Convergence of discrete snakes

Svante Janson
Department of Mathematics
Uppsala University
PO Box 480
SE-751 06 Uppsala
Sweden

Jean-François Marckert
Laboratoire de Mathématiques
Université de Versailles Saint-Quentin
45 avenue des Etats-Unis
78000 Versailles
France

Abstract

The discrete snake is an arborescent structure built with the help of a conditioned Galton–Watson tree and random i.i.d. increments Y. In this paper, we show that if $\mathbb{E}Y = 0$ and $\mathbb{P}(|Y| > y) = o(y^{-4})$, then the discrete snake converges weakly to the Brownian snake (this result was known under the hypothesis $\mathbb{E}Y^{8+\varepsilon} < +\infty$). Moreover, if this condition fails, and the tails of Y are sufficiently regular, we show that the discrete snake converges weakly to an object that we name jumping snake. In both case, the limit of the occupation measure is shown to be the integrated super-Brownian excursion. The proofs rely on the convergence of the codings of discrete snake with the help of two processes, called tours.

1 Introduction

1.1 Limits of arborescent structures

In the last years, numerous arborescent structures related to Brownian snake (BS) (see the works of Le Gall [16] and Duquesne & Le Gall [9] for a complete overview of the subject) or to the ISE (defined by Aldous [3]) have been published. ISE, acronym for the integrated super-Brownian excursion, is the occupation measure of the Brownian snake with lifetime the Brownian excursion. These objects (BS and ISE) now seem to be natural limits for discrete branching structures, and they are expected to appear often; we can cite the following:
– Aldous [3]: ISE is (defined as) a simple asymptotic model of repartition of mass.
– Chassaing & Schaeffer [6]: relation between the support of ISE and the diameter of quadrangulations
– Marckert & Mokkadem [19], Gittenberger [12] and Chassaing & Schaeffer [6]: convergence of discrete snakes to the Brownian snake (see Section 1.4 for more precisions)
– Slade [21] (and several other works of the same authors and coauthors): percolation and trees drawn on lattices
– Marckert [17]: about the difference of two related trees.

In the present paper, we search the limit of discrete snakes under more general assumptions than in [6, 12, 19]. The Brownian snake appears as the limit of discrete snakes (after suitable normalizations), when the increments Y are centered and have moments of order 4; to be precise, we need $\mathbb{P}(|Y| > y) = o(y^{-4})$. The convergence to the Brownian snake takes place in the set of continuous functions (in $C[0,1]^2$ for the tour, in $C([0,1] \times [0,+\infty[) \times C[0,1]$ for the snake).

When the tails of distribution of Y are less concentrated than that, the discrete snake does not converge anymore to the Brownian snake. The reason is that some of the displacements

October 14, 2003

are too large (in some branches, a large Y is of the same order of magnitude as the sum of the other displacements on this branch). The consequence of these *spikes* is that the limit snake is no more continuous. Moreover, if Y is not centered, the discrete snake drifts. Actually, the limit of the discrete snake is a combination (in some sense) of three parts: the first part being the Brownian snake, the second one being a random set of jumps and the third one being the drift. A kind of competition between these three parts occurs. According to the properties of Y, the limit snake, will present a combinations of one or two of these parts. In some case, the limit snake is christened the *jumping snake*. Since, in general, the limit processes are no more continuous, we choose to work in a topological space which is not a functional space. Actually, we show that the graph of the discrete snake (as well as the graph of the tour of the discrete snake) converges as a closed set.

The limit of the occupation measure of discrete snake is shown to converge to ISE (if $\mathbb{E}Y = 0$ and $0 < \operatorname{Var} Y < +\infty$).

Here is the plan of the paper: Section 1 contains the definition of conditioned branching random walks, of discrete snake and of its tour; known results about the convergence of discrete snakes (and tour) are recalled. The homeomorphism theorem of [19] that allows to deduce the convergence of the snake via the convergence of the tour (in the case of continuous functions) is recalled.

Section 2 is devoted to the study of the convergence of the tour of discrete snakes. Two main limit objects are obtained (the tour of the Brownian snake, and the hairy tour); as consequences we find again the asymptotic behavior of the maximum of the branching random walks in the different settings. The limit of the occupation measure of branching random walk is also shown to converge to ISE under quite general assumptions.

In Section 3, we investigate the limit of the discrete snake according to the properties of the distribution of the displacements. We define a new snake (the jumping snake) that is the limit of the discrete snake when the displacements have heavy tails.

Section 4 contains the proofs of the theorems of the previous sections.

1.2 Finite branching random walks

Consider a rooted ordered finite tree T in which each node is marked by a real number called *value*. Let u be a node, $y(u)$ its value and $h(u)$ its depth. Consider $(u_0 = root, u_1, \cdots, u_{h(u)} = u)$ the path from the root to u in T. We associate to u a trajectory of a killed walk $\Phi_u = (\Phi_u(j))_{j\in[\![0,h(u)]\!]}$ defined by

$$\Phi_u(j) = \sum_{i=0}^{j} y(u_i), \text{ for } j \in [\![0, h(u)]\!],$$

where $[\![a, b]\!] = [a, b] \cap \mathbb{Z}$. The branching walk associated to this marked tree is the multiset B of the trajectories Φ_u. The tree T is called the skeleton (or the underlying tree) of the branching walk. We construct a branching *random* walk by choosing the skeleton T and the values at random.

We consider a Galton–Watson branching process with offspring distribution $p = (p_i)_{i\geq 0}$, starting with 1 individual in generation 0. Consider a p-distributed random variable ξ. We

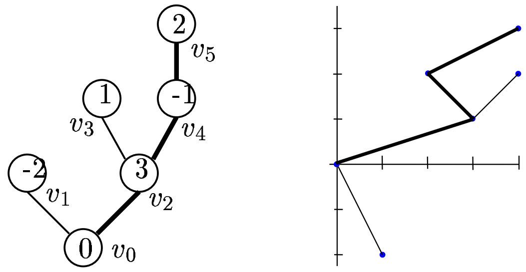

Figure 1: A valued tree and its associated BRW

assume that ξ satisfies the following condition

$$(A) : \begin{cases} \mathbb{E}\xi = 1, \quad 0 < \mathrm{Var}(\xi) = \sigma_\xi^2 < +\infty \\ \text{there exists a constant } \alpha > 0 \text{ s.t. } \mathbb{E}e^{\alpha\xi} < +\infty. \end{cases}$$

We write τ for the random family tree of this branching process and we let τ_n be τ conditioned by $|\tau| = n+1$; this is a random tree with $n+1$ nodes. Recall that by choosing appropriate ξ, we obtain for example random ordered trees, random labelled (Cayley) trees, random binary trees, ..., and in general any "simply generated tree" (with a minor technical assumption corresponding to $\mathbb{E}e^{\alpha\xi} < +\infty$), see e.g. [1] or [8].

We assume now that $T = \tau_n$. For the displacements, we set $y(root) = 0$ and the other values $y(u)$ are assumed to be i.i.d. random variables, independent of T. (One could assume with few changes that the values can be dependent only for brothers as in [19]. Also, it is obviously equivalent to have values assigned to the edges instead.) Hence, for each node u, Φ_u is a random walk with $h(u)$ steps, where the increments are i.i.d.; such a random walk is usually called a killed random walk.

We denote by Q the law of $y(u)$ (for $u \neq root$), and let Y be a generic random variable with this distribution. The two distributions p and Q determine the law of the marked trees. The random multiset of trajectories B is then called a branching random walk. The discrete snake will be defined as a normalized, interpolated, version of B, endowed with an order. An aim of this paper is to examine the limit of the discrete snake as a function of Q (for ξ satisfying (A), p will have influence on the limit only through its variance); in particular, we find necessary and sufficient conditions for convergence to the Brownian snake.

1.3 The discrete snake and the tour of the discrete snake

Depth first traversal of the BRW

We recall the depth first procedure. Let τ be an ordered tree with $n+1$ nodes. We define a function (see Aldous [1, p. 260]):

$$f : [\![0, 2n]\!] \longrightarrow \{\text{nodes of } \tau\},$$

which we regard as a walk around τ, as follows:

$$f(0) = \text{root}.$$

Given $f(i) = v$, choose, if possible, the leftmost child w of v which has not already been visited, and set $f(i+1) = w$. If no child is left unvisited, let $f(i+1)$ be the parent of v.

The function f is only defined for integer arguments. It is natural to map intermediate values into the corresponding edges; we will however instead map into nodes and define

$$\tilde{f} : [0, 2n] \longrightarrow \{\text{nodes of } \tau\} \tag{1}$$

by taking $\tilde{f}(t)$ to be $f(\lfloor t \rfloor)$ or $f(\lceil t \rceil)$, whichever is most distant from the root. Note that each node except the root then is the image of two intervals of unit length. Thus, if we choose t uniformly at random in $[0, 2n]$, $\tilde{f}(t)$ will be uniformly distributed over the non-root nodes.

The depth first walk (DFW)

The DFW of τ is the process V_n defined by:

$$V_n(i) = d(root, f(i)), \quad 0 \le i \le 2n.$$

The DFW is also called the Harris walk or the tour of the tree.

For i from 0 to n, let v_i be the ith new node visited by the depth first procedure on τ ($v_0 = root$).

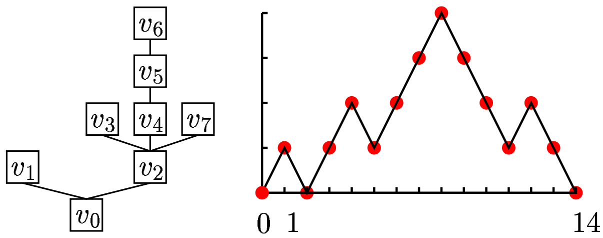

Figure 2: A tree and its associated DFW

In order to encode the branching random walk in keeping the historical branching structure, we use the depth first traversal of the tree. We denote by

$$R_n(k) = \Phi_{f(k)}(h(f(k)))$$

the terminal points of the random walk $\Phi_{f(k)}$. We call the process R_n the discrete head process. The two dimensional process $\{(R_n(k), V_n(k)), k \in [\![0, 2n]\!]\}$ determines T and all the values.

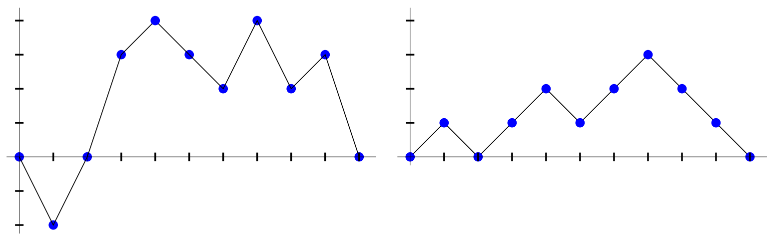

Figure 3: Interpolation of processes R_n and V_n associated to the marked tree of Figure 1

Indeed, the process V_n alone determines the skeleton of the marked tree T. The value of the node u, whose father is v, is obtained by the difference $\Phi_u(h(u)) - \Phi_v(h(v))$.

Remark 1. V_n is sometimes described as follows: $V_n(k)$ is the height at time k of a fly that is (clockwise) walking around the tree, 1 edge per time unit (the fly is on the root at time 0 and at time $2n$). Consider the successive values $y(u)$ of the visited nodes as abscissa displacements; then $R_n(k)$ is the abscissa of the fly at time k.

We will assume that $\tau = \tau_n$ and the values $y(u)$, $u \in \tau$, are random as above; thus V_n and R_n are random processes.

