/2, 1/2)$; Theorem 1.1 then yields the following.

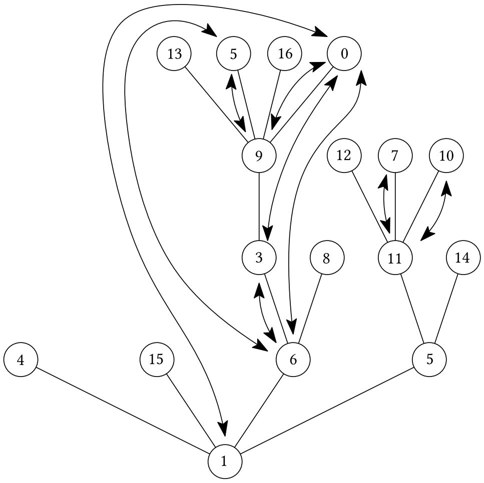

Fig. 5. A tree labelled by a permutation; there are 9 inversions, indicated by arrows.

Corollary 1.3 (Inversions on trees). *We have the convergences in distribution*

$$\frac{2B_n}{n^2}\mathbf{E}\big[I(T_n) \mid T_n\big] \xrightarrow[n\to\infty]{(d)} \int_0^1 \mathscr{H}_t \,\mathrm{d}t, \quad \textit{and} \quad \left(\frac{12B_n}{n^3}\right)^{1/2}\big(I(T_n) - \mathbf{E}\big[I(T_n) \mid T_n\big]\big) \xrightarrow[n\to\infty]{(d)} \int_0^1 \mathscr{S}_t \,\mathrm{d}t,$$

where $(\mathscr{H}, \mathscr{S})$ *is the Brownian snake driven by the excursion* $\mathscr{H}$.

We stress that the appearance of the Brownian snake really arises from the aforementioned coupling, where the constant 12 is $1/\sigma_Y^2$. If one wanted to view the labels given by the permutation as spatial positions on the tree (the Y_u's would not be independent anymore), then the corresponding spatial height process would be very different: it would not depend on the tree and would only be the interpolation of the points $(i, \sigma(i))_{0\le i\le n}$, which is much more irregular.

When $\alpha = 2$, recall that $\mathscr{H}$ is $\sqrt{2}$ times the standard Brownian excursion, then the law of $2\int_0^1 \mathscr{H}_t \,\mathrm{d}t$ is known as the Airy distribution; further, if $\mathscr{N}$ is a standard Gaussian random variable independent of $\mathscr{H}$, then $\int_0^1 \mathscr{S}_t \,\mathrm{d}t$ is distributed as $(\int_{0\le s<t\le 1} \min_{r\in[s,t]} \mathscr{H}_r \,\mathrm{d}s\,\mathrm{d}t)^{1/2}\mathscr{N}$. We refer to [14] for more information on this random variable.

The main idea to prove tightness of spatial processes is to appeal to Kolmogorov's criterion, which enables one to avoid dealing with all the correlations due to the genealogy of the trees. This requires a strong control on the geometry of the trees. Precisely, although the convergence (1) implies that the sequence $(\frac{B_n}{n}H_n(n\cdot))_{n\ge 1}$ is tight in $\mathscr{C}([0,1],\mathbf{R})$, we need the following more precise estimate on the geometry of the trees which, we believe, is of independent interest.

Lemma 1.4 (Hölder norm of the height process). *For every* $\gamma \in (0, (\alpha-1)/\alpha)$, *it holds that*

$$\lim_{C\to\infty} \liminf_{n\to\infty} \mathbf{P}\left(\sup_{0\le s\ne t\le 1} \frac{B_n}{n} \cdot \frac{|H_n(nt) - H_n(ns)|}{|t-s|^\gamma} \le C\right) = 1,$$

and the same holds when $H_n(n\cdot)$ *is replaced by* $C_n(2n\cdot)$.

By very different means, Gittenberger [13] proved a similar statement in the case $\alpha = 2$, when the offspring distribution admits finite exponential moments[3] and Janson and Marckert [15] built upon this result. Note that the maximal exponent $(\alpha-1)/\alpha$ corresponds to the maximal exponent for which the limit process $\mathscr{H}$ is Hölder continuous, see Duquesne and Le Gall [10, Theorem 1.4.4].

1.3. *More general models and random maps*

The initial motivation for studying spatial trees comes from the theory of *random planar maps*. Indeed, the *Schaeffer bijection* relates uniformly random *quadrangulations* of the sphere with n faces and such a model of spatial trees, when μ is the geometric distribution with parameter $1/2$ – in which case T_n has the uniform distribution amongst plane trees of size $n+1$ – and when Y has the uniform distribution on $\{-1, 0, 1\}$. The convergence of this particular spatial tree has been obtained Chassaing and Schaeffer [6]. More general models of random maps are also related to spatial trees, via the *Bouttier–Di Francesco–Guitter* bijection [4] and the *Janson–Stefánsson* bijection [16]; however, in this case, the displacements are neither independent nor identically distributed. Analogous convergences to Theorem 1.1 in this case have been proved by Marckert and Mokkadem [27] still for the uniform random trees, but for general displacements, under an '$(8+\varepsilon)$-moment' assumption; Gittenberger [13] extended this result to the case where μ has finite exponential moments, and then Marckert and Miermont [26] reduced the assumption on the displacements to a '$(4+\varepsilon)$-moment'; see also Miermont [29] for similar results on multi-type Bienaymé–Galton–Watson trees, Marckert [25] for 'globally centred' displacements, and finally [28] for trees (more general than size-conditioned Bienaymé–Galton–Watson trees) with finite variance, but only for the very particular displacements associated with maps. Appealing to Lemma 1.4, it seems that the '$(4+\varepsilon)$-moment' assumption suffices in the case where μ belongs to the domain of attraction of a Gaussian law to ensure the convergence towards $(\mathscr{H}, \mathscr{S})$. However in the α-stable case with $\alpha < 2$, the limit may be different and depend more precisely on the displacements, see Le Gall and Miermont [24], again for the very particular displacements associated with maps.

---

[3] Even if the assumption is written as 'finite variance' in [13], the proof requires exponential moments.

1.4. *Techniques*

The rest of this paper is organised as follows: In Section 2, we first recall the coding of plane trees by paths and define the limit object of interest $(\mathscr{H}, \mathscr{S})$; after recalling a few results on slowly varying functions and well-known results on Bienaymé–Galton–Watson trees, we prove Lemma 1.4. The idea is to rely on the Łukasiewicz path of the tree, since height of vertices corresponds to positive records of the latter, which is an excursion of a left-continuous random walk in the domain of attraction of a stable law, so it already has attracted a lot of attention and we may use several existing results, such as those due to Doney [7]. In Section 3, we prove Theorem 1.1, Theorem 1.2 and Corollary 1.3. The proof of the two theorems follows the ideas of Janson and Marckert [15] which are quite general once we have Lemma 1.4. However, several technical adaptations are needed here to deal with the heavier tails for the offspring distribution. Finally, in Section 4, we state and prove results similar to Theorem 1.1 when Y has heavier, regularly varying tails. Again, the proof scheme follows that of [15] but requires technical adaptation.

2. Geometry of large Bienaymé–Galton–Watson trees

2.1. *Discrete and continuum snakes*

Following the notation of [30], we view discrete trees as words. Let $\mathbf{N} = \{1, 2, \dots\}$ be the set of all positive integers and set $\mathbf{N}^0 = \{\varnothing\}$. Then a (plane) *tree* is a non-empty subset $T \subset \bigcup_{n\geq 0} \mathbf{N}^n$ such that: $\varnothing \in T$, it is called the *root* of T, and for every $v = (v_1, \dots, v_n) \in T$, we have $\mathrm{pr}(v) := (v_1, \dots, v_{n-1}) \in T$ and there exists an integer $k_v \geq 0$ such that $vi := (v_1, \dots, v_n, i) \in T$ if and only if $1 \leq i \leq k_v$. We shall view each vertex v of a tree T as an individual of a population for which T is the genealogical tree. For every $v = (v_1, \dots, v_n) \in T$, the vertex $\mathrm{pr}(v)$ is its *parent*, k_v is the number of *children* of v (if $k_v = 0$, then v is called a *leaf*, otherwise, v is called an *internal vertex*), and $|v| = n$ is its *generation*. We shall denote by $[\![u, v]\!]$ the unique non-crossing path between u and v.

Fix a tree T with $n + 1$ vertices, listed $\varnothing = u_0 < u_1 < \cdots < u_n$ in lexicographical order. We describe three discrete paths which each encode T. First, its *Łukasiewicz path* $W = (W(j); 0 \leq j \leq n + 1)$ is defined by $W(0) = 0$ and for every $0 \leq j \leq n$,

$$W(j + 1) = W(j) + k_{u_j} - 1.$$

One easily checks that $W(j) \geq 0$ for every $0 \leq j \leq n$ but $W(n + 1) = -1$. Next, we define the *height process* $H = (H(j); 0 \leq j \leq n)$ by setting for every $0 \leq j \leq n$,

$$H(j) = |u_j|.$$

Finally, define the *contour sequence* $(c_0, c_1, \dots, c_{2n})$ of T as follows: $c_0 = \varnothing$ and for each $i \in \{0, \dots, 2n - 1\}$, c_{i+1} is either the first child of c_i which does not appear in the sequence $(c_0, \dots, c_i)$, or the parent of c_i if all its children already appear in this sequence. The lexicographical order on the tree corresponds to the *depth-first search order*, whereas the contour order corresponds to 'moving around the tree in clockwise order'. The *contour process* $C = (C(j); 0 \leq j \leq 2n)$ is defined by setting for every $0 \leq j \leq 2n$,

$$C(j) = |c_j|.$$

We refer to Figure 6 for an illustration of these functions.

A *spatial tree* $(T, (S_u; u \in T))$ is a tree T in which each individual u is assigned a spatial position $S_u \in \mathbf{R}$, with $S_\varnothing = 0$. We encode these positions via the *spatial height* and *spatial contour* processes H^{sp} and C^{sp} respectively, defined by $H^{\mathrm{sp}}(j) = S_{u_j}$ for every $0 \leq j \leq n$ and $C^{\mathrm{sp}}(j) = S_{c_j}$ for every $0 \leq j \leq 2n$, where n is the number of edges of the tree. See Figure 3 for an illustration of H^{sp}.

Without further notice, throughout this work, every Łukasiewicz path shall be viewed as a step function, jumping at integer times, whereas height and contour processes, as well as their spatial versions, shall be viewed as continuous functions after interpolating linearly between integer times.

The analogous continuum objects we shall consider are the *stable Lévy tree* of Duquesne, Le Gall and Le Jan [8,23] which generalise Aldous' Brownian Continuum Random Tree [2] in the case $\alpha = 2$. Recall that $\mathscr{H} = (\mathscr{H}_t; t \in [0, 1])$ denotes the excursion of the *height process* associated with the α-stable Lévy process with no negative jump; we shall not need the precise definition of this process but we refer the reader to [8, Section 3.1 and 3.2]. For every $s, t \in [0, 1]$, set

$$d_{\mathscr{H}}(s, t) = \mathscr{H}_s + \mathscr{H}_t - 2 \min_{r\in[s\wedge t, s\vee t]} \mathscr{H}_r.$$

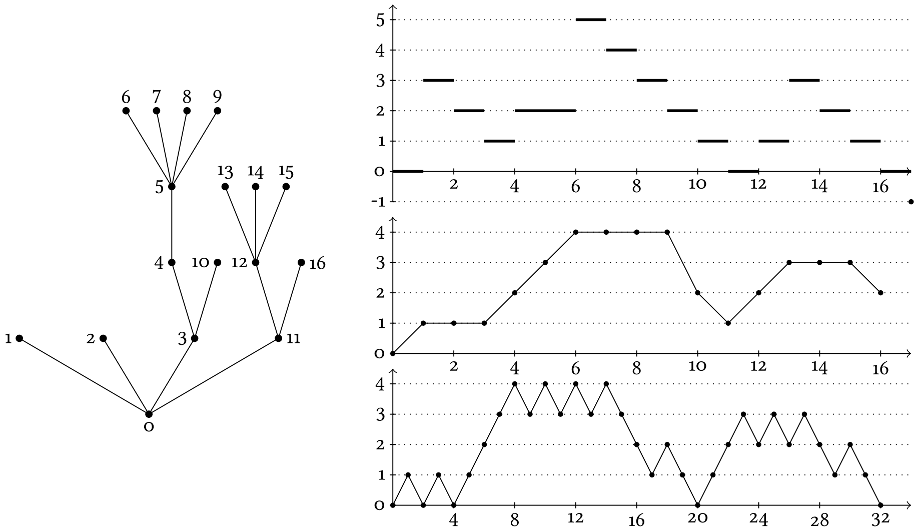

Fig. 6. A tree on the left with the lexicographical order of the vertices, and on the right, from top to bottom: its Łukasiewicz path W, its height process H, and its contour process C.

