ANNALES
HENRI LEBESGUE

CYRIL MARZOUK

ON SCALING LIMITS OF RANDOM TREES AND MAPS WITH A PRESCRIBED DEGREE SEQUENCE

SUR LES LIMITES D'ÉCHELLE D'ARBRES ET DE CARTES ALÉATOIRES À DEGRÉS PRESCRITS

ABSTRACT. — We study a configuration model on bipartite planar maps in which, given n even integers, one samples a planar map with n faces uniformly at random with these face degrees. We prove that when suitably rescaled, such maps always admit nontrivial subsequential limits as $n \to \infty$ in the Gromov–Hausdorff–Prokhorov topology. Further, we show that they converge in distribution towards the celebrated Brownian sphere, and more generally a Brownian disk for maps with a boundary, if and only if there is no inner face with a macroscopic degree, or, if the perimeter is too big, the maps degenerate and converge to the Brownian tree. By first sampling the degrees at random with an appropriate distribution, this model recovers that of size-conditioned Boltzmann maps associated with critical weights in the domain of attraction of a stable law with index $\alpha \in [1, 2]$. The Brownian tree and disks then appear respectively in the case $\alpha = 1$ and $\alpha = 2$, whereas in the case $\alpha \in (1, 2)$ our results partially

recover previous known ones. Our proofs rely on known bijections with labelled plane trees, which are similarly sampled uniformly at random given n outdegrees. Along the way, we obtain some results on the geometry of such trees, such as a convergence to the Brownian tree but only in the weaker sense of subtrees spanned by random vertices, which are of independent interest.

RÉSUMÉ. — Nous étudions un modèle de configuration sur les cartes planaires biparties dans lequel on se donne n entiers pairs et l'on tire une carte planaire avec n faces uniformément au hasard parmi celles avec ces degrés. Nous démontrons que ces cartes, convenablement remises à l'échelle, admettent toujours des limites non triviales le long de sous-suites pour la topologie de Gromov–Hausdorff–Prokhorov. De plus nous montrons qu'elles convergent vers la célèbre sphère brownienne, et plus généralement un disque brownien pour des cartes à bord, si et seulement si elles ne contiennent pas de face de degré macroscopique, ou bien, si le périmètre est trop grand, les cartes dégénèrent et convergent vers l'arbre brownien. En choisissant les degrés eux-mêmes de façon aléatoire, ce modèle recouvre celui déjà étudié des cartes de Boltzmann conditionnées par la taille associées à une suite de poids critique et dans le bassin d'attraction d'une loi stable d'indice $\alpha \in [1,2]$. L'arbre et les disques browniens apparaissent alors respectivement dans les cas $\alpha = 1$ et $\alpha = 2$, tandis que dans le cas $\alpha \in]1,2[$ nous retrouvons partiellement des résultats connus. Les preuves reposent sur des bijections connues avec des arbres plans étiquetés qui sont de la même manière tirés uniformément au hasard avec n degrés donnés. Nous obtenons ainsi plusieurs résultats sur la géométrie de ces arbres intéressants par ailleurs, notamment leur convergence en loi vers l'arbre brownien, mais uniquement dans un sens faible de sous-arbres engendrés par un nombre fini de sommets.

1. Introduction

This paper deals with continuum limits of random planar maps as their size tends to infinity and their edge-length tends to zero appropriately. The most notable result in this theory has been obtained by Le Gall [LG13] and Miermont [Mie13] who proved the convergence of large random quadrangulations towards a limit called the *Brownian map* or the *Brownian sphere*. Building on these works, the Brownian sphere has then been shown to be a universal limit of many models of discrete maps and this paper continues with a large class of distributions introduced and studied previously in [Mar18b] in a restricted case.

