



**HAL**  
open science

## Scaling limits of discrete snakes with stable branching

Cyril Marzouk

► **To cite this version:**

Cyril Marzouk. Scaling limits of discrete snakes with stable branching. Annales de l'Institut Henri Poincaré (B) Probabilités et Statistiques, 2020, 56 (1), pp.502-523. 10.1214/19-AIHP970. hal-01718881

**HAL Id: hal-01718881**

**<https://hal.science/hal-01718881>**

Submitted on 27 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# SCALING LIMITS OF DISCRETE SNAKES WITH STABLE BRANCHING

by

Cyril MARZOUK

---

**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  $\alpha$ -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.

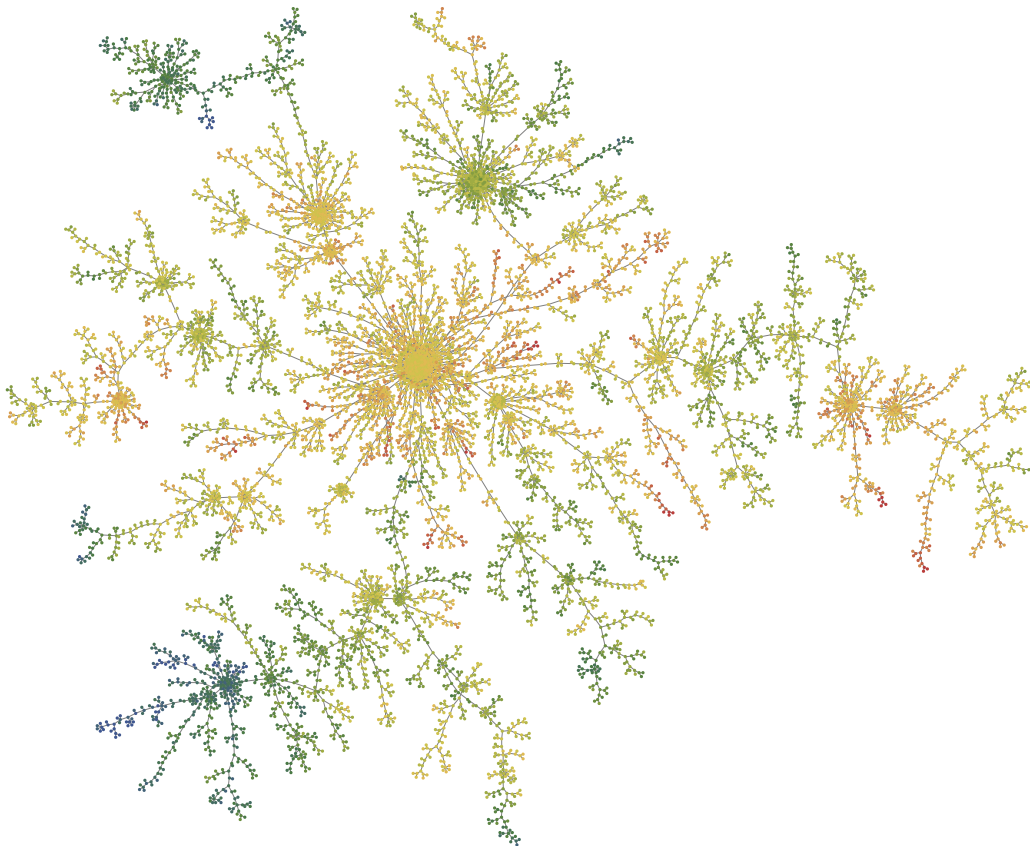


FIGURE 1. A spatial stable Lévy tree with index  $\alpha = 1,2$ ; colours indicate the spatial position of each vertex, from blue for the lowest (negative) ones to green then yellow and finally red for the highest ones. It corresponds to Theorem 1 with  $Y$  uniformly distributed on  $\{-1, 0, 1\}$ .

## 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 & Marckert [JM05] 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.

**1.1. Large Bienaymé–Galton–Watson trees.** — Throughout this work, we fix a probability measure  $\mu$  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  $\mu$  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  $\mu$  and conditioned to have  $n + 1$  vertices,<sup>(1)</sup> which is well defined for every  $n$  large enough from the aperiodicity of  $\mu$ . It is well known that for every  $a, b > 0$ ,  $T_n$  has the same law as a random *simply generated* tree with  $n + 1$  vertices, associated with the weight sequence  $(ab^k\mu(k))_{k \geq 0}$ , so there is almost no loss of generality to assume that  $\mu$  has mean 1.

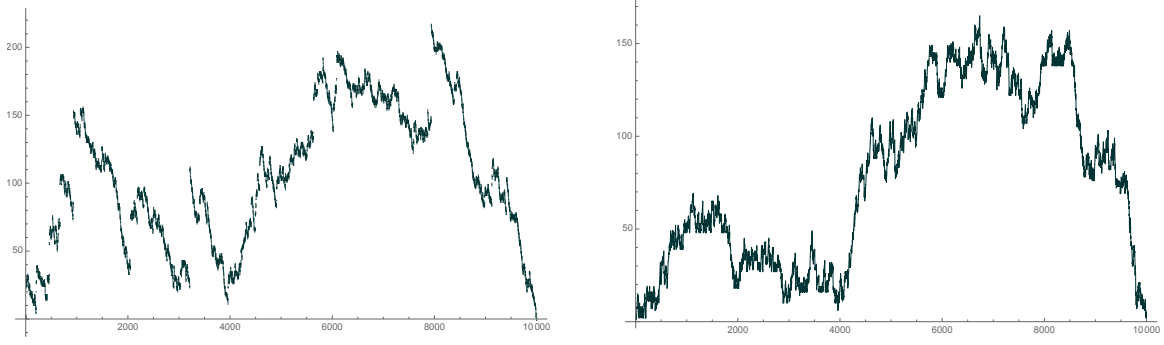
Finally, we assume that there exists  $\alpha \in (1, 2]$  such that  $\mu$  belongs to the domain of attraction of an  $\alpha$ -stable law, which means that there exists an increasing sequence  $(B_n)_{n \geq 1}$  such that if  $(\xi_n)_{n \geq 1}$  is a sequence of i.i.d. random variables sampled from  $\mu$ , then  $B_n^{-1}(\xi_1 + \dots + \xi_n - n)$  converges in distribution to a random variable  $X^{(\alpha)}$  whose law is given by the Laplace exponent  $\mathbb{E}[\exp(-\lambda X^{(\alpha)})] = \exp(\lambda^\alpha)$  for every  $\lambda \geq 0$ . Recall that  $n^{-1/\alpha}B_n$  is *slowly varying* at infinity and that if  $\mu$  has variance  $\sigma^2 \in (0, \infty)$ , then this falls in the case  $\alpha = 2$  and we may take  $B_n = \sqrt{n\sigma^2/2}$ .

It is well-known that a planar tree can be encoded by discrete paths; in Section 2.1, we recall the definition of the *Lukasiewicz path*  $W_n$ , the *height process*  $H_n$  and the *contour process*  $C_n$  associated with the tree  $T_n$ . Duquesne [Duq03] (see also Kortchemski [Kor13, Kor12]) has proved that

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

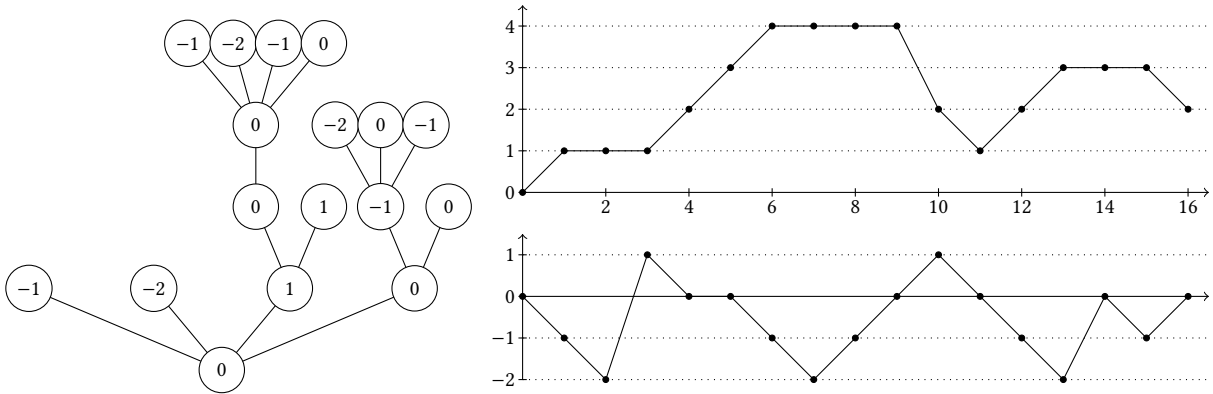
in the Skorokhod space  $\mathcal{D}([0, 1], \mathbf{R}^3)$ , where  $\mathcal{X}$  is the normalised excursion of the  $\alpha$ -stable Lévy process with no negative jump, whose value at time 1 has the law of  $X^{(\alpha)}$ , and  $\mathcal{H}$  is the associated continuous *height function*; see the references above for definitions and Figure 2 below 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 ‘continuous tree’ called the  $\alpha$ -stable Lévy tree  $\mathcal{T}_\alpha$  of Duquesne, Le Gall & Le Jan [Duq03, LGLJ98], which generalises the celebrated Brownian tree of Aldous [Ald93] in the case  $\alpha = 2$ . The convergence (1) implies that the tree  $T_n$ , viewed as a metric space by endowing its vertex-set by 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 & Le Gall [DLG05].

<sup>(1)</sup>Our results also hold when the tree is conditioned to have  $n + 1$  vertices with out-degree in a fixed set  $A \subset \mathbf{Z}_+$ , appealing to [Kor12].


 FIGURE 2. The Łukasiewicz path and the height process of  $T_{10000}$  with  $\alpha = 1,3$ .

**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  $\emptyset$  of the tree has position  $S_\emptyset = 0$  by convention so the spatial positions  $(S_u)_{u \in T}$  are entirely characterised by the *displacements*  $(Y_u)_{u \in T \setminus \{\emptyset\}}$ . 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  $(Y_u)_{u \in T \setminus \{\emptyset\}}$  are i.i.d. copies of  $Y$ .

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*  $H_n^{\text{sp}}$  and the *spatial contour process*  $C_n^{\text{sp}}$ . We consider scaling limits of these processes as  $n \rightarrow \infty$ . The most general such results are due to Janson & Marckert [JM05] 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 below for a formal definition and Figure 4 for two simulations.


 FIGURE 3. A spatial tree, its height process  $H$  on top and its spatial height process  $H^{\text{sp}}$  below.

**THEOREM 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  $\Sigma^2 := \mathbf{E}[Y^2] \in (0, \infty)$ , then the following convergence in distribution holds in the sense of finite-dimensional marginals:

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

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  $B_n = \sqrt{n\sigma^2/2}$ , the last assumption is equivalent to  $\mathbf{P}(|Y| \geq y) = o(y^{-4})$ , which is slightly weaker than  $\mathbf{E}[Y^4] < \infty$ ; otherwise, when the tree is less regular, 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. Duquesne & Le Gall [DLGo2, Section 2.5], or Marckert & Mokkadem [MMo3]. The same arguments apply for their spatial counterparts  $H_n^{\text{sp}}$  and  $C_n^{\text{sp}}$  so we concentrate only on the joint convergence of  $H_n$  and  $H_n^{\text{sp}}$ .

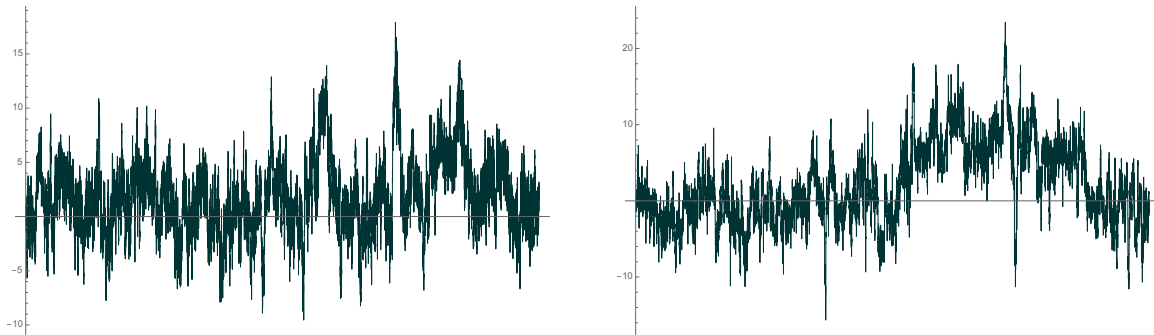


FIGURE 4. Two instances of the spatial height process  $\Sigma^{-1}H_n(n\cdot)$  associated with the height process of Figure 2: on the left,  $Y$  is uniformly distributed on  $[-1, 1]$  and on the right,  $Y$  is symmetric and such that  $\mathbf{P}(Y > y) = \frac{1}{2}(1+y)^{-10}$  so both satisfy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

Janson & Marckert [JM05] 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 below. Let us only mention the next result, which extends Theorem 8 in [JM05].