1.3.1 The discrete snake

We interpolate V_n linearly between integral points:

$$V_n(x) = V_n(\lfloor x \rfloor) + \{x\}\big(V_n(\lceil x \rceil) - V_n(\lfloor x \rfloor)\big), \text{ for } x \in [0, 2n]. \tag{2}$$

By analogy with the Brownian snake, we define the discrete snake as the process

$$\big(W_n(x,t), V_n(x)\big)_{x \in [0,2n],\, t \in [0,\infty)}$$

where for each $x \in [0, 2n]$, $W_n(x, \cdot)$ is a stopped continuous process defined as follows:
– For $k \in [\![0, 2n]\!]$ and $t \in [0, V_n(k)]$, $W_n(k, \cdot)$ is the process that interpolates piecewise the random walk $\Phi_{f(k)}$:

$$W_n(k,t) = \Phi_{f(k)}(\lfloor t \rfloor) + \{t\}\big(\Phi_{f(k)}(\lceil t \rceil) - \Phi_{f(k)}(\lfloor t \rfloor)\big).$$

– For $x \in [0, 2n] \setminus \mathbb{Z}$ and $t \in [0, V_n(x)]$,

$$W_n(x,t) = \begin{cases} W_n(\lfloor x \rfloor, t) & \text{if} \quad V_n(\lceil x \rceil) < V_n(\lfloor x \rfloor) \\ W_n(\lceil x \rceil, t) & \text{if} \quad V_n(\lceil x \rceil) > V_n(\lfloor x \rfloor). \end{cases}$$

– For $x \in [0, 2n]$ and $t \in]V_n(x), +\infty[$, we set

$$W_n(x,t) = W_n(x, V_n(x)). \tag{3}$$

Note that

$$W_n(\lfloor x \rfloor, t) = W_n(\lceil x \rceil, t) \text{ for } 0 \le t \le V_n(\lfloor x \rfloor) \wedge V_n(\lceil x \rceil).$$

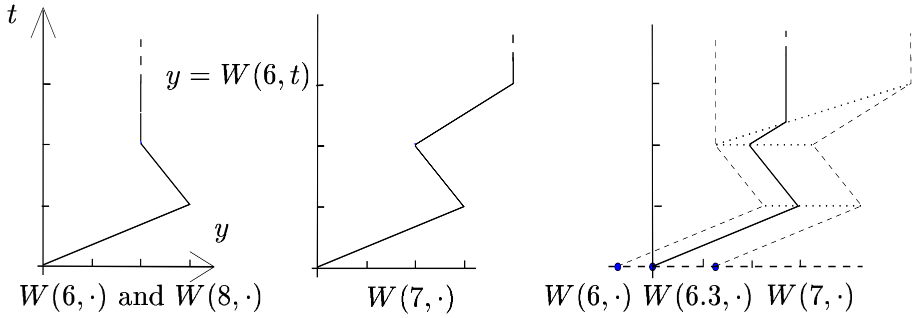

Figure 4: Three stopped paths associated to the BRW of Figure 1.

1.3.2 The tour of the discrete snake

We define the tour of the discrete snake as the process $\big(R_n(x), V_n(x)\big)_{x\in[0,2n]}$ where V_n is given in (2) and, similarly,

$$R_n(x) = R_n(\lfloor x\rfloor) + \{x\}\big(R_n(\lceil x\rceil) - R_n(\lfloor x\rfloor)\big), \text{ for } x \in [0, 2n].$$

1.3.3 Normalizations

We set,

$$\begin{aligned} v_n(s) &= \frac{V_n(2ns)}{\sqrt{n}} && \text{for } s \in [0,1] \\ r_n(s) &= \frac{R_n(2ns)}{n^{1/4}} && \text{for } s \in [0,1] \\ w_n(s,t) &= \frac{W_n(2ns, t\sqrt{n})}{n^{1/4}} && \text{for } (s,t) \in [0,1] \times [0,+\infty[. \end{aligned}$$

We call the processes (w_n, v_n) and (r_n, v_n) the normalized discrete snake and the normalized tour of the discrete snake respectively. The process r_n is called the normalized discrete head process. Note that w_n, by (3), enjoys the property

$$w_n(s,t) = w_n(s, v_n(s)) \text{ for } t \geq v_n(s).$$

1.4 Convergence of discrete snakes: known results

We say that the discrete snake converges weakly (with the normalizations chosen above) if, for a suitable topology, w_n converges weakly to a limit process. We say that the head of the discrete snake converges (or the tour of the snake converges) weakly if (r_n, v_n) converges. Whenever possible, we will use the space $C(A)$ of bounded continuous functions on a suitable metric space A, equipped with the topology of uniform convergence.

Marckert & Mokkadem [19] establish that the weak convergence of (r_n, v_n) in $C([0,1])^2$ is equivalent to the weak convergence of (w_n, v_n) in $C\big([0,1]\times[0,+\infty[\big) \times C[0,1]$.

The convergence of the second marginal of the tour is ensured by Aldous' theorem [1, 2] (see also [18]):

$$v_n \xrightarrow{(d)} v \overset{\text{def}}{=} \frac{2}{\sigma_\xi} e \quad \text{in } C[0,1], \tag{4}$$

where e is a normalized Brownian excursion. (We use $\xrightarrow{(d)}$ and $\xrightarrow{(p)}$ to denote convergence in distribution and in probability, respectively. All unspecified limits are as $n\to\infty$.) Further, assuming that $0 < \sigma_Y^2 = \operatorname{Var} Y < \infty$, let $r = \sigma_Y \bar{r}$, where $\bar{r}$ is the process such that conditionally given v, $\bar{r}$ is a centered Gaussian process with

$$\operatorname{Cov}(\bar{r}(s), \bar{r}(t) \mid v) = \check{v}(s,t) \overset{\text{def}}{=} \min_{s\leq u\leq t} v(u), \qquad s \leq t. \tag{5}$$

Note that we, for notational convenience, have hidden σ_ξ and σ_Y inside the definitions of v and r; if we want variables not depending on the distributions of ξ and Y we should change the normalization to $(\sigma_Y^{-1}\sigma_\xi^{1/2} r, \sigma_\xi v) = (\sigma_\xi^{1/2}\bar{r}, \sigma_\xi v)$.

In the case when p is geometric, that is when T is chosen uniformly among the rooted plane trees with $n+1$ vertices, $\mathbb{E}Y = 0$ and $\mathbb{E}|Y|^{6+\varepsilon} < +\infty$, Marckert & Mokkadem [19] show that $(r_n, v_n) \xrightarrow{(d)} (r, v)$. A consequence of that result is that the discrete snake converges weakly to the Brownian snake with lifetime process $v = \sqrt{2}e$. Chassaing & Schaeffer [6] prove independently the convergence of the head of the discrete snake to the head of the Brownian snake in the case p geometric and Y uniform in $\{-1, 0, 1\}$; the aim of their work was to compute the asymptotic diameter of rooted quadrangulations. Gittenberger [12] extended the result by [19] and showed that $(r_n, v_n) \xrightarrow{(d)} (r, v)$ holds for all conditioned Galton–Watson tree and Y satisfying $\mathbb{E}Y = 0$ and $\mathbb{E}|Y|^{8+\varepsilon} < +\infty$. One of our aims is to weaken this moment condition on Y as far as possible. (Theorem 2 below.)

2 Convergence of the tour of the discrete snake

2.1 Case $\mathbb{E}Y = 0$

This section is devoted to the study of the head of the snake when Y is centered.

We first observe that for finite-dimensional convergence, the existence of the second moment of Y is enough.

Theorem 1. *If $\mathbb{E}Y = 0$ and $\mathbb{E}Y^2 < +\infty$, then the finite-dimensional distributions of (r_n, v_n) converge to those of (r, v).*

Weak convergence in $C[0,1]$ is equivalent to the convergence of the finite distributions together with tightness (see [4, Theorem 8.1]). Often, the tightness is a technical nuisance, that can be verified with more or less work. Here, that is not the case and we need a stronger condition on Y in order to obtain convergence.

Theorem 2. *Assume $\mathbb{E}Y = 0$. Then the following are equivalent*

(i) $\mathbb{P}(|Y| \geq y) = o(y^{-4})$,

(ii) $r_n \xrightarrow{(d)} r$ *in* $C[0,1]$,

(iii) $(r_n, v_n) \xrightarrow{(d)} (r, v)$ *in* $C[0,1] \times C[0,1]$.

In particular (ii) and (iii) hold if $\mathbb{E}Y^4 < +\infty$, and no weaker moment condition suffices. We have thus extended as far as possible the result by Gittenberger [12] (where $\mathbb{E}|Y|^{8+\varepsilon} < \infty$).

When the condition $\mathbb{P}(|Y| \geq y) = o(y^{-4})$ of Theorem 2 is not satisfied, the conclusion fails because some values $y(u)$ are too large; these exceptionally large values of the $y(u)$ will give raise to narrow spikes in r_n. We still can obtain $r_n \xrightarrow{(d)} r$ in a weaker topology than $C[0,1]$, for example in the L^2 topology.

Theorem 3. *If $\mathbb{E}Y = 0$ and $\mathbb{E}Y^2 < +\infty$, then $r_n \xrightarrow{(d)} r$ in $L^2[0,1]$ and $(r_n, v_n) \xrightarrow{(d)} (r, v)$ in $L^2[0,1] \times C[0,1]$.*

Remark 2. One can similarly show that if $\mathbb{E}Y = 0$ and $\mathbb{E}|Y|^p < \infty$, with $p > 2$, then $r_n \xrightarrow{(d)} r$ in $L^p[0,1]$.

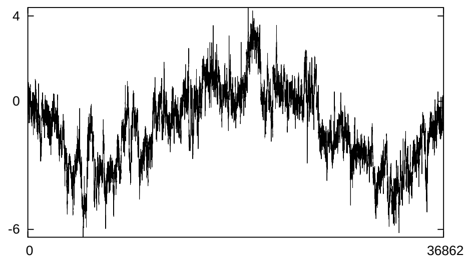

Figure 5: Simulation of R_n associated with a random plane tree with 18432 nodes, where Y is symmetric and satisfies $\mathbb{P}(|Y| > y) = (1+y)^{-6}$ for y non-negative.

The convergence in $L^2[0,1]$ is weaker than in $C[0,1]$. For example, Theorem 3 does not imply the convergence in distribution of $\max r_n$ to $\max r$; in fact, $\max r_n$ does not converge in distribution to $\max r$ because of the large spikes evoked above (see Theorem 5 and its proof). The convergence in $L^2[0,1]$ is enough, however, to imply the convergence of the corresponding empirical measures, as stated by Aldous [3].

Theorem 4. *Assume* $\mathbb{E}Y = 0$ *and* $0 < \mathbb{E}Y^2 < \infty$, *and let*

$$\nu_n = \frac{1}{n+1} \sum_{i=0}^{n} \delta\left(\sigma_Y^{-1} \sigma_\xi^{1/2} n^{-1/4} \Phi_{v_i}(h(v_i))\right) \tag{6}$$

be the normalized empirical measure of the terminal points of the branching random walk. Then $\nu_n \xrightarrow{(\mathrm{d})} \mu$ *in the space of probability measures on* $\mathbb{R}$, *where* μ *is the random probability measure ISE (in dimension one) defined by Aldous [3].*

The random measure μ has a representation using the process r; if g is a continuous bounded function on $\mathbb{R}$, one has [16, 9, 18]

$$\int g d\mu = \int_0^1 g\left(\sigma_\xi^{1/2} \sigma_Y^{-1} r(t)\right) dt. \tag{7}$$

To state a limit theorem that describes, rather than ignores, the spikes in r_n, we let $\mathcal{K} = \mathcal{K}([0,1] \times \mathbb{R})$ be the space of non-empty compact subsets of $[0,1] \times \mathbb{R}$ equipped with the Hausdorff metric

$$d(K_1, K_2) = \max\left(\max_{x \in K_1} d(x, K_2), \max_{x \in K_2} d(x, K_1)\right). \tag{8}$$

We identify a continuous function from $[0,1]$ to $\mathbb{R}$ with its graph, which is an element of $\mathcal{K}$, and note that for functions $f_n, f \in C[0,1]$, $f_n \longrightarrow f$ in $C[0,1]$ if and only if $f_n \longrightarrow f$ in $\mathcal{K}$.