1.1. Model and main results

A rooted planar map is a finite connected multigraph embedded in the two-dimensional sphere, in which one oriented root edge is distinguished, and viewed up to orientation-preserving homeomorphisms. The embedding allows to define the faces of the map, which are the connected components of the complement of the graph in the sphere; the degree of a face is then the number of edges incident to it, counted with multiplicity. We shall consider maps with a, non-necessarily simple, boundary, given by the face incident to the right of the root edge; the degree of this
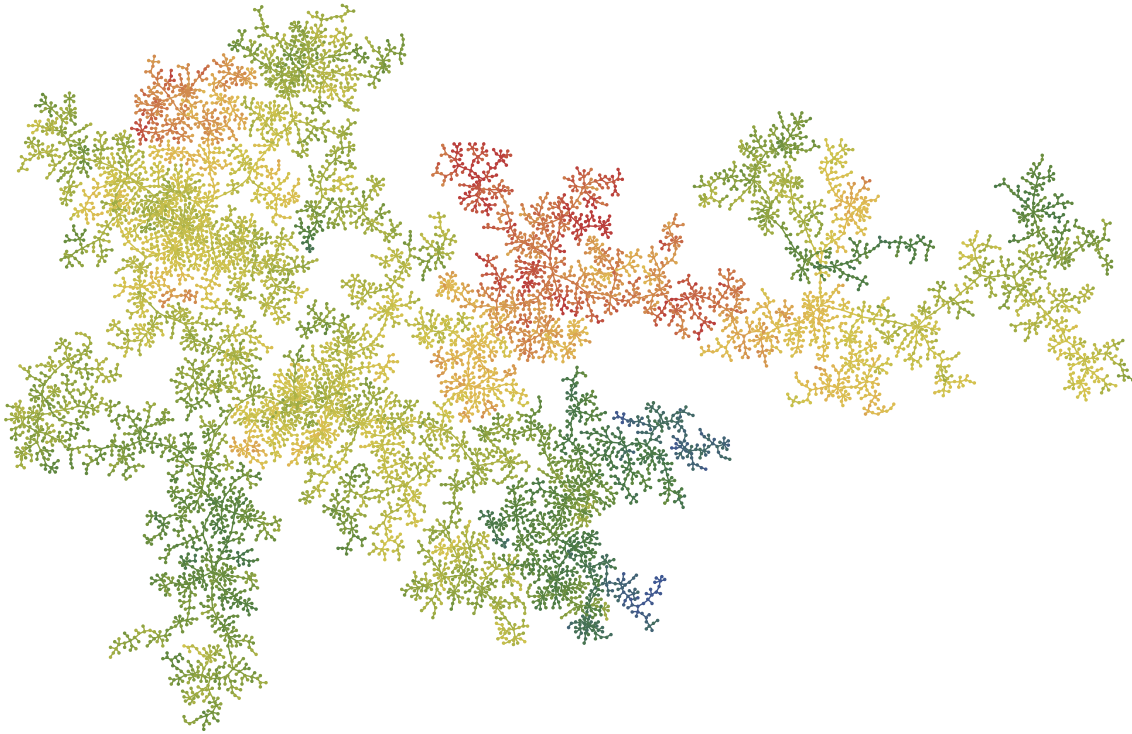
Figure 0.1. A Brownian CRT with Brownian labels encoding the Brownian sphere: labels are indicated by colours (red for the highest values and blue for the lowest).

face is called the perimeter of the map. As usual in this field, for technical reasons briefly discussed later in this section, we restrict ourselves to *bipartite* maps, in which all faces have even degree.

Random maps provide discrete models of random geometry on the sphere and an important question is their convergence towards continuum random geometries when their size tends to infinity and one rescales the edge-lengths properly. In this respect, in the case of uniformly random quadrangulations with n faces, i.e. n faces each with degree 4, Le Gall [LG07] proved that such graphs admit subsequential limits at the scaling $n^{1/4}$, and that these limits all have the same topology (later identified as the sphere [LGP08, Mie08]) and the same Hausdorff dimension 4. The problem of uniqueness of the subsequential limits was solved simultaneously by Le Gall [LG13] and Miermont [Mie13] who proved that rescaled random quadrangulations converge in distribution towards a limit called the Brownian map, or sphere, which is a compact metric measured space $M^0 = (M^0, D^0, p^0)$ thus almost surely homeomorphic to the sphere and with Hausdorff dimension 4. In the case of quadrangulations with a boundary, Bettinelli & Miermont [BM17] proved that when the perimeter behaves like $\varrho n^{1/2}$ with $\varrho \in (0, \infty)$ fixed, they converge to the *Brownian disk* with perimeter ϱ, denoted by $M^\varrho = (M^\varrho, D^\varrho, p^\varrho)$, which now has the topology of a disk, with Hausdorff dimension 4, and its boundary has Hausdorff dimension 2 [Bet15].