**THEOREM 2 (NON CENTRED SNAKES).** — Suppose that  $m := \mathbf{E}[Y] \neq 0$ . Then each process  $\frac{B_n}{n}H_n^{\text{sp}}(n\cdot)$  and  $\frac{B_n}{n}C_n^{\text{sp}}(2n\cdot)$  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  $\mathcal{C}([0, 1], \mathbf{R}^4)$

$$\left( \frac{B_n}{n}H_n(nt), \frac{B_n}{n}C_n(2nt), \frac{B_n}{n}H_n^{\text{sp}}(nt), \frac{B_n}{n}C_n^{\text{sp}}(2nt) \right)_{t \in [0, 1]} \xrightarrow[n \rightarrow \infty]{(d)} (\mathcal{H}_t, \mathcal{H}_t, m \cdot \mathcal{H}_t, m \cdot \mathcal{H}_t)_{t \in [0, 1]}.$$

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  $m > 0$ . Then 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  $H_n^{\text{sp}}$  and  $C_n^{\text{sp}}$  can be interpreted as the height and contour processes of the tree  $T_n$  with such random edge-lengths and Theorem 2 shows that this tree is close to the one obtained by assigning deterministic length  $m$  to each edge of  $T_n$ , and it converges towards  $m$  times the stable tree for the Gromov–Hausdorff topology, jointly with the original tree.

The main result of [JM05] has been used very recently by Cai *et al.* [CHJ<sup>+</sup>17] 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,  $\sigma$ , 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

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

This extends the classical definition of the number of inversions of a permutation, when the tree contains a single branch. We refer to [CHJ<sup>+</sup>17] 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.* [CHJ<sup>+</sup>17, Theorem 5] study the fluctuations of  $I(T_n)$  when  $T_n$  is a size-conditioned Bienaymé–Galton–Watson tree whose offspring distribution admits exponential moments. Their argument is based on the convergence of snakes from [JM05] and extends readily as follows thanks to Theorem 1.

**COROLLARY 1 (INVERSIONS ON TREES).** — *We have the convergences in distribution*

$$\frac{2B_n}{n^2} \mathbf{E}[I(T_n) \mid T_n] \xrightarrow[n \rightarrow \infty]{(d)} \int_0^1 \mathcal{H}_t dt, \quad \text{and} \quad \left( \frac{B_n}{12n^3} \right)^{1/2} (I(T_n) - \mathbf{E}[I(T_n) \mid T_n]) \xrightarrow[n \rightarrow \infty]{(d)} \int_0^1 \mathcal{S}_t dt.$$

Note that the scaling factors are respectively of order  $n^{-3/2}$  and  $n^{-5/4}$  in the finite-variance regime  $B_n = \sqrt{n\sigma^2/2}$ . When  $\alpha = 2$ , recall that  $\mathcal{H}$  is  $\sqrt{2}$  times the standard Brownian excursion; then the distribution of  $2 \int_0^1 \mathcal{H}_t dt$  is known as the *Airy distribution*; further, the random variable  $\int_0^1 \mathcal{S}_t dt$  is distributed as

$$\left( \int_{0 \leq s < t \leq 1} \min_{r \in [s, t]} \mathcal{H}_r ds dt \right)^{1/2} \cdot \mathcal{N},$$

where  $\mathcal{N}$  is standard Gaussian random variable independent of  $\mathcal{H}$ ; we refer to [JCo4] 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 between vertices. This requires a strong control on the geometry of the trees. Precisely, although the convergence (1) implies that the sequence  $(\frac{B_n}{n} H_n(n))_{n \geq 1}$  is tight in  $\mathcal{C}([0, 1], \mathbf{R})$ , we need the following more precise estimate on the geometry of the trees.

**LEMMA 1 (HÖLDER NORM OF THE HEIGHT PROCESS).** — *For every  $\gamma \in (0, (\alpha - 1)/\alpha)$ , it holds that*

$$\lim_{C \rightarrow \infty} \liminf_{n \rightarrow \infty} \mathbf{P} \left( \sup_{0 \leq s \neq t \leq 1} \frac{B_n}{n} \cdot \frac{|H_n(nt) - H_n(ns)|}{|t - s|^\gamma} \leq C \right) = 1.$$

By very different means, Gittenberger [Gito3] proved a similar statement for the contour function  $C_n$ , in the case  $\alpha = 2$ , when the offspring distribution admits finite exponential moments<sup>(2)</sup> and Janson & Marckert [JM05] built upon this result. Note that the maximal exponent  $(\alpha - 1)/\alpha$  corresponds to the maximal exponent for which the limit process  $\mathcal{H}$  is Hölder continuous, see [DLGo2, 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  $\mu$  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 & Schaeffer [CS04]. More general models of random maps are also related to spatial trees, via the *Bouttier–Di Francesco–Guitter* bijection [BDFGo4] and the *Janson–Stefánsson* bijection [JS15]; however, in this case, the displacements are neither independent nor identically distributed. Analogous convergences to Theorem 1 in this case have been proved by Marckert & Mokkadem [MM03] still for the uniform random trees, but for general displacements, under a ‘ $(8 + \varepsilon)$ -moment’ assumption; Gittenberger [Gito3] extended this result to the case where  $\mu$

<sup>(2)</sup>Even if the assumption is written as ‘finite variance’ in [Gito3], the proof actually requires exponential moments.

has finite exponential moments, and then Marckert & Miermont [MM07] reduced the assumptions on the displacements to a ‘ $(4 + \varepsilon)$ -moment’, see also Miermont [Mie08] for similar results on multi-type Bienaymé–Galton–Watson trees, Marckert [Mar08] for ‘globally centred’ displacements, and finally [Mar16] 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, it seems that the ‘ $(4 + \varepsilon)$ -moment’ assumption suffices in the case where  $\mu$  belongs to the domain of attraction of a Gaussian law to ensure the convergence towards  $(\mathcal{H}, \mathcal{S})$ . However in the  $\alpha$ -stable case with  $\alpha < 2$ , the limit may be different and depend more precisely on the displacements, see Le Gall & Miermont [LGM11], again for the very particular displacements associated with maps.

**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  $\mathcal{S}$ ; after recalling a few results on slowly varying functions and well-known results on Bienaymé–Galton–Watson trees, we prove Lemma 1. 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 [Don82].

In Section 3, we prove Theorem 1, Theorem 2 and Corollary 1. The proof of the two theorems follows the ideas of Janson & Marckert [JM05] which are quite general once we have Lemma 1. 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 on the convergence of similar to Theorem 1 when  $Y$  has heavier, regularly varying tails. Again, the proof scheme follows that of [JM05] but requires technical adaptation.

**Acknowledgment.** — I wish to thank Nicolas Curien for a stimulating discussion on the proof of Lemma 1 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.

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

## 2. Geometry of large Bienaymé–Galton–Watson trees

**2.1. Discrete and continuous snakes.** — We follow the notation of Neveu [Nev86] and view discrete trees as words. Let  $\mathbb{N} = \{1, 2, \dots\}$  be the set of all positive integers, set  $\mathbb{N}^0 = \{\emptyset\}$  and consider the set  $\mathbb{U} = \bigcup_{n \geq 0} \mathbb{N}^n$ . For every  $u = (u_1, \dots, u_n) \in \mathbb{U}$ , we denote by  $|u| = n$  the length of  $u$ ; if  $n \geq 1$ , we define its *prefix*  $pr(u) = (u_1, \dots, u_{n-1})$  and we let  $\chi_u = u_n$ ; for  $v = (v_1, \dots, v_m) \in \mathbb{U}$ , we let  $uv = (u_1, \dots, u_n, v_1, \dots, v_m) \in \mathbb{U}$  be the concatenation of  $u$  and  $v$ . We endow  $\mathbb{U}$  with the *lexicographical order*: given  $u, v \in \mathbb{U}$ , let  $w \in \mathbb{U}$  be their longest common prefix, that is  $u = w(u_1, \dots, u_n)$ ,  $v = w(v_1, \dots, v_m)$  and  $u_1 \neq v_1$ , then  $u < v$  if  $u_1 < v_1$ .

A (plane) *tree* is a non-empty, finite subset  $T \subset \mathbb{U}$  such that:

- (i)  $\emptyset \in T$ ;
- (ii) if  $u \in T$  with  $|u| \geq 1$ , then  $pr(u) \in T$ ;
- (iii) if  $u \in T$ , then there exists an integer  $k_u \geq 0$  such that  $ui \in T$  if and only if  $1 \leq i \leq k_u$ .

We shall view each vertex  $u$  of a tree  $T$  as an individual of a population for which  $T$  is the genealogical tree. The vertex  $\emptyset$  is called the *root* of the tree and for every  $u \in T$ ,  $k_u$  is the number of *children* of  $u$  (if  $k_u = 0$ , then  $u$  is called a *leaf*, otherwise,  $u$  is called an *internal vertex*) and  $u_1, \dots, u_{k_u}$  are these children from left to right,  $\chi_u$  is the relative position of  $u$  among its siblings,  $|u|$  is its *generation*,  $pr(u)$

is its *parent* and more generally, the vertices  $u, pr(u), pr \circ pr(u), \dots, pr^{|u|}(u) = \emptyset$  are its *ancestors*; the longest common prefix of two elements is their *last common ancestor*. We shall denote by  $\llbracket u, v \rrbracket$  the unique non-crossing path between  $u$  and  $v$ .

Fix a tree  $T$  with  $n + 1$  vertices, listed  $\emptyset = u_0 < u_1 < \dots < u_n$  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:  $c_0 = \emptyset$  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 5 for an illustration of these functions.

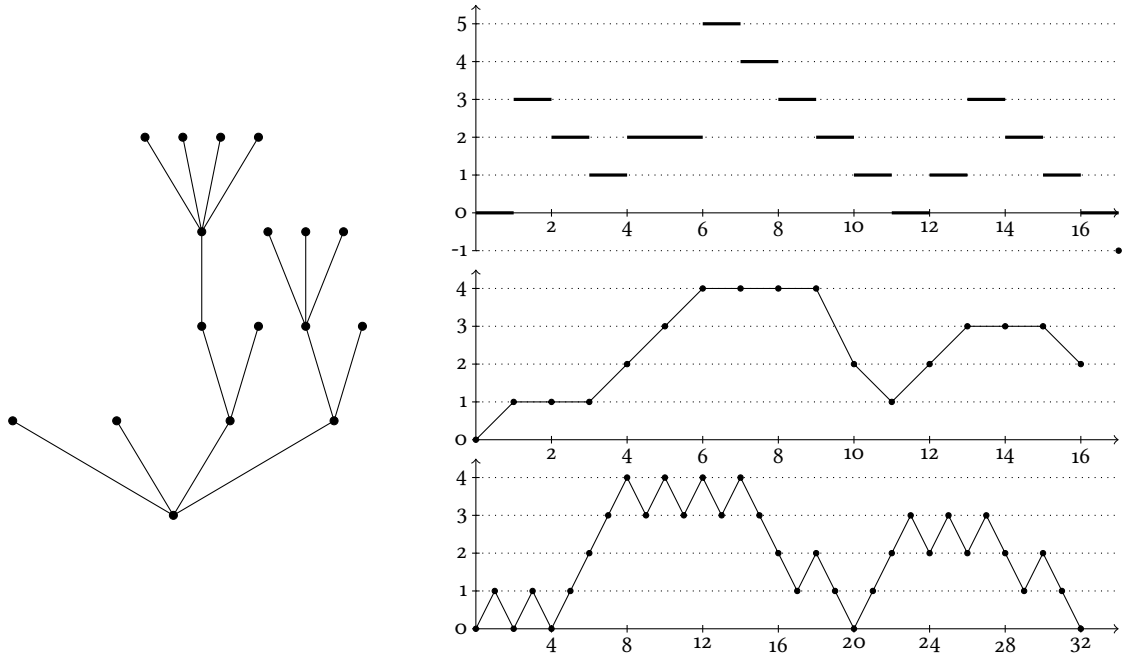


FIGURE 5. A tree on the left, and on the right, from top to bottom: its Łukasiewicz path  $W$ , its height process  $H$ , and its contour process  $C$ .

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  $S_\emptyset = 0$ . We encode these positions via the *spatial height* and *spatial contour* processes  $H^{\text{sp}}$  and  $C^{\text{sp}}$  respectively, defined by  $H^{\text{sp}}(j) = S_{u_j}$  for every  $0 \leq j \leq n$  and  $C^{\text{sp}}(j) = S_{c_j}$  for every  $0 \leq j \leq 2n$ , where  $n$  is the number of edges of the tree. See Figure 3 for an illustration of  $H^{\text{sp}}$ .



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 continuous objects we shall consider are the *stable Lévy tree* of Duquesne, Le Gall and Le Jan [Duq03, LGLJ98] which generalise Aldous' Brownian Continuum Random Tree [Ald93] in the case  $\alpha = 2$ . Recall that  $\mathcal{H} = (\mathcal{H}_t; t \in [0, 1])$  denotes the excursion of the *height process* associated with the  $\alpha$ -stable Lévy process with no negative jump; we shall not need the precise definition of this process but we refer the reader to [Duq03, Section 3.1 and 3.2]. For every  $s, t \in [0, 1]$ , set

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

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  $[0, 1] \rightarrow \mathcal{T}_{\alpha}$ ; 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  $0 \leq s \leq t \leq 1$ ,

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

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 continuous analog of spatial plane trees.

The *Brownian snake* driven by  $\mathcal{H}$  [LG99, DLGo2] is a path-valued process which associates with each time  $t \in [0, 1]$  the hole 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 than the entire snakes; nevertheless, Marckert & Mokkadem [MM03] proved a homeomorphism theorem which translates one into the other. Theorem 1 then implies the convergence of the whole snake towards the Brownian snake, see [JM05, Corollary 2].