Moreover, if $f \in C[0,1]$ and $\Xi \subset [0,1] \times (\mathbb{R} \setminus \{0\})$ is a set such that $\Xi \cap \big([0,1] \times (\mathbb{R} \setminus [-a,a])\big)$ is finite for every $a > 0$, define $H(f, \Xi)$ to be the union of the graph of f and the vertical line segments $[(x, f(x)), (x, f(x) + y)]$ for $(x, y) \in \Xi$. In other words, $H(f, \Xi)$ is the graph of f with added *hairs*, whose positions, lengths and directions (up or down) are described by Ξ.

It is easy to see that $H(f, \Xi)$ is compact, so it is an element of $\mathcal{K}$. Moreover, the mapping $(f, \Xi) \mapsto H(f, \Xi)$ is measurable for the natural σ-fields, so we may let f and Ξ be random and obtain a random element of $\mathcal{K}$. In the case $\mathbb{P}(|Y| \geq y) = \Theta(y^{-4})$, we have the following extension of Theorem 2, which shows the reason why Theorem 2 fails when $\mathbb{P}(|Y| \geq y) \neq o(y^{-4})$. Note that the limit (a "hairy tour", or a "millipede") is not the graph of a function unless $\Xi = \emptyset$ (which happens only when $\mathbb{P}(|Y| \geq y) = o(y^{-4})$); this is the reason for considering $\mathcal{K}$ rather than a function space.

As an illustration of the convergence in $\mathcal{K}$, consider the function g_n defined in $[0,1]$ by $g_n(x) = nx$ for $x \in [0, 1/n]$, $g_n(x) = 2 - nx$ for $x \in [1/n, 2/n]$, and 0 elsewhere on $[0,1]$. The function g_n is the archetype of functions that converge simply to $g \equiv 0$ but not uniformly. The sequence of functions g_n converges in $\mathcal{K}$ since the sequence of graphs of g_n converges to the union of the two segments $[(0,0),(0,1)]$ and $[(0,0),(1,0)]$. Note, to end this example, that g_n does not converge to g in $\mathcal{K}$.

Theorem 5. *Suppose that* $\mathbb{E}Y = 0$ *and* $\mathbb{P}(Y \geq y) = (a_+ + o(1))y^{-4}$ *and* $\mathbb{P}(Y \leq -y) = (a_- + o(1))y^{-4}$ *for some* $a_+, a_- \geq 0$. *Let* Ξ *be a Poisson process in* $[0,1] \times (\mathbb{R} \setminus \{0\})$ *with intensity* $4a_+ y^{-5} dx\, dy$ *for* $y > 0$ *and* $4a_- |y|^{-5} dx\, dy$ *for* $y < 0$, *which is independent of* (r, v). *Then* $r_n \xrightarrow{(d)} H(r, \Xi)$ *in* $\mathcal{K}$ *and* $(r_n, v_n) \xrightarrow{(d)} \big(H(r,\Xi), v\big)$ *in* $\mathcal{K} \times C[0,1]$.

Remark 3. If $\mathbb{P}(|Y| > y) = O(y^{-4})$ but the tails are less regular than in Theorem 5, similar results hold for suitable subsequences n_k.

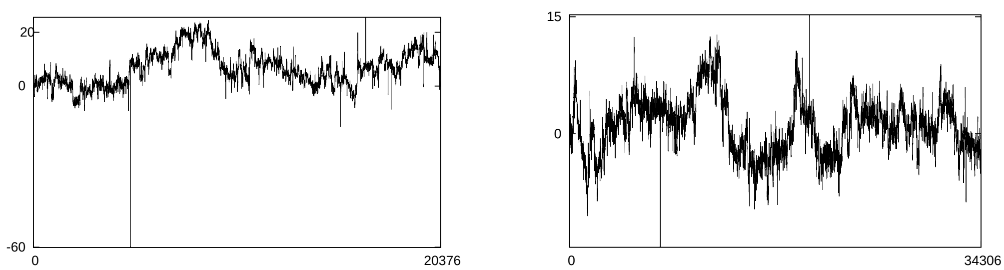

Figure 6: Simulations of R_n associated with a plane tree with 10189 nodes (respectively 17154 nodes), where Y is symmetric and satisfies $\mathbb{P}(|Y| > y) = (1+y)^{-a}$ for y non-negative and $a = 3$ for the first figure and $a = 4$ for the second one.

One can note that the convergence in $\mathcal{K}$ implies convergence of the maximal value of the second coordinate. Hence Theorem 5 implies convergence in distribution of $\max r_n$, but not to $\max r$ unless $a_+ = 0$.

If the tails of Y are even larger, the spikes dominate the Brownian part, and with a different normalization, one may get convergence in $\mathcal{K}$ to a flat millipede.

Theorem 6. *Suppose that* $\mathbb{E}Y = 0$ *and* $\mathbb{P}(Y \geq y) = (a_+ + o(1))y^{-q}$ *and* $\mathbb{P}(Y \leq -y) = (a_- + o(1))y^{-q}$ *for some* $q < 4$ *and* $a_+, a_- \geq 0$. *Let* Ξ *be a Poisson process in* $[0,1] \times (\mathbb{R} \setminus \{0\})$ *with intensity* $qa_+ y^{-q-1} dx\, dy$ *for* $y > 0$ *and* $qa_- |y|^{-q-1} dx\, dy$ *for* $y < 0$ *which is independent of* v. *Then* $n^{-1/q} R_n(2n\cdot) \xrightarrow{(d)} H(0, \Xi)$ *in* $\mathcal{K}$, *and this holds jointly with* $v_n \xrightarrow{(d)} v$.

Remark 4. Theorem 6 is easily extended to regularly varying tails with exponent < 4 (see [5] for definitions); we leave the details to the reader.

A consequence of Theorem 6 is that

$$\max n^{-1/q} R_n \xrightarrow{(d)} \max \Xi, \tag{9}$$

the distribution of $\max \Xi$ being given by

$$\mathbb{P}(\max \Xi \le x) = \exp(-a_+ x^{-q}) \text{ for } x \ge 0 \tag{10}$$

since $\#\Big(\Xi \cap [0,1] \times [x, +\infty[\Big)$ is a $\mathrm{Po}(\int_x^{+\infty} q a_+ y^{-q-1} dy)$ distributed random variable.

2.2 Case $\mathbb{E}Y \neq 0$

We assume now that $m = \mathbb{E}Y \neq 0$. The variable $Y - m$ is centered and hence we can apply the results of the previous subsection to $R_n(t, Y - m)$. Note that $n^{-1/2} R_n(2n\cdot, m) = m v_n$ converges in $C[0,1]$ to mv. The convergence of $R_n(t,Y) = R_n(t, Y-m) + R_n(t,m)$ depends then of the tails of Y. For example, suppose that Y have regular tails as in Theorem 6 for some q, with $a_+ + a_- \neq 0$. The critical parameter that makes $R_n(t,m)$ and $R_n(t, Y-m)$ to be of the same order then is $q = 2$. More precisely, we have

Theorem 6′. *If $q < 2$, Theorem 6 holds also when $\mathbb{E}Y \neq 0$.*

Theorem 7. *If the assumptions of Theorem 6 hold with $q = 2$, but $\mathbb{E}Y = m$ may be non-zero, then $n^{-1/2}(R_n(2n\cdot), V_n(2n\cdot))$ converges in $\mathcal{K} \times C[0,1]$ to $(H(mv, \Xi), v)$ where Ξ is described in Theorem 6.*

Theorem 8. *If $\mathbb{P}(|Y| > y) = o(y^{-2})$, then $n^{-1/2}(R_n(2n\cdot), V_n(2n\cdot)) \xrightarrow{(d)} (mv, v)$ in $C[0,1]^2$.*

As consequences, we recover (assuming (A)) the main results of Durrett & al. [10] on the maximum of R_n. If $\mathbb{P}(|Y| > y) = o(y^{-2})$, then Theorem 8 implies that $n^{-1/2} R_n \xrightarrow{(d)} \frac{2m}{\sigma} \max e$ (Theorem 1 of [10]). If the tails of Y are regular as in Theorem 6 with $q < 2$ and $a_+ > 0$, then, by Theorem 6′, the maximum of R_n is described by formulas (9) and (10) (Theorem 2 of [10], which also allows regular varying tails); it can be seen from the proofs that we can weaken the condition on $\mathbb{P}(Y \le -y)$. In the case $q = 2$, we obtain from Theorem 7 convergence to the law of the maximum of $H(mv, \Xi)$ (the maximum of the projection on the second coordinate of $\mathbb{R}^2$); this distribution seems difficult to describe [10, Remark 3]. (See also Kesten [14, 15] for related results.)

Note also that if $Y \ge 0$, we can regard $y(u)$ as the length of the edge leading up to u; R_n then is the DFW on a tree with random edge lengths. Theorem 8 shows that if $\mathbb{E}Y = 1$ and $\mathbb{P}(|Y| > y) = o(y^{-2})$, then this walk converges as when the edge lengths are constant. By Aldous [2], Theorem 20 and Corollary 19, this implies that set representations of the tree (after scaling by $\sigma_\xi n^{-1/2}$) converge to the Brownian continuum random tree, thus answering a question by Aldous [2, Remark in 5.3] whether a second moment is enough for this.

2.3 Case $\mathbb{E}(|Y|) = +\infty$

With regular tails as in Theorem 6, this implies $q \le 1$. In this case, the result still holds.

Theorem 6″. *For $q \le 1$, Theorem 6 holds also when $\mathbb{E}Y$ does not exist.*

3 Convergence of the discrete snake

Consider (r_n, v_n) the tour of the discrete snake and (w_n, v_n) the discrete snake. It is clear that knowledge of (r_n, v_n) implies knowledge of (w_n, v_n) and *vice versa*. As said above, if the convergence of (r_n, v_n) holds in $C[0,1]^2$ then the convergence of (w_n, v_n) holds in $C([0,1]\times\mathbb{R}^+)\times C[0,1]$ and *vice versa* (this is a consequence of the homeomorphism theorem of [19]). But, it is not clear that convergence of the tour of the discrete snake in some other functional space will imply the convergence of the snake. An immediate corollary of Theorem 2 is that

Corollary 1. *If $\mathbb{E}Y = 0$ and if $\mathbb{P}(|Y| \geq y) = o(y^{-4})$ then (w_n, v_n) converges in $C\big([0,1]\times[0,+\infty[\big)\times C[0,1]$ to the Brownian snake with lifetime process $(2/\sigma_\xi)e$.*

To have an image of what happens, consider the random walk $W_n(2ns,\cdot)$, for $2ns$ integer. The evolution of $W_n(2ns,\cdot)$ is the one of a simple random walk on the interval $[0, V_n(2ns)]$ and so, at the limit, $W_n(2ns,\cdot)$ normalized as said in Subsection 1.3.3 will converge to a killed Brownian motion (with variance given by the variance of Y and lifetime $v(s)$). The normalization by $n^{1/4}$ comes from the height of the skeleton which is about $n^{1/2}$.