These works raise then the question of the universality of the limits which should not depend too much on the details of the discrete models. The present work proposes

an answer by considering a more general model that we now introduce. For every integer $n \geqslant 1$, let us give ourselves an integer $\varrho_n \geqslant 1$ and a sequence $(d_n(k))_{k \geqslant 1}$ of nonnegative integers such that $\sum_{k \geqslant 1} d_n(k) = n$ and $d_n(1) < n$ to avoid trivialities; then let $\mathbf{M}_{d_n}^{\varrho_n}$ denote the set of all (rooted planar bipartite) maps with perimeter $2\varrho_n$ and n inner faces, amongst which exactly $d_n(k)$ have degree $2k$ for every $k \geqslant 1$. A key quantity in this work is

$$\sigma_n^2 := \sum_{k \geqslant 1} k(k-1) d_n(k).$$

We stress that σ_n^2 really depends on the sequence d_n, not only on n, but we chose this lighter notation. We sample $M_{d_n}^{\varrho_n}$ uniformly at random in $\mathbf{M}_{d_n}^{\varrho_n}$ and consider the asymptotic behaviour as $n \to \infty$ of its vertex set $V(\mathbf{M}_{d_n}^{\varrho_n})$ endowed with the graph distance d_{gr} and the uniform probability measure p_{unif}.

THEOREM 1.1. — *Fix any sequence of perimeters $(\varrho_n)_{n \geqslant 1}$ and any degree sequence $(d_n)_{n \geqslant 1}$. From every increasing sequence of integers, one can extract a subsequence along which the sequence of metric measured spaces*

$$\left(V(M_{d_n}^{\varrho_n}), (\sigma_n + \varrho_n)^{-1/2} \, d_{\mathrm{gr}}, p_{\mathrm{unif}} \right)_{n \geqslant 1}$$

converges in distribution in the Gromov–Hausdorff–Prokhorov topology to a limit with a nonzero diameter.

Our first main theorem thus identifies the correct scale of the random maps and provides a general tightness result; on the other hand, without any assumptions, it cannot give any further information on the subsequential limits. In our second main theorem, we prove that the Brownian sphere and disks appear when there is no macroscopic (face-)degree, in the sense that none of them dominates the others. We let

$$\Delta_n = \max \{ k \geqslant 1 : d_n(k) \neq 0 \}$$

be the largest half-degree of an inner face in $\mathbf{M}_{d_n}^{\varrho_n}$, which we assume is always larger than or equal to 2 in order to avoid trivialities.

THEOREM 1.2. — *Assume that $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ for some $\varrho \in [0, \infty)$. Then we have the convergence in distribution in the Gromov–Hausdorff–Prokhorov topology*

$$\left(V\left(M_{d_n}^{\varrho_n} \right), \left(\frac{3}{2\sigma_n} \right)^{1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} (M^\varrho, D^\varrho, p^\varrho)$$

if and only if

$$\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0.$$

Theorem 1.2 fully recovers [Mar18b, Theorem 1] obtained under additional technical assumptions and in a restricted "finite variance" regime when $n^{-1} \sigma_n^2$ converges to a positive and finite limit. An important point is that, since the characterisation of the limit spaces M^ϱ, showing convergence of discrete models to these objects has become much easier and we stress that, although the statement extends the results of [BM17, LG13, Mie13], we rely on these works and we do not claim any independent proof of the convergence of quadrangulations.

The "only if" part of the statement is simple to prove: we shall see that an inner face with degree of order σ_n has a diameter, in the whole map, of order $\sigma_n^{1/2}$, which either creates a pinch-point or a hole, so the space cannot converge in distribution towards a limit which has the topology of the sphere or the disk. This argument also shows that in the presence of such a macroscopic face, the diameter of the rescaled map is bounded away from zero, which proves that the subsequential limits in Theorem 1.1 are indeed not reduced to a single point (and without such a large face the maps converge to the Brownian sphere, which is nontrivial as well).

The behaviour drastically changes if ϱ_n is much larger than σ_n; indeed as shown by Bettinelli [Bet15, Theorem 5] for quadrangulations, in this case, the boundary takes over the rest of the map and we obtain in the limit $T_{X^0} = (T_{X^0}, d_{X^0}, p_{X^0})$ the Brownian Continuum Random Tree of Aldous [Ald93] encoded by the standard Brownian excursion X^0.

THEOREM 1.3. — *Suppose that* $\lim_{n\to\infty} \sigma_n^{-1}\varrho_n = \infty$. *Then the convergence in distribution*

$$\left(V(M_{d_n}^{\varrho_n}), (2\varrho_n)^{-1/2}\, d_{\mathrm{gr}}, p_{\mathrm{unif}}\right) \xrightarrow[n\to\infty]{(d)} (T_{X^0}, d_{X^0}, p_{X^0})$$

holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion.