It is known, see, e.g. [LG99, 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  $\mu$  is a probability measure on  $\mathbf{Z}_+$  satisfying a few assumptions given in the introduction. 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 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 & Miermont [LGM11, 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 domains of attraction.** — Let us present a few prerequisites on slowly varying functions. First, recall that a function  $l : [0, \infty) \rightarrow \mathbf{R}$  is said to be slowly varying (at infinity) when for every  $c > 0$ , it holds that

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

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

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

see e.g. Seneta's book [Sen76] 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  $\mu$  belongs to the domain of attraction of a stable law with index  $\alpha \in (1, 2]$ , there exists two slowly varying functions  $L$  and  $L_1$  such that for every  $n \geq 1$ ,

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

The two functions are related by

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

see Feller [Fel71, Chapter XVII, Equation 5.16]. We shall need a third slowly varying function  $L^*$  (see Doney [Don82, Equation 2.2]), defined uniquely up to asymptotic equivalence as the *conjugate* of  $1/L$  by the following equivalent asymptotic relations:

$$L(x)^{-1/\alpha} L^*(x^\alpha L(x)^{-1}) \xrightarrow{x \rightarrow \infty} 1 \quad \text{and} \quad L^*(x)^{-\alpha} L(x^{1/\alpha} L^*(x)) \xrightarrow{x \rightarrow \infty} 1.$$

We refer to [Sen76, 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.i.d. copies of  $X$ , then  $B_n^{-1} S(n)$  converges in distribution to some  $\alpha$ -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 [Don82], but it extends to the case  $\alpha = 2$ , see the remark between Equation 2.2 and Theorem 1 in [Don82]: the function  $L$  there is  $1/L$  here. By comparing the preceding asymptotic relations between  $L$  and  $L^*$  to [Kor17, Equation 7], one gets precisely

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

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

Doney [Don82, 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.i.d. and according to [Don82, Theorem 1], it holds that

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

with a constant  $C > 0$  which shall not be important here. By a Tauberian theorem, see e.g. [Fel71, Chapter XVII, Theorem 5.5] it follows that

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

for some constant  $C_\alpha > 0$ , where we recall that  $\ell$  is a slowly varying function at infinity such that  $B_n = n^{1/\alpha} \ell(n)$ , so, taking  $\lambda = 1/N$  with  $N \in \mathbf{N}$ , we obtain

$$(3) \quad 1 - \mathbf{E} \left[ e^{-\tau_1/N} \right] \underset{N \rightarrow \infty}{\sim} C_\alpha \cdot N^{-1} \cdot B_N.$$

**2.2.2. Łukasiewicz paths and random walks.** — Recall that  $S = (S(i))_{i \geq 0}$  denotes a random walk started from 0 with steps  $(X_i)_{i \geq 1}$  given by i.i.d. random variables with law  $\mathbf{P}(X_1 = k) = \mu(k+1)$  for every  $k \geq -1$ . Let  $(X_n(i))_{1 \leq i \leq n+1}$  have the law of  $(X_i)_{1 \leq i \leq n+1}$  conditioned to satisfy  $X_1 + \dots + X_{n+2} = -1$  and let  $S_n = (S_n(i))_{0 \leq i \leq n+2}$  be the associated path. For every  $1 \leq j \leq n+2$ , put

$$X_n^{(j)}(k) = X_n(k + j \bmod n + 2), \quad 1 \leq k \leq n + 2.$$

We say that  $X_n^{(j)}$  is the  $j$ -th cyclic shift of  $X_n$ . Obviously, for every  $1 \leq j \leq n+2$ , we have  $X_{n,1}^{(j)} + \dots + X_{n,n+2}^{(j)} = -1$ , but it turns out there is a unique  $j$  such that  $X_{n,1}^{(j)} + \dots + X_{n,k}^{(j)} \geq 0$  for every  $1 \leq k \leq n$ . This index is the least time at which the path  $S_n$  achieves its minimum overall value:

$$(4) \quad j = \inf \left\{ 1 \leq k \leq n + 2 : S_n(k) = \inf_{1 \leq i \leq n+2} S_n(i) \right\}.$$

Moreover, it is a standard fact that this time  $j$  has the uniform distribution on  $\{1, \dots, n+2\}$  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. [Pit06, 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 also invariant in law under time and space *reversal*, so by combining these observations, we obtain the following property: let  $(X_n(i))_{1 \leq i \leq n+2}$  be the increments of  $S_n$  and for a given  $1 \leq i \leq n+2$ , let  $\widehat{X}_n^{(i)}(k) = X_n(i+1-k)$  for  $1 \leq k \leq i$  and  $\widehat{X}_n^{(i)}(k) = X_n(n+i+2-k)$  for  $i+1 \leq k \leq n+2$ ; 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 event  $A_n$  depending only on the first  $\lceil n/2 \rceil$  steps of the path, we have

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

see e.g. [Kor17], 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  $\emptyset = u_0 < u_1 < \dots < u_n$ . It is well-known that the processes  $H_n$  and  $W_n$  are related as follows (see e.g. Le Gall & Le Jan [LGLJ98]): for every  $0 \leq j \leq n$ ,

$$H_n(j) = \# \left\{ k \in \{0, \dots, j-1\} : W_n(k) \leq \inf_{[k+1, j]} W_n \right\}.$$

Indeed, for  $k < j$ , we have  $W_n(k) \leq \inf_{[k+1, j]} W_n$  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  $W_n(i) - \min_{i \leq k \leq j} W_n(k)$  counts the number of vertices branching-off of the ancestral line  $\llbracket u_{\overline{ij}}, u_i \rrbracket$  which lie between  $u_i$  and  $u_j$ , i.e. all the vertices visited between time  $i$  and  $j$  whose parent belongs to  $\llbracket u_{\overline{ij}}, u_i \rrbracket$ . 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 ancestor 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  $\llbracket u_{\overline{ij}}, u_i \rrbracket$ , until the last one which is  $u_{j'}$ . We conclude that

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

It follows that the length of the path  $\llbracket u_{j'}, u_j \rrbracket$  is

$$\begin{aligned} H_n(j) - H_n(j') &= \# \left\{ k \in \{j', \dots, j\} : W_n(k) = \min_{k \leq l \leq j} W_n(l) \right\} \\ &= \# \left\{ k \in \{i, \dots, j\} : W_n(k) = \min_{k \leq l \leq j} W_n(l) \right\}. \end{aligned}$$

We can now prove Lemma 1 appealing to the preceding subsections.

**2.3. Proof of Lemma 1.** — 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  $c_1, c_2 > 0$  such that for every  $n$  large enough, every  $0 \leq s \leq t \leq 1$ , and every  $x \geq 0$ , we have

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

and

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

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, 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). Note that we may, and shall, restrict to times  $0 \leq s < t \leq 1$  such that  $t - s \leq 1/2$  and both  $ns$  and  $nt$  are integers.

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

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

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, this 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.i.d. and let  $\tau = \tau(1)$ . The exponential Markov inequality shows that the preceding probability is bounded by

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

From (3), we get that

$$\begin{aligned} \left[ x \frac{n}{B_n} |t-s|^\gamma \right] \ln \left( 1 - \left( 1 - \mathbf{E} \left[ \exp \left( - \frac{\tau}{n|t-s|} \right) \right] \right) \right) &= \left[ x \frac{n}{B_n} |t-s|^\gamma \right] \ln \left( 1 - C_\alpha \frac{B_n |t-s|}{n|t-s|} (1 + o(1)) \right) \\ &= -x \frac{n}{B_n} |t-s|^\gamma C_\alpha \frac{B_n |t-s|}{n|t-s|} (1 + o(1)) \\ &= -C_\alpha x \frac{(n|t-s|)^{-1/\alpha} B_n |t-s|}{n^{-1/\alpha} B_n} |t-s|^{\gamma-1+\frac{1}{\alpha}} (1 + o(1)), \end{aligned}$$

where the  $o(1)$  does not depend on  $s$  and  $t$ . 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. [BS15, Lemma 4.2] or [Kor17, Equation 9]) asserts that there exists a constant  $c$ , depending on  $\varepsilon$  (and so on  $\gamma$ ), such that for every  $n$  large enough,

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

We conclude that

$$\mathbf{P} \left( \# \left\{ k \in \{ns, \dots, nt\} : S_n(k) = \min_{k \leq l \leq nt} S_n(l) \right\} \geq x \frac{n}{B_n} |t-s|^\gamma \right) \leq C \cdot \exp(1 - C_\alpha c x (1 + o(1))),$$

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

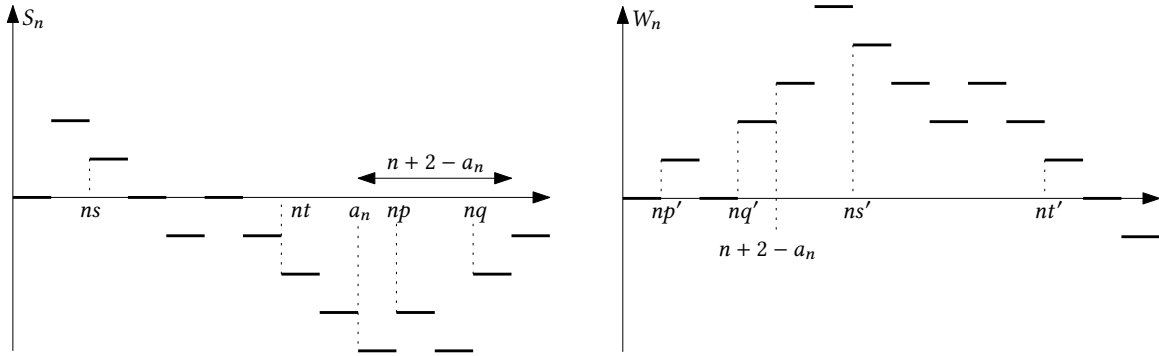


FIGURE 6. A bridge  $S_n$  and its shifted excursion  $W_n$ ; the times  $s, t$  fall into the first case, whereas  $p, q$  fall into the second case and  $s, p$  into the third case.

We next prove (7) by relating  $W_n$  and  $S_n$ , as depicted in Figure 6. Recall that these paths have length  $n+2$ . 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  $ns < nt \leq a_n$ , in which case  $ns' = ns + (n+2-a_n) < nt + (n+2-a_n) = nt'$ ;
- (ii) Either  $a_n \leq ns < nt$ , in which case  $ns' = ns - a_n < nt - a_n = nt'$ ;
- (iii) Or  $ns < a_n < nt$ , in which case  $nt' = nt - a_n < ns + (n+2-a_n) = ns'$ .

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

$$\mathbf{P} \left( \# \left\{ k \in \{ns', \dots, nt'\} : W_n(k) = \min_{k \leq l \leq nt'} W_n(l) \right\} \geq x \frac{n}{B_n} |t' - s'|^Y \right) \leq c_1 e^{-c_2 x}.$$

In the third case above, we have to be a little more careful; by cutting  $W_n$  at time  $n + 2 - a_n$  (which corresponds to  $n + 2$  for  $S_n$ ), we observe that

$$\begin{aligned} & \# \left\{ k \in \{nt', \dots, ns'\} : W_n(k) = \min_{k \leq l \leq ns'} W_n(l) \right\} \\ & \leq \# \left\{ k \in \{nt', \dots, n + 2 - a_n\} : W_n(k) = \min_{k \leq l \leq n + 2 - a_n} W_n(l) \right\} \\ & \quad + \# \left\{ k \in \{n + 2 - a_n, \dots, ns'\} : W_n(k) = \min_{k \leq l \leq ns'} W_n(l) \right\} \\ & = \# \left\{ k \in \{nt, \dots, n + 2\} : S_n(k) = \min_{k \leq l \leq n + 2} S_n(l) \right\} + \# \left\{ k \in \{0, \dots, ns\} : S_n(k) = \min_{k \leq l \leq ns} S_n(l) \right\}. \end{aligned}$$

A union bound then yields

$$\begin{aligned} & \mathbf{P} \left( \# \left\{ k \in \{nt', \dots, ns'\} : W_n(k) = \min_{k \leq l \leq nt'} W_n(l) \right\} \geq x \frac{n}{B_n} |t' - s'|^Y \right) \\ & \leq \mathbf{P} \left( \# \left\{ k \in \{nt, \dots, n + 2\} : S_n(k) = \min_{k \leq l \leq n + 2} S_n(l) \right\} \geq \frac{x}{2} \frac{n}{B_n} |t' - s'|^Y \right) \\ & \quad + \mathbf{P} \left( \# \left\{ k \in \{0, \dots, ns\} : S_n(k) = \min_{k \leq l \leq ns} S_n(l) \right\} \geq \frac{x}{2} \frac{n}{B_n} |t' - s'|^Y \right) \\ & \leq \mathbf{P} \left( \# \left\{ k \in \{nt, \dots, n + 2\} : S_n(k) = \min_{k \leq l \leq n + 2} S_n(l) \right\} \geq \frac{x}{2} \frac{n}{B_n} |1 - t|^Y \right) \\ & \quad + \mathbf{P} \left( \# \left\{ k \in \{0, \dots, ns\} : S_n(k) = \min_{k \leq l \leq ns} S_n(l) \right\} \geq \frac{x}{2} \frac{n}{B_n} |s|^Y \right) \\ & \leq c_1 e^{-c_2 x}, \end{aligned}$$

which concludes the proof of (7).  $\square$

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 7. Put  $i = ns$  and  $j = nt$ . Let us denote by  $\tilde{T}_n$  the image of  $T_n$  by the following two operations: first exchange the subtrees of the progeny of the  $i$ -th and the  $j$ -th vertices of  $T_n$  and then take the mirror image of the whole tree, the resulting tree is  $\tilde{T}_n$ . Observe that  $T_n$  and  $\tilde{T}_n$  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  $\tilde{T}_n$  correspond to the  $i$ -th and the  $j$ -th 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  $i$ -th vertex, then all the vertices that lie strictly between the two ancestral lines between the  $i$ -th and  $j$ -th 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  $\tilde{T}_n$ , the Łukasiewicz path  $\tilde{W}_n$  visits all the progeny of the  $\tilde{j}$ -th vertex, which is the same as that of the  $i$ -th 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  $j$ -th vertex of  $T_n$  and not those of its  $i$ -th vertex, whereas  $\tilde{W}_n$  visits the ancestors of the  $i$ -th vertex of  $T_n$  and not those of its  $j$ -th 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  $\tilde{W}_n$  (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.