One easily checks that $d_{\mathcal{H}}$ is a random pseudo-metric on $[0, 1]$, we then define an equivalence relation on $[0, 1]$ by setting $s \sim_{\mathcal{H}} t$ whenever $d_{\mathcal{H}}(s, t) = 0$. Consider the quotient space $\mathcal{T}_\alpha = [0, 1]/{\sim_{\mathcal{H}}}$, we let $\pi_{\mathcal{H}}$ be the canonical projection $[0, 1] \to \mathcal{T}_\alpha$; then $d_{\mathcal{H}}$ induces a metric on $\mathcal{T}_\alpha$ that we still denote by $d_{\mathcal{H}}$. The space $(\mathcal{T}_\alpha, d_{\mathcal{H}})$ is a so-called compact real-tree, naturally rooted at $\pi_{\mathcal{H}}(0) = \pi_{\mathcal{H}}(1)$, called the *stable tree* coded by $\mathcal{H}$.

We construct another process $\mathcal{S} = (\mathcal{S}_t; t \in [0, 1])$ on the same probability space as $\mathcal{H}$ which, conditional on $\mathcal{H}$, is a centred Gaussian process satisfying for every $0 \le s \le t \le 1$,

$$\mathbf{E}\big[|\mathcal{S}_s - \mathcal{S}_t|^2 \mid \mathcal{H}\big] = d_{\mathcal{H}}(s, t) \quad \text{or, equivalently,} \quad \mathbf{E}[\mathcal{S}_s \mathcal{S}_t \mid \mathcal{H}] = \min_{r \in [s,t]} \mathcal{H}_r.$$

Observe that, almost surely, $\mathcal{S}_0 = 0$ and $\mathcal{S}_s = \mathcal{S}_t$ whenever $s \sim_{\mathcal{H}} t$ so $\mathcal{S}$ can be seen as a Brownian motion indexed by $\mathcal{T}_\alpha$ by setting $\mathcal{S}_{\pi_{\mathcal{H}}(t)} = \mathcal{S}_t$ for every $t \in [0, 1]$. We interpret $\mathcal{S}_x$ as the spatial position of an element $x \in \mathcal{T}_\alpha$; the pair $(\mathcal{T}_\alpha, (\mathcal{S}_x; x \in \mathcal{T}_\alpha))$ is a continuum analogue of spatial plane trees.

The *Brownian snake* driven by $\mathcal{H}$ [10,21] is a path-valued process which associates with each time $t \in [0, 1]$ the whole path of values $\mathcal{S}_x$ where x ranges over all the ancestors of $\pi_{\mathcal{H}}(t)$ in $\mathcal{T}_\alpha$, from the root to $\pi_{\mathcal{H}}(t)$, so the process $\mathcal{S}$ that we consider is only its 'tip', which is called the *head of the Brownian snake*. In this work we only consider the head of the snakes, which is in principle different from the entire snakes; nevertheless, Marckert and Mokkadem [27] proved a homeomorphism theorem which translates one into the other. Theorem 1.1 then implies the convergence of the whole snake towards the Brownian snake, see [15, Corollary 2].

It is known, see, e.g. [21, Chapter IV.4] on the whole Brownian snake, that the pair $(\mathcal{H}, \mathcal{S})$ admits a continuous version and, without further notice, we shall work throughout this paper with this version.

2.2. *Bienaymé–Galton–Watson trees and random walks*

Recall that μ is an aperiodic probability measure on $\mathbf{Z}_+$ with mean one, in the domain of attraction of an α-stable law for some $\alpha \in (1, 2]$. The Bienaymé–Galton–Watson distribution is the law on the set of all finite plane trees, which gives mass $\prod_{u \in T} \mu(k_u)$ to every such tree T. We then denote by T_n such a random tree conditioned to have $n + 1$ vertices.

The key to prove Lemma 1.4 is a well-known relation between the height process H_n and the Łukasiewicz path W_n, as well as a representation of the latter from a random walk. Our argument is inspired by the work of Le Gall and Miermont [24, Proof of Lemma 6 and 7] who consider an infinite forest of unconditioned trees, which is slightly easier thanks to the fact that the Łukasiewicz path is then a non-conditioned random walk; furthermore, there it is supposed that $\mu([k, \infty)) \sim ck^{-\alpha}$ for some constant $c > 0$, which is a stronger assumption than ours, and several arguments do not carry over.

2.2.1. *On slowly varying functions and stable domains of attraction*

Let us present a few prerequisites on slowly varying functions. First, recall that a measurable function $l : [0, \infty) \to \mathbf{R}$ is said to be slowly varying (at infinity) when for every $c > 0$, it holds that

$$\lim_{x\to\infty} \frac{l(cx)}{l(x)} = 1.$$

A property of slowly varying functions that we shall used repeatedly in Sections 3 and 4 is that for every $\varepsilon > 0$, it holds that

$$\lim_{x\to\infty} x^{-\varepsilon} l(x) = 0 \quad \text{and} \quad \lim_{x\to\infty} x^{\varepsilon} l(x) = \infty,$$

see e.g. Seneta's book [33] for more information on slowly varying functions (see Chapter 1.5 there for this property).

Let us fix a random variable X on $\{-1, 0, 1, \dots\}$ with law $\mathbf{P}(X = k) = \mu(k+1)$ for every $k \geq -1$, so $\mathbf{E}[X] = 0$. Since μ belongs to the domain of attraction of a stable law with index $\alpha \in (1, 2]$, there exist two slowly varying functions L and L_1 such that for every $n \geq 1$,

$$\mathbf{E}\big[X^2 \mathbf{1}_{\{X\leq n\}}\big] = n^{2-\alpha} L(n) \quad \text{and} \quad \mathbf{P}(X \geq n) = n^{-\alpha} L_1(n).$$

The two functions are related by

$$\lim_{n\to\infty} \frac{L_1(n)}{L(n)} = \lim_{n\to\infty} \frac{n^2 \mathbf{P}(X \geq n)}{\mathbf{E}[X^2 \mathbf{1}_{\{X\leq n\}}]} = \frac{2-\alpha}{\alpha},$$

see Feller [12, Chapter XVII, Equation 5.16]. We shall need a third slowly varying function L^* (see Doney [7, Equation 2.2]), defined uniquely up to asymptotic equivalence as the *conjugate* of $1/L$ by the following equivalent asymptotic relations:

$$\lim_{x\to\infty} L(x)^{-1/\alpha} L^*\big(x^\alpha L(x)^{-1}\big) = 1 \quad \text{and} \quad \lim_{x\to\infty} L^*(x)^{-\alpha} L\big(x^{1/\alpha} L^*(x)\big) = 1.$$

We refer to [33, Chapter 1.6] for more information about conjugation of slowly varying functions. Let $S = (S(n))_{n\geq 0}$ be a random walk started from 0 with step distribution X. As recalled in the introduction, there exists an increasing sequence $(B_n)_{n\geq 1}$ such that if $(X_n)_{n\geq 1}$ are i.d.d. copies of X, then $B_n^{-1} S(n)$ converges in distribution to some α-stable random variable. The sequence $\ell(n) = n^{-1/\alpha} B_n$ is slowly varying at infinity and in fact, the ratio $L^*(n)/\ell(n)$ converges to some positive and finite limit. For $\alpha < 2$, this was observed by Doney [7], but it extends to the case $\alpha = 2$, see the remark between Equation 2.2 and Theorem 1 in [7]: the function L there is $1/L$ here. By comparing the preceding asymptotic relations between L and L^* to [20, Equation 7], one gets precisely

$$\lim_{x\to\infty} \frac{L^*(x)}{\ell(x)} = \frac{1}{(2-\alpha)\Gamma(-\alpha)},$$

where, by continuity, the limit is interpreted as equal to 2 if $\alpha = 2$.

Doney [7, Theorem 1] studies the behaviour of the strict record times of the walk S, but his work extends *mutatis mutandis* to weak record times: let $\tau_0 = 0$ and for every $i \geq 1$, let $\tau_i = \inf\{k > \tau_{i-1} : S(k) \geq S(\tau_{i-1})\}$; in other words, the times $(\tau_n)_{n\geq 0}$ list those $k \geq 0$ such that $S(k) = \max_{0\leq i\leq k} S(i)$. Then the random variables $(\tau_{n+1} - \tau_n)_{n\geq 0}$ are i.d.d. and according to [7, Theorem 1], it holds that

$$\mathbf{P}(\tau_1 \geq n) \underset{n\to\infty}{\sim} C \cdot n^{-\frac{\alpha-1}{\alpha}} L^*(n), \tag{2}$$

with a constant $C > 0$ which shall not be important here. By a Tauberian theorem, see e.g. [12, Chapter XVII, Theorem 5.5] it follows that

$$\begin{aligned} 1 - \mathbf{E}\big[\mathrm{e}^{-\lambda\tau_1}\big] &= \big(1 - \mathrm{e}^{-\lambda}\big) \sum_{n\geq 0} \mathrm{e}^{-\lambda n} \mathbf{P}(\tau_1 > n) \\ &\underset{\lambda\downarrow 0}{\sim} C \cdot \Gamma(1/\alpha) \cdot \big(1 - \mathrm{e}^{-\lambda}\big)^{\frac{\alpha-1}{\alpha}} L^*\big(\big(1 - \mathrm{e}^{-\lambda}\big)^{-1}\big) \\ &\underset{\lambda\downarrow 0}{\sim} C_\alpha \cdot \lambda^{\frac{\alpha-1}{\alpha}} \ell\big(\lambda^{-1}\big), \end{aligned}$$

for some constant $C_\alpha > 0$, where we recall that ℓ is a slowly varying function at infinity such that $B_n = n^{1/\alpha}\ell(n)$. Taking $\lambda = N^{-1}$ with $N \in \mathbf{N}$, we obtain in particular that there exists a constant $K > 0$, which depend on μ, such that

$$\mathbf{E}\big[\mathrm{e}^{-N^{-1}\tau_1}\big] \le 1 - K \cdot N^{-1} B_N \quad \text{for every } N \in \mathbf{N}. \tag{3}$$

2.2.2. *Łukasiewicz paths and random walks*

Recall that $S = (S(i))_{i\ge 0}$ denotes a random walk started from 0 with steps $(X_i)_{i\ge 1}$ given by i.i.d. random variables with law $\mathbf{P}(X_1 = k) = \mu(k+1)$ for every $k \ge -1$. Let $X_n = (X_n(i))_{1\le i\le n+1}$ have the law of $(X_i)_{1\le i\le n+1}$ conditioned to satisfy $X_1 + \cdots + X_{n+1} = -1$ and let $S_n = (S_n(i))_{0\le i\le n+1}$ be the associated path. For every $1 \le j \le n+1$, put

$$X_n^j(k) = X_n(k + j \bmod n+1), \quad 1 \le k \le n+1.$$

We say that X_n^j is the jth *cyclic shift* of X_n. Obviously, for every $1 \le j \le n+1$, we have $X_n^j(1) + \cdots + X_n^j(n+1) = -1$, but it turns out there is a unique j such that $X_n^j(1) + \cdots + X_n^j(k) \ge 0$ for every $1 \le k \le n$. This index is the least time at which the path S_n achieves its minimum overall value:

$$j = \inf\Big\{k \in \{1, \dots, n+1\} : S_n(k) = \inf_{1\le i\le n+1} S_n(i)\Big\}. \tag{4}$$

Moreover, it is a standard fact that this time j has the uniform distribution on $\{1, \dots, n+1\}$ and furthermore $X_n^* = X_n^j$ has the same law as the increments of the Łukasiewicz path W_n of the tree T_n and it is independent of j. See e.g. [32, Chapter 6.1] for details.

We see that cyclicly shifting the path W_n at a fixed time, we obtain a random walk bridge S_n. The latter is invariant in law under time and space reversal, so by combining these observations, we obtain the following property: let $(X_n(i))_{1\le i\le n+1}$ be the increments of S_n and for a given $1 \le i \le n+1$, let $\widehat{X}_n^{(i)}(k) = X_n(i+1-k)$ for $1 \le k \le i$ and $\widehat{X}_n^{(i)}(k) = X_n(n+2+i-k)$ for $i+1 \le k \le n+1$; let $\widehat{S}_n^{(i)}$ be the associated path started from 0, then it has the same distribution as S_n.

Let us finally note that the bridge conditioning is not important: an argument based on the Markov property of S applied at time $\lceil n/2\rceil$ and the local limit theorem shows that there exists a constant $C > 0$ such that for every $n \in \mathbf{N}$ and every event A_n depending only on the first $\lceil n/2\rceil$ steps of the path, we have

$$\mathbf{P}\big(A_n \mid S(n) = -1\big) \le C \cdot \mathbf{P}(A_n),$$

see e.g. [20], near the end of the proof of Theorem 9 there.