Theorem 8 provides a second case where the convergence holds in $C[0,1]$. But the limit of the discrete snake in this case is not the Brownian snake. Indeed, fix again s such that $2ns$ is an integer and examine the evolution of $W(2ns,\cdot)$. We condition on V_n. Consider two real numbers t_1 and t_2, $t_1 < t_2$ such that $\sqrt{n}t_2 < V_n(s)$. The difference $D = W(2ns,\sqrt{n}t_1) - W(2ns,\sqrt{n}t_2)$ is again the evolution of a random walk with generic increment Y, but in this case, since Y is not centered, D is concentrated around $\sqrt{n}(t_2 - t_1)\mathbb{E}Y$ (there is a drift). Hence, for any fixed s and t, $n^{-1/2}W_n(2ns,\sqrt{n}t)$ converges to $t\mathbb{E}Y$ for $0 \leq t < v(s)$ and to $v(s)\mathbb{E}Y$ if $t \geq v(s)$. In this case, the limit snake is flat. In other words, knowing v, $n^{-1/2}W_n(2ns,\sqrt{n}t)$ converges to $w(s,t) = (t \wedge v(s))\mathbb{E}Y$.

A natural question arising now is the description of a snake having for tour the hairy tour, and the convergence to the *jumping snake.*

3.1 Hairy tour and jumping snake

Let (f,ζ) be an element of $C([0,1])^2$ such that f is the head of a snake with lifetime process ζ. Hence, f is assumed to be *compatible* with ζ, that is, for $0 \leq s \leq t \leq 1$, if $\zeta(s) = \zeta(t) = \min_{u\in[s,t]}\zeta(u)$ then $f(s) = f(t)$ (this translates the fact that in the tree coded by ζ, the two real numbers s and t code the same node, and so the head of the snake coded by f should take one unique value on these points). The snake with tour (f,ζ) is the function $(\nu,\zeta) \in C([0,1]\times[0,+\infty])\times C[0,1]$ defined as follows: For any $s\in[0,1]$ and $t\in[0,+\infty]$, define

$$\rho(s,t,\zeta) = \begin{cases} \sup\{\alpha \leq s,\ \zeta(\alpha) = t\} & \text{if } 0 \leq t \leq \zeta(s), \\ s & \text{if } t \geq \zeta(s). \end{cases}$$

The function ρ plays an important role in the homeomorphism theorem of [19]; it allows us to express the snake with the tour of the snake. The snake is

$$\nu(s,t) = f(\rho(s,t,\zeta)) \quad \text{for} \quad (s,t) \in [0,1]\times[0,+\infty].$$

All nodes of the path from the root to the node coded by $v(s)$ are also terminal points in the branching random walk. The function ρ allows us to find their positions in the tree (r only gives the values of terminal points); we refer to [19] for additional explanations.

Since we can not work in the set of continuous functions (because of the hairs), we denote by $\Gamma(f,\zeta)$ the graph of ν:

$$\Gamma(f,\zeta) = \{\, (s,t,\nu(s,t))\,,\ (s,t) \in [0,1] \times [0,+\infty]\}.$$

We let $\mathcal{K}' = \mathcal{K}([0,1]\times[0,+\infty]\times\mathbb{R})$ be the space of non-empty compact subsets of $[0,1]\times[0,+\infty]\times\mathbb{R}$ equipped with the Hausdorff metric (8). We allow here $t=+\infty$ so that $\Gamma(f,\zeta)$ is a compact set; since $(s,t) = \nu(s,\zeta(s))$ for $t \geq \zeta(s)$, this is just a technical convenience. (Alternatively, we could use the space of closed sets in $[0,1] \times [0,+\infty[\times\mathbb{R})$ with the Fell topology [13, Appendix A.2].) The function $(f,\zeta) \mapsto \Gamma(f,\zeta)$ is continuous from $C[0,1]^2$ into $\mathcal{K}'$.

Let $A \subset [0,1] \times [0,+\infty] \times \mathbb{R}$, $\zeta \in C[0,1]$, and $B \subset [0,1] \times \mathbb{R}$ with $B \cap \big([0,1] \times (\mathbb{R} \setminus [-a,a])\big)$ finite for every $a > 0$. We set

$$(A,\zeta) \oplus B := \left\{ (x,y,z+a) : (x,y,z) \in A \text{ and } \begin{cases} a = 0 & \text{if } y < \zeta(x) \\ a \in [0,a^\star] \text{ or } [a^\star,0] & \text{if } y \geq \zeta(x),\ (x,a^\star) \in B \end{cases} \right\}$$

We define the jumping snake $J(f,\zeta,X)$ with set of jumps $X \subset [0,1]\times\mathbb{R}$, lifetime process ζ, and continuous head process f (compatible with ζ), as the compact set

$$J(f,\zeta,X) = (\Gamma(f,\zeta),\zeta) \oplus X.$$

A simulation

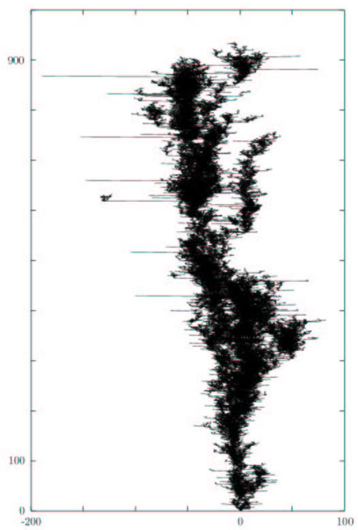

Figure 7: Discrete branching random walk (size 180 000) where Y is symmetric and satisfies $\mathbb{P}(|Y| > y) = (1+y)^{-5/2}$ for $y \geq 0$.

Figure 7 represents a non-normalized branching random walk. The set of points represented in this picture is the union of the graphs $(W_n(k,t),t)$, $t \leq V_n(k)$ for all k. It is a projection of the discrete snake *alive* (under the lifetime process). One can see that all big jumps are done at leaves or displace only small subtrees. It is also clear that the occupation measure of this branching random walk is almost the same as that without the displaced subtrees (this illustrates Theorems 1 and 4). Finally, one guesses (a little bit) that the jumps are larger that the "continuous part". The limit picture when the size goes to $+\infty$ is a vertical line with horizontal hairs. To finish, to have a good picture of a branching random walk with displacements having moment of order 4, just shave the hairs in Figure 7.

Remark 5 (About the name "jumping snake"). At first, recall that the name Brownian snake comes from the interpretation of $(w(s,t), 0 \le t \le v(s))$ as the body of a snake at time s; the head of the snake is at point $w(s, v(s))$. The head of the snake in the time interval $[s, s+ds]$ moves forward in a new direction if v increases in $s+$ (the remaining of its body does not move) and comes back in its tracks if $v(s)$ decreases in $s+$ (of course, since v is a Brownian excursion, v is not really increasing or decreasing in any interval of positive length). In the case of the jumping snake, at any point (s,t) such that $0 \le t < v(s)$, $w(s,t) = w(s_-, t) = w(s_+, t)$ but for the point $(s, a(s))$ in Ξ (that correspond of the hairs of the tour), we have

$$w(s, v(s)) = w(s_-, v(s-)) + a(s) = w(s_+, v(s)) + a(s).$$

In our interpretation in term of snake, we see that the snake makes a jump of size $a(s)$ at time s and come back instantaneously to its position before the jump.

Theorem 9. *If the hypotheses of Theorem 5 are fulfilled, then $\Gamma(r_n, v_n)$ converges weakly in $\mathcal{K}'$ to $J(r, v, \Xi)$ where Ξ is described in Theorem 5.*

Theorem 10. *If the hypotheses of Theorem 6, 6′ or 6″ are fulfilled, then $\Gamma(n^{-1/q} R_n(2n\cdot), v_n)$ converges weakly in $\mathcal{K}'$ to $J(0, v, \Xi)$, where Ξ is described in Theorem 6.*

4 Proofs

Lemma 1. *There exists $\alpha > 0$ such that for every $\varepsilon > 0$ there exists a finite real number C_ε such that*

$$\mathbb{P}(|v_n(s) - v_n(t)| \le C_\varepsilon |t-s|^\alpha \textit{ for all } s,t \in [0,1]) \ge 1 - \varepsilon.$$

In other words, the sequence $\|v_n\|_{H_\alpha}$ of Hölder norms is tight.

Proof. Gittenberger [12] proved (in a stronger form) that for all s, t, $\varepsilon > 0$

$$\mathbb{P}(|v_n(s) - v_n(t)| \ge \varepsilon) \le C_1 |s-t|^{-1} \exp\bigl(-C_2 \varepsilon |s-t|^{-1/2}\bigr),$$

which immediately gives, for any $p > 0$,

$$\mathbb{E}|v_n(s) - v_n(t)|^p \le C(p) |s-t|^{p/2-1}.$$

Taking $p = 6$, the result follows by Kolmogorov's continuity criterion [13, Theorem 3.23] (and its proof, to get uniformity in n), cf. [4, Theorem 12.3]. (Indeed, taking p large enough, we obtain the result for any $\alpha < 1/2$.)

An alternative, more probabilistic, proof runs as follows. (We only sketch the details.) We may assume $\alpha < 1/4$, and then [18, Theorem 2] implies that it is equivalent to prove the statement for the depth first queue process instead of the DFW V. The depth first queue process is a random walk $\{S_k\}_0^{n+1}$ with $S_0 = 0$ and i.i.d. increments distributed as $\xi - 1$, conditioned on $S_{n+1} = -1$ and $S_j \ge 0$ for $0 \le j \le n$. It is well known [11] (and easy to prove) that if

$\sum_1^{n+1} \xi_i = -1$ (with $\xi_i \geq -1$ integers), exactly one conjugation (i.e. a cyclic shift) of $\{\xi_i\}_1^{n+1}$ yields a non-negative sequence of partial sums $\{S_j\}_1^n$. Hence we can construct our $\{S_j\}$ by conditioning on $S_{n+1} = -1$ only, followed by the proper conjugation. Since conjugation changes the Hölder norm by at most a factor 2 (because of wrapping around), it is sufficient to prove the corresponding result for (normalized) partial sums conditioned on $S_{n+1} = -1$. It is, by exchangeability, enough to take $0 \leq s \leq t \leq 1/2$, which means that we only have to look at the first half of the walk $\{S_j\}_1^n$. But it is easy to see, by conditioning on $S_{\lceil n/2 \rceil}$, that the probability of any event for this part of the walk is bounded by a constant times the same probability for a random walk without conditioning. For such walks, the result is a well-known consequence of Kolmogorov's continuity criterion [13, Theorem 3.23]. □

Remark 6. This is the only place where we use the assumption $\mathbb{E}e^{\alpha\xi} < \infty$. (It is used in both proofs, although it probably could be replaced by existence of a sufficiently high moment in [18].) Hence, a proof of this lemma for all (or at least some other) ξ with $\mathbb{E}\xi^2 < \infty$ would extend our results to that case. Note that Aldous [2] proves tightness of v_n assuming only $\mathbb{E}\xi^2 < \infty$; it is not easy to see whether his proof can be strengthen to yield Hölder continuity.