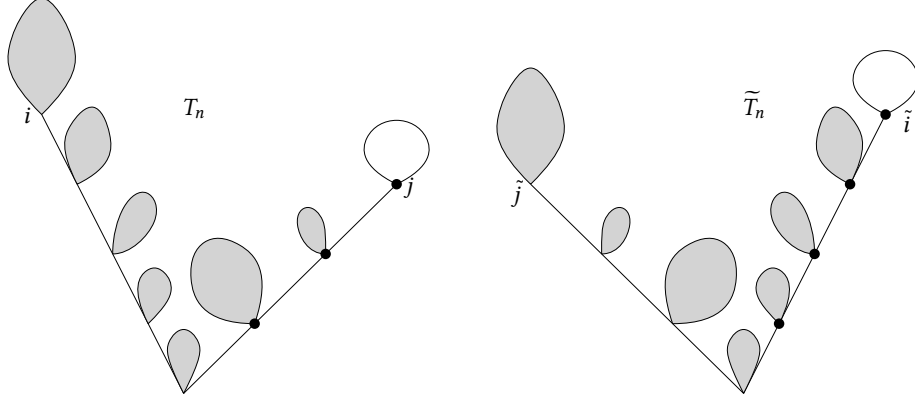


FIGURE 7. On the left: a portion of the tree  $T_n$  and two vertices  $u_i$  and  $u_j$ ; on the right: the ‘mirror’ images  $\tilde{T}_n$ ,  $\tilde{i}$  and  $\tilde{j}$ . The vertices visited by  $W_n$  (resp.  $\tilde{W}_n$ ) between time  $i$  and  $j$  (resp.  $\tilde{j}$  and  $\tilde{i}$ ) are those black dots on the right branch as well as all the vertices strictly inside the grey trees.

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

LEMMA 2. — Let  $C = \frac{10}{\mu(0)^2}$ , then the probability of the event

$$\left\{ \frac{\#\{w \in \llbracket u, v \rrbracket : \chi_w = 1\}}{\#\llbracket u, v \rrbracket} \leq 1 - \frac{\mu(0)}{2} \text{ for every } u, v \in T_n \text{ such that } u \in \llbracket \emptyset, v \rrbracket \text{ and } \#\llbracket u, v \rrbracket > C \ln n \right\}$$

tends to 1 as  $n \rightarrow \infty$ .

We can now finish the proof of Lemma 1.

*Proof of (6).* — From the preceding lemma, we deduce that there exists some  $p \in (0, 1)$  such that with high probability, on all ancestral paths in  $T_n$  of length at least logarithmic, there is a proportion at least  $p$  of individuals which are not the first child of their parent; symmetrically, there is the same proportion of individuals which are not the last child of their parent. Consequently, the length of such a path, multiplied by  $p$ , is bounded below by the number of vertices whose parent belongs to this path, and which themselves lie strictly to its right. With the notation of the discussion preceding the lemma, on the event described in this lemma, the lexicographical distance in  $\tilde{T}_n$  between the images of the  $i$ -th and  $j$ -th vertex of  $T_n$  is

$$\begin{aligned} |\tilde{j} - \tilde{i}| &= |i - j| - |H_n(j) - \inf_{i \leq k \leq j} H_n(k)| + |H_n(i) - \inf_{i \leq k \leq j} H_n(k)| \\ &\leq |i - j| + p^{-1}|i - j|, \end{aligned}$$

where the second (very rough) bound holds only if  $|H_n(i) - \inf_{i \leq k \leq j} H_n(k)| > C \ln n$ , with  $C$  as in Lemma 2. Note that (6) is trivial otherwise since  $|t - s| \geq \frac{1}{n}$  as we restricted to integer times, so  $x \frac{n}{B_n} |t - s|^\gamma \geq x \frac{n^{1-\gamma}}{B_n}$  which tends to infinity like a power of  $n$ . We then conclude from the bound (5) applied to the ‘mirror’ Łukasiewicz path  $\tilde{W}_n$ .  $\square$

It remains to prove Lemma 2. A similar statement was proved in [Mar16, 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.* — Fix  $\varepsilon > 0$  and let  $T$  be an unconditioned Bienaymé–Galton–Watson tree with offspring distribution  $\mu$ . Define the set

$$A_T = \left\{ u, v \in T : u \in \llbracket \emptyset, v \rrbracket \text{ and } \# \llbracket u, v \rrbracket > C \ln n \text{ and } \frac{\#\{w \in \llbracket u, v \rrbracket : \chi_w = 1\}}{\# \llbracket u, v \rrbracket} > 1 - \frac{\mu(0)}{2} \right\}.$$

Note that the maximal height of a tree cannot exceed its total size. Then by sampling  $v$  uniformly at random in  $T$ , we obtain

$$(8) \quad \begin{aligned} \mathbf{P}(\exists(u, v) \in A_T \text{ and } \#T = n + 1) &= \frac{1}{n + 1} \mathbf{E} \left[ \sum_{h=1}^{n+1} \sum_{\substack{v \in T \\ |v|=h}} \mathbb{1}_{\{\exists u \in T \text{ such that } (u, v) \in A_T\}} \mathbb{1}_{\{\#T = n+1\}} \right] \\ &\leq \frac{1}{n + 1} \sum_{h=1}^{n+1} \mathbf{E} \left[ \sum_{\substack{v \in T \\ |v|=h}} \mathbb{1}_{\{\exists u \in T \text{ such that } (u, v) \in A_T\}} \right]. \end{aligned}$$

We then use a spinal decomposition due to Duquesne [Duq09, Equation 24] which results in an absolute continuity relation between the tree  $T$  and the tree  $T_\infty$  ‘conditioned to survive’, which is the infinite tree which arise as the *local limit* of  $T_n$ . It was introduced by Kesten [Kes86] and the most general results on such convergences are due to Abraham & Delmas [AD14]. The tree  $T_\infty$  contains a unique infinite simple path called the *spine*, starting from the root, and the vertices which belong to this spine reproduce according to the *size-biased* law  $(k\mu(k))_{k \geq 1}$ , whereas the other vertices reproduce according to  $\mu$ , and all the vertices reproduce independently. For a tree  $\tau$  and a vertex  $v \in \tau$ , let  $\theta_v(\tau)$  be the subtree consisting of  $v$  and all its progeny, and let  $\text{Cut}_v(\tau) = \{v\} \cup (\tau \setminus \theta_v(\tau))$  be its complement (note that  $v$  belongs to the two parts). Then for every non-negative measurable functions  $G_1, G_2$ , for every  $h \geq 0$ , we have

$$\mathbf{E} \left[ \sum_{\substack{v \in T \\ |v|=h}} G_1(\text{Cut}_v(T), v) \cdot G_2(\theta_v(T)) \right] = \mathbf{E} \left[ G_1(\text{Cut}_{v_h^*}(T_\infty), v_h^*) \right] \cdot \mathbf{E} [G_2(T)],$$

where  $v_h^*$  is the only vertex on the spine of  $T_\infty$  at height  $h$ . Then the expectation in (8) equals

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

Now recall that on the spine, the vertices reproduce according to the size-biased law  $(i\mu(i))_{i \geq 1}$ , and furthermore, conditional on the number of children of its parent, the position of a vertex amongst its sibling is uniformly chosen, this means that for every vertex  $w$  on the spine, the probability  $\mathbf{P}(\chi_w = 1)$  equals  $\sum_{i \geq 1} (i\mu(i))/i = 1 - \mu(0)$  and these events are independent. Therefore, if  $\text{Bin}(N, p)$  denotes a random variable with the binomial law with parameters  $N$  and  $p$ , then the preceding probability is bounded by

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



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

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

It is well-known that  $nB_n\mathbf{P}(\#T = n + 1) \rightarrow p_1(0)$  as  $n \rightarrow \infty$ , where  $p_1$  is the density of the stable random variable  $X^{(\alpha)}$  from the introduction; this follows e.g. from the fact that the event  $\mathbf{P}(\#T = n + 1)$  is the probability that the random walk  $S$  first hits  $-1$  at time  $n + 2$ , which equals by cyclic shift  $(n + 1)^{-1}$  times the probability that  $S_{n+2} = -1$  and the asymptotic behaviour of this probability is dictated by the local limit theorem, see e.g. [Kor13, Lemma 1]. We conclude that for  $n$  large enough

$$\begin{aligned} \mathbf{P}(\exists(u, v) \in A_T \mid \#T = n + 1) &\leq \frac{n^2 B_n}{2p_1(0)} e^{-C \ln n \mu(0)^2/2} (1 + o(1)) \\ &\leq n^3 e^{-C \ln n \mu(0)^2/2}, \end{aligned}$$

which converges to 0 from our choice of  $C$ .  $\square$

### 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  $H_n^{\text{sp}}$ .

**3.1. Proof of Theorem 1 for centred snakes.** — Let us first focus on the case  $\mathbf{E}[Y] = 0$ ; we aim at showing the convergence in distribution in  $\mathcal{C}([0, 1], \mathbf{R}^2)$

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

where  $\Sigma^2 := \mathbf{E}[Y^2] \in (0, \infty)$ ; 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  $(B_n/n)^{1/2} H_n^{\text{sp}}(n \cdot)$ . The idea is to apply Kolmogorov's criterion but our assumption does not give us sufficiently large moments. We therefore adapt the argument from [JM05] 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  $(Y_i)_{i \geq 1}$  be i.i.d. copies of  $Y$ , independent of  $T_n$ , then 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

$$\mathbf{P} \left( \bigcup_{i=1}^{n+1-\lambda(T_n)} \{|Y_i| \geq (n/B_n)^{1/2}\} \mid T_n \right) = 1 - \mathbf{P} \left( \bigcap_{i=1}^{n+1-\lambda(T_n)} \{|Y_i| < (n/B_n)^{1/2}\} \mid T_n \right) \geq 1 - \left( 1 - \frac{\delta}{n} \right)^{n+1-\lambda(T_n)},$$

where  $\lambda(T_n)$  denotes the number of leaves of the tree. Since  $\limsup_{n \rightarrow \infty} \lambda(T_n)/n < 1$  with high probability, and indeed  $\lambda(T_n)/n$  converges to  $\mu(0)$ , see e.g. [Kor12, 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  $n \in \mathbb{N}$ , there exists  $0 \leq i < n$  such that  $(n/B_n)^{1/2} |H_n^{\text{SP}}(i+1) - H_n^{\text{SP}}(i)| \geq 1$  so the sequence of continuous processes  $((n/B_n)^{1/2} H_n^{\text{SP}}(n))_{n \geq 1}$  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  $\varepsilon$  and  $\delta$  small. The idea is to take into account separately the large increments. For every vertex  $u \in T_n$ , let  $Y'_u = Y_u \mathbb{1}_{\{|Y_u| \leq b_n\}}$  and  $Y''_u = Y_u \mathbb{1}_{\{|Y_u| > b_n\}}$ , define then  $H_n^{\text{SP}'}$  and  $H_n^{\text{SP}''}$  the spatial processes in which the increments  $Y_u$  are replaced by  $Y'_u$  and  $Y''_u$  respectively, so  $H_n^{\text{SP}} = H_n^{\text{SP}'} + H_n^{\text{SP}''}$ .

**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  $|Y_v| > b_n$ . Then  $\mathbf{P}(E_n | T_n) \leq \Lambda(T_n) \mathbf{P}(|Y| > b_n)^2$ , where  $\Lambda(T_n) = \sum_{u \in T_n} |u|$  is called the *total path length* of  $T_n$ . It is a simple matter to prove the following well-known integral representation: if  $C_n$  denotes the contour process of  $T_n$ , then

$$\Lambda(T_n) = \frac{n}{2} + \frac{1}{2} \int_0^{2n} C_n(t) dt = \frac{n^2}{B_n} \left( \frac{B_n}{2n} + \int_0^1 \frac{B_n}{n} C_n(2nt) dt \right).$$

We deduce that if  $T_n$  satisfies **1**, then we have the convergence in distribution:

$$(9) \quad \frac{B_n}{n^2} \Lambda(T_n) \xrightarrow[n \rightarrow \infty]{(d)} \int_0^1 \mathcal{H}_t dt.$$

We then write for every  $K > 0$ ,

$$\limsup_{n \rightarrow \infty} \mathbf{P}(E_n) \leq \limsup_{n \rightarrow \infty} \mathbf{P}(\Lambda(T_n) > Kn^2/B_n) + K \limsup_{n \rightarrow \infty} \frac{n^2}{B_n} \mathbf{P}(|Y| > b_n)^2.$$

The first term on the right tends 0 when  $K \rightarrow \infty$ , and as for the second term, from our choice of  $b_n$ , we have for every  $\delta > 0$ ,

$$\frac{n^2}{B_n} \mathbf{P}(|Y| > b_n)^2 \ll \left( \frac{n^2}{B_n} \right)^{1 - (\frac{\alpha-1}{4\alpha} + \varepsilon) (\frac{4\alpha}{\alpha(1+\delta)-1})},$$

and the exponent is negative for  $\delta$  sufficiently small. Now on the 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  $\max_{0 \leq i \leq n} |H_n^{\text{SP}''}(i)|$  simply equals  $\max_{u \in T_n} |Y''_u|$  and so for every  $\delta > 0$ ,

$$\begin{aligned} \mathbf{P} \left( \left\{ \max_{0 \leq t \leq 1} |H_n^{\text{SP}''}(2nt)| > \delta(n/B_n)^{1/2} \right\} \cap E_n^c \right) &\leq \mathbf{P} \left( \max_{u \in T_n} |Y_u| > \delta(n/B_n)^{1/2} \right) \\ &\leq n \mathbf{P}(|Y| > (n/B_n)^{1/2}), \end{aligned}$$

which converges to 0 as  $n \rightarrow \infty$ . Thus  $(B_n/n)^{1/2} H_n^{\text{SP}''}(n \cdot)$  converges to 0 so it only remains to prove that  $(B_n/n)^{1/2} H_n^{\text{SP}'}(n \cdot)$  is tight.