2.2.3. *The height process as local times*

Let us list the vertices of T_n in lexicographical order as $\varnothing = u_0 < u_1 < \cdots < u_n$. It is well-known that the processes H_n and W_n are related as follows (see e.g. Le Gall and Le Jan [23]): for every $0 \le j \le n$,

$$H_n(j) = \#\Big\{k \in [0, j-1] : W_n(k) \le \inf_{k+1\le l\le j} W_n(l)\Big\}.$$

Indeed, for $k < j$, we have $W_n(k) \le \inf_{k+1\le l\le j} W_n(l)$ if and only if u_k is an ancestor of u_j; moreover, the inequality is an equality if and only if the last child of u_k is also an ancestor of u_j. Fix $i < j$ and suppose that u_i is not an ancestor of u_j (this case is treated similarly); denote by $\overline{ij} < i$ the index of the last common ancestor of u_i and u_j, and $j' \in (i, j]$ the index of the child of $u_{\overline{ij}}$ which is an ancestor of u_j. It follows from the preceding identity that the quantity $W_n(i) - \min_{i\le k\le j} W_n(k)$ counts the number of vertices branching-off of the ancestral line $[\![u_{\overline{ij}}, u_i[\![$ which lie between u_i and u_j, i.e. all the vertices visited between time i and j whose parent belongs to $[\![u_{\overline{ij}}, u_i[\![$. Indeed, started from i, the path W_n will take only values larger than or equal to $W_n(i)$ until it visits the last descendant of u_i, in which case it takes value exactly $W_n(i)$. Then W_n will decrease by one exactly at every time it visits a vertex whose parent belongs to $[\![u_{\overline{ij}}, u_i[\![$, until the last one which is $u_{j'}$. We conclude that

$$W_n\big(j'\big) = \inf_{i\le k\le j} W_n(k) \quad \text{and} \quad H_n\big(j'\big) = \inf_{i\le k\le j} H_n(k).$$

It follows that the length of the path $[\![u_{j'}, u_j]\!]$ is

$$H_n(j) - H_n\big(j'\big) = \#\Big\{k \in \big[j', j\big] : W_n(k) = \min_{k\le l\le j} W_n(l)\Big\} = \#\Big\{k \in [i, j] : W_n(k) = \min_{k\le l\le j} W_n(l)\Big\}.$$

We can now prove Lemma 1.4 appealing to the preceding subsections.

2.3. *Proof of Lemma* 1.4

Fix $\gamma \in (0, (\alpha-1)/\alpha)$. We claim that there exists a sequence of events $(E_n)_{n\geq 1}$ whose probability tends to 1 such that the following holds. There exists $c_1, c_2 > 0$ which depend on μ such that for every n large enough, every $0 \leq s \leq t \leq 1$, and every $x \geq 0$, we have

$$\mathbf{P}\left(\left|H_n(nt) - \inf_{r\in[s,t]} H_n(nr)\right| \geq x\frac{n}{B_n}|t-s|^{\gamma}\right) \leq c_1 \mathrm{e}^{-c_2 x}, \tag{5}$$

and

$$\mathbf{P}\left(\left|H_n(ns) - \inf_{r\in[s,t]} H_n(nr)\right| \geq x\frac{n}{B_n}|t-s|^{\gamma} \;\middle|\; E_n\right) \leq c_1 \mathrm{e}^{-c_2 x}. \tag{6}$$

This shows that under the conditional probability $\mathbf{P}(\cdot \mid E_n)$, the moments of $\frac{B_n}{n}\frac{|H_n(nt)-H_n(ns)|}{|t-s|^{\gamma}}$ are bounded uniformly in n and $s, t \in [0,1]$, so Lemma 1.4, first under $\mathbf{P}(\cdot \mid E_n)$, but then under the unconditioned law, follows from Kolmogorov's tightness criterion. Let us start by considering the right branch and prove (5).

Proof of (5). According to the discussion closing Section 2.2.3, our claim (5) reads as follows: for every pair $s < t$,

$$\mathbf{P}\left(\#\left\{k \in [ns, nt] : W_n(k) = \min_{k\leq l\leq nt} W_n(l)\right\} \geq x\frac{n}{B_n}|t-s|^{\gamma}\right) \leq c_1 \mathrm{e}^{-c_2 x}. \tag{7}$$

Let us first consider the random walk bridge S_n and prove that (7) holds when W_n is replaced by S_n. Note that we may, and shall, restrict to times such that $t-s \leq 1/2$ and both ns and nt are integers. By shifting the path at time nt and then taking its time and space reversal, the cardinal of the set in this probability has the same law as the number of weak records of S_n up to time $n|t-s|$. Let $(\tau_n(i))_{i\geq 0}$ be the weak record times of S_n, we therefore aim at bounding the probability

$$\mathbf{P}\left(\tau_n\left(\left\lfloor x\frac{n}{B_n}|t-s|^{\gamma}\right\rfloor\right) \leq n|t-s|\right).$$

Since $n|t-s| \leq n/2$, as explained in Section 2.2.2, this probability is bounded by some constant $C > 0$ times

$$\mathbf{P}\left(\tau\left(\left\lfloor x\frac{n}{B_n}|t-s|^{\gamma}\right\rfloor\right) \leq n|t-s|\right),$$

where $(\tau(i))_{i\geq 0}$ are the weak record times of the unconditioned walk S. Recall that $(\tau(i+1)-\tau(i))_{i\geq 0}$ are i.d.d. and let $\tau = \tau(1)$. The exponential Markov inequality shows that the preceding probability is bounded by

$$\mathrm{e}\cdot\mathbf{E}\left[\exp\left(-\frac{\tau(\lfloor x\frac{n}{B_n}|t-s|^{\gamma}\rfloor)}{n|t-s|}\right)\right] = \exp\left(1 + \left\lfloor x\frac{n}{B_n}|t-s|^{\gamma}\right\rfloor \ln\mathbf{E}\left[\exp\left(-\frac{\tau}{n|t-s|}\right)\right]\right).$$

From (3), we get that

$$\ln\mathbf{E}\left[\exp\left(-\frac{\tau}{n|t-s|}\right)\right] \leq \ln\left(1 - K\frac{B_{n|t-s|}}{n|t-s|}\right) \leq -K\frac{B_{n|t-s|}}{n|t-s|}.$$

It follows that

$$\begin{aligned}\mathrm{e}\cdot\mathbf{E}\left[\exp\left(-\frac{\tau(\lfloor x\frac{n}{B_n}|t-s|^{\gamma}\rfloor)}{n|t-s|}\right)\right] &\leq \exp\left(1 - x\frac{n}{B_n}|t-s|^{\gamma}K\frac{B_{n|t-s|}}{n|t-s|}\right)\\ &= \exp\left(1 - Kx\frac{(n|t-s|)^{-\frac{1}{\alpha}}B_{n|t-s|}}{n^{-\frac{1}{\alpha}}B_n}|t-s|^{\gamma-1+\frac{1}{\alpha}}\right).\end{aligned}$$

Let $\varepsilon = 1 - \frac{1}{\alpha} - \gamma > 0$, since the sequence $(n^{-1/\alpha}B_n)_{n\geq 1}$ is slowly varying, the so-called Potter bound (see e.g.[3, Lemma 4.2] or [20, Equation 9]) asserts that there exists a constant c, depending on ε (and so on γ), such that for every n large enough,

$$\frac{(n|t-s|)^{-1/\alpha}B_{n|t-s|}}{n^{-1/\alpha}B_n} \geq c\cdot|t-s|^{\varepsilon}.$$

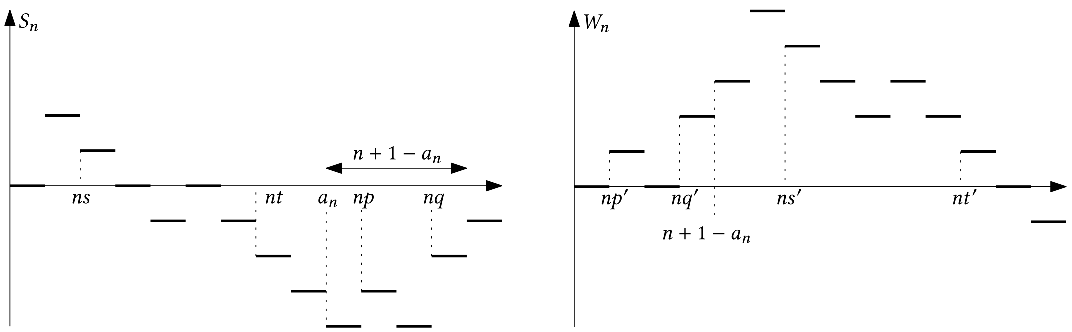

Fig. 7. A bridge S_n and its shifted excursion W_n; the times s, t fall into the first case, whereas p, q fall into the second case and s, p into the third case.

We conclude that

$$\mathbf{P}\left(\#\left\{k \in [ns, nt] : S_n(k) = \min_{k \le l \le nt} S_n(l)\right\} \ge x\frac{n}{B_n}|t-s|^\gamma\right) \le C \cdot \exp(1 - cKx),$$

for every pair $s < t$, which indeed corresponds to (7) with S_n instead of W_n.

We next prove (7) by relating W_n and S_n, as depicted in Figure 7. Recall that these paths have length $n+1$. Let us denote by a_n the time j in (4) so the path S_n shifted at time a_n has the law of W_n. Fix two times $s < t$ such that ns and nt are integers and denote by s' and t' their respective image after the shift. We distinguish three cases:

(i) Either $ns < nt \le a_n$, in which case $ns' = ns + (n+1-a_n) < nt + (n+1-a_n) = nt'$;
(ii) Either $a_n \le ns < nt$, in which case $ns' = ns - a_n < nt - a_n = nt'$;
(iii) Or $ns < a_n < nt$, in which case $nt' = nt - a_n < ns + (n+1-a_n) = ns'$.

In the first two cases, the parts of the two paths $(S_n(k))_{ns \le k \le nt}$ and $(W_n(k))_{ns' \le k \le nt'}$ are identical (up to a vertical shift), and $t' - s' = t - s$ so, according to (5), we have

$$\mathbf{P}\left(\#\left\{k \in \left[ns', nt'\right] : W_n(k) = \min_{k \le l \le nt'} W_n(l)\right\} \ge x\frac{n}{B_n}\left|t'-s'\right|^\gamma\right) \le c_1 \mathrm{e}^{-c_2 x}.$$

In the third case above, we have to be a little more careful; by cutting W_n at time $n+1-a_n$ (which corresponds to $n+1$ for S_n), we observe that $\#\{k \in [nt', ns'] : W_n(k) = \min_{k \le l \le ns'} W_n(l)\}$ is smaller than or equal to

$$\begin{aligned}
&\#\left\{k \in \left[nt', n+1-a_n\right] : W_n(k) = \min_{k \le l \le n+1-a_n} W_n(l)\right\} + \#\left\{k \in \left[n+1-a_n, ns'\right] : W_n(k) = \min_{k \le l \le ns'} W_n(l)\right\} \\
&\quad = \#\left\{k \in [nt, n+1] : S_n(k) = \min_{k \le l \le n+1} S_n(l)\right\} + \#\left\{k \in [0, ns] : S_n(k) = \min_{k \le l \le ns} S_n(l)\right\}.
\end{aligned}$$

Therefore, if we have

$$\#\left\{k \in \left[nt', ns'\right] : W_n(k) = \min_{k \le l \le ns'} W_n(l)\right\} \ge x\frac{n}{B_n}\left|t'-s'\right|^\gamma,$$

then either

$$\#\left\{k \in [nt, n+1] : S_n(k) = \min_{k \le l \le n+1} S_n(l)\right\} \ge \frac{x}{2}\frac{n}{B_n}\left|t'-s'\right|^\gamma \ge \frac{x}{2}\frac{n}{B_n}|1-t|^\gamma,$$

or (and both can occur)

$$\#\left\{k \in [0, ns] : S_n(k) = \min_{k \le l \le ns} S_n(l)\right\} \ge \frac{x}{2}\frac{n}{B_n}\left|t'-s'\right|^\gamma \ge \frac{x}{2}\frac{n}{B_n}|s|^\gamma.$$

A union bound thus yields

$$\mathbf{P}\left(\#\left\{k \in \left[nt', ns'\right] : W_n(k) = \min_{k \le l \le nt'} W_n(l)\right\} \ge x\frac{n}{B_n}\left|t'-s'\right|^\gamma\right) \le 2c_1 \mathrm{e}^{-c_2 x/2},$$

which concludes the proof of (7). □

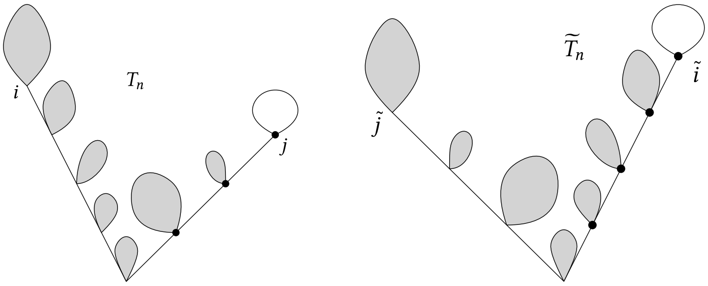

Fig. 8. On the left: a portion of the tree T_n and two vertices u_i and u_j; on the right: the 'mirror' images $\widetilde{T}_n$, $\tilde{i}$ and $\tilde{j}$. The vertices visited by W_n (resp. $\widetilde{W}_n$) between time i and j (resp. $\tilde{j}$ and $\tilde{i}$) are those black dots on the right branch as well as all the vertices strictly inside the grey trees.