Proof of Theorem 1. Let $x_1, \ldots, x_k$ be fixed with $0 \leq x_1 < x_2 < \cdots < x_k \leq 1$. We know by Aldous [2] that $v_n \xrightarrow{(d)} v$ in $C[0,1]$. By the Skohorod coupling theorem (see e.g. [13, Theorem 4.30]), we may assume that the random trees τ_n are realized on the same probability space and that $v_n \longrightarrow v$ in $C[0,1]$ almost surely. Let T_n be the subtree of τ_n spanned by the root and the nodes $(f(\lfloor 2nx_i \rfloor))_{i \in [\![1,k]\!]}$, and let T be the corresponding subtree of the continuum random tree defined by v (see Aldous [1, 2]). Now, T is determined by the vector

$$W(v; x_1, \ldots, x_k) = (v(x_1), \check{v}(x_1, x_2), v(x_2), \ldots, \check{v}(x_{k-1}, x_k), v(x_k))$$

and T_n by the corresponding vector $W(v_n; x_1^{(n)}, \ldots, x_k^{(n)})$ where $x_i^{(n)} = \lfloor 2nx_i \rfloor / 2n$. It follows from $v_n \longrightarrow v$ a.s. that $W(v_n; x_1^{(n)}, \ldots, x_k^{(n)})$ converges a.s. to $W(v; x_1, \ldots, x_k)$. Thus, a.s., for large n, T_n has the same shape as T, and the lengths of its branches divided by $\sqrt{n}$ converge to the corresponding lengths of the branches of T. By the standard central limit theorem applied to each branch, it follows that conditioned on v, the increments of Φ along the branches, divided by $\sigma_Y n^{1/4}$, converge jointly to independent normal variables that can be regarded as the increments of a Brownian motion on T along the branches. This implies that unconditionally,

$$(r_n(x_1), \ldots, r_n(x_k), v_n(x_1), \ldots, v_n(x_k)) \xrightarrow{(d)} (r(x_1), \ldots, r(x_k), v(x_1), \ldots, v(x_k)).$$ □

Truncation

Fix a sequence $\{b_n\}$ and truncate Y_i as follows:

$$\begin{aligned} Y'_{ni} &= Y_i \, \mathbb{1}_{|Y_i| \leq b_n}, \\ Y''_{ni} &= Y_i - Y'_i = Y_i \, \mathbb{1}_{|Y_i| > b_n}. \end{aligned}$$

We similarly truncate the generic variable Y to Y'_n and Y''_n.

We write $R_n(t; \{Z_i\})$ for the head of the discrete snake obtained by replacing the values $y(u)$ (with the vertices in depth-first order, say) by $Z_1, \ldots, Z_n$. When all Z_i are equal to the same constant β, we also write $R_n(t;\beta)$. We similarly write $r_n(t;\{Z_i\}) = n^{-1/4}R_n(2nt;\{Z_i\})$ and $r_n(t;\beta)$ for the normalized head. We further write

$$\begin{aligned} R'_n(t) &= R_n(t;\{Y'_{ni}\}), & R''_n(t) &= R_n(t;\{Y''_{ni}\}),\\ r'_n(t) &= r_n(t;\{Y'_{ni}\}) = n^{-1/4}R'_n(2nt), & r''_n(t) &= r_n(t;\{Y''_{ni}\}) = n^{-1/4}R''_n(2nt), \end{aligned}$$

and with centered variables (when the means exist)

$$\begin{aligned} \widetilde{R}'_n(t) &= R_n(t;\{Y'_{ni} - \mathbb{E}Y'_{ni}\}), & \widetilde{R}''_n(t) &= R_n(t;\{Y''_{ni} - \mathbb{E}Y''_{ni}\}),\\ \widetilde{r}'_n(t) &= n^{-1/4}\widetilde{R}'_n(2nt), & \widetilde{r}''_n(t) &= n^{-1/4}\widetilde{R}''_n(2nt). \end{aligned}$$

Note that one has always $r_n(t,\{Z_i+Y_i\}) = r_n(t,\{Z_i\}) + r_n(t,\{Y_i\})$ and thus $r_n = r'_n + r''_n$.

We begin with the versions with centering. We will use the following special case of Rosenthal's inequality (see Petrov [20, Theorem 2.9]):

Lemma 2. *Suppose $Z, Z_1, Z_2, \ldots$ are i.i.d. with $\mathbb{E}(Z) = 0$. Let $S_n = \sum_{i=1}^n Z_i$. Then for each $p \geq 2$,*

$$\mathbb{E}(|S_n|^p) \leq C_p\left(n\mathbb{E}(|Z|^p) + n^{p/2}(\mathbb{E}(Z^2))^{p/2}\right)$$

for a constant C_p that depends only on p. □

We will also frequently use (without further comment) the fact that if Z, Z_n and W_n are random variables in some metric space, $Z_n \xrightarrow{(d)} Z$ and $d(Z_n, W_n) \xrightarrow{(p)} 0$, then $W_n \xrightarrow{(d)} Z$ [4, Theorem 4.1].

Lemma 3. (i) *Suppose $b_n = O(n^{-\varepsilon}a_n)$ for some $\varepsilon > 0$ and $\mathbb{E}(|Y'_n|^2) = O(n^{-1/2}a_n^2)$. Then $\{a_n^{-1}\widetilde{R}'_n(2n\cdot)\}_n$ is tight in $C[0,1]$.*

(ii) *If further $\mathbb{E}(|Y'_n|^2) = o(n^{-1/2}a_n^2)$, then $a_n^{-1}\widetilde{R}'_n(2n\cdot) \xrightarrow{(p)} 0$ in $C[0,1]$.*

Proof. Let $g_n(x) = a_n^{-1}\widetilde{R}'_n(2nx)$.

(i): By Lemma 1, there exist $C > 0$ and $\alpha > 0$ such that

$$|v_n(s) - v_n(t)| \leq C|s-t|^\alpha \text{ for } s,t \in [0,1] \tag{11}$$

with probability arbitrarily close to 1. We condition on v_n, assumed to satisfy (11). Given $s,t \in [0,1]$, with $2ns$ and $2nt$ integers, let $u \in [s,t]$ with $v_n(u) = \check{v}_n(s,t)$. The path from $f(2ns)$ to $f(2nt)$ consists of two parts, one descending from $f(2ns)$ to $f(2nu)$ and one increasing from $f(2nu)$ to $f(2nt)$. The total length of the path is, using (11),

$$\begin{aligned} V_n(2ns) + V_n(2nt) - 2V_n(2nu) &= n^{1/2}\big(v_n(s) + v_n(t) - 2v_n(u)\big)\\ &\leq Cn^{1/2}(|s-u|^\alpha + |t-u|^\alpha) \leq 2Cn^{1/2}|s-t|^\alpha. \end{aligned}$$

Hence, by Lemma 2, for any fixed $p \geq 2$,

$$\begin{aligned} \mathbb{E}|g_n(s) - g_n(t)|^p &\leq c_1 a_n^{-p}\left(n^{1/2}|s-t|^\alpha\mathbb{E}|Y' - \mathbb{E}Y'|^p + n^{p/4}|s-t|^{p\alpha/2}(\mathbb{E}|Y' - \mathbb{E}Y'|^2)^{p/2}\right)\\ &\leq c_2 a_n^{-p}\left(n^{1/2}|s-t|^\alpha\mathbb{E}|Y'|^p + n^{p/4}|s-t|^{p\alpha/2}(\mathbb{E}|Y'|^2)^{p/2}\right)\\ &\leq c_2(b_n/a_n)^p n^{1/2}|s-t|^\alpha + c_3|s-t|^{p\alpha/2}\\ &\leq c_4 n^{1/2-p\varepsilon} + c_3|s-t|^{p\alpha/2}. \end{aligned}$$

Choosing p such that $p\varepsilon \geq 5/2$ and $p\alpha \geq 4$, this is at most, since $|s-t| \geq 1/(2n)$ except in the trivial case $s=t$,

$$c_4 n^{-2} + c_3|s-t|^2 \leq c_5|s-t|^2.$$

We have shown that, conditioned on v satisfying (11),

$$\mathbb{E}|g_n(s) - g_n(t)|^p \leq c_5|s-t|^2 \tag{12}$$

for all s and t such that $2ns$ and $2nt$ are integers. Since g_n is defined by linear interpolation between these points, (12) holds for all $s, t \in [0,1]$. The tightness now follows by [4, Theorem 12.3].

(ii): Again conditioning on (11), we have $|v_n(s)| \leq C|s|^\alpha \leq C$, and thus by the same argument

$$\mathbb{E}|g_n(s)|^2 \leq c_6 a_n^{-2} n^{1/2} \mathbb{E}|Y' - \mathbb{E}Y'|^2 \longrightarrow 0.$$

Hence the finite-dimensional distributions tend to 0, and the result follows by part (i). □

Lemma 4. *Suppose that* $\mathbb{E}Y^2 < \infty$ *and* $b_n \longrightarrow \infty$. *Then* $\widetilde{r}''_n(x) \overset{(p)}{\longrightarrow} 0$ *for each fixed* $x \in [0,1]$, *and* $\widetilde{r}''_n \overset{(p)}{\longrightarrow} 0$ *in* $L^2[0,1]$.

Proof. If $2nx$ is an integer, $\widetilde{R}''_n(2nx)$ is a sum of $n^{1/2}v_n(x)$ independent copies of $Y''_n - \mathbb{E}Y''_n$. It follows that, also if $2nx$ is not an integer,

$$\mathbb{E}\big(\widetilde{R}''_n(2nx)^2 \mid v_n\big) \leq n^{1/2}v_n(x)\mathbb{E}(Y''_n - \mathbb{E}Y''_n)^2 \leq n^{1/2}v_n(x)\mathbb{E}(Y''_n)^2$$

and thus

$$\mathbb{E}\big(\widetilde{r}''_n(x)^2 \mid v_n\big) \leq v_n(x)\mathbb{E}(Y''_n)^2. \tag{13}$$

Let us again condition on (11) and thus $|v_n(x)| \leq C$. We then have, since $\mathbb{E}(Y''_n)^2 \longrightarrow 0$,

$$\mathbb{E}\big(\widetilde{r}''_n(x)^2\big) \leq C\mathbb{E}(Y''_n)^2 \longrightarrow 0 \tag{14}$$

and thus $\widetilde{r}''_n(x) \overset{(p)}{\longrightarrow} 0$.

Similarly, integrating (14),

$$\mathbb{E}\big(\|\widetilde{r}''_n\|^2_{L^2[0,1]}\big) = \int_0^1 \mathbb{E}\big(\widetilde{r}''_n(x)^2\big)\,dx \leq C\mathbb{E}(Y''_n)^2 \longrightarrow 0$$

and thus $\|\widetilde{r}''_n\|_{L^2[0,1]} \overset{(p)}{\longrightarrow} 0$. □

Lemma 5. *Suppose that* $\mathbb{E}Y^2 < \infty$, *and let* $b_n = n^{1/4-\varepsilon}$ *for some* $\varepsilon > 0$ *with* $\varepsilon < 1/4$. *Then* $(\widetilde{r}'_n, v_n) \overset{(d)}{\longrightarrow} (r, v)$ *in* $C([0,1])^2$.

Proof. Since $\mathbb{E}|Y'_n|^2 \leq \mathbb{E}|Y|^2 < \infty$, the conditions of Lemma 3 are satisfied with $a_n = n^{1/4}$, and thus $\{\widetilde{r}'_n\}_n$ is tight in $C[0,1]$. Hence also $\{(\widetilde{r}'_n, v_n)\}_n$ is tight.

Since $\widetilde{r}'_n(x) = r_n(x; \{Y_i - \mathbb{E}Y_i\}) - \widetilde{r}''_n(x)$, and $\widetilde{r}''_n(x) \overset{(p)}{\longrightarrow} 0$ for every fixed x by Lemma 4, Theorem 1 (applied to $Y - \mathbb{E}Y$) implies that the finite-dimensional distributions of $(\widetilde{r}'_n, v_n)$ converge to those of (r, v), and the result follows. □

We now study r'_n and r''_n, defined without centering.