**3.1.4. Average contribution of small jumps.** — The process  $(B_n/n)^{1/2} H_n^{\text{SP}'}(n \cdot)$  is simpler to analyse than  $(B_n/n)^{1/2} H_n^{\text{SP}}(n \cdot)$  since all the increments are bounded in absolute value by  $(n/B_n)^{-\varepsilon}$ . Note nonetheless that  $H_n^{\text{SP}'}$  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  $\mathbf{E}[H_n^{\text{SP}'}(n \cdot) | T_n] = m_n H_n(n \cdot)$ . Recall that  $(B_n/n) H_n(n \cdot)$  converges in distribution (to  $\mathcal{H}$ ); from the tail behaviour of  $Y$  we get:

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

Note that for  $\varepsilon$  and  $\delta$  sufficiently small, we have

$$\left(\frac{\alpha-1}{4\alpha} + \varepsilon\right) \left(1 - \frac{2\alpha}{\alpha(1+\delta)-1}\right) > -\frac{\alpha+1}{4\alpha},$$

and this bound gets tighter as  $\varepsilon$  and  $\delta$  get closer to 0. Then, since

$$\frac{\alpha+1}{4\alpha} < \frac{1}{2} < \frac{\alpha+1}{2\alpha},$$

we conclude that for  $\varepsilon$  and  $\delta$  sufficiently small,

$$\left(\frac{n}{B_n}\right)^{1/2} \left(b_n^{1-\frac{2\alpha}{\alpha(1+\delta)-1}}\right) = n^{2(\frac{\alpha-1}{4\alpha}+\varepsilon)(1-\frac{2\alpha}{\alpha(1+\delta)-1})+\frac{1}{2}} \cdot B_n^{-(\frac{\alpha-1}{4\alpha}+\varepsilon)(1-\frac{2\alpha}{\alpha(1+\delta)-1})-\frac{1}{2}}$$

converges to 0 as  $n \rightarrow \infty$  since both exponents are negative. Therefore the process  $(B_n/n)^{1/2} \mathbf{E}[H_n^{\text{SP}'}(n\cdot) | T_n] = (n/B_n)^{1/2} m_n \cdot (B_n/n) H_n(n\cdot)$  converges in probability to 0 and we focus for the rest of the proof on the centred process  $H_n^{\text{SP}'}(n\cdot) = H_n^{\text{SP}'}(n\cdot) - \mathbf{E}[H_n^{\text{SP}'}(n\cdot) | T_n]$ .

**3.1.5. Re-centred small jumps are tight.** — It only remains to prove that  $(B_n/n)^{1/2} H_n^{\text{SP}'}(n\cdot)$  is tight. Fix  $\gamma \in (0, (\alpha-1)/\alpha)$ , let  $\eta > 0$  arbitrary and let us fix  $C > 0$  such that, according to Lemma 1, for every  $n$  large enough,

$$\mathbf{P}\left(\sup_{0 \leq s \neq t \leq 1} \frac{B_n \cdot |H_n(nt) - H_n(ns)|}{n \cdot |t - s|^\gamma} \leq C\right) \geq 1 - \eta.$$

We shall denote by  $A_n$  the event in the preceding probability. Our aim is to apply Kolmogorov's tightness criterion to  $(B_n/n)^{1/2} H_n^{\text{SP}'}(n\cdot)$  on the event  $A_n$ . Let us enumerate the vertices of  $T_n$  in lexicographical order as  $u_0 < u_1 < \dots < u_n$ . Fix  $0 \leq s < t \leq 1$  such that  $ns$  and  $nt$  are both integers. Then  $H_n^{\text{SP}'}(nt) - H_n^{\text{SP}'}(ns)$  is the sum of  $\#\llbracket u_{ns}, u_{nt} \rrbracket$  i.i.d. random variables distributed as  $Y' = Y' - \mathbf{E}[Y']$ . 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

$$\#\llbracket u_{ns}, u_{nt} \rrbracket \leq 2 + H_n(ns) + H_n(nt) - 2H_n(nr).$$

On the event  $A_n$ , the right-hand side is bounded by

$$C \frac{n}{B_n} (|t-r|^\gamma + |r-s|^\gamma) \leq 2C \frac{n}{B_n} |t-s|^\gamma.$$

Fix any  $q \geq 2$  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  $\mathbf{E}[|Y'|^2] = \text{Var}(Y') \leq \mathbf{E}[|Y'|^2] \leq \mathbf{E}[|Y|^2] < \infty$  and  $|Y'|^q \leq 2^q (n^2/B_n)^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)}$ . Appealing to [Pet95, Theorem 2.9] (sometimes called the Rosenthal inequality), we obtain

$$\begin{aligned} \mathbf{E}\left[\left(\frac{|H_n^{\text{SP}'}(nt) - H_n^{\text{SP}'}(ns)|}{(n/B_n)^{1/2}}\right)^q \middle| A_n\right] &\leq C_q \left(\frac{B_n}{n}\right)^{q/2} \left(\frac{n}{B_n} |t-s|^\gamma \mathbf{E}[|Y'|^q] + \left(\frac{n}{B_n} |t-s|^\gamma\right)^{q/2} \mathbf{E}[|Y'|^2]^{q/2}\right) \\ &\leq C_q \left(\left(\frac{B_n}{n}\right)^{q/2-1} \left(\frac{n^2}{B_n}\right)^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)} |t-s|^\gamma + |t-s|^{q\gamma/2}\right). \end{aligned}$$

Recall that for every  $\delta > 0$ , we have  $B_n \ll n^{\frac{1}{\alpha}+\delta}$  so

$$\frac{1}{\ln n} \ln \left( \left(\frac{B_n}{n}\right)^{q/2-1} \left(\frac{n^2}{B_n}\right)^{q(\frac{\alpha-1}{4\alpha}+\varepsilon)} \right) \leq \left(\frac{1}{\alpha} + \delta\right) \left(\frac{q}{2} - 1 - q\left(\frac{\alpha-1}{4\alpha} + \varepsilon\right)\right) + 2q\left(\frac{\alpha-1}{4\alpha} + \varepsilon\right) - \left(\frac{q}{2} - 1\right).$$

Taking  $\varepsilon = \delta = 0$ , the right-hand side reads

$$\frac{1}{\alpha} \left( \frac{q}{2} - 1 - q \frac{\alpha - 1}{4\alpha} \right) + q \frac{\alpha - 1}{2\alpha} - \left( \frac{q}{2} - 1 \right) = \frac{\alpha - 1}{\alpha} \left( 1 - \frac{q}{4\alpha} \right),$$

which tends to  $-\infty$  as  $q \rightarrow \infty$ . Now if  $\varepsilon, \delta > 0$  are small, one obtains instead the exponent

$$\frac{\alpha - 1}{\alpha} \left( 1 - \frac{q}{4\alpha} \left[ 1 - \delta \left( \frac{4\alpha^2}{\alpha - 1} + \frac{1}{2} \right) + 2\alpha\varepsilon \left( 2 \left( \frac{1}{\alpha} + \delta \right) - 1 \right) \right] \right) - \delta,$$

which still tends to  $-\infty$  as  $q \rightarrow \infty$ . Notice also that  $n^{-1} \leq |s - t| \leq 1$ , so we may choose  $q$  large enough so that

$$\sup_{n \rightarrow \infty} \mathbf{E} \left[ \left( \frac{|H_n^{\text{SP}'}(nt) - H_n^{\text{SP}'}(ns)|}{(n/B_n)^{1/2}} \right)^q \middle| A_n \right] \leq C_q |t - s|^2.$$

This bound holds whenever  $s, t \in [0, 1]$  are such that  $ns$  and  $nt$  are both integers. Since  $H_n^{\text{SP}'}$  is defined by linear interpolation between such times, then it also holds for every  $s, t \in [0, 1]$ . The standard Kolmogorov criterion then implies the following bound for the Hölder norm of  $H_n^{\text{SP}'}$  for some  $\theta > 0$ : for every  $\delta > 0$ , there exists  $C > 0$  such that for every  $n$  large enough,

$$\mathbf{P} \left( \sup_{0 \leq s \neq t \leq 1} \frac{|H_n^{\text{SP}}(nt) - H_n^{\text{SP}}(ns)|}{(n/B_n)^{1/2} |t - s|^\theta} \leq C \middle| A_n \right) \geq 1 - \delta.$$

Since  $\mathbf{P}(A_n) \rightarrow 1$  as  $n \rightarrow \infty$ , the same result holds under the unconditional probability, which shows that the sequence  $(B_n/n)^{1/2} H_n^{\text{SP}'}(n \cdot)$  is indeed tight and the proof is complete.

**3.2. Proof of Theorem 2 for non-centred snakes.** — We next assume that  $\mathbf{E}[Y] = m \neq 0$  and prove Theorem 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  $\frac{B_n}{n} \mathbf{E}[H_n^{\text{SP}}(n \cdot) \mid T_n] = m \frac{B_n}{n} H_n(n \cdot)$  which converges to  $m \cdot \mathcal{H}$  so it is equivalent to consider the centred version of  $Y$ . For the rest of the proof, we thus assume that  $\mathbf{E}[Y] = 0$  and  $\mathbf{P}(|Y| \geq n/B_n) = o(n^{-1})$ , and we prove that the corresponding scaled spatial process  $\frac{B_n}{n} H_n^{\text{SP}}(n \cdot)$  converges to the null process.

The fact that our assumption is necessary for tightness of this process goes exactly as for Theorem 1, in the first step: Now the tails of  $Y$  are so that  $\mathbf{P}(|Y| \geq y) = o(y^{-\frac{\alpha}{\alpha(1+\delta)-1}})$  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  $\delta, \varepsilon$ , 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*: we have

$$\lim_{n \rightarrow \infty} \frac{n^2}{B_n} \mathbf{P}(|Y| > b_n)^2 = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \left( \frac{n}{B_n} \right)^{1/2} b_n \mathbf{P}(|Y| > b_n) = 0,$$

so both processes  $\frac{B_n}{n} H_n^{\text{SP}''}(n \cdot)$  and  $\frac{B_n}{n} \mathbf{E}[H_n^{\text{SP}'}(n \cdot) \mid T_n]$  converge to the null process. Similarly, with the preceding notations, for  $s, t \in [0, 1]$ , we have

$$\begin{aligned} \mathbf{E} \left[ \left( \frac{|H_n^{\text{SP}'}(nt) - H_n^{\text{SP}'}(ns)|}{n/B_n} \right)^q \middle| A_n \right] &\leq C_q \left( \frac{B_n}{n} \right)^q \left( \frac{n}{B_n} |t - s|^\gamma \mathbf{E}[|Y'|^q] + \left( \frac{n}{B_n} |t - s|^\gamma \right)^{q/2} \mathbf{E}[|Y'|^2]^{q/2} \right) \\ &\leq C_q \left( \left( \frac{B_n}{n} \right)^{q-1} \left( \frac{n^2}{B_n} \right)^{q(\frac{\alpha-1}{2\alpha} + \varepsilon)} |t - s|^\gamma + \left( \frac{B_n}{n} \mathbf{E}[|Y'|^2] |t - s|^\gamma \right)^{q/2} \right). \end{aligned}$$

The first term in the last line is controlled as previously: the factor  $1/2$  in the exponent in  $b_n$  compensate the fact that we now rescale by  $n/B_n$  instead of  $(n/B_n)^{1/2}$  and similar calculations as in the preceding section show that this first term is bounded by  $|t - s|^\gamma$  times  $n$  raised to a power which converges to  $-\infty$

as  $q \rightarrow \infty$ . The only change compared to the proof of Theorem 1 is that we may not have  $\mathbf{E}[|Y|^2] < \infty$ . Still,

$$\mathbf{E}[|Y'|^2] = 2 \int_0^{b_n} y \mathbf{P}(|Y| > y) dy = O\left(\int_1^{b_n} y^{1-\frac{\alpha}{\alpha(1+\delta)-1}} dy\right).$$

Note that if  $\alpha < 2$ , then for  $\delta$  sufficiently small, the exponent is smaller than  $-1$  so the integral converges. If  $\alpha = 2$ , then since  $B_n$  is at least of order  $n^{1/2}$  (and it is exactly of this order if and only if  $\mu$  has finite variance), then we do not need any  $\delta$ : we have  $\mathbf{P}(|Y| \geq y) = o(y^{-2})$  and so

$$\mathbf{E}[|Y'|^2] = O\left(\int_1^{b_n} y^{-1} dy\right) = O(\ln b_n) = O(\ln n).$$

In both cases,  $\frac{B_n}{n} \mathbf{E}[|Y'|^2]$  is bounded above by  $n^{-\eta}$  for some  $\eta > 0$  and we may conclude as in the preceding proof that for  $q$  large enough,

$$\sup_{n \rightarrow \infty} \mathbf{E}\left[\left(\frac{|H_n^{\text{sp}'}(nt) - H_n^{\text{sp}'}(ns)|}{n/B_n}\right)^q \middle| A_n\right] \leq C_q |t - s|^2,$$

and so the process  $\frac{B_n}{n} H_n^{\text{sp}'}(n \cdot)$  is tight. Moreover, the preceding bounds applied with  $s = 0$  and  $t \in [0, 1]$  fixed show that the one-dimensional marginals converge in distribution 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 to prove Corollary 1, following the argument of Cai *et al.* [CHJ<sup>+</sup>17, Section 5].