The idea to control the left branch $|H_n(ns) - \inf_{r\in[s,t]} H_n(nr)|$ is to consider the 'mirror tree' obtained from T_n by flipping the order of the children of every vertex. There is one subtlety though, let us explain how to make this argument rigorous, with the help of Figure 8. Put $i = ns$ and $j = nt$. Let us denote by $\widetilde{T}_n$ the image of T_n by the following two operations: first exchange the subtrees of the progeny of the ith and the jth vertices of T_n and then take the mirror image of the whole tree, the resulting tree is $\widetilde{T}_n$. Observe that T_n and $\widetilde{T}_n$ have the same law. Let $\tilde{i} > \tilde{j}$ be the indices such that the $\tilde{i}$th and the $\tilde{j}$th vertices of $\widetilde{T}_n$ correspond to the ith and the jth vertices of T_n respectively. Then between times i and j, in T_n, the Łukasiewicz path W_n visits all the progeny of the ith vertex, then all the vertices that lie strictly between the two ancestral lines between the ith and jth vertices and their last common ancestor, and also all the vertices on this ancestral line leading to j. Similarly, between times $\tilde{j}$ and $\tilde{i}$, in $\widetilde{T}_n$, the Łukasiewicz path $\widetilde{W}_n$ visits all the progeny of the $\tilde{j}$th vertex, which is the same as that of the ith vertex of T_n, then all the vertices that lie strictly between the two ancestral lines between the $\tilde{j}$th and $\tilde{i}$th vertices and their last common ancestor, which again are the same as in T_n, and also all the vertices on this ancestral line leading to $\tilde{i}$. So the two Łukasiewicz paths visit the same vertices, except that W_n visits the ancestors of the jth vertex of T_n and not those of its ith vertex, whereas $\widetilde{W}_n$ visits the ancestors of the ith vertex of T_n and not those of its jth vertex. In principle, the lexicographical distance $|\tilde{j} - \tilde{i}|$ may thus be much larger than $|i - j|$ so we cannot directly apply the bound (5) to $\widetilde{W}_n$ (note that it could also be smaller, but this is not an issue for us, it actually helps). The following lemma shows that this difference is indeed not important.

Recall that for a vertex v of a tree T different from its root, we denote by $\mathrm{pr}(v)$ its parent and by $k_{\mathrm{pr}(v)}$ the number of children of the latter; denote further by χ_v the relative position of v among the children of $\mathrm{pr}(v)$: formally, the index $\chi_v \in \{1, \dots, k_{\mathrm{pr}(v)}\}$ satisfies $v = \mathrm{pr}(v)\chi_v$.

Lemma 2.1. *Let $C = 5/\mu(0)^2$ and define the set*

$$A_{T_n} = \left\{(u, v) \in T_n : u \in [\![\varnothing, v[\![\text{ and } \#]\!]u, v]\!] > C \ln n \text{ and } \frac{\#\{w \in]\!]u, v]\!] : \chi_w = k_{\mathrm{pr}(w)}\}}{\#]\!]u, v]\!]} > 1 - \frac{\mu(0)}{2}\right\}.$$

Then the probability that there exists a pair $(u, v) \in A_{T_n}$ converges to 0 as $n \to \infty$.

We can now finish the proof of Lemma 1.4.

Proof of (6). From the preceding lemma, we deduce that with high probability, on all ancestral paths in T_n of length at least logarithmic, there is a proportion at least $\mu(0)/2$ of individuals which are not the last child of their parent. Consequently, the length of such a path is at most $2/\mu(0)$ times the number of individuals branching-off strictly to the right of the path. With the notation of the discussion preceding the lemma, with high probability, the lexicographical distance in $\widetilde{T}_n$ between the images of the ith and jth vertex of T_n is

$$|\tilde{j} - \tilde{i}| = |i - j| - \left|H_n(j) - \inf_{i\le k\le j} H_n(k)\right| + \left|H_n(i) - \inf_{i\le k\le j} H_n(k)\right| \le |i - j| + \frac{2}{\mu(0)}|i - j|,$$

where the second (very rough) bound holds when $|H_n(i) - \inf_{i\le k\le j} H_n(k)| > C \ln n$, with C as in Lemma 2.1. Note that when this length is smaller than $C \ln n$, then (6) is trivial since $|t - s| \ge \frac{1}{n}$ as we restricted to integer times, so $x \frac{n}{B_n}|t - s|^\gamma \ge x \frac{n^{1-\gamma}}{B_n}$ which tends to infinity like a power of n. We then conclude from the bound (5) applied to the 'mirror' Łukasiewicz path $\widetilde{W}_n$. □

It remains to prove Lemma 2.1. A similar statement was proved in [28, Corollary 3] in the context of trees 'with a prescribed degree sequence'. The argument may be extended to our present case but we chose to modify it in order to directly use the existing references on Bienaymé–Galton–Watson trees.

Proof of Lemma 2.1. Fix $\varepsilon > 0$ and let T be an unconditioned Bienaymé–Galton–Watson tree with offspring distribution μ. Let us consider the set A_T. Note that the maximal height of a tree cannot exceed its total size. Then we have

$$\mathbf{P}\big(\exists(u,v)\in A_T \text{ and } \#T=n+1\big) \le \frac{1}{n+1}\sum_{h=1}^{n+1}\mathbf{E}\Bigg[\sum_{\substack{v\in T\\ |v|=h}} \mathbf{1}_{\{\exists u\in T \text{ such that } (u,v)\in A_T\}}\Bigg]. \tag{8}$$

We then use a spinal decomposition due to Duquesne [9, Equation 24] which results in an absolute continuity relation between the tree T and the tree T_∞ 'conditioned to survive', which is the infinite tree which arises as the *local limit* of T_n. It was introduced by Kesten [17] and the most general results on such convergences are due to Abraham and Delmas [1]. The tree T_∞ contains a unique infinite simple path called the *spine*, starting from the root: the vertices which belong to this spine reproduce according to the *size-biased* law $(i\mu(i))_{i\ge1}$, and the unique child which remains on the spine is chosen uniformly at random, whereas the other vertices reproduce according to μ, and all the vertices reproduce independently. Let v_h^* be the only vertex on the spine of T_∞ at height h, then, by [9, Equation 24], the expectation in (8) equals

$$\begin{aligned}
&\mathbf{P}\left(\exists u\in T_\infty : u\in \llbracket\varnothing, v_h^*\llbracket \text{ and } \#\rrbracket u, v_h^*\rrbracket > C\ln n \text{ and } \frac{\#\{w\in\rrbracket u, v_h^*\rrbracket : \chi_w = k_{\mathrm{pr}(w)}\}}{\#\rrbracket u, v_h^*\rrbracket} > 1-\frac{\mu(0)}{2}\right)\\
&\quad\le \sum_{k=C\ln n}^{h} \mathbf{P}\left(\frac{\#\{w\in\rrbracket v_{h-k}^*, v_h^*\rrbracket : \chi_w = k_{\mathrm{pr}(w)}\}}{k} > 1-\frac{\mu(0)}{2}\right).
\end{aligned}$$

Now for every vertex w on the spine, we have $\mathbf{P}(\chi_w = k_{\mathrm{pr}(w)}) = \sum_{i\ge1} i^{-1}(i\mu(i)) = 1-\mu(0)$ and these events are independent. Therefore, if $\mathsf{Bin}(N,p)$ denotes a random variable with the binomial law with parameters N and p, then the expectation in (8) is bounded by

$$\sum_{k=C\ln n}^{h} \mathbf{P}\left(k^{-1}\,\mathsf{Bin}\big(k, 1-\mu(0)\big) > 1-\frac{\mu(0)}{2}\right) \le \sum_{k=C\ln n}^{h} \mathrm{e}^{-k\mu(0)^2/2},$$

where we have used the celebrated Chernoff bound. Putting things together, we obtain the bound

$$\begin{aligned}
\mathbf{P}\big(\exists(u,v)\in A_T \mid \#T=n+1\big) &\le \frac{1}{(n+1)\mathbf{P}(\#T=n+1)}\sum_{h=1}^{n+1}\sum_{k=C\ln n}^{h} \mathrm{e}^{-k\mu(0)^2/2}\\
&\le \frac{1}{\mathbf{P}(\#T=n+1)}\frac{\mathrm{e}^{-C\ln n\mu(0)^2/2}}{1-\mathrm{e}^{-\mu(0)^2/2}}.
\end{aligned}$$

It is well-known that $nB_n\mathbf{P}(\#T=n+1)\to p_1(0)$ as $n\to\infty$, where p_1 is the density of the stable random variable $X^{(\alpha)}$ from the introduction; this follows e.g. from the fact that $\mathbf{P}(\#T=n+1)$ is the probability that the random walk S first hits -1 at time $n+1$, which equals by cyclic shift $(n+1)^{-1}$ times the probability that $S(n+1)=-1$ and the asymptotic behaviour of this probability is dictated by the local limit theorem, see e.g. [19, Lemma 1]. We conclude that for n large enough

$$\mathbf{P}\big(\exists(u,v)\in A_T \mid \#T=n+1\big) \le \frac{n^{1-C\mu(0)^2/2}B_n}{p_1(0)(1-\mathrm{e}^{-\mu(0)^2/2})}\big(1+o(1)\big),$$

which converges to 0 from our choice of C since $B_n = o(n)$. □

3. Convergence of snakes

We prove in this section the results presented in the introduction when we add to T_n spatial positions given by i.i.d. increments with law Y. Recall that we concentrate only on the joint convergence of H_n and H_n^{sp}.

3.1. *Proof of Theorem 1.1 for centred snakes*

Let us first focus on the case $\mathbf{E}[Y]=0$; we aim at showing the convergence in distribution in $\mathcal{C}([0,1],\mathbf{R}^2)$

$$\left(\frac{B_n}{n}H_n(nt), \left(\frac{B_n}{n\sigma_Y^2}\right)^{1/2} H_n^{\mathrm{sp}}(nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} (\mathscr{H}_t, \mathscr{S}_t)_{t\in[0,1]},$$

where $\sigma_Y^2 := \mathbf{E}[Y^2] \in (0,\infty)$; this convergence in the sense of finite-dimensional marginals follows easily from (1) appealing e.g. to Skorohod's representation theorem and Donsker's invariance principle applied to finitely many branches. We thus only focus on the tightness of the rescaled process $(\frac{B_n}{n})^{1/2}H_n^{\mathrm{sp}}(n\cdot)$. The idea is to apply Kolmogorov's criterion but our assumption does not give us sufficiently large moments. We therefore adapt the argument from [15] and treat separately the large and small values of Y's: the large ones are too rare to contribute much and the small ones now have sufficiently large moments. The proof takes five steps.

3.1.1. *Necessity of the assumption*

Suppose first that the assumption $\mathbf{P}(|Y| \geq (n/B_n)^{1/2}) = o(n^{-1})$ does not hold. Then there exists $\delta > 0$ such that for infinitely many indices $n \in \mathbf{N}$, we have $\mathbf{P}(|Y| \geq (n/B_n)^{1/2}) \geq \delta n^{-1}$; let us implicitly restrict ourselves to such indices. Let us denote by $\iota(T_n)$ the number of internal vertices of T_n. The conditional probability given T_n that there exists an internal vertex u such that its first child satisfies $|Y_{u1}| \geq (n/B_n)^{1/2}$ equals

$$1 - \mathbf{P}\big(|Y| < (n/B_n)^{1/2}\big)^{\iota(T_n)} \geq 1 - \left(1 - \frac{\delta}{n}\right)^{\iota(T_n)}.$$

Since $\iota(T_n)/n$ converges to $1-\mu(0)$, see e.g. [18, Lemma 2.5], the right-most term is bounded away from 0 uniformly in n. We conclude that with a probability bounded away from 0, for infinitely many indices $n \in \mathbf{N}$, there exists $0 \leq i < n$ such that $(\frac{B_n}{n})^{1/2}|H_n^{\mathrm{sp}}(i+1) - H_n^{\mathrm{sp}}(i)| \geq 1$ so the sequence of continuous processes $((\frac{B_n}{n})^{1/2}H_n^{\mathrm{sp}}(n\cdot))_{n\geq 1}$ cannot be tight.