Lemma 6. *Suppose that* $\mathbb{E}Y = 0$ *and* $\mathbb{E}|Y|^q < \infty$ *for some* $q > 2$*, and let* $b_n = n^{1/4-\varepsilon}$ *for some* $\varepsilon > 0$ *with* $(1-4\varepsilon)(q-1) > 1$*. Then* $(r'_n, v_n) \xrightarrow{(d)} (r, v)$ *in* $C[0,1]^2$*.*

Proof. Let $\beta_n = \mathbb{E}(Y'_n) = -\mathbb{E}(Y''_n)$. We have $\mathbb{P}(|Y| > t) = O(t^{-q})$ and thus

$$\begin{aligned}|\beta_n| = |\mathbb{E}Y''_n| \le \mathbb{E}|Y''_n| &= \int_0^\infty \mathbb{P}(|Y''_n| > t)\,dt = b_n\mathbb{P}(|Y| > b_n) + \int_{b_n}^\infty \mathbb{P}(|Y| > t)\,dt \\ &= O(b_n^{1-q}) = o(n^{-1/4}).\end{aligned} \tag{15}$$

Since β_n is non-random,

$$r_n(x;\beta_n) = n^{-1/4}R(2nx;\beta_n) = n^{-1/4}\beta_n V_n(2nx) = n^{-1/4}\beta_n n^{1/2} v_n(x) \xrightarrow{(p)} 0$$

in $C[0,1]$, because $n^{1/4}\beta_n \to 0$ by (15) and $v_n \xrightarrow{(d)} v$. Since $r'_n = \tilde{r}'_n(x) + r_n(x;\beta_n)$, the result follows from Lemma 5. ☐

We let $\mathcal{E}$ be the event that some path from the root in the tree contains more than one value $y(u)$ with $|y(u)| > b_n$.

Lemma 7. *If* $\mathbb{P}(|Y| > b_n) = o(n^{-3/4})$*, then* $\mathbb{P}(\mathcal{E}) \longrightarrow 0$*.*

Proof. A node v in T has $h(v)$ ancestors. Hence, the number of pairs (u,v) of nodes in T with u an ancestor of v equals $L = \sum_{v\in T} h(v)$, the total path length.

If $\mathcal{E}$ happens, there is such a pair (u,v) with $|y(u)|, |y(v)| > b_n$, so given T, $\mathbb{P}(\mathcal{E} \mid T) \le L\mathbb{P}(|Y| > b_n)^2$. Hence, for any $A > 0$,

$$\mathbb{P}(\mathcal{E}) \le \mathbb{P}(L > An^{3/2}) + An^{3/2}\big(\mathbb{P}(|Y| > b_n)\big)^2 = \mathbb{P}(n^{-3/2}L > A) + o(1). \tag{16}$$

It is well known that $n^{-3/2}L$ converges in distribution [1, 2]; indeed,

$$L = \frac{1}{2}\int_0^{2n} V_n(t)\,dt + \frac{n}{2} = n\int_0^1 n^{1/2}v_n(x)\,dx + \frac{n}{2},$$

so

$$n^{-3/2}L \xrightarrow{(d)} \int_0^1 v(x)\,dx. \tag{17}$$

Hence (16) yields

$$\limsup_{n\to\infty} \mathbb{P}(\mathcal{E}) \le \mathbb{P}\Big(\int_0^1 v \ge A\Big),$$

and the result follows by letting $A \to \infty$. ☐

Proof of Theorem 2. Suppose first that (i) does not hold. Then, for some $\delta > 0$, at least for n in a subsequence n_k,

$$\mathbb{P}(|Y| > n^{1/4}) \ge \delta n^{-1}.$$

Consequently, if N is the number of $y(u)$ with $|y(u)| > n^{1/4}$, then

$$\mathbb{P}(N = 0) \le (1 - \delta n^{-1})^n \longrightarrow e^{-\delta}.$$

Hence, with probability at least $1 - e^{-\delta} + o(1)$, at least one $y(u)$ satisfies $|y(u)| > n^{1/4}$, which means that $|R_n(j) - R_n(j+1)| > n^{1/4}$ for some j and thus $|r_n(s) - r_n(t)| > 1$ for some s and t with $|s - t| = 1/2n$. It follows from [4, Theorem 8.2] that $(r_n)_n$ cannot be tight. Hence (ii) and (iii) do not hold.

Conversely, suppose that (i) holds. Choose $b_n = n^{1/4-\varepsilon}$ for a small $\varepsilon > 0$ such that $1 - 4\varepsilon > 3/4$. Lemma 6 applies, with $q = 3$, say, so to prove (ii) and (iii) it suffices to show

$$r''_n \xrightarrow{\text{(p)}} 0 \quad \text{in } C[0,1]. \tag{18}$$

To show this, note that $\mathbb{P}(|Y| > b_n) = O(b_n^{-4}) = o(n^{-3/4})$, so Lemma 7 applies. If the event $\mathcal{E}$ does not happen, every sum $R''_n(k) = R_n(k; \{Y''_{ni}\})$ contains at most one non-zero term, and thus

$$\max_x |R_n(x; \{Y''_{ni}\})| = \max_{i \le n} |Y''_{ni}|.$$

Hence, for every $\delta > 0$,

$$\begin{aligned}\mathbb{P}\big(\sup |r''_n| > \delta\big) &\le \mathbb{P}(\mathcal{E}) + \mathbb{P}\big(n^{-1/4} \max_{i \le n} |Y''_{ni}| > \delta\big) \le \mathbb{P}(\mathcal{E}) + n\mathbb{P}\big(|Y''_n| > \delta n^{1/4}\big) \\ &\le \mathbb{P}(\mathcal{E}) + n\mathbb{P}\big(|Y| > \delta n^{1/4}\big) \longrightarrow 0,\end{aligned}$$

by Lemma 7 and the assumption.

Consequently, (18) holds, which completes the proof. ∎

Proof of Theorem 3. Take $b_n = n^{1/8}$. By Lemma 5, $(\tilde{r}'_n, v_n) \longrightarrow (r, v)$ in $C[0,1] \times C[0,1]$, and thus in $L^2[0,1] \times C[0,1]$. Moreover, Lemma 4 shows that $\tilde{r}''_n \xrightarrow{\text{(p)}} 0$ in $L^2[0,1]$, and the result follows because $r_n = r'_n + r''_n$. ∎

Proof of Theorem 4. To prove $\nu_n \xrightarrow{\text{(d)}} \mu$, it is sufficient to prove $\int g \, d\nu_n \xrightarrow{\text{(d)}} \int g \, d\mu$ for every bounded, uniformly continuous function g on $\mathbb{R}$, see [13, Theorem 16.16]; by approximating g by a Lipschitz function (using [4, Theorem 4.2]), we can further assume that $|g(x) - g(y)| \le C|x - y|$ for some C and all x, y. By (6) and (7), we thus have to show

$$\frac{1}{n} \sum_{i=1}^{n} g\Big(\sigma_\xi^{1/2} \sigma_Y^{-1} n^{-1/4} \Phi_{v_i}(h(v_i))\Big) \xrightarrow{\text{(d)}} \int_0^1 g\Big(\sigma_\xi^{1/2} \sigma_Y^{-1} r(t)\Big) dt. \tag{19}$$

Let us for convenience write $\bar{g}(x) = g\big(\sigma_\xi^{1/2} \sigma_Y^{-1} x\big)$ and $\Psi(u) = \Phi_u(h(u))$. Using the fact that $\tilde{f}(t)$ defined in (1) is uniformly distributed over $v_1, \dots, v_n$, (19) can be written

$$\frac{1}{2n} \int_0^{2n} \bar{g}\big(n^{-1/4} \Psi(\tilde{f}(t))\big) \, dt \xrightarrow{\text{(d)}} \int_0^1 \bar{g}\big(r(t)\big) \, dt. \tag{20}$$

Now, $\Psi(\tilde{f}(t))$ is either $R_n(\lfloor t \rfloor)$ or $R_n(\lceil t \rceil)$, and $R_n(t)$ is defined by linear interpolation between these values. Hence

$$\big|\Psi(\tilde{f}(t)) - R_n(t)\big| \le \big|R_n(\lfloor t \rfloor) - R_n(\lceil t \rceil)\big| = 3 \int_{\lfloor t \rfloor}^{\lceil t \rceil} \int_{\lfloor t \rfloor}^{\lceil t \rceil} |R_n(s) - R_n(u)| \, ds \, du$$

and thus, with $\bar{C} = \sigma_\xi^{1/2}\sigma_Y^{-1}C$, the Lipschitz constant of $\bar{g}$,

$$\begin{aligned}\frac{1}{2n}\int_0^{2n}\Big|\bar{g}\big(n^{-1/4}\,\Psi(\tilde{f}(t))\big) - \bar{g}\big(n^{-1/4}R_n(t)\big)\Big|\,dt &\le \frac{3\bar{C}}{2n}\iint_{|s-u|\le 1} n^{-1/4}|R_n(s)-R_n(u)|\,ds\,du \\ &= 6\bar{C}n\iint_{|x-y|\le 1/2n}|r_n(x)-r_n(y)|\,dx\,dy.\end{aligned} \tag{21}$$

We know by Theorem 3 that $r_n \xrightarrow{(d)} r$ in $L^2[0,1]$, and thus in $L^1[0,1]$; By the Skorohod coupling theorem [13, Theorem 4.30], we may assume that $\|r_n - r\|_{L^1} = \int |r_n - r| \longrightarrow 0$.

Let $\varepsilon > 0$. Since r is continuous, we have $|r(x)-r(y)| \le \varepsilon$ if $|x-y| \le 1/2n$ when n is large enough. For such n,

$$\begin{aligned} n\iint\limits_{|x-y|\le 1/2n}|r_n(x)-r_n(y)|\,dx\,dy &\le n\iint\limits_{|x-y|\le 1/2n}|r(x)-r(y)|\,dx\,dy + n\iint\limits_{|x-y|\le 1/2n}|r_n(x)-r(x)|\,dx\,dy \\ &\quad + n\iint\limits_{|x-y|\le 1/2n}|r_n(y)-r(y)|\,dx\,dy \\ &\le \varepsilon + 2\int_0^1 |r_n(x)-r(x)|\,dx \longrightarrow \varepsilon.\end{aligned}$$

Since ε was arbitrary, this shows that the right hand side in (21) tends to 0, so to show (20), it suffices to show that

$$\frac{1}{2n}\int_0^{2n}\bar{g}\big(n^{-1/4}R_n(t)\big)dt = \int_0^1 \bar{g}\big(r_n(t)\big)dt \longrightarrow \int_0^1 \bar{g}\big(r(t)\big)dt.$$

This is immediate consequence of $r_n \to r$ and the Lipschitz condition on g. □

Lemma 8. *Let u be a random node in T, and let T_u be the subtree consisting of u and its descendants. Then $|T_u|/n \xrightarrow{(p)} 0$.*

Proof. Clearly, $v \in T_u$ if and only if $u = v$ or u is an ancestor of v, so, see the proof of Lemma 7,

$$\sum_{u\in T}|T_u| = |T| + L.$$

Hence, conditioned on T,

$$\mathbb{E}(|T_u| \mid T) = \frac{|T|+L}{|T|} = 1 + \frac{L}{n+1}$$

and, by Markov's inequality,

$$\mathbb{P}(|T_u| > \varepsilon n \mid T) \le \frac{1}{\varepsilon n} + \frac{L}{\varepsilon n^2}.$$

Consequently, for any $\varepsilon > 0$,

$$\mathbb{P}(|T_u| > \varepsilon n) \le \frac{1}{\varepsilon n} + \frac{n^{7/4}}{\varepsilon n^2} + \mathbb{P}(L > n^{7/4}) \longrightarrow 0,$$

because $n^{-7/4}L \xrightarrow{(p)} 0$ by (17). □

Lemma 9. *Suppose that* $\mathbb{P}(Y \geq y) = (a_+ + o(1))y^{-q}$ *and* $\mathbb{P}(Y \leq -y) = (a_- + o(1))y^{-q}$ *for some* $q > 0$ *and* $a_+, a_- \geq 0$. *If* $b_n = n^{1/q-\varepsilon}$ *with* $0 < \varepsilon < 1/(4q)$, *then*

$$n^{-1/q} R''_n(2n\cdot) \xrightarrow{\text{(d)}} H(0, \Xi) \quad \text{in } \mathcal{K}, \tag{22}$$

where Ξ *is a Poisson process in* $[0,1] \times (\mathbb{R} \setminus \{0\})$ *with intensity* $qa_+ y^{-q-1} dx\, dy$ *for* $y > 0$ *and* $q_-|y|^{-q-1} dx\, dy$ *for* $y < 0$. *Moreover,* (22) *holds jointly with* $v_n \xrightarrow{\text{(d)}} v$ *(in* $C[0,1]$*), and further with* $r_n(\cdot\,; \{Y_i^{(n)}\}) \xrightarrow{\text{(d)}} r$, *with* (r, v) *independent of* Ξ, *whenever* $\{Y_i^{(n)}\}$ *are random variables independent of* $\{Y''_{ni}\}$ *such that this limit holds.*

Proof. We truncate further. We let $Y''_{ni} = \overline{Y}^\delta_{ni} + Y^\delta_{ni}$ with

$$\overline{Y}^\delta_{ni} = Y''_{ni}\, \mathbb{1}_{|Y''_{ni}| \leq \delta n^{1/q}}, \qquad\qquad Y^\delta_{ni} = Y''_{ni}\, \mathbb{1}_{|Y''_{ni}| > \delta n^{1/q}},$$

and let Ξ^δ be Ξ restricted to $[0,1] \times (\mathbb{R} \setminus [-\delta, \delta])$.