First note that for a given tree  $T$  with  $n + 1$  vertices listed  $\emptyset = u_0 < u_1 < \dots < u_n$  in lexicographical order, we have

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

where we recall the notation  $\Lambda(T)$  for the total path length of  $T$ . Therefore the convergence of the conditional expectation of  $I(T_n)$  follows from (9). 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  $(-1/2, 1/2)$ , with variance  $\Sigma^2 = 1/12$ . The main idea, see the discussion around Equation (5.1) in [CHJ<sup>+</sup>17], 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:

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

where  $J(T_n) = \sum_{v \in T_n} S_v$  and we recall that  $S_v$  is the spatial position of the vertex  $v$ . Still following [CHJ<sup>+</sup>17, Section 5], let us define a process  $\widehat{R}_n$  on  $[0, 2n]$  as follows: recall that  $C_n^{\text{sp}}$  is the spatial process in *contour order*, then for every  $t \in [0, 2n]$ , if  $t$  is an integer, then set  $\widehat{R}_n(t) = C_n^{\text{sp}}(t)$ , otherwise set

$$\widehat{R}_n(t) = \begin{cases} C_n^{\text{sp}}(\lfloor t \rfloor), & \text{if } C_n(\lfloor t \rfloor) > C_n(\lceil t \rceil), \\ C_n^{\text{sp}}(\lceil t \rceil), & \text{if } C_n(\lfloor t \rfloor) < C_n(\lceil t \rceil). \end{cases}$$

In other words,  $\widehat{R}_n$  is a step function on  $[0, 2n]$  which is constant on each interval  $[i, i + 1)$  with  $0 \leq i \leq 2n - 1$ , on which it takes the value of the position of either the vertex visited at time  $i$  or  $i + 1$  in the contour order, whichever is the farthest (in graph distance) from the root. It then readily follows that

$$J(T_n) = \frac{1}{2} \int_0^{2n} \widehat{R}_n(t) dt.$$

Observe that  $|\widehat{R}_n(t) - C_n^{\text{sp}}(t)| \leq 1/2$  for every  $t \in [0, 2n]$  so,

$$I(T_n) - \frac{\Lambda(T_n)}{2} = \frac{1}{2} \int_0^{2n} C_n^{\text{sp}}(t) dt + O(n) = n \int_0^1 C_n^{\text{sp}}(2nt) dt + O(n).$$

Notice that  $(n^3/B_n)^{1/2} \gg n$ ; since  $\Sigma^2 = 1/12$ , we conclude from Theorem 1 that

$$\left(\frac{B_n}{12n^3}\right)^{1/2} \left(I(T_n) - \frac{\Lambda(T_n)}{2}\right) = \int_0^1 \left(\frac{B_n}{12n}\right)^{1/2} C_n^{\text{sp}}(2nt) dt + o(1) \xrightarrow[n \rightarrow \infty]{(d)} \int_0^1 \mathcal{S}_t dt.$$

#### 4. Heavy-tailed snakes

We investigate more precisely in this section the behaviour of  $H_n^{\text{sp}}$  and  $C_n^{\text{sp}}$  when the assumption  $\mathbf{P}(|Y| \geq (n/B_n)^{1/2}) = o(n^{-1})$  of Theorem 1 fails. In this case, 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  $H_n^{\text{sp}}$  and  $C_n^{\text{sp}}$  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 & Marckert [JM05] they still converge in distribution in the following weaker sense.

In this section, we identify continuous functions from  $[0, 1]$  to  $\mathbf{R}$  with their graph, which belong 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

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

where  $A^{(r)} = \{x \in \mathbf{R}^2 : d(x, A) \leq r\}$ . Then a sequence of functions  $(f_n)_{n \geq 1}$  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  $\Xi$  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  $\delta > 0$ , the set  $\Xi \cap ([0, 1] \times (\mathbf{R} \setminus [-\delta, \delta]))$  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 so we may take a random function  $f$  and a random set  $\Xi$  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 [JM05, Section 3.1].

**4.1. The intermediate regime.** — In the next result, we investigate the case where  $n \cdot \mathbf{P}(|Y| \geq (n/B_n)^{1/2})$  is uniformly bounded. Extracting a subsequence if necessary, we may assume in fact that both tails  $n \cdot \mathbf{P}(Y \geq (n/B_n)^{1/2})$  and  $n \cdot \mathbf{P}(-Y \geq (n/B_n)^{1/2})$  converge.

*THEOREM 3 (CONVERGENCE TO A ‘HAIRY SNAKE’).* — Suppose that  $\mathbf{E}[Y] = 0$ , that  $\Sigma^2 := \mathbf{E}[Y^2] \in (0, \infty)$  and that there exists  $a_+, a_- \in [0, \infty)$  such that  $a_+ + a_- > 0$  and

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

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

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

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

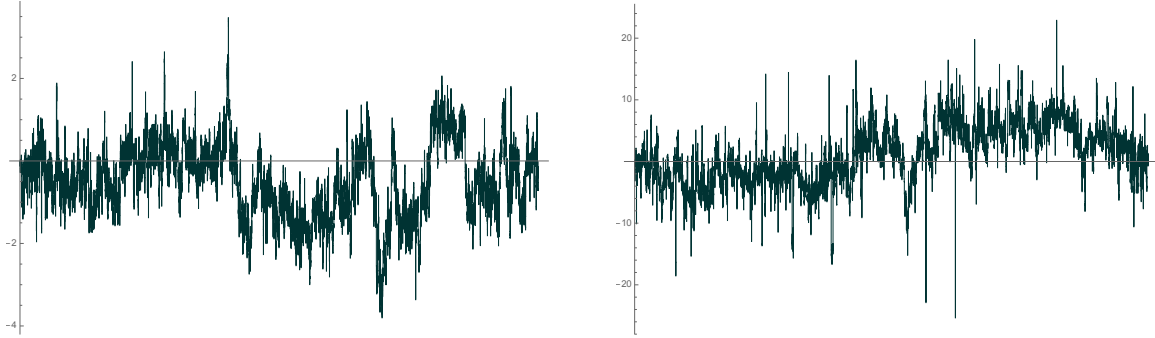


FIGURE 8. Two instances of the spatial height process  $H_n(n \cdot)$  associated with the height process of Figure 2 where in both cases,  $Y$  is symmetric and such that  $\lim_{n \rightarrow \infty} n \cdot \mathbf{P}(Y \geq (n/B_n)^{1/p}) = 1$ ; on the left,  $p = 2$  and on the right,  $p = 0,6$ .

The intuition behind this result is that, as opposed to Theorem 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  $\Xi$  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 at infinity. Therefore, for every  $y > 0$  fixed, we have

$$y \left( \frac{n}{B_n} \right)^{1/2} \underset{n \rightarrow \infty}{\sim} \left( \frac{y^{\frac{2\alpha}{\alpha-1}} n}{B_{y^{\frac{2\alpha}{\alpha-1}} n}} \right)^{1/2}.$$

Observe that under the assumption of Theorem 3, we have

$$\mathbf{P} \left( Y > \left( \frac{y^{\frac{2\alpha}{\alpha-1}} n}{B_{y^{\frac{2\alpha}{\alpha-1}} n}} \right)^{1/2} \right) \underset{n \rightarrow \infty}{\sim} a_+ y^{-\frac{2\alpha}{\alpha-1}} n^{-1}.$$

For every  $\varepsilon > 0$ , the same relation holds when each occurrence (on both sides) of  $y^{\frac{2\alpha}{\alpha-1}}$  is multiplied by  $1 + \varepsilon$  or by  $1 - \varepsilon$ , which enables to conclude (the second limit is obtained by a symmetric argument) that under the assumption of Theorem 3, we have

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

Since, 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  $\mathbf{P}(Y > y(n/B_n)^{1/2})$ , this number is asymptotically Poisson distributed with rate  $a_+ y^{-\frac{2\alpha}{\alpha-1}}$ , which indeed corresponds to  $\Xi$ , provided that the locations are asymptotically uniformly distributed in the tree.

As for Theorem 1, let us decompose the proof into several steps, following closely the argument of the proof of Theorem 5 in [JM05].

**4.1.1. Contribution of the small jumps.** — As in the proof of Theorem 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  $Y'_u = Y_u \mathbb{1}_{\{|Y_u| \leq b_n\}}$  and  $Y''_u = Y_u \mathbb{1}_{\{|Y_u| > b_n\}}$ , define then  $H_n^{\text{sp}'}$  and  $H_n^{\text{sp}''}$  the spatial processes in which the increments  $Y_u$  are replaced by  $Y'_u$  and  $Y''_u$  respectively, so  $H_n^{\text{sp}} = H_n^{\text{sp}'} + H_n^{\text{sp}''}$ . Observe that the only difference with Theorem 1 is that our assumption now implies  $\mathbf{P}(|Y| \geq y) = O(y^{-\frac{2\alpha}{\alpha(1+\delta)-1}})$ , whereas the big  $O$  was a small  $o$  there. Actually, the arguments used to control the small jumps in Section 3.1.4 and 3.1.5 only requires a big  $O$  so we conclude that, as there, we have

$$\left( \frac{B_n}{n} H_n(nt), \left( \frac{B_n}{n\Sigma^2} \right)^{1/2} H_n^{\text{sp}'}(nt) \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{(d)} (\mathcal{H}_t, \mathcal{S}_t)_{t \in [0,1]}.$$

**4.1.2. Contribution of the medium-large jumps.** — We next claim that we have

$$(10) \quad \left\{ \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{sp}''}(nt); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} 0 \bowtie \Xi,$$

in  $\mathcal{H}$ . Let us approximate both sides. First note that the intensity measure of  $\Xi$  explodes at the  $x$  axis, but for every  $\eta > 0$ , the restricted Poisson random measure  $\Xi^\eta = \Xi \cap ([0, 1] \times (\mathbf{R} \setminus [-\eta, \eta]))$  is almost surely finite. Clearly, the Hausdorff distance between  $0 \bowtie \Xi$  and  $0 \bowtie \Xi^\eta$  is at most  $\eta$ . Similarly, let us truncate further  $Y''_u$  by setting  $Y''_\eta = Y_u \mathbb{1}_{\{|Y_u| > \eta(n/B_n)^{1/2}\}}$  and then define  $H_n^{\text{sp}\eta}$  as the spatial process in which the increments  $Y_u$  are replaced by  $Y''_\eta$ . 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 since the argument used in Section 3.1.3 only requires a big  $O$  in the tail of  $Y$ . Furthermore, on the complement event, we have  $\max_{0 \leq i \leq n} |H_n^{\text{sp}''}(i)| = \max_{u \in T_n} |Y''_u|$  and  $\max_{0 \leq i \leq n} |H_n^{\text{sp}\eta}(i)| = \max_{u \in T_n} |Y''_\eta|$ . The latter is either  $\max_{u \in T_n} |Y''_\eta|$  or 0 in the case  $\max_{u \in T_n} |Y''_\eta| \leq \eta(n/B_n)^{1/2}$  so we conclude that on the event  $E_n^c$  whose probability tends to 1, we have

$$\max_{0 \leq i \leq n} |H_n^{\text{sp}''}(i)| - \max_{0 \leq i \leq n} |H_n^{\text{sp}\eta}(i)| \leq \eta \left( \frac{n}{B_n} \right)^{1/2}.$$

Therefore, in order to prove (10), it suffices to prove that for every  $\eta > 0$ , it holds that

$$(11) \quad \left\{ \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{sp}\eta}(nt); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} 0 \bowtie \Xi^\eta.$$

**4.1.3. A discrete Poisson random measure.** — Let  $\emptyset = u_0 < u_1 < \dots < u_n$  be the vertices of  $T_n$  listed in lexicographical order, and let  $1 \leq k_1 < \dots < k_{N_n} \leq n$  be the indices of those vertices  $u$  of  $T_n$  for which  $Y''_\eta$  is non-zero, or otherwise said  $|Y_u| > \eta(n/B_n)^{1/2}$ . Let  $\nu_\eta(dy) = \frac{2\alpha}{\alpha-1} y^{-1-\frac{2\alpha}{\alpha-1}} (a_+ \mathbb{1}_{\{y > \eta\}} + a_- \mathbb{1}_{\{y < -\eta\}}) dy$  so  $dx \nu_\eta(dy)$  is the intensity of  $\Xi^\eta$ . Then the discussion just after the statement of the theorem shows that  $N_n$  has the binomial distribution with parameters  $n$  and  $\mathbf{P}(|Y| > \eta(n/B_n)^{1/2})$  which converges to the Poisson distribution with parameter  $\nu_\eta(\mathbf{R}) \in (0, \infty)$ . Furthermore, conditional on  $N_n$ , the values  $(Y''_{u_{k_i}})_{1 \leq i \leq N_n}$  have the same distribution as i.i.d. copies of  $Y$  conditioned on  $|Y| > \eta(n/B_n)^{1/2}$ , the indices  $(k_i)_{1 \leq i \leq N_n}$  have the uniform distribution amongst the ranked  $N_n$ -tuples in  $\{1, \dots, n\}$ , and they are independent of the values  $(Y''_{u_{k_i}})_{1 \leq i \leq N_n}$ . All the other  $Y_u$ 's are null.