3.1.2. *A cut-off argument*

We assume for the rest of the proof that $\mathbf{P}(|Y| \geq (n/B_n)^{1/2}) = o(n^{-1})$. Recall that for every $\delta > 0$, we have $n^{\frac{1}{\alpha}-\delta} \ll B_n \ll n^{\frac{1}{\alpha}+\delta}$ so this assumption implies $\mathbf{P}(|Y| \geq y) = o(y^{-\frac{2\alpha}{\alpha(1+\delta)-1}})$. Set $b_n = (n^2/B_n)^{\frac{\alpha-1}{4\alpha}+\varepsilon}$ for some $\varepsilon > 0$; we shall tune ε and δ small. For example, we have $\mathbf{P}(|Y| > b_n) \ll (n^2/B_n)^{-(\frac{\alpha-1}{4\alpha}+\varepsilon)(\frac{2\alpha}{\alpha(1+\delta)-1})}$, and the exponent is smaller than $-1/2$ if δ is sufficiently small. It follows that

$$\frac{n^2}{B_n}\mathbf{P}\big(|Y| > b_n\big)^2 \underset{n\to\infty}{\longrightarrow} 0. \tag{9}$$

The idea is to take into account separately the large increments. For every vertex $u \in T_n$, let $Y'_u = Y_u \mathbf{1}_{\{|Y_u| \leq b_n\}}$ and $Y''_u = Y_u \mathbf{1}_{\{|Y_u| > b_n\}}$, define then $H_n^{\mathrm{sp}\prime}$ and $H_n^{\mathrm{sp}\prime\prime}$ as the spatial processes in which the increments Y_u are replaced by Y'_u and Y''_u respectively, so $H_n^{\mathrm{sp}} = H_n^{\mathrm{sp}\prime} + H_n^{\mathrm{sp}\prime\prime}$.

3.1.3. *Contribution of the large jumps*

Let E_n be the event that T_n contains two vertices, say u and v, such that u is an ancestor of v and both $|Y_u| > b_n$ and $|Y_v| > b_n$; we claim that its probability tends to 0. First note that $\mathbf{P}(E_n \mid T_n) \leq \Lambda(T_n) \cdot \mathbf{P}(|Y| > b_n)^2$, where $\Lambda(T_n) = \sum_{u\in T_n} |u| = \sum_{i\leq n} H_n(i)$ is called the *total path length* of T_n. The convergence (1) implies that

$$\frac{B_n}{n^2}\Lambda(T_n) = \int_0^1 \frac{B_n}{n} H_n\big(\lceil nt \rceil\big)\,\mathrm{d}t \xrightarrow[n\to\infty]{(d)} \int_0^1 \mathscr{H}_t\,\mathrm{d}t. \tag{10}$$

We then write for every $K > 0$,

$$\limsup_{n\to\infty} \mathbf{P}(E_n) \leq \limsup_{n\to\infty} \mathbf{P}\left(\Lambda(T_n) > K\frac{n^2}{B_n}\right) + K \limsup_{n\to\infty} \frac{n^2}{B_n}\mathbf{P}\big(|Y| > b_n\big)^2.$$

According to (9), the second term on the right vanishes for any $K > 0$, and the first term further tends 0 when $K \to \infty$ according to (10). Now on the complement event E_n^c, there is at most one edge on each branch along which the spatial displacement is in absolute value larger than b_n, therefore $\max_{0 \le i \le n} |H_n^{\mathrm{sp}\prime\prime}(i)|$ simply equals $\max_{u \in T_n} |Y''_u|$ and so for every $\gamma > 0$, we have

$$\mathbf{P}\left(\left\{\max_{0 \le t \le 1} \left|H_n^{\mathrm{sp}\prime\prime}(2nt)\right| > \gamma \left(\frac{n}{B_n}\right)^{1/2}\right\} \cap E_n^c\right) \le \mathbf{P}\left(\max_{u \in T_n} |Y_u| > \gamma \left(\frac{n}{B_n}\right)^{1/2}\right) \le n \cdot \mathbf{P}\left(|Y| > \gamma \left(\frac{n}{B_n}\right)^{1/2}\right),$$

which converges to 0 as $n \to \infty$. Thus $(\frac{B_n}{n})^{1/2} H_n^{\mathrm{sp}\prime\prime}(n\cdot)$ converges to 0 so it only remains to prove that $(\frac{B_n}{n})^{1/2} H_n^{\mathrm{sp}\prime}(n\cdot)$ is tight.

3.1.4. *Average contribution of small jumps*

The process $(\frac{B_n}{n})^{1/2} H_n^{\mathrm{sp}\prime}(n\cdot)$ is simpler to analyse than $(\frac{B_n}{n})^{1/2} H_n^{\mathrm{sp}}(n\cdot)$ since its increments are bounded. Note nonetheless that it is non centred in general, we next prove that its conditional expectation given T_n is negligible. Let $m_n = \mathbf{E}[Y'] = -\mathbf{E}[Y'']$ and observe that $\mathbf{E}[H_n^{\mathrm{sp}\prime}(n\cdot) \mid T_n] = m_n H_n(n\cdot)$. Recall that $\frac{B_n}{n} H_n(n\cdot)$ converges in distribution (to $\mathscr{H}$); from the tail behaviour of Y we get:

$$|m_n| \le \mathbf{E}\left[\left|Y''\right|\right] \le b_n \mathbf{P}\left(|Y| > b_n\right) + \int_{b_n}^{\infty} \mathbf{P}\left(|Y| > y\right) \mathrm{d}y = O\left(b_n^{1 - \frac{2\alpha}{\alpha(1+\delta)-1}}\right).$$

By definition, we have $(\frac{n}{B_n})^{1/2} b_n^{1 - \frac{2\alpha}{\alpha(1+\delta)-1}} = n^{2(\frac{\alpha-1}{4\alpha}+\varepsilon)(1-\frac{2\alpha}{\alpha(1+\delta)-1})+\frac{1}{2}} \cdot B_n^{-(\frac{\alpha-1}{4\alpha}+\varepsilon)(1-\frac{2\alpha}{\alpha(1+\delta)-1})-\frac{1}{2}}$ and we claim that both exponents are negative, i.e. $-\frac{1}{2} < (\frac{\alpha-1}{4\alpha} + \varepsilon)(1 - \frac{2\alpha}{\alpha(1+\delta)-1}) < -\frac{1}{4}$. For $\varepsilon = \delta = 0$, it reduces to $-\frac{1}{2} < -\frac{\alpha+1}{4\alpha} < -\frac{1}{4}$ which is easily checked, and the bounds still hold by continuity for ε and δ sufficiently small. Therefore

$$\lim_{n \to \infty} \left(\frac{n}{B_n}\right)^{1/2} m_n = 0, \tag{11}$$

and thus $(B_n/n)^{1/2} \mathbf{E}[H_n^{\mathrm{sp}\prime}(n\cdot) \mid T_n]$ converges in probability to the null process; we finally focus on the centred process $\widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot) = H_n^{\mathrm{sp}\prime}(n\cdot) - \mathbf{E}[H_n^{\mathrm{sp}\prime}(n\cdot) \mid T_n]$.

3.1.5. *Re-centred small jumps are tight*

It only remains to prove that $(B_n/n)^{1/2} \widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot)$ is tight. Fix $\gamma \in (0, (\alpha - 1)/\alpha)$. According to Lemma 1.4, we may fix $C > 0$ such that the probability of the event

$$A_n = \left\{\sup_{0 \le s \ne t \le 1} \frac{B_n}{n} \cdot \frac{|H_n(nt) - H_n(ns)|}{|t - s|^{\gamma}} \le C\right\}$$

is arbitrarily close to 1 when n is large. Our aim is to apply Kolmogorov's tightness criterion to $(B_n/n)^{1/2} \widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot)$ on the event A_n. Let us enumerate the vertices of T_n in lexicographical order as $u_0 < u_1 < \cdots < u_n$. Fix $0 \le s < t \le 1$ such that ns and nt are both integers. Then $\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)$ is the sum of $\#[\![u_{ns}, u_{nt}[\![$ i.i.d. random variables distributed as $\widetilde{Y}' = Y' - \mathbf{E}[Y']$. Let $r \in [s, t]$ be as follows: set $r = s$ if u_{ns} is an ancestor of u_{nt}; otherwise, nr is an integer and u_{nr} is the ancestor of u_{nt} whose parent is the last common ancestor of u_{ns} and u_{nt}. In this way, r satisfies $H_n(nr) = \inf_{[s,t]} H_n(n\cdot)$ and it holds that $\#[\![u_{ns}, u_{nt}[\![\le 2 + H_n(ns) + H_n(nt) - 2H_n(nr)$ which, on the event A_n, is bounded by $C\frac{n}{B_n}(|t - r|^{\gamma} + |r - s|^{\gamma}) \le 2C\frac{n}{B_n}|t - s|^{\gamma}$. Fix any $q \ge 2$ and let us write C_q for a constant which will vary from one line to the other, and which depends on q and the law of Y, but not on s, t nor n.

Note that $\mathbf{E}[|\widetilde{Y}'|^2] = \mathrm{Var}(Y') \le \mathbf{E}[|Y'|^2] \le \mathbf{E}[|Y|^2] < \infty$ and $|\widetilde{Y}'|^q \le 2^q (\frac{n^2}{B_n})^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)}$. Appealing to [31, Theorem 2.9] (sometimes called the Rosenthal inequality), we obtain

$$\begin{aligned}\mathbf{E}\left[\left(\frac{|\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)|}{(n/B_n)^{1/2}}\right)^q \,\middle|\, A_n\right] &\le C_q \left(\frac{B_n}{n}\right)^{\frac{q}{2}} \left(\frac{n}{B_n}|t - s|^{\gamma} \mathbf{E}\left[\left|\widetilde{Y}'\right|^q\right] + \left(\frac{n}{B_n}|t - s|^{\gamma}\right)^{\frac{q}{2}} \mathbf{E}\left[\left|\widetilde{Y}'\right|^2\right]^{\frac{q}{2}}\right) \\ &\le C_q \left(\left(\frac{B_n}{n}\right)^{\frac{q}{2}-1} \left(\frac{n^2}{B_n}\right)^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)} |t - s|^{\gamma} + |t - s|^{\frac{q\gamma}{2}}\right).\end{aligned}$$

Recall that for every $\delta > 0$, we have $n^{\frac{1}{\alpha}-\delta} \ll B_n \ll n^{\frac{1}{\alpha}+\delta}$, then one can check that $(\frac{B_n}{n})^{\frac{q}{2}-1}(\frac{n^2}{B_n})^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)}$ is bounded above by some power of n which converges to $-\infty$ as $q \to \infty$. Indeed, taking $\varepsilon = \delta = 0$, we obtain $\frac{1}{\ln n} \ln((n^{\frac{1}{\alpha}-1})^{\frac{q}{2}-1}(n^{2-\frac{1}{\alpha}})^{q\frac{\alpha-1}{4\alpha}}) = -q\frac{\alpha-1}{4\alpha^2} + \frac{\alpha-1}{\alpha}$, and the result still hold by continuity for ε and δ sufficiently small.

Since we assume that ns and nt are integers, it holds that $n^{-1} \le |s-t| \le 1$, and therefore our expectation is bounded by some constant C_q times a power of $|t-s|$ which converges to ∞ as $q \to \infty$; in particular, for q large enough,

$$\mathbf{E}\biggl[\biggl(\frac{|\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)|}{(n/B_n)^{1/2}}\biggr)^q \biggm| A_n\biggr] \le C_q|t-s|^2.$$

This bound holds whenever $s, t \in [0, 1]$ are such that ns and nt are both integers. Since $\widetilde{H}_n^{\mathrm{sp}\prime}$ is defined by linear interpolation between such times, then it also holds for every $s, t \in [0, 1]$. The standard Kolmogorov criterion then shows that the sequence $(B_n/n)^{1/2}\widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot)$ is tight under the conditional probability $\mathbf{P}(\cdot \mid A_n)$ but then also unconditionally and the proof is complete.

3.2. *Proof of Theorem 1.2 for non-centred snakes*

We next assume that $\mathbf{E}[Y] = m_Y \neq 0$ and prove Theorem 1.2. The intuition behind the result is that the fluctuations are small and disappear after scaling, only the contribution of the expected displacement remains. Indeed, as in the preceding proof, we have $\frac{B_n}{n}\mathbf{E}[H_n^{\mathrm{sp}}(n\cdot) \mid T_n] = m_Y \frac{B_n}{n} H_n(n\cdot)$ which converges to $m_Y \cdot \mathscr{H}$ so it is equivalent to consider the centred version of Y. For the rest of the proof, we thus assume instead that $\mathbf{E}[Y] = 0$ and $\mathbf{P}(|Y| \ge n/B_n) = o(n^{-1})$, and we prove that the corresponding scaled spatial process $\frac{B_n}{n} H_n^{\mathrm{sp}}(n\cdot)$ converges to the null process.