Theorems 1.2 and 1.3 apply directly to models of random maps when all faces have the same degree. For an integer $p \geqslant 2$, a $2p$-*angulation* denotes a map in which all faces have degree $2p$. In this case $\Delta_n = p$ and $\sigma_n^2 = p(p-1)n$. Then Le Gall [LG13, Theorem 1] actually proved that for any p fixed, such a $2p$-angulation with n faces sampled uniformly at random converges in distribution towards the Brownian sphere; this was extended to $2p$-angulations with a boundary by Bettinelli & Miermont [BM17, Corollary 6]. Our results allow to let p vary with n.

COROLLARY 1.4. — *Let $\varrho \in [0, \infty]$ and let $(p_n)_{n\geqslant 1} \in \{2, 3, \ldots\}^{\mathbf{N}}$ and $(\varrho_n)_{n\geqslant 1} \in \mathbf{N}^{\mathbf{N}}$ be any sequences such that $\lim_{n\to\infty}(p_n(p_n-1)n)^{-1/2}\varrho_n = \varrho$. For every $n \geqslant 1$, let $M_{n,p_n}^{\varrho_n}$ be a uniformly chosen random $2p_n$-angulation with n inner faces and with perimeter $2\varrho_n$.*

(1) *Suppose that $\varrho < \infty$, then the convergence in distribution*

$$\left(V\left(M_{n,p_n}^{\varrho_n}\right), \left(\frac{9}{4p_n(p_n-1)n}\right)^{1/4} d_{\mathrm{gr}}, p_{\mathrm{unif}}\right) \xrightarrow[n\to\infty]{(d)} (M^\varrho, D^\varrho, p^\varrho)$$

holds in the Gromov–Hausdorff–Prokhorov topology.

(2) *Suppose that $\varrho = \infty$, then the convergence in distribution*

$$\left(V\left(M_{n,p_n}^{\varrho_n}\right), (2\varrho_n)^{-1/2}\, d_{\mathrm{gr}}, p_{\mathrm{unif}}\right) \xrightarrow[n\to\infty]{(d)} (T_{X^0}, d_{X^0}, p_{X^0})$$

holds in the Gromov–Hausdorff–Prokhorov topology.

An important point about our main theorems is that the assumptions are very simple to check. We shall illustrate this with so-called size-conditioned *critical α-stable Boltzmann maps* in Section 6. This model can be seen as a mixture of our present one, where (n itself and) the sequence d_n is first sampled at random, informally as

a conditioned version of some i.i.d. random variables, and then one samples a map uniformly at random given these degrees. The size of a random Boltzmann map is not fixed but one can sample such a map conditioned to have either n faces, or n edges, or n vertices and let $n \to \infty$.

When the index α lies in $(1, 2)$, we shall prove that for some deterministic sequence a_n of order $n^{1/\alpha}$ (up to a *slowly varying function*), the ratio $a_n^{-1}\sigma_n$ converges in distribution as $n \to \infty$, so we deduce from Theorem 1.1 tightness of these maps once rescaled by $a_n^{-1/2} \approx n^{-1/(2\alpha)}$. This model was first studied by Le Gall & Miermont [LGM11] who obtained more information in addition to tightness, although proving the uniqueness of the subsequential limit is still open.

When $\alpha = 2$, we shall see that $a_n^{-1}\Delta_n \to 0$ in probability and $a_n^{-1}\sigma_n$ converges in probability to an explicit constant which depends on the law of the degrees and the notion of size (faces, edges, or vertices) where a_n is of order $n^{1/2}$ (again up to a slowly varying function). Then Theorem 1.2 implies the convergence of these maps, once rescaled by the factor $a_n^{-1/2} \approx n^{-1/4}$, to a Brownian disk. In the case of maps without boundary, this recovers [Mar18a, Theorem 1], whereas for maps with a boundary, this extends [BM17, Theorem 5] which assumes small exponential moments.

Finally, in the other extreme case $\alpha = 1$, relying on the recent work by Kortchemski & Richier [KR19], we prove that there is a unique giant face, of order some other deterministic sequence $|b_n| \approx n$ (still up to a slowly varying function), and this falls into the framework of Theorem 1.3 so the maps converge, once rescaled by a factor $|b_n|^{-1/2} \approx n^{-1/2}$ to the Brownian CRT. A similar result holds for *subcritical maps*, at the scaling exactly $n^{-1/2}$, as first shown by Janson & Stefánsson [JS15]. We refer to Section 6 for precise statements.