Since  $N_n$  is bounded in probability, the  $N_n$ -tuple  $(k_i)_{1 \leq i \leq N_n}$  is well approximated as  $n \rightarrow \infty$  by i.i.d. random times on  $\{1, \dots, n\}$ . Furthermore, one easily checks from their tail behaviour that the random variables  $(B_n/n)^{1/2} Y''_{u_{k_i}}$  converge in distribution as  $n \rightarrow \infty$  towards i.i.d. random variables sampled from the probability  $\nu_\eta(\cdot)/\nu_\eta(\mathbf{R})$ . Therefore, the set  $\xi_n^\eta = \{(n^{-1}k_i, (B_n/n)^{1/2} Y''_{u_{k_i}}); 1 \leq i \leq N_n\} \subset [0, 1] \times \mathbf{R}$  converges in law in  $\mathcal{H}$  towards the random set  $\xi_\infty^\eta = \{(U_i, X_i); 1 \leq i \leq N\}$  where  $(U_i)_{i \geq 1}$  are i.i.d. uniformly distributed on  $[0, 1]$ ,  $(X_i)_{i \geq 1}$  are i.i.d. with law  $\nu_\eta(\cdot)/\nu_\eta(\mathbf{R})$ ,  $N$  has the Poisson distribution



with parameter  $\nu_\eta(\mathbf{R})$ , and all are independent. Then  $\xi_\infty^\eta$  has the law of  $\Xi^\eta$ . One easily check that the mapping  $\xi \mapsto 0 \bowtie \xi$  is continuous  $\Xi$ -almost surely, so we conclude that

$$0 \bowtie \xi_n \xrightarrow[n \rightarrow \infty]{(d)} 0 \bowtie \Xi^\eta.$$

in  $\mathcal{K}$ .

**4.1.4. Contribution of the very large jumps.** — In order to prove that (11), and therefore (10), holds, in only remains to prove that

$$d_H \left( \left\{ \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{sp}\eta}(nt); t \in [0, 1] \right\}, 0 \bowtie \xi_n^\eta \right) \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0.$$

We implicitly work conditional on the event  $E_n^c$  so there is at most one non zero value of  $Y_u^\eta$  along each branch of  $T_n$ . In this case, the process  $(B_n/n) \cdot H_n^{\text{sp}\eta}$  can be described 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  $k_2$  and so on. On the other hand,  $0 \bowtie \xi_n^\eta$  is constructed by putting value 0 for every time  $t \in [0, 1]$  except the  $k_i$ 's where we place vertical peaks given by the  $Y_{u_{k_i}}$ 's. Then the previous convergence is an easy consequence of the following lemma which extends [JM05, Lemma 8]:

*LEMMA 3.* — *Let  $\nu_n$  be uniformly distributed in  $T_n$ , and let  $D(\nu_n)$  be its number of descendants, then  $D(\nu_n)/n$  converges in probability to 0.*

Indeed, since  $N_n$  is bounded in probability and our vertices  $(u_{k_i})_{1 \leq i \leq N_n}$  are uniformly distributed (and conditioned to be different and to lie on different branches, but this occurs with high probability), it follows that their progeny are all small compared to  $n$ , so as  $n \rightarrow \infty$ , the process  $(B_n/n) \cdot H_n^{\text{sp}\eta}$  indeed goes back almost immediately to 0 after reaching a high value, as for  $0 \bowtie \xi_n^\eta$ .

*Proof of Lemma 3.* — Let us condition on  $T_n$ :

$$\mathbf{E}[D(\nu_n) \mid T_n] = \frac{1}{n+1} \sum_{v \in T_n} \sum_{w \in T_n} \mathbb{1}_{\{w \text{ is an ancestor of } v\}} = \frac{1}{n+1} \sum_{w \in T_n} |v| = \frac{1}{n+1} \Lambda(T_n),$$

where we recall the notation  $\Lambda(T_n)$  for the total path length of  $T_n$ , which is of order  $n^2/B_n$  according to (9). We conclude from the Markov inequality that for every  $\varepsilon, C > 0$ , we have

$$\mathbf{P}(D(\nu_n) > \varepsilon n) \leq \mathbf{P}(\Lambda(T_n) > C(n+1)^2/B_n) + C/B_n,$$

which converges to 0 when letting first  $n \rightarrow \infty$  and then  $C \rightarrow \infty$ .  $\square$

**4.1.5. Combining small and large jumps.** — The proof is not finished! We have shown that

$$\left( \frac{B_n}{n} H_n(nt), \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{sp}'\eta}(nt) \right)_{t \in [0, 1]} \xrightarrow[n \rightarrow \infty]{(d)} (\mathcal{H}_t, \Sigma^2 \cdot \mathcal{S}_t)_{t \in [0, 1]},$$

and

$$\left\{ \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{sp}''\eta}(nt); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} 0 \bowtie \Xi,$$

and yet, although, if for  $f \in \mathcal{C}([0, 1], \mathbf{R})$  and  $B \in \mathcal{K}$ , we put  $A \dot{+} B = \{(x, y_A + y_B); (x, y_A) \in A, (x, y_B) \in B\}$ , then this addition is continuous, we cannot directly conclude that

$$\left\{ \left( \frac{B_n}{n} \right)^{1/2} (H_n^{\text{sp}'\eta}(nt) + H_n^{\text{sp}''\eta}(nt)); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} (\Sigma^2 \cdot \mathcal{S}_t) \bowtie \Xi = (\Sigma^2 \cdot \mathcal{S}_t) \dot{+} (0 \bowtie \Xi),$$

because the previous convergences in distribution may not hold simultaneously. Indeed, the processes  $H_n^{\text{sp}'\eta}$  and  $H_n^{\text{sp}''\eta}(nt)$  are not independent since each  $Y_u$  either contributes to one or to the other.

We create independence by resampling the  $Y_u$ 's which contribute to  $H_n^{\text{SP}''}(nt)$  as follows: let  $(Z_i)_{i \geq 1}$  be i.i.d. copies of  $Y \mathbb{1}_{\{|Y| \leq b_n\}}$  independent of the rest and put

$$\tilde{Y}_i = Y_i \mathbb{1}_{\{|Y| \leq b_n\}} + Z_i \mathbb{1}_{\{|Y| > b_n\}},$$

for each  $1 \leq i \leq n$ . Now the processes  $\tilde{H}_n^{\text{SP}}$  and  $H_n^{\text{SP}''}(nt)$  are independent, and furthermore, the error between  $H_n^{\text{SP}'}$  and  $\tilde{H}_n^{\text{SP}}$  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  $\max_{0 \leq i \leq n} |\tilde{H}_n^{\text{SP}}(i) - H_n^{\text{SP}'}(i)| \leq b_n = o(n/B_n)^{1/2}$  since each  $\tilde{Y}_i$  and each  $Y'_i$  belongs to  $[0, b_n]$ . We conclude that

$$\left( \left( \frac{B_n}{n} \right)^{1/2} H_n^{\text{SP}'}(nt), \left( \frac{B_n}{n} \right)^{1/2} \tilde{H}_n^{\text{SP}}(nt) \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{(d)} (\Sigma^2 \cdot \mathcal{S}_t, \Sigma^2 \cdot \mathcal{S}_t)_{t \in [0,1]},$$

and the proof of Theorem 3 is now complete.

**4.2. The strong heavy tail regime.** — We finally investigate the regime where the tails of  $Y$  are much stronger than what requires Theorem 1. In this case, the extreme values dominate the small ones and the snake disappears at the limit, and only the vertical peaks remain, see Figure 8 for a comparison with the previous case.

*THEOREM 4 (CONVERGENCE TO A 'FLAT HAIRY SNAKE').* — Fix  $p \in (0, 2]$  and suppose that  $\mathbf{E}[Y] = 0$ . Assume that there exists  $\varrho \in [0, 1]$  and two slowly varying functions at infinity  $L^+$  and  $L^-$  such that if  $L = L^+ + L^-$ , then as  $x \rightarrow \infty$ , the ratios  $L^+(x)/L(x)$  and  $L^-(x)/L(x)$  converge respectively to  $\varrho$  and  $1 - \varrho$ , and furthermore

$$n \cdot \mathbf{P}(Y \geq (n/B_n)^{1/p} L^+(n/B_n)) \xrightarrow[n \rightarrow \infty]{} 1 \quad \text{and} \quad n \cdot \mathbf{P}(-Y \geq (n/B_n)^{1/p} L^-(n/B_n)) \xrightarrow[n \rightarrow \infty]{} 1,$$

If  $p = 2$ , assume also that the function  $L$  tends to infinity. Let  $\Xi$  be a Poisson random measure on  $[0, 1] \times \mathbf{R}$  independent of the pair  $\mathcal{S}$ , with intensity

$$\frac{p\alpha}{\alpha - 1} y^{-1 - \frac{p\alpha}{\alpha - 1}} \left( \varrho^{1 + \frac{p\alpha}{\alpha - 1}} \mathbb{1}_{\{y > 0\}} + (1 - \varrho)^{1 + \frac{p\alpha}{\alpha - 1}} \mathbb{1}_{\{y < 0\}} \right) dx dy.$$

Then the convergence in distribution of the sets

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

holds in  $\mathcal{K}$ , jointly with (1). The same holds (jointly) when  $H_n^{\text{SP}}(n \cdot)$  is replaced by  $C_n^{\text{SP}}(2n \cdot)$ .

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

$$n \cdot \mathbf{P}(Y \geq (n/B_n)^{1/p}) \xrightarrow[n \rightarrow \infty]{} a_+ \quad \text{and} \quad n \cdot \mathbf{P}(-Y \geq (n/B_n)^{1/p}) \xrightarrow[n \rightarrow \infty]{} a_-,$$

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

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

where  $\Xi$  has intensity  $\frac{p\alpha}{\alpha - 1} y^{-1 - \frac{p\alpha}{\alpha - 1}} (a_+ \mathbb{1}_{\{y > 0\}} + a_- \mathbb{1}_{\{y < 0\}}) dx dy$ , which recovers [JM05, Theorem 6].

*REMARK 1.* — Recall if  $\mathbf{E}[Y] = m \neq 0$ , then  $(B_n/n) \cdot \mathbf{E}[H_n^{\text{SP}}(n \cdot) | T_n]$  converges to  $m \cdot \mathcal{H}$  so the previous result still holds in this case for  $p < 1$ , and when  $p = 1$  and  $L^+$  and  $L^-$  both converge, then

$$\left\{ \left( \frac{B_n}{n} \right)^{1/p} H_n^{\text{SP}}(nt); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} m \cdot \mathcal{H} \bowtie \Xi,$$

in  $\mathcal{K}$ , jointly with (1), where  $\mathcal{H}$  and  $\Xi$  are independent, and the same holds (jointly) when  $H_n^{\text{SP}}(n\cdot)$  is replaced by  $C_n^{\text{SP}}(2n\cdot)$ .

*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 to be tuned. For every vertex  $u \in T_n$ , let  $Y'_u = Y_u \mathbb{1}_{\{|Y_u| \leq b_n\}}$  and  $Y''_u = Y_u \mathbb{1}_{\{|Y_u| > b_n\}}$ , define then  $H_n^{\text{SP}'}$  and  $H_n^{\text{SP}''}$  the spatial processes in which the increments  $Y_u$  are replaced by  $Y'_u$  and  $Y''_u$  respectively, so  $H_n^{\text{SP}} = H_n^{\text{SP}'} + H_n^{\text{SP}''}$ . Similarly as in the proof of Theorem 1 and Theorem 2, the exponent in  $b_n$  matches that in the tails of  $Y$ . Therefore, again, it holds that

$$\frac{n^2}{B_n} \mathbf{P}(|Y| > b_n)^2 \ll \left(\frac{n^2}{B_n}\right)^{1 - \left(\frac{\alpha-1}{2p\alpha} + \varepsilon\right) \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}},$$

and the exponent is negative for  $\delta, \theta$  small enough. 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 Section 4.1.2, 4.1.3 and 4.1.4 extend readily to prove that

$$\left\{ \frac{B_n^{1/p}}{n^{1/p}L(n/B_n)} H_n^{\text{SP}''}(nt); t \in [0, 1] \right\} \xrightarrow[n \rightarrow \infty]{(d)} 0 \bowtie \Xi,$$

and it only remains to prove the convergence

$$\left( \frac{B_n^{1/p}}{n^{1/p}L(n/B_n)} H_n^{\text{SP}'}(nt) \right)_{t \in [0, 1]} \xrightarrow[n \rightarrow \infty]{\mathbf{P}} 0.$$

We first consider the average displacement induced by these small jumps. Let  $m_n = \mathbf{E}[Y'] = -\mathbf{E}[Y'']$ , so

$$\frac{B_n^{1/p}}{n^{1/p}L(n/B_n)} \mathbf{E} \left[ H_n^{\text{SP}'}(n\cdot) \mid T_n \right] = \frac{B_n^{-(p-1)/p}}{n^{-(p-1)/p}L(n/B_n)} m_n \cdot \frac{B_n}{n} H_n(n\cdot).$$

From the tail behaviour of  $Y$  we get:

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

Exactly as in the proof of Theorem 1, note that for  $\varepsilon, \delta, \theta$  sufficiently small, we have

$$\left(\frac{\alpha-1}{2p\alpha} + \varepsilon\right) \left(1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}\right) > -\frac{\alpha+1}{2\alpha},$$

and this bound gets tighter as  $\varepsilon, \delta, \theta$  get closer to 0. Then, since

$$\frac{\alpha+1}{2\alpha} < \frac{p-1}{p} < \frac{\alpha+1}{\alpha},$$

we conclude that for  $\varepsilon, \delta, \theta$  sufficiently small,

$$\frac{n^{(p-1)/p}}{B_n^{(p-1)/p}L(n/B_n)} \left( b_n^{1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}} \right) \leq n^{2\left(\frac{\alpha-1}{2p\alpha} + \varepsilon\right)\left(1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}\right) + \frac{p-1}{p-\theta}} \cdot B_n^{-\left(\frac{\alpha-1}{2p\alpha} + \varepsilon\right)\left(1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}\right) - \frac{p-1}{p-\theta}}$$

converges to 0 as  $n \rightarrow \infty$  since both exponents are negative. Therefore the process  $\mathbf{E}[H_n^{\text{SP}'}(n\cdot) \mid T_n]$  suitably rescaled converges in probability to 0 and we focus for the rest of the proof on the centred process  $H_n^{\text{SP}'}(n\cdot) = H_n^{\text{SP}'}(n\cdot) - \mathbf{E}[H_n^{\text{SP}'}(n\cdot) \mid T_n]$ .