Since $H(0, \Xi^\delta)$ is obtained from $H(0, \Xi)$ by adding hairs of lengths at most δ, we have, in $\mathcal{K}$,

$$d\big(H(0, \Xi^\delta), H(0, \Xi)\big) \leq \delta,$$

and thus $H(0, \Xi^\delta) \longrightarrow H(0, \Xi)$ a.s. as $\delta \to 0$.

Similarly, if the event $\mathcal{E}$ does not happen,

$$\max_x |n^{-1/q} R_n(2nx; \{\overline{Y}^\delta_{ni}\})| = \max_{i \leq n} n^{-1/q} |\overline{Y}^\delta_{ni}| \leq \delta$$

and hence, in $C[0,1]$ and thus in $\mathcal{K}$,

$$d\big(n^{-1/q} R_n(2nx; \{Y^\delta_{ni}\}), n^{-1/q} R_n(2nx; \{Y''_{ni}\})\big) \leq \delta. \tag{23}$$

Note that Lemma 7 applies because $\mathbb{P}(|Y| > b_n) = O(b_n^{-q}) = o(n^{-3/4})$. Consequently, for any $\varepsilon > 0$ and every $\delta < \varepsilon$, (23) implies

$$\limsup_{n\to\infty} \mathbb{P}\Big(d\big(n^{-1/q} R_n(2nx; \{Y^\delta_{ni}\}), n^{-1/q} R_n(2nx; \{Y''_{ni}\})\big) \geq \varepsilon\Big) \leq \lim_{n\to\infty} \mathbb{P}(\mathcal{E}) = 0.$$

By [4, Theorem 4.2], (22) now follows if we can show

$$n^{-1/q} R_n(2nt; \{Y^\delta_{ni}\}) \xrightarrow{\text{(d)}} H(0, \Xi^\delta) \quad \text{in } \mathcal{K} \tag{24}$$

for every fixed $\delta > 0$.

Thus fix $\delta > 0$, and assume that n is so large that $b_n < \delta n^{1/q}$. Let ν be the measure on $\mathbb{R}$ with density $qa_+ y^{-q-1} dy$ for $y > 0$ and $q_-|y|^{-q-1} dy$ for $y < 0$, and let ν_δ be the restriction of ν to $\mathbb{R} \setminus [-\delta, \delta]$. Let N be the number of non-zero Y^δ_{ni}, $1 \leq i \leq n$. Then N has a binomial distribution with parameters n and

$$\mathbb{P}(Y^\delta_{n1} \neq 0) = \mathbb{P}(|Y| > \delta n^{1/q}) \sim a\delta^{-q} n^{-1},$$

where $a = a_+ + a_-$. Consequently,

$$N \xrightarrow{\text{(d)}} \operatorname{Po}(a\delta^{-q}) = \operatorname{Po}\big(\nu_\delta(\mathbb{R})\big). \tag{25}$$

The case $\nu_\delta(\mathbb{R}) = 0$, i.e. $a_+ = a_- = 0$, is now trivial, because then $N \xrightarrow{\text{(p)}} 0$ by (25). We may thus assume $\nu_\delta(\mathbb{R}) > 0$.

Let the non-zero values of $n^{-1/q}Y_{ni}^\delta$ be $X_1, \dots, X_N$ (in random order), and let the corresponding vertices be $v_1, \dots, v_N$. Suppose that $\mathcal{E}$ does not happen. Let f be the depth first traversal defined in the introduction and let $h_n(t) = n^{-1/q}R_n(2nt; \{Y_{ni}^\delta\})$. Then, for integers j, $h_n(j/2n) = X_i$ if $f(j)$ is in the subtree T_i consisting of v_i and its descendants, and $h_n(j/2n) = 0$ otherwise. By the properties of the depth first traversal, $\{j : f(j) \in T_i\}$ is an interval, say $[l_i, m_i]$, for each i, and $m_i - l_i = 2|T_i| - 2 \geq 0$. When R_n is extended to non-integer values, we see that $h_n(t)$ is 0 except on the N disjoint intervals $J_i = [(l_i - 1)/2n, (m_i + 1)/2n]$; it is 0 at the endpoints of J_i, constant X_i on the subinterval $[l_i/2n, m_i/2n]$, and linear in between.

Conditioned on N, we can construct $\{v_i\}$ and $\{X_i\}$, and thus $\{Y_{ni}^\delta\}$ and h_n by choosing N independent, uniformly distributed points t_i in $[0,1]$, taking $v_i = \tilde{f}(2nt_i)$, with $\tilde{f}$ defined in (1), and $X_i = Z_i$ where $Z_1, \dots, Z_N$ are independent copies of $n^{-1/q}Y$ conditioned on $|n^{-1/q}Y| > \delta$; note that this makes the v_i independent and uniformly distributed among the non-root vertices. Actually, to get the correct distribution we should condition on the v_i being distinct. Since the probability of a collision is $< N^2/n \to 0$ (for N fixed), the error we make can be ignored in the limit.

Let $S \subset [0,1] \times \mathbb{R}$ be the set of the N points (t_i, Z_i). Thus $H(0, S)$ is the set consisting of the horizontal interval $[(0,0), (1,0)]$ and the N vertical segments $[(t_i, 0), (t_i, Z_i)]_{i \in [\![1,N]\!]}$. Noting that $t_i \in J_i$, it follows that

$$d\big(h_n, H(0,S)\big) \leq \max_i |J_i|.$$

By Lemma 8, $|J_i| = |T_i|/n \xrightarrow{\text{(p)}} 0$ for each i, and thus

$$d\big(h_n, H(0,S)\big) \xrightarrow{\text{(p)}} 0. \tag{26}$$

We have conditioned on N, but since N converges in distribution, (26) holds also unconditionally. It is convenient to assume that (t_i, Z_i) are defined as above for all $i \geq 1$, independent of each other and of N, with $S = \{(t_i, Z_i)\}_1^N$.

Recall that the variables N, t_i and Z_i depend on n, although we do not show that in the notation, and that the distribution of t_i is uniform on $[0,1]$ and the distribution of Z_i is the conditional distribution of $n^{-1/q}Y$ given $|Y| > \delta n^{1/q}$. For $y \geq \delta$,

$$\mathbb{P}(Z_i > y) = \frac{\mathbb{P}(Y > yn^{1/q})}{\mathbb{P}(|Y| > \delta n^{1/q})} \longrightarrow \frac{a_+ y^{-q}}{a\delta^{-q}} = \frac{\nu_\delta]y, \infty[}{\nu_\delta(\mathbb{R})},$$

and similarly for $\mathbb{P}(Z_i < -y)$. It follows that Z_i converges in distribution to a random variable Z_i' with the distribution $\nu_\delta/\nu_\delta(\mathbb{R})$. By the Skorohod coupling theorem (see e.g. [13, Theorem 4.30]) we can assume that actually $Z_i \to Z_i'$ a.s. for every i as $n \to \infty$, and similarly, using (25), that $N \to N' \sim \mathrm{Po}(\nu_\delta(\mathbb{R}))$; we can also assume that we use the same t_i for every n. Then, a.s.,

$$S = \{(t_i, Z_i)\}_1^N \longrightarrow S' = \{(t_i, Z_i')\}_1^{N'} \quad \text{in } \mathcal{K},$$

and it is easily seen that this implies

$$H(0,S) \longrightarrow H(0,S') \quad \text{in } \mathcal{K}.$$

Together with (26), this shows

$$n^{-1/q} R_n(2nt; \{Y_{ni}^{\delta}\}) = h_n \xrightarrow{\text{(p)}} H(0, S') \quad \text{in } \mathcal{K}.$$

However, S' is constructed as a $\mathrm{Po}(\nu_\delta(\mathbb{R}))$ number of independent points (t_i, Z_i') with the distribution $dt \times \nu_\delta/\nu_\delta(\mathbb{R})$ on $[0,1] \times \mathbb{R}$; this means that S' is a Poisson process with intensity $dt \times \nu_\delta$, and thus equals (in distribution) Ξ^δ (see also [13, Proposition 16.17]). Consequently, we have shown (24) and thus (22).

The final claim on joint convergence holds because we can choose all t_i and Z_i independent of v_n and $\{Y_i^{(n)}\}$. □

Proof of Theorem 5. Let $b_n = n^{1/4-1/100}$. Then Lemma 6 (with $q = 3$) yields

$$r_n' \xrightarrow{\text{(d)}} r \quad \text{in } C[0,1] \tag{27}$$

and Lemma 9 (with $q = 4$) yields

$$r_n'' \xrightarrow{\text{(d)}} H(0, \Xi) \quad \text{in } \mathcal{K}.$$

Unfortunately, r_n' and r_n'' are not independent. We therefore modify the truncation as follows. Let Z be a random variable whose distribution equals the conditional distribution of Y given $|Y| \le b_n$. We let $Z_1, Z_2, \dots$ be copies of Z, independent of each other, $\{Y_i\}$, and τ_n, and define

$$\widehat{Y}_{ni} = \begin{cases} Y_{ni}' = Y_i & \text{if } |Y_i| \le b_n, \\ Z_i & \text{if } |Y_i| > b_n. \end{cases}$$

(Another way to describe this is that if Y_i is too large, we resample until we get an acceptable value.) Note that each $\widehat{Y}_{ni}$ has the same distribution as Z, and that they are independent of $\{Y_{ni}''\}$. Define

$$\hat{r}_n(t) = n^{-1/4} R_n(2nt; \{\widehat{Y}_{ni}\}),$$

and note that

$$\hat{r}_n(t) - r_n'(t) = n^{-1/4} R_n(2nt; \{\widehat{Y}_{ni} - Y_{ni}'\}).$$

Now $\widehat{Y}_{ni} - Y_{ni}'$ is non-zero only if $|Y_i| > b_n$, and if $\mathcal{E}$ does not happen, this occurs at most once on each path. Further, $|Z| \le b_n$ and thus $|\widehat{Y}_{ni} - Y_{ni}'| \le b_n$. Consequently, using Lemma 7,

$$\mathbb{P}(\sup |\hat{r}_n - r_n'| > n^{-1/4} b_n) = \mathbb{P}\big(\sup_t |R_n(t; \{\widehat{Y}_{ni} - Y_{ni}'\})| > b_n\big) \le \mathbb{P}(\mathcal{E}) \longrightarrow 0,$$

and thus

$$\hat{r}_n - r_n' \xrightarrow{\text{(p)}} 0 \quad \text{in } C[0,1]. \tag{28}$$