The fact that our assumption is necessary for tightness of this process goes exactly as for Theorem 1.1, in the first step: Now the tails of Y are so that $\mathbf{P}(|Y| \ge y) = o(y^{-\frac{\alpha}{\alpha(1+\delta)-1}})$ for every $\delta > 0$ and we may proceed as previously, with the sequence $b_n = (n^2/B_n)^{\frac{\alpha-1}{2\alpha}+\varepsilon}$ instead: up to δ, ε, both exponents in the tails of Y and in b_n are half what they were in the preceding section, so these changes compensate each other. Then the previous arguments apply *mutatis mutandis* and the limits (9) and (11) now become

$$\lim_{n\to\infty} \frac{n^2}{B_n}\mathbf{P}\bigl(|Y| > b_n\bigr)^2 = 0, \quad \text{and} \quad \lim_{n\to\infty} m_n = 0,$$

so both processes $\frac{B_n}{n} H_n^{\mathrm{sp}\prime\prime}(n\cdot)$ and $\frac{B_n}{n}\mathbf{E}[H_n^{\mathrm{sp}\prime}(n\cdot) \mid T_n] = m_n \frac{B_n}{n} H_n(n\cdot)$ converge to the null process. Similarly, for $s, t \in [0, 1]$ such that $ns, nt \in \mathbf{N}$, we have

$$\begin{aligned}\mathbf{E}\biggl[\biggl(\frac{|\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)|}{n/B_n}\biggr)^q \biggm| A_n\biggr] &\le C_q\biggl(\frac{B_n}{n}\biggr)^q\biggl(\frac{n}{B_n}|t-s|^\gamma\mathbf{E}\bigl[\bigl|\widetilde{Y}'\bigr|^q\bigr] + \biggl(\frac{n}{B_n}|t-s|^\gamma\biggr)^{\frac{q}{2}}\mathbf{E}\bigl[\bigl|\widetilde{Y}'\bigr|^2\bigr]^{\frac{q}{2}}\biggr) \\ &\le C_q\biggl(\biggl(\frac{B_n}{n}\biggr)^{q-1}\biggl(\frac{n^2}{B_n}\biggr)^{q(\frac{\alpha-1}{2\alpha}+\varepsilon)}|t-s|^\gamma + \biggl(\frac{B_n}{n}\mathbf{E}\bigl[\bigl|Y'\bigr|^2\bigr]|t-s|^\gamma\biggr)^{\frac{q}{2}}\biggr).\end{aligned}$$

The first term in the last line is controlled as previously: the factor $1/2$ in the exponent in b_n compensates the fact that we now rescale by $\frac{n}{B_n}$ instead of $(\frac{n}{B_n})^{1/2}$ and similar calculations as in the preceding section show that there exists $K_q \to \infty$ as $q \to \infty$ such that $(\frac{B_n}{n})^{q-1}(\frac{n^2}{B_n})^{q(\frac{\alpha-1}{2\alpha}+\varepsilon)} \le n^{-K_q} \le |t-s|^{K_q}$. The only change compared to the proof of Theorem 1.1 is that we may not have $\mathbf{E}[|Y'|^2] < \infty$.

If $\alpha = 2$, observe that $n^{-1/2}B_n$ either converges to $(\sigma_\mu^2/2)^{1/2}$, when μ variance $\sigma_\mu^2 < \infty$, or it converges to ∞ otherwise. Therefore, one can fix $K > 0$ such that we have eventually $K \ge n^{1/2}B_n^{-1}$, and so $\mathbf{P}(|Y| \ge Kn^{1/2}) \le \mathbf{P}(|Y| \ge n/B_n) = o(n^{-1})$ by our assumption. It easily follows that $\mathbf{P}(|Y| \ge y) = o(y^{-2})$ and so

$$\mathbf{E}\bigl[\bigl|Y'\bigr|^2\bigr] = 2\int_0^{b_n} y\mathbf{P}\bigl(|Y| > y\bigr)\,\mathrm{d}y = O(\ln b_n) = O(\ln n).$$

If $\alpha < 2$, then $y\mathbf{P}(|Y| \ge y)$ is small compared to $y^{1-\frac{\alpha}{\alpha(1+\delta)-1}}$ and the exponent is smaller than -1 for δ sufficiently small so the integral above converges. In any case, there exists $\eta > 0$ such that $\frac{B_n}{n}\mathbf{E}[|Y'|^2] \le n^{-\eta} \le |t-s|^\eta$ and we may conclude as in the preceding proof that for q large enough,

$$\mathbf{E}\biggl[\biggl(\frac{|\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)|}{n/B_n}\biggr)^q \biggm| A_n\biggr] \le C_q|t-s|^2,$$

and so the process $\frac{B_n}{n}\widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot)$ is tight. Moreover, the preceding bounds applied with $s = 0$ and $t \in [0, 1]$ fixed show that the moments of the one-dimensional marginals converge to 0 so the whole process converges in distribution to the null process, which completes the proof.

3.3. *Application to the number of inversions*

Before discussing heavy-tailed snakes, let us apply Theorem 1.1 to prove Corollary 1.3, following the argument of Cai *et al.* [5, Section 5]. First note that for a given tree T with $n+1$ vertices listed $\varnothing = u_0 < u_1 < \cdots < u_n$ in lexicographical order, we have

$$\mathbf{E}\big[I(T)\big] = \frac{1}{2} \sum_{0 \le i < j \le n} \mathbf{1}_{\{u_i \text{ is a ancestor of } u_j\}} = \frac{1}{2} \sum_{u \in T} |u| = \frac{1}{2} \Lambda(T),$$

where we recall the notation $\Lambda(T)$ for the total path length of T. Therefore the convergence of the conditional expectation of $I(T_n)$ in Corollary 1.3 follows from (10). We focus on the fluctuations.

Let $(Y_u)_{u \in T_n}$ be i.i.d. spatial increments on the tree T_n, where each Y_u has the uniform distribution on the interval $(-1/2, 1/2)$. The main idea, see the discussion around Equation (5.1) in [5], is the introduction of a coupling between an inversion I on T_n and $(Y_u)_{u \in T_n}$, which yields the following comparison:

$$\left| J(T_n) - \left(I(T_n) - \frac{\Lambda(T_n)}{2} \right) \right| \le 2n,$$

where $J(T_n) = \sum_{v \in T_n} S_v = \sum_{1 \le k \le n} H_n^{\mathrm{sp}}(k)$ and we recall that S_v is the spatial position of the vertex v associated with i.d.d. increments of law Y. Since $\sigma_Y^2 = 1/12$, Theorem 1.1 yields

$$\left(\frac{12 B_n}{n^3} \right)^{1/2} J(T_n) = \int_0^1 \left(\frac{12 B_n}{n} \right)^{1/2} H_n^{\mathrm{sp}}\big(\lceil nt \rceil\big)\, \mathrm{d}t \xrightarrow[n \to \infty]{(d)} \int_0^1 \mathcal{S}_t \, \mathrm{d}t.$$

Since $(n^3/B_n)^{1/2} \gg n$, the coupling above finally implies

$$\left(\frac{12 B_n}{n^3} \right)^{1/2} \left(I(T_n) - \frac{\Lambda(T_n)}{2} \right) \xrightarrow[n \to \infty]{(d)} \int_0^1 \mathcal{S}_t \, \mathrm{d}t,$$

which is the second convergence in Corollary 1.3.

4. Heavy-tailed snakes

We investigate more precisely in this section the behaviour of H_n^{sp} and C_n^{sp} when the assumption $\mathbf{P}(|Y| \ge (\frac{n}{B_n})^{1/2}) = o(n^{-1})$ of Theorem 1.1 fails. In this case, we have seen that these processes cannot converge to continuous function since they admit large increments. In fact, they do not converge to functions at all; indeed, with high probability as n becomes large, we may find in the tree T_n vertices, say, u, which have a microscopic descendance and such that $|Y_u|$ is very large so the processes H_n^{sp} and C_n^{sp} have a macroscopic increment, almost immediately followed by the opposite increment, which gives rise at the limit to a vertical peak. Nonetheless, as proved by Janson and Marckert [15] they still converge in distribution in the following weaker sense.

Throughout this section, we identify continuous functions from $[0, 1]$ to $\mathbf{R}$ with their graph, which belongs to the space $\mathcal{K}$ of compact subsets of $[0, 1] \times \mathbf{R}$, which is a Polish space when equipped with the Hausdorff distance: the distance between two compact sets A and B is

$$d_H(A, B) = \inf\big\{r > 0 : A \subset B^{(r)} \text{ and } B \subset A^{(r)}\big\},$$

where $A^{(r)} = \{x \in \mathbf{R}^2 : d(x, A) \le r\}$. Then a sequence of functions $(f_n)_{n \ge 1}$ in $\mathcal{C}([0, 1], \mathbf{R})$ may converge in $\mathcal{K}$ to a limit K which is not the graph of a function; note that if K is the graph of a continuous function, then this convergence is equivalent to the uniform convergence considered previously. The type of limits we shall consider are constructed as follows. Take $f \in \mathcal{C}([0, 1], \mathbf{R})$ and Ξ a collection of points in $[0, 1] \times \mathbf{R}$ such that for every $x \in [0, 1]$ there exists at most one element $y \in \mathbf{R}$ such that $(x, y) \in \Xi$, and for every $\eta > 0$, the set $\Xi \cap ([0, 1] \times (\mathbf{R} \setminus [-\eta, \eta]))$ is finite. We then define a subset $f \bowtie \Xi \subset [0, 1] \times \mathbf{R}$ as the union of the graph of f and the following collection of vertical segments: for every point $(x, y) \in \Xi$, we place a vertical segment of length $|y|$ at $(x, f(x))$, directed up or down according to the sign of y. Then $f \bowtie \Xi$ belongs to $\mathcal{K}$ and the map $(f, \Xi) \mapsto f \bowtie \Xi$ is measurable (even continuous) so we may take a random function f and a random set Ξ and obtain a random set $f \bowtie \Xi$.

Again, our results focus on the head of the snakes, but they imply the convergence of the entire snakes towards 'jumping snakes', see [15, Section 3.1].

4.1. *The intermediate regime*

In the next result, we investigate the case where $n \cdot \mathbf{P}(|Y| \geq (\frac{n}{B_n})^{1/2})$ is uniformly bounded. Extracting a subsequence if necessary, we may assume that both tails converge.

Theorem 4.1 (Convergence to a 'hairy snake'). *Suppose that* $\mathbf{E}[Y] = 0$, *that* $\sigma_Y^2 := \mathbf{E}[Y^2] \in (0, \infty)$, *and that there exist* $a_+, a_- \in [0, \infty)$ *such that* $a_+ + a_- > 0$ *and*

$$\lim_{n\to\infty} n \cdot \mathbf{P}\big(Y \geq (n/B_n)^{1/2}\big) = a_+ \quad \textit{and} \quad \lim_{n\to\infty} n \cdot \mathbf{P}\big(-Y \geq (n/B_n)^{1/2}\big) = a_-.$$

Let Ξ *be a Poisson random measure on* $[0, 1] \times \mathbf{R}$ *with intensity* $\frac{2\alpha}{\alpha-1} y^{-1-\frac{2\alpha}{\alpha-1}} (a_+ \mathbf{1}_{\{y>0\}} + a_- \mathbf{1}_{\{y<0\}}) \,\mathrm{d}x \,\mathrm{d}y$ *which is independent of the pair* $(\mathscr{H}, \mathscr{S})$. *Then the convergence in distribution of the sets*

$$\left\{\left(\frac{B_n}{n}\right)^{1/2} H_n^{\mathrm{sp}}(nt); t \in [0, 1]\right\} \xrightarrow[n\to\infty]{(d)} (\sigma_Y \cdot \mathscr{S}) \bowtie \Xi,$$

holds in $\mathscr{K}$, *jointly with* (1). *The same holds* (*jointly*) *when* $H_n^{\mathrm{sp}}(n\cdot)$ *is replaced by* $C_n^{\mathrm{sp}}(2n\cdot)$.

The intuition behind this result is that, as opposed to Theorem 1.1, we can find here vertices u of T_n such that $|Y_u|$ is macroscopic, and these points lead to the peaks given by Ξ at the limit. Indeed, for every $c > 0$, we have

$$\left(\frac{cn}{B_{cn}}\right)^{1/2} = c^{\frac{\alpha-1}{2\alpha}} \left(\frac{n}{B_n}\right)^{1/2} \left(\frac{n^{-1/\alpha} B_n}{(cn)^{-1/\alpha} B_{cn}}\right)^{1/2},$$

and the very last term converges to 1 since $n^{-1/\alpha} B_n$ is slowly varying. Taking $y = c^{\frac{\alpha-1}{2\alpha}}$, we deduce that the assumption of Theorem 4.1 implies that

$$\lim_{n\to\infty} n \cdot \mathbf{P}\big(Y > y(n/B_n)^{1/2}\big) = a_+ y^{-\frac{2\alpha}{\alpha-1}}, \quad \text{and} \quad \lim_{n\to\infty} n \cdot \mathbf{P}\big(-Y < y(n/B_n)^{1/2}\big) = a_- y^{-\frac{2\alpha}{\alpha-1}}.$$

Conditional on T_n, the cardinal $\#\{u \in T_n : Y_u > y(n/B_n)^{1/2}\}$ has the binomial distribution with parameters n and $\mathbf{P}(Y > y(n/B_n)^{1/2})$, so this number is asymptotically Poisson distributed with rate $a_+ y^{-\frac{2\alpha}{\alpha-1}}$, which indeed corresponds to Ξ, since furthermore the locations are uniformly distributed in the tree.

As for Theorem 1.1, let us decompose the proof into several steps; here the argument is more straightforwardly adapted from the proof of Theorem 5 in [15] so we shall only give the main ideas.