Similarly, with the notation from Section 3, for  $s, t \in [0, 1]$ , we have

$$\begin{aligned} & \mathbf{E} \left[ \left( \frac{B_n^{1/p}}{n^{1/p}L(n/B_n)} |H_n^{\text{sp}'}(nt) - H_n^{\text{sp}'}(ns)| \right)^q \middle| A_n \right] \\ & \leq C_q \frac{B_n^{q/p}}{n^{q/p}L(n/B_n)^q} \left( \frac{n}{B_n} |t-s|^\gamma \mathbf{E}[|Y'|^q] + \left( \frac{n}{B_n} |t-s|^\gamma \right)^{q/2} \mathbf{E}[|Y'|^2]^{q/2} \right) \\ & \leq C_q \left( L \left( \frac{n}{B_n} \right)^{-q} \left( \frac{B_n}{n} \right)^{\frac{q}{p}-1} \left( \frac{n^2}{B_n} \right)^{q(\frac{\alpha-1}{p}+\varepsilon)} |t-s|^\gamma + \left[ L \left( \frac{n}{B_n} \right)^{-2} \left( \frac{B_n}{n} \right)^{\frac{2}{p}-1} \mathbf{E}[|Y'|^2] \right]^{q/2} |t-s|^{q\gamma/2} \right). \end{aligned}$$

The first term in the last line is controlled as in the proof of Theorem 1 and Theorem 2: the slowly varying function will not cause any trouble, and the factor  $1/p$  in the exponent in  $b_n$  compensates the fact that we now rescale by  $(n/B_n)^{1/p}$  instead of  $(n/B_n)^{1/2}$  for Theorem 1 and similar calculations as in the proof of the latter show that this first term is bounded by  $|t-s|^\gamma$  times  $n$  raised to a power which converges to  $-\infty$  as  $q \rightarrow \infty$ . The only change compared to the proof of Theorem 1 is that, as for Theorem 2, we may not have  $\mathbf{E}[|Y|^2] < \infty$ . Still,

$$\mathbf{E}[|Y'|^2] = 2 \int_0^{b_n} y \mathbf{P}(|Y| > y) dy = O \left( \int_1^{b_n} y^{1-\frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}} dy \right).$$

Either  $\frac{\alpha(p-\theta)}{\alpha-1} > 2$ , in which case  $1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1} < -1$  for  $\delta$  small enough and the integral converges, or  $\frac{\alpha(p-\theta)}{\alpha-1} \leq 2$ , in which case,  $1 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1} > -1$ , so  $\mathbf{E}[|Y'|^2] = O(b_n^{2-\frac{\alpha(p-\theta)}{\alpha(1+\delta)-1}})$  and then

$$\left( \frac{B_n}{n} \right)^{\frac{2}{p}-1} \mathbf{E}[|Y'|^2] = O \left( B_n^{\frac{2}{p}-1 - (\frac{\alpha-1}{2p\alpha} + \varepsilon)(2 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1})} n^{1 - \frac{2}{p} + 2(\frac{\alpha-1}{2p\alpha} + \varepsilon)(2 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1})} \right).$$

Now for every  $\eta > 0$ , it holds that  $B_n = o(n^{\frac{1+\eta}{\alpha}})$ , so finally

$$\left( \frac{B_n}{n} \right)^{\frac{2}{p}-1} \mathbf{E}[|Y'|^2] = o \left( n^{\frac{1+\eta}{\alpha} (\frac{2}{p}-1 - (\frac{\alpha-1}{2p\alpha} + \varepsilon)(2 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1})) + 1 - \frac{2}{p} + 2(\frac{\alpha-1}{2p\alpha} + \varepsilon)(2 - \frac{\alpha(p-\theta)}{\alpha(1+\delta)-1})} \right).$$

Note that we got rid of the term  $L(n/B_n)^{-2}$  but its contribution is negligible. For  $\eta = \varepsilon = \delta = \theta = 0$ , the preceding exponent reduces to

$$\frac{1}{\alpha} \left( \frac{2}{p} - 1 - \frac{\alpha-1}{2p\alpha} \left( 2 - \frac{\alpha p}{\alpha-1} \right) \right) + 1 - \frac{2}{p} + \frac{\alpha-1}{p\alpha} \left( 2 - \frac{\alpha p}{\alpha-1} \right) = \frac{2 - \alpha(p+2)}{4\alpha^2 p},$$

which is negative for any  $p > 0$  and  $\alpha > 1$ . Then the exponent is still negative for  $\eta, \varepsilon, \delta, \theta > 0$  small enough. In almost all cases,  $(B_n/n)^{\frac{2}{p}-1} \mathbf{E}[|Y'|^2]$  is thus bounded by a negative power of  $n$ , the only case where it fails is for  $p = 2$ , in which case  $\frac{\alpha(p-\theta)}{\alpha-1} > 2$  for  $\theta$  small enough, so  $(B_n/n)^{\frac{2}{p}-1} \mathbf{E}[|Y'|^2]$  is uniformly bounded, but this is not an issue. We conclude as in the proof of Theorem 1 and Theorem 2 that for  $q$  large enough,

$$\sup_{n \rightarrow \infty} \mathbf{E} \left[ \left( \frac{B_n^{1/p}}{n^{1/p}L(n/B_n)} |H_n^{\text{sp}'}(nt) - H_n^{\text{sp}'}(ns)| \right)^q \middle| A_n \right] \leq C_q |t-s|^2,$$

so the process  $\frac{B_n}{n} H_n^{\text{sp}'}(n \cdot)$  is tight, and the bounds applied with  $s = 0$  and  $t \in [0, 1]$  fixed show that the one-dimensional marginals converge in distribution to 0 so the whole process converges in distribution to the null process, which completes the proof.  $\square$

## References

- [AD14] Romain Abraham and Jean-François Delmas. Local limits of conditioned Galton–Watson trees: the infinite spine case. *Electron. J. Probab.*, 19:no. 2, 19, 2014.
- [Ald93] David Aldous. The continuum random tree III. *Ann. Probab.*, 21(1):248–289, 1993.
- [BDFG04] Jérémie Bouttier, Philippe Di Francesco, and Emmanuel Guitter. Planar maps as labeled mobiles. *Electron. J. Combin.*, 11(1):Research Paper 69, 27, 2004.
- [BS15] Jakob E. Björnberg and Sigurdur Örn Stefánsson. Random walk on random infinite looptrees. *J. Stat. Phys.*, 158(6):1234–1261, 2015.
- [CHJ<sup>+</sup>17] Xing Shi Cai, Cecilia Holmgren, Svante Janson, Tony Johansson, and Fiona Skerman. Inversions in split trees and conditional Galton–Watson trees. Preprint available at [arXiv:1709.00216](https://arxiv.org/abs/1709.00216), 2017.
- [CS04] Philippe Chassaing and Gilles Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probab. Theory Related Fields*, 128(2):161–212, 2004.
- [Don82] Ronald A. Doney. On the exact asymptotic behaviour of the distribution of ladder epochs. *Stochastic Process. Appl.*, 12(2):203–214, 1982.
- [Duq03] Thomas Duquesne. A limit theorem for the contour process of conditioned Galton–Watson trees. *Ann. Probab.*, 31(2):996–1027, 2003.
- [Duq09] Thomas Duquesne. An elementary proof of Hawkes’s conjecture on Galton–Watson trees. *Electron. Commun. Probab.*, 14:151–164, 2009.
- [DLG02] Thomas Duquesne and Jean-François Le Gall. Random trees, Lévy processes and spatial branching processes. *Astérisque*, (281):vi+147, 2002.
- [DLG05] Thomas Duquesne and Jean-François Le Gall. Probabilistic and fractal aspects of Lévy trees. *Probab. Theory Related Fields*, 131(4):553–603, 2005.
- [Fel71] William Feller. *An introduction to probability theory and its applications. Vol. II*. Second edition. John Wiley & Sons, Inc., New York-London-Sydney, 1971.
- [Gito3] Bernhard Gittenberger. A note on: “State spaces of the snake and its tour—convergence of the discrete snake” [J. Theoret. Probab. **16** (2003), no. 4, 1015–1046; mr2033196] by J.-F. Marckert and A. Mokkadem. *J. Theoret. Probab.*, 16(4):1063–1067 (2004), 2003.
- [JCo4] Svante Janson and Philippe Chassaing. The center of mass of the ISE and the Wiener index of trees. *Electron. Comm. Probab.*, 9:178–187, 2004.
- [JM05] Svante Janson and Jean-François Marckert. Convergence of discrete snakes. *J. Theoret. Probab.*, 18(3):615–647, 2005.
- [JS15] Svante Janson and Sigurður Örn Stefánsson. Scaling limits of random planar maps with a unique large face. *Ann. Probab.*, 43(3):1045–1081, 2015.
- [Kes86] Harry Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(4):425–487, 1986.
- [Kor12] Igor Kortchemski. Invariance principles for Galton–Watson trees conditioned on the number of leaves. *Stochastic Process. Appl.*, 122(9):3126–3172, 2012.
- [Kor13] Igor Kortchemski. A simple proof of Duquesne’s theorem on contour processes of conditioned Galton–Watson trees. In *Séminaire de Probabilités XLV*, volume 2078 of *Lecture Notes in Math.*, pages 537–558. Springer, Cham, 2013.
- [Kor17] Igor Kortchemski. Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees. *Probab. Theory Related Fields*, 168(1-2):1–40, 2017.
- [LG99] Jean-François Le Gall. *Spatial branching processes, random snakes and partial differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1999.
- [LGLJ98] Jean-François Le Gall and Yves Le Jan. Branching processes in Lévy processes: the exploration process. *Ann. Probab.*, 26(1):213–252, 1998.
- [LGM11] Jean-François Le Gall and Grégory Miermont. Scaling limits of random planar maps with large faces. *Ann. Probab.*, 39(1):1–69, 2011.
- [Mar08] Jean-François Marckert. The lineage process in Galton–Watson trees and globally centered discrete snakes. *Ann. Appl. Probab.*, 18(1):209–244, 2008.
- [MM07] Jean-François Marckert and Grégory Miermont. Invariance principles for random bipartite planar maps. *Ann. Probab.*, 35(5):1642–1705, 2007.

- [MM03] Jean-François Marckert and Abdelkader Mokrakadem. States spaces of the snake and its tour—convergence of the discrete snake. *J. Theoret. Probab.*, 16(4):1015–1046, 2003.
- [Mar16] Cyril Marzouk. Scaling limits of random bipartite planar maps with a prescribed degree sequence. To appear in *Random Struct. Alg.* Preprint available at [arXiv:1612.08618](https://arxiv.org/abs/1612.08618), 2016.
- [Mie08] Grégory Miermont. Invariance principles for spatial multitype Galton-Watson trees. *Ann. Inst. Henri Poincaré Probab. Stat.*, 44(6):1128–1161, 2008.
- [Nev86] Jacques Neveu. Arbres et processus de Galton-Watson. *Ann. Inst. H. Poincaré Probab. Statist.*, 22(2):199–207, 1986.
- [Pet95] Valentin V. Petrov. *Limit theorems of probability theory*, volume 4 of *Oxford Studies in Probability*. The Clarendon Press, Oxford University Press, New York, 1995. Sequences of independent random variables, Oxford Science Publications.
- [Pit06] Jim Pitman. *Combinatorial stochastic processes*, volume 1875 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2006. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard.
- [Sen76] Eugene Seneta. *Regularly varying functions*. Lecture Notes in Mathematics, Vol. 508. Springer-Verlag, Berlin-New York, 1976.

---

CYRIL MARZOUK, Laboratoire de Mathématiques d’Orsay, Univ. Paris-Sud, CNRS, Université Paris-Saclay, 91405 Orsay, France.  
E-mail : [cyril.marzouk@u-psud.fr](mailto:cyril.marzouk@u-psud.fr)