By (27) and (28), $\hat{r}_n \xrightarrow{\text{(d)}} r$. Since $\{\widehat{Y}_{ni}\}$ and $\{Y_{ni}''\}$ are independent, Lemma 9 yields

$$(\hat{r}_n, r_n'') \xrightarrow{\text{(d)}} \big(r, H(0, \Xi)\big) \quad \text{in } C[0,1] \times \mathcal{K},$$

and (28) again yields

$$(r_n', r_n'') \xrightarrow{\text{(d)}} \big(r, H(0, \Xi)\big) \quad \text{in } C[0,1] \times \mathcal{K}. \tag{29}$$

We define an addition in $\mathcal{K}$ by

$$K_1 \dotplus K_2 = \{(x, y_1 + y_2) : (x, y_1) \in K_1,\ (x, y_2) \in K_2\}, \tag{30}$$

that is, Minkowski addition in the second coordinate. It is easy to see that $K_1 \dotplus K_2$ is compact (but possibly empty). It is also easy to see that $\dotplus$ is a continuous map $C[0,1] \times \mathcal{K} \to \mathcal{K}$. (It is *not* continuous $\mathcal{K} \times \mathcal{K} \to \mathcal{K}$.) Finally, for (graphs of) continuous functions, $\dotplus$ equals the usual addition. Hence, (29) implies

$$r_n = r_n' + r_n'' = r_n' \dotplus r_n'' \xrightarrow{\text{(d)}} r \dotplus H(0, \Xi) = H(r, \Xi) \quad \text{in } \mathcal{K}.$$

The same argument shows joint convergence of r_n and v_n. □

Proof of Theorems 6, 6′, 6″. Let $b_n = n^{1/q-\varepsilon}$ with $\varepsilon = 1/100$. We use the decomposition

$$n^{-1/q} R_n(2n\cdot) = n^{-1/q} R_n''(2n\cdot) + n^{-1/q} \widetilde{R}_n'(2n\cdot) + n^{-1/q} R_n(2n\cdot; \mathbb{E}Y_n'). \tag{31}$$

The first term converges to $H(0, \Xi)$ in $\mathcal{K}$ by Lemma 9, so it suffices to show that the other two terms converge to 0 in $C[0,1]$. For the second term, this is Lemma 3(ii), with $a_n = n^{-1/q}$, provided we can verify the condition

$$\mathbb{E}|Y_n'|^2 = o\big(n^{-1/2+2/q}\big). \tag{32}$$

We have, by the assumptions, $\mathbb{P}(|Y| > t) = O(t^{-q})$, and thus

$$\mathbb{E}|Y_n'|^2 = 2\int_0^{b_n} t\mathbb{P}(|Y| > t)\,dt = O\left(1 + \int_1^{b_n} t^{1-q}\,dt\right) = \begin{cases} O\big(b_n^{2-q}\big) = o\big(n^{2/q-1}\big), & q < 2, \\ O\big(\log b_n\big) = O\big(\log n\big), & q = 2, \\ O(1), & q > 2. \end{cases}$$

In all three case, (32) holds, and thus the second term in (31) tends to 0.

The final term in (31) equals

$$n^{-1/q}\mathbb{E}Y_n'\, V_n(2n\cdot) = n^{1/2-1/q}\mathbb{E}Y_n'\, v_n.$$

Since $v_n \xrightarrow{\text{(d)}} v$, it thus remains to show that

$$n^{1/2-1/q}\mathbb{E}Y_n' \longrightarrow 0. \tag{33}$$

Similarly to the estimate of $\mathbb{E}|Y_n'|^2$ above, we have

$$\mathbb{E}|Y_n'| = \int_0^{b_n} \mathbb{P}(|Y| > t)\,dt = O\left(1 + \int_1^{b_n} t^{-q}\,dt\right) = \begin{cases} O\big(b_n^{1-q}\big) = o\big(n^{1/q-1}\big), & q < 1, \\ O\big(\log b_n\big) = O\big(\log n\big), & q = 1, \\ O(1), & q > 1. \end{cases}$$

Hence (33) holds when $0 < q < 2$. When $q \geq 2$, we have the additional assumption $\mathbb{E}Y = 0$, and thus $\mathbb{E}Y_n' = -\mathbb{E}Y_n''$. Hence, as in (15),

$$|\mathbb{E}Y_n'| = |\mathbb{E}Y_n''| \leq \mathbb{E}|Y_n''| = b_n\mathbb{P}(|Y| > b_n) + \int_{b_n}^{\infty} \mathbb{P}\big(|Y| > t\big)\,dt = O\left(b_n^{1-q}\right) = o\Big(n^{1/q-1+q\varepsilon}\Big),$$

and again (33) holds. □

Proof of Theorem 7. By Theorem 6 applied to $Y - \mathbb{E}Y$,

$$\big(n^{-1/2}R_n(2n\cdot;\{Y_i - m\}), n^{-1/2}V_n(2n\cdot)\big) \xrightarrow{\text{(d)}} \big(H(0,\Xi), v\big)$$

in $\mathcal{K} \times C[0,1]$. Since $R_n(x) = R_n(x;\{Y_i - m\}) + mV_n(x)$ and $\dot{+}$ defined in (30) is continuous $\mathcal{K} \times C[0,1] \to \mathcal{K}$, it follows that

$$\big(n^{-1/2}R_n(2n\cdot), n^{-1/2}V_n(2n\cdot)\big) \xrightarrow{\text{(d)}} \big(H(0,\Xi) \dot{+} mv, v\big) = \big(H(mv,\Xi), v\big).$$

in $\mathcal{K} \times C[0,1]$. □

Proof of Theorem 8. We apply Theorem 7 with $a_+ = a_- = 0$ and thus $\Xi = \emptyset$ and $H(mv,\Xi) = mv$. This shows convergence to (mv, v) in $\mathcal{K} \times C[0,1]$. Since, as remarked in Section 2, $g_n \longrightarrow g$ in $\mathcal{K}$ is equivalent to $g_n \longrightarrow g$ in $C[0,1]$ when $g_n, g \in C[0,1]$, we also have convergence in $C[0,1]^2$. □

Proof of Theorem 9. Thanks to (29), joint convergence with $v_n \longrightarrow v$, and the Skorohod coupling theorem, we may assume that $r_n' \longrightarrow r$ in $C[0,1]$, $r_n'' \longrightarrow H(0,\Xi)$ in $\mathcal{K}$, and $v_n \longrightarrow v$ in $C[0,1]$. By the homeomorphism theorem of [19], one then has

$$\Gamma(r_n', v_n) \longrightarrow \Gamma(r, v) \quad \text{in } C([0,1]\times[0,+\infty]) \subset \mathcal{K}'.$$

A simple analysis shows that

$$\Gamma(r_n'', v_n) \longrightarrow (\Gamma(0,v), v) \oplus \Xi \quad \text{in } \mathcal{K}',$$

and that

$$(\Gamma(0,v), v) \oplus \Xi = ([0,1]\times\mathbb{R}^+ \times \{0\}, v) \oplus \Xi.$$

Now, if f_1 and f_2 are continuous functions, then

$$\Gamma(f_1 + f_2, \zeta) = \Gamma(f_1, \zeta) \ddot{+} \Gamma(f_2, \zeta)$$

where $\ddot{+}$ is the Minkowski addition in the third coordinate, cf. (30). So, we have

$$\Gamma(r_n, v_n) = \Gamma(r_n', v_n) \ddot{+} \Gamma(r_n'', v_n)$$

converging to

$$\Gamma(r, v) \ddot{+} \Big(([0,1]\times\mathbb{R}^+ \times \{0\}, v) \oplus \Xi\Big) = J(r, v, \Xi)$$

since the operator $C([0,1]\times[0,+\infty]) \times \mathcal{K}' \longrightarrow \mathcal{K}'$, $(K_1, K_2) \mapsto K_1 \ddot{+} K_2$, is continuous. □

Proof of Theorem 10. By Theorems 6, 6′, 6″ and the Skorohod coupling theorem, we may assume that $\big(n^{-1/q}R_n(2n\cdot), v_n\big) \longrightarrow \big(H(0,\Xi), v\big)$ in $\mathcal{K} \times C[0,1]$ a.s. It is easy to see that this implies $\Gamma\big(n^{-1/q}R_n(2n\cdot), v_n\big) \longrightarrow J(0, v, \Xi)$ in $\mathcal{K}'$. □

Acknowledgement. This research was done while the first author visited Université de Versailles Saint-Quentin.

References

[1] D. Aldous (1991), *The continuum random tree. II: An overview.* Stochastic analysis, Proc. Symp., Durham/UK 1990, Lond. Math. Soc. Lect. Note Ser. 167, 23–70.

[2] D. Aldous (1991), *The continuum random tree. III.* Ann. Probab. 21, No.1, 248–289.

[3] D. Aldous (1993), *Tree-based models for random distribution of mass.* J. Statist. Phys. 73 625–641.

[4] P. Billingsley (1968), *Convergence of Probability Measures*, Wiley, New York.

[5] N. H. Bingham, C. M. Goldie and J. L. Teugels (1987), *Regular Variation.* Encyclopedia of Mathematics and its Applications, 27. Cambridge University Press, Cambridge.

[6] P. Chassaing, G. Schaeffer (2002), *Random planar lattices and integrated superBrownian excursion.* To appear in Prob. Theo. Rel. Fields.

[7] T. Cox, R. Durrett, E.A. Perkins (1999), *Rescaled particle systems converging to super-Brownian motion.* Perplexing problems in probability, Festschrift in honor of Harry Kesten, eds. M. Bramson et al., Prog. Probab. 44, 269–284.

[8] L. Devroye (1998), *Branching processes and their applications in the analysis of tree structures and tree algorithms.* Probabilistic methods for algorithmic discrete mathematics, eds. M. Habib et al., Algorithms Combin. 16, Springer, Berlin, 249–314.

[9] T. Duquesne, J.F. Le Gall (2002), *Random trees, Lévy processes and spatial branching processes.* Astérisque 281. Soc. Math. de France, Paris.

[10] R. Durrett, H. Kesten, E. Waymire (1991), *On weighted heights of random trees.* J. Theor. Probab. 4, No.1, 223–237.

[11] A. Dvoretzky, Th. Motzkin (1947), *A problem of arrangements*, Duke Math. J. 14, 305–313.

[12] B. Gittenberger (2003), *A note on "State spaces of the snake and its tour – Convergence of the discrete snake" by J.-F. Marckert and A. Mokkadem.* To appear in J. Theo. Prob.

[13] O. Kallenberg (2002), *Foundations of Modern Probability.* 2nd ed., Probability and Its Applications. Springer, New York, NY.

[14] H. Kesten (1994), *A limit theorem for weighted branching process trees.* The Dynkin Festschrift. Markov processes and their applications, ed. Freidlin, Birkhäuser, 153–166.

[15] H. Kesten (1995), *Branching random walk with a critical branching part.* J. Theo. Prob 8, No. 4, 921–962.

[16] J.F. Le Gall (1999), *Spatial branching processes, random snakes and partial differential equations.* Lectures in Mathematics, Birkhäuser.

[17] J.F. Marckert (2003), *The rotation correspondence is asymptotically a dilatation*, Preprint.

[18] J.F. Marckert, A. Mokkadem (2003), *The depth first processes of Galton–Watson trees converge to the same Brownian excursion.* Ann. Probab. 31, no. 3, 1655–1678.

[19] J.F. Marckert, A. Mokkadem (2002), *States spaces of the snake and of its tour – Convergence of the discrete snake.* To appear in J. Theo. Prob.

[20] V.V. Petrov (1995), *Limit Theorems of Probability Theory. Sequences of Independent Random Variables.* Oxford Studies in Probability, 4. Oxford University Press, New York.

[21] G. Slade (1999), *Lattice trees, percolation and super-Brownian motion.* Perplexing problems in probability, Festschrift in honor of Harry Kesten, eds. M. Bramson et al., Prog. Probab. 44, 35–51.