4.1.1. *Contribution of the small jumps*

As in the proof of Theorem 1.1, let us treat separately the large and small increments: put $b_n = (n^2/B_n)^{\frac{\alpha-1}{4\alpha}+\varepsilon}$ for some $\varepsilon > 0$ small to be tuned. For every vertex $u \in T_n$, let $Y_u' = Y_u \mathbf{1}_{\{|Y_u| \leq b_n\}}$ and $Y_u'' = Y_u \mathbf{1}_{\{|Y_u| > b_n\}}$, define then $H_n^{\mathrm{sp}\prime}$ and $H_n^{\mathrm{sp}\prime\prime}$ as the spatial processes in which the increments Y_u are replaced by Y_u' and Y_u'' respectively, so $H_n^{\mathrm{sp}} = H_n^{\mathrm{sp}\prime} + H_n^{\mathrm{sp}\prime\prime}$. Our assumption implies that $\mathbf{P}(|Y| \geq y) = o(y^{-\frac{2\alpha}{\alpha(1+\delta)-1}})$ for every $\delta > 0$, and we have seen in Section 3.1.4 and 3.1.5 that

$$\left(\frac{B_n}{n} H_n(nt), \left(\frac{B_n}{n}\right)^{1/2} H_n^{\mathrm{sp}\prime}(nt)\right)_{t\in[0,1]} \xrightarrow[n\to\infty]{(d)} \big(\mathscr{H}_t, \sigma_Y^2 \cdot \mathscr{S}_t\big)_{t\in[0,1]}. \tag{12}$$

We next claim that we have the convergence in $\mathscr{K}$

$$\left\{\left(\frac{B_n}{n}\right)^{1/2} H_n^{\mathrm{sp}\prime\prime}(nt); t \in [0, 1]\right\} \xrightarrow[n\to\infty]{(d)} 0 \bowtie \Xi. \tag{13}$$

4.1.2. *Contribution of the large jumps*

The convergence (13) is basically [15, Lemma 9], let us describe the argument and refer to this reference for details. The idea is to truncate further Y_u'' by setting $Y_u^\eta = Y_u \mathbf{1}_{\{|Y_u| > \eta(n/B_n)^{1/2}\}}$ for some $\eta > 0$ and then define $H_n^{\mathrm{sp}\eta}$ as the spatial process in which the increments Y_u are replaced by Y_u^η. Recall the event E_n that T_n contains two vertices, say u and v, such that u is an ancestor of v and both $|Y_u| > b_n$ and $|Y_v| > b_n$. Then again, $\mathbf{P}(E_n)$ converges to 0, and on E_n^c, we have that $0 \leq \max_{0\leq i\leq n} |H_n^{\mathrm{sp}\prime\prime}(i)| - \max_{0\leq i\leq n} |H_n^{\mathrm{sp}\eta}(i)| \leq \eta(\frac{n}{B_n})^{1/2}$. Let us then replace $H_n^{\mathrm{sp}\prime\prime}$ by $H_n^{\mathrm{sp}\eta}$ in the left-hand side of (13). We implicitly work conditional on the event E_n^c so there is at most one non-zero value of Y_u^η along each branch of T_n. Let $\varnothing = u_0 < u_1 < \cdots < u_n$ be the vertices of T_n listed in lexicographical order, and let $1 \le k_1 < \cdots < k_{N_n} \le n$ be the indices of those vertices u of T_n for which $Y_u^\eta \neq 0$. Define then the sets $\xi_n^\eta = \{(n^{-1}k_i, (B_n/n)^{1/2} Y_{u_{k_i}}); 1 \le i \le N_n\}$ and $\Xi^\eta = \Xi \cap ([0,1] \times (\mathbf{R} \setminus [-\eta, \eta]))$ which has a finite intensity measure. Then in this context, it is easy to check (recall the discussion just after the statement of the theorem on the asymptotic behaviour of N_n and the k_i's) that ξ_n^η converges in distribution to Ξ^η. By continuity, this yields the convergence in law $0 \bowtie \xi_n^\eta \to 0 \bowtie \Xi^\eta$ and the latter is at Hausdorff distance at most η from $0 \bowtie \Xi$ by construction.

Now on the event E_n^c, the process $H_n^{\mathrm{sp}\eta}$ goes as follows: it is null until time $k_1 - 1$, then it moves to a random value $Y_{u_{k_1}}$ at time k_1, it stays at this value for a time given by the total progeny of u_{k_1} before going back to zero where it stays until time $k_2 - 1$ and so on. As shown in [15, Lemma 8], and the argument is easily adapted using (10), the total progeny of a uniformly random vertex of T_n is $o(n)$ in probability, so as $n \to \infty$, the process $H_n^{\mathrm{sp}\eta}(n\cdot)$ goes back almost immediately to 0 after reaching a high value, leading at the limit to vertical peaks as in $0 \bowtie \xi_n^\eta$. It follows that the Hausdorff distance between $\{(\frac{B_n}{n})^{1/2} H_n^{\mathrm{sp}\eta}(nt); t \in [0,1]\}$ and $0 \bowtie \xi_n^\eta$ converges to zero in probability, which, combined with the preceding paragraph, yields (13).

4.1.3. *Combining small and large jumps*

The proof is not finished! We cannot directly conclude from (12) and (13) that

$$\left\{\left(\frac{B_n}{n}\right)^{1/2}\left(H_n^{\mathrm{sp}\prime}(nt) + H_n^{\mathrm{sp}\prime\prime}(nt)\right); t \in [0,1]\right\} \xrightarrow[n\to\infty]{(d)} \left(\sigma_Y^2 \cdot \mathscr{S}_t\right) \bowtie \Xi, \tag{14}$$

because (12) and (13) may not hold simultaneously. Indeed, the processes $H_n^{\mathrm{sp}\prime}$ and $H_n^{\mathrm{sp}\prime\prime}$ are not independent since each Y_u contributes either to one or to the other. As in [15, Proof of Theorem 5], we create independence by re-sampling the Y_u's which contribute to $H_n^{\mathrm{sp}\prime\prime}$ as follows: let $(Z_i)_{i\ge1}$ be i.i.d. copies of $Y\mathbf{1}_{\{|Y|\le b_n\}}$ independent of the rest and put $\widehat{Y}_i = Y_i \mathbf{1}_{\{|Y|\le b_n\}} + Z_i \mathbf{1}_{\{|Y|>b_n\}}$ for each $1 \le i \le n$. Now the processes $\widehat{H}_n^{\mathrm{sp}}$ and $H_n^{\mathrm{sp}\prime\prime}(nt)$ are independent, and furthermore, the error between $H_n^{\mathrm{sp}\prime}$ and $\widehat{H}_n^{\mathrm{sp}}$ comes from those Y_u's for which $|Y_u| > b_n$; on the event E_n^c, there exists at most one such u on each branch and therefore $\max_{0\le i\le n} |\widehat{H}_n^{\mathrm{sp}}(i) - H_n^{\mathrm{sp}\prime}(i)| \le b_n = o(\frac{n}{B_n})^{1/2}$ since each $\widehat{Y}_i$ and each Y_i' belongs to $[0, b_n]$. Therefore the left-hand side of (14) is close to the similar set when $H_n^{\mathrm{sp}\prime}$ is replaced by $\widehat{H}_n^{\mathrm{sp}}$ which satisfies (12) and is independent of $H_n^{\mathrm{sp}\prime\prime}$. We conclude that (14) does hold and the proof of Theorem 4.1 is now complete.

4.2. *The heavy tail regime*

We finally investigate the regime where the tails of Y are much heavier than what requires Theorem 1.1. In this case, the extreme values dominate the small ones and the snake disappears at the limit, only the vertical peaks remain, see Figure 9 for a comparison with the previous case.

Theorem 4.2 (Convergence to a 'flat hairy snake'). *Fix $p \in (0,2]$ and suppose that $\mathbf{E}[Y] = 0$. Assume that there exist $\varrho \in [0,1]$ and two slowly varying functions at infinity L^+ and L^- such that if $L = L^+ + L^-$, then as $x \to \infty$, the ratios $L^+(x)/L(x)$ and $L^-(x)/L(x)$ converge respectively to ϱ and $1 - \varrho$, and furthermore*

$$n \cdot \mathbf{P}\big(Y \ge (n/B_n)^{1/p} L^+(n/B_n)\big) \underset{n\to\infty}{\longrightarrow} 1 \quad \textit{and} \quad n \cdot \mathbf{P}\big(-Y \ge (n/B_n)^{1/p} L^-(n/B_n)\big) \underset{n\to\infty}{\longrightarrow} 1.$$

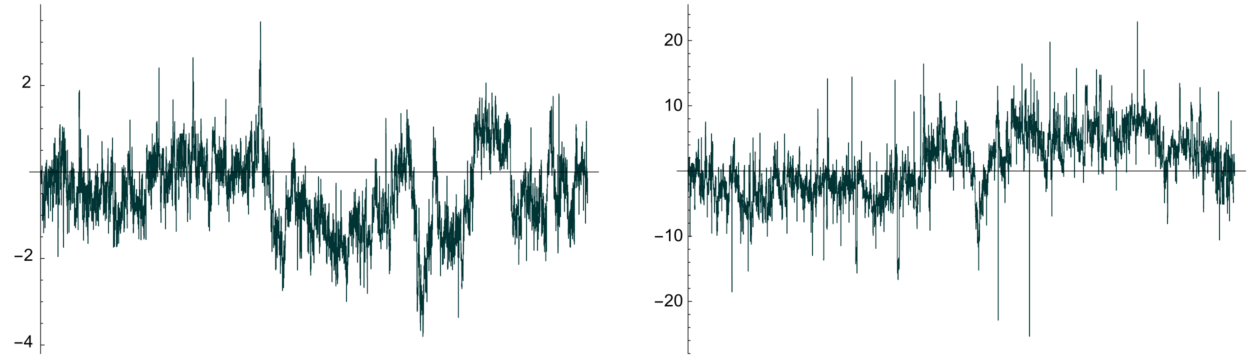

Fig. 9. Two instances of the spatial height process $H_n^{\mathrm{sp}}(n\cdot)$ associated with the height process of Figure 2 where in both cases, Y is symmetric and such that $\lim_{n\to\infty} n \cdot \mathbf{P}(Y \ge (n/B_n)^{1/p}) = 1$; on the left, $p = 2$ and on the right, $p = 0.6$.

If $p = 2$, *assume also that the function* L *tends to infinity. Let* Ξ *be a Poisson random measure on* $[0,1] \times \mathbf{R}$ *independent of the pair* $(\mathscr{H}, \mathscr{S})$, *with intensity* $\frac{p\alpha}{\alpha-1} y^{-1-\frac{p\alpha}{\alpha-1}} (\varrho^{\frac{p\alpha}{\alpha-1}} \mathbf{1}_{\{y>0\}} + (1-\varrho)^{\frac{p\alpha}{\alpha-1}} \mathbf{1}_{\{y<0\}}) \,\mathrm{d}x\,\mathrm{d}y$. *Then the convergence in distribution of the sets*

$$\left\{ \frac{B_n^{1/p}}{n^{1/p} L(\frac{n}{B_n})} H_n^{\mathrm{sp}}(nt); t \in [0,1] \right\} \xrightarrow[n\to\infty]{(d)} 0 \bowtie \Xi,$$

holds in $\mathscr{K}$, *jointly with* (1). *The same holds (jointly) when* $H_n^{\mathrm{sp}}(n\cdot)$ *is replaced by* $C_n^{\mathrm{sp}}(2n\cdot)$.

In the case $L^+(x) \to c_+ \in [0, \infty)$ and $L^-(x) \to c_- \in [0, \infty)$, the assumption reads

$$n \cdot \mathbf{P}\big(Y \geq (n/B_n)^{1/p}\big) \underset{n\to\infty}{\longrightarrow} a_+ \quad \text{and} \quad n \cdot \mathbf{P}\big(-Y \geq (n/B_n)^{1/p}\big) \underset{n\to\infty}{\longrightarrow} a_-,$$

where $a_+ = (c_+)^{\frac{p\alpha}{\alpha-1}}$ and $a_- = (c_-)^{\frac{p\alpha}{\alpha-1}}$, and then the conclusion reads

$$\left\{ \left(\frac{B_n}{n}\right)^{1/p} H_n^{\mathrm{sp}}(nt); t \in [0,1] \right\} \xrightarrow[n\to\infty]{(d)} 0 \bowtie \Xi,$$

where Ξ has intensity $\frac{p\alpha}{\alpha-1} y^{-1-\frac{p\alpha}{\alpha-1}} (a_+ \mathbf{1}_{\{y>0\}} + a_- \mathbf{1}_{\{y<0\}}) \,\mathrm{d}x\,\mathrm{d}y$, which recovers [15, Theorem 6].

Remark 4.3. Recall from Theorem 1.2 that if $\mathbf{E}[Y] = m_Y \neq 0$, then $\frac{B_n}{n} \mathbf{E}[H_n^{\mathrm{sp}}(n\cdot) \mid T_n]$ converges to $m_Y \cdot \mathscr{H}$. Therefore the previous result still holds in this case for $p < 1$; when $p = 1$ and L^+ and L^- both converge, then

$$\left\{ \left(\frac{B_n}{n}\right)^{1/p} H_n^{\mathrm{sp}}(nt); t \in [0,1] \right\} \xrightarrow[n\to\infty]{(d)} (m_Y \cdot \mathscr{H}) \bowtie \Xi,$$

in $\mathscr{K}$, jointly with (1), where $\mathscr{H}$ and Ξ are independent, and the same holds (jointly) when $H_n^{\mathrm{sp}}(n\cdot)$ is replaced by $C_n^{\mathrm{sp}}(2n\cdot)$.

Proof. Since $L^{+/-}$ are slowly varying, we have $(n/B_n)^{-\theta} \ll L^{+/-}(n/B_n) \ll (n/B_n)^{\theta}$ for every $\theta > 0$ so the tails of Y satisfy now $\mathbf{P}((Y)_{+/-} > y) = o(y^{-\frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}})$ for every $\delta, \theta > 0$. As usual, let us cut the increments: put $b_n = (n^2/B_n)^{\frac{\alpha-1}{2p\alpha}+\varepsilon}$ for some $\varepsilon > 0$ small. For every vertex $u \in T_n$, let $Y'_u = Y_u \mathbf{1}_{\{|Y_u| \leq b_n\}}$ and $Y''_u = Y_u \mathbf{1}_{\{|Y_u| > b_n\}}$, define then $H_n^{\mathrm{sp}\prime}$ and $H_n^{\mathrm{sp}\prime\prime}$ as the spatial processes in which the increments Y_u are replaced by Y'_u and Y''_u respectively, finally set $m_n = \mathbf{E}[Y'] = -\mathbf{E}[Y'']$ and $\widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot) = H_n^{\mathrm{sp}\prime}(n\cdot) - m_n H_n(n\cdot)$. Then we have $H_n^{\mathrm{sp}} = H_n^{\mathrm{sp}\prime\prime} + m_n H_n + \widetilde{H}_n^{\mathrm{sp}\prime}$. Similarly as in the preceding proofs, the exponent in b_n matches that in the tails of Y. Therefore, taking $\varepsilon, \delta, \theta$ small enough, similarly to (9) and (11), it holds that

$$\lim_{n\to\infty} \frac{n^2}{B_n} \mathbf{P}\big(|Y| > b_n\big)^2 = 0, \quad \text{and} \quad \lim_{n\to\infty} \frac{n^{(p-1)/p}}{B_n^{(p-1)/p} L(n/B_n)} m_n = 0.$$

The event E_n that T_n contains two vertices, say u and v, such that u is an ancestor of v and both $|Y_u| > b_n$ and $|Y_v| > b_n$, thus has a probability tending to 0. Then the arguments used in the proof of Theorem 4.1 extend readily to prove that

$$\left\{ \frac{B_n^{1/p}}{n^{1/p} L(n/B_n)} H_n^{\mathrm{sp}\prime\prime}(nt); t \in [0,1] \right\} \xrightarrow[n\to\infty]{(d)} 0 \bowtie \Xi,$$

furthermore it holds that

$$\frac{B_n^{1/p}}{n^{1/p} L(n/B_n)} m_n H_n(n\cdot) = \frac{n^{(p-1)/p}}{B_n^{(p-1)/p} L(n/B_n)} m_n \cdot \frac{B_n}{n} H_n(n\cdot) \xrightarrow[n\to\infty]{\mathbf{P}} 0.$$

Finally, with the notation from Section 3, for $s, t \in [0,1]$ such that $ns, nt \in \mathbf{N}$, we have

$$\mathbf{E}\left[\left(\frac{B_n^{1/p}}{n^{1/p} L(n/B_n)} \big|\widetilde{H}_n^{\mathrm{sp}\prime}(nt) - \widetilde{H}_n^{\mathrm{sp}\prime}(ns)\big|\right)^q \;\Big|\; A_n\right]$$

$$\leq C_q \frac{B_n^{q/p}}{n^{q/p} L(n/B_n)^q} \left(\frac{n}{B_n} |t-s|^\gamma \mathbf{E}\big[\big| \widetilde{Y}' \big|^q \big] + \left(\frac{n}{B_n} |t-s|^\gamma \right)^{\frac{q}{2}} \mathbf{E}\big[\big| \widetilde{Y}' \big|^2 \big]^{\frac{q}{2}} \right)$$

$$\leq C_q \left(L\left(\frac{n}{B_n}\right)^{-q} \left(\frac{B_n}{n}\right)^{\frac{q}{p}-1} \left(\frac{n^2}{B_n}\right)^{q(\frac{\alpha-1}{p\alpha}+\varepsilon)} |t-s|^\gamma + \left[L\left(\frac{n}{B_n}\right)^{-2} \left(\frac{B_n}{n}\right)^{\frac{2}{p}-1} \mathbf{E}\big[\big| Y' \big|^2 \big] \right]^{\frac{q}{2}} |t-s|^{q\gamma/2} \right).$$

The first term in the last line is controlled exactly as in the proof of Theorem 1.2: the slowly varying function will not cause any trouble, and the factor $1/p$ here acts as $1/2$ there. On the other hand, exactly as in the proof of Theorem 1.2, one can check from the tail behaviour of Y that if $p < 2$, then there exists $\eta > 0$ such that $(\frac{B_n}{n})^{\frac{2}{p}-1}\mathbf{E}[|Y'|^2] \leq n^{-\eta} \leq |t-s|^\eta$; this fails when $p = 2$ but in this case $(\frac{B_n}{n})^{\frac{2}{p}-1}\mathbf{E}[|Y'|^2]$ is uniformly bounded, so this is not an issue. We conclude as in the proof of Theorem 1.2 that the process $\frac{B_n}{n}\widetilde{H}_n^{\mathrm{sp}\prime}(n\cdot)$ is tight and even converges to the null process, which completes the proof. □

Acknowledgement

I wish to thank Nicolas Curien for a stimulating discussion on the proof of Lemma 1.4 when I started to have some doubts on the strategy used below. Many thanks are also due to Igor Kortchemski who spotted a mistake in a first draft as well as two anonymous referees for their helpful remarks and encouragements to refine Section 4.

References

[1] R. Abraham and J.-F. Delmas. Local limits of conditioned Galton–Watson trees: The infinite spine case. *Electron. J. Probab.* **19** (2) (2014) 2. MR3164755 https://doi.org/10.1214/ejp.v19-2747
[2] D. Aldous. The continuum random tree. III. *Ann. Probab.* **21** (1) (1993) 248–289. MR1207226
[3] J. E. Björnberg and S. O. Stefánsson. Random walk on random infinite looptrees. *J. Stat. Phys.* **158** (6) (2015) 1234–1261. MR3317412 https://doi.org/10.1007/s10955-014-1174-9
[4] J. Bouttier, P. Di Francesco and E. Guitter. Planar maps as labeled mobiles. *Electron. J. Combin.* **11** (1) (2004) 27. Research Paper 69. MR2097335
[5] X. Cai, C. Holmgren, S. Janson, T. Johansson and F. Skerman. Inversions in split trees and conditional Galton–Watson trees. *Combin. Probab. Comput.* **28** (3) (2019) 335–364. MR3933780 https://doi.org/10.1017/S0963548318000512
[6] P. Chassaing and G. Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probab. Theory Related Fields* **128** (2) (2004) 161–212. MR2031225 https://doi.org/10.1007/s00440-003-0297-8
[7] R. A. Doney. On the exact asymptotic behaviour of the distribution of ladder epochs. *Stochastic Process. Appl.* **12** (2) (1982) 203–214. MR0651904 https://doi.org/10.1016/0304-4149(82)90042-4
[8] T. Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Probab.* **31** (2) (2003) 996–1027. MR1964956 https://doi.org/10.1214/aop/1048516543
[9] T. Duquesne. An elementary proof of Hawkes's conjecture on Galton–Watson trees. *Electron. Commun. Probab.* **14** (2009) 151–164. MR2497323 https://doi.org/10.1214/ECP.v14-1454
[10] T. Duquesne and J.-F. Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque* **281** (2002) vi+147. MR1954248
[11] T. Duquesne and J.-F. Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields* **131** (4) (2005) 553–603. MR2147221 https://doi.org/10.1007/s00440-004-0385-4
[12] W. Feller. *An Introduction to Probability Theory and Its Applications. Vol. II*, 2nd edition. Wiley, New York–London–Sydney, 1971. MR0270403
[13] B. Gittenberger. A note on: "State spaces of the snake and its tour – convergence of the discrete snake" by J.-F. Marckert and A. Mokkadem. *J. Theoret. Probab.* **16** (4) (2003) 1063–1067. MR2033198 https://doi.org/10.1023/B:JOTP.0000012006.91251.e8
[14] S. Janson and P. Chassaing. The center of mass of the ISE and the Wiener index of trees. *Electron. Commun. Probab.* **9** (2004) 178–187. MR2108865 https://doi.org/10.1214/ECP.v9-1088
[15] S. Janson and J.-F. Marckert. Convergence of discrete snakes. *J. Theoret. Probab.* **18** (3) (2005) 615–647. MR2167644 https://doi.org/10.1007/s10959-005-7252-9
[16] S. Janson and S. Ö. Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.* **43** (3) (2015) 1045–1081. MR3342658 https://doi.org/10.1214/13-AOP871
[17] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** (4) (1986) 425–487. MR0871905
[18] I. Kortchemski. Invariance principles for Galton–Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.* **122** (9) (2012) 3126–3172. MR2946438 https://doi.org/10.1016/j.spa.2012.05.013
[19] I. Kortchemski. A simple proof of Duquesne's theorem on contour processes of conditioned Galton–Watson trees. In *Séminaire de Probabilités XLV* 537–558. *Lecture Notes in Math.* **2078**. Springer, Cham, 2013. MR3185928 https://doi.org/10.1007/978-3-319-00321-4_20
[20] I. Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees. *Probab. Theory Related Fields* **168** (1–2) (2017) 1–40. MR3651047 https://doi.org/10.1007/s00440-016-0704-6
[21] J.-F. Le Gall. *Spatial Branching Processes, Random Snakes and Partial Differential Equations*. *Lectures in Mathematics ETH Zürich*. Birkhäuser Verlag, Basel, 1999. MR1714707 https://doi.org/10.1007/978-3-0348-8683-3
[22] J.-F. Le Gall. Random trees and applications. *Probab. Surv.* **2** (2005) 245–311. MR2203728 https://doi.org/10.1214/154957805100000140
[23] J.-F. Le Gall and Y. Le Jan. Branching processes in Lévy processes: The exploration process. *Ann. Probab.* **26** (1) (1998) 213–252. MR1617047 https://doi.org/10.1214/aop/1022855417

[24] J.-F. Le Gall and G. Miermont. Scaling limits of random planar maps with large faces. *Ann. Probab.* **39** (1) (2011) 1–69. MR2778796 https://doi.org/10.1214/10-AOP549

[25] J.-F. Marckert. The lineage process in Galton–Watson trees and globally centered discrete snakes. *Ann. Appl. Probab.* **18** (1) (2008) 209–244. MR2380897 https://doi.org/10.1214/07-AAP450

[26] J.-F. Marckert and G. Miermont. Invariance principles for random bipartite planar maps. *Ann. Probab.* **35** (5) (2007) 1642–1705. MR2349571 https://doi.org/10.1214/009117906000000908

[27] J.-F. Marckert and A. Mokkadem. States spaces of the snake and its tour – convergence of the discrete snake. *J. Theoret. Probab.* **16** (4) (2003) 1015–1046. MR2033196 https://doi.org/10.1023/B:JOTP.0000012004.73051.f3

[28] C. Marzouk. Scaling limits of random bipartite planar maps with a prescribed degree sequence. *Random Structures Algorithms* **53** (3) (2018) 448–503. MR3854042 https://doi.org/10.1002/rsa.20773

[29] G. Miermont. Invariance principles for spatial multitype Galton–Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.* **44** (6) (2008) 1128–1161. MR2469338 https://doi.org/10.1214/07-AIHP157

[30] J. Neveu. Arbres et processus de Galton–Watson. *Ann. Inst. Henri Poincaré Probab. Stat.* **22** (2) (1986) 199–207. MR0850756

[31] V. V. Petrov. *Limit Theorems of Probability Theory*. *Sequences of Independent Random Variables*. *Oxford Studies in Probability.* **4**. The Clarendon Press, Oxford University Press, New York, 1995. Oxford Science Publications. MR1353441

[32] J. Pitman. Combinatorial stochastic processes. In *Lectures from the 32nd Summer School on Probability Theory Held in Saint-Flour*. *Lecture Notes in Mathematics* **1875**. Springer-Verlag, Berlin, 2006. July 7–24, 2002, with a foreword by Jean Picard. MR2245368

[33] E. Seneta. *Regularly Varying Functions*. *Lecture Notes in Mathematics* **508**. Springer-Verlag, Berlin–New York, 1976. MR0453936