1.2. Random trees with a prescribed degree sequence

This model of random maps was inspired by a similar model of random trees, whose scaling limits were first studied by Broutin & Marckert [BM14] and extended to forests recently by Lei [Lei19]. Let $(\varrho_n)_{n \geqslant 1}$ and $(d_n)_{n \geqslant 1}$ be as above, then for each $n \geqslant 1$ we let $\mathbf{T}_{d_n}^{\varrho_n}$ be the set of ordered plane forests with ϱ_n trees and n internal vertices, amongst which $d_n(k)$ have k offspring for every $k \geqslant 1$. Such a forest always has

$$E_n \coloneqq \sum_{k \geqslant 1} k d_n(k), \qquad d_n(0) \coloneqq \varrho_n + \sum_{k \geqslant 1}(k-1)d_n(k), \qquad V_n \coloneqq \sum_{k \geqslant 0} d_n(k),$$

edges, leaves, and vertices respectively.

Under technical assumptions on the degree sequences $(d_n)_{n \geqslant 1}$, in particular that $\sigma_n^2 \sim \sigma^2 E_n$ for some $\sigma \in (0, \infty)$ and $\lim_{n \to \infty} \sigma_n^{-1}\Delta_n = 0$, Broutin & Marckert [BM14] showed that if T_{d_n} is a tree sampled uniformly at random in $\mathbf{T}_{d_n}^1$, then the convergence in distribution

$$(1.1) \qquad \left(V(T_{d_n}), \frac{\sigma_n}{2E_n} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} (T_{X^0}, d_{X^0}, p_{X^0})$$

holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion. More precisely, the *contour* or *height* process of the rescaled tree converges in distribution towards X^0.

Under the assumption of no macroscopic degree only, we prove a weaker convergence, in the sense of subtrees spanned by finitely many random vertices, as depicted on Figure 1.1. Fix $q \geqslant 1$ and let $u_1, \ldots, u_q$ be q i.i.d. uniform random vertices of T_{d_n} and keep only these vertices and their ancestors and remove all the other ones; further merge each chain of vertices with only one child in this new tree into a single edge with a length given by the number of edges of the former chain. The resulting tree $R_{d_n}(q)$ is called a discrete tree with edge-lengths; its combinatorial structure is that of a plane tree with at most q leaves and no vertex with outdegree 1, so there are only finitely many possibilities, and thus there are a bounded number of edge-lengths to record. We equip the space of trees with edge-lengths with the natural product topology. For $x_1, \ldots, x_q$ i.i.d. random points of T_{X^0} sampled from the mass measure p_{X^0}, one can construct similarly a discrete tree with edge-lengths $R_{X^0}(q)$; its law is described by Aldous [Ald93, Section 4.3].
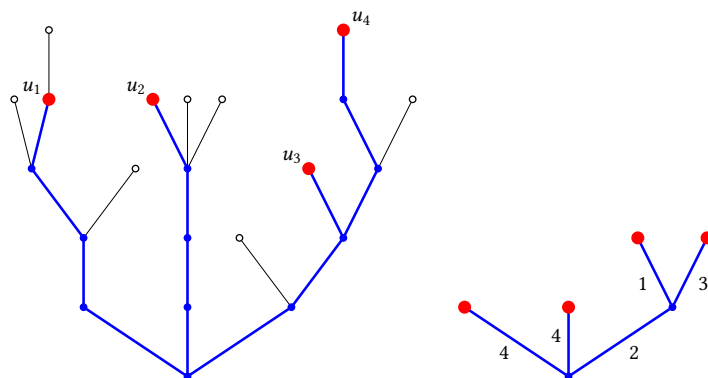

Figure 1.1. Left: a plane tree and four distinguished vertices in red. Right: the associated reduced tree with edge-lengths.

THEOREM 1.5. — *If* $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$, *then for every* $q \geqslant 1$ *we have the convergence in distribution*

$$\frac{\sigma_n}{2E_n} R_{d_n}(q) \xrightarrow[n \to \infty]{(d)} R_{X^0}(q).$$

where X^0 *is the standard Brownian excursion.*

This result holds more generally when $\varrho_n \sim \varrho \sigma_n$ for some $\varrho \in [0, \infty)$, for the forest $T_{d_n}^{\varrho_n}$, viewed as a single tree by attaching all the roots to an extra root vertex. The limit is a different continuum tree which involves a Brownian first-passage bridge from 0 to $-\varrho$; we refer to Theorem 2.4 for a statement involving the height process of the forest.

Theorem 1.5 characterises the possible scaling limits of the trees T_{d_n} in this regime with no macroscopic degrees; in order to extend the convergence from (1.1) and thus obtain an analogue of Theorem 1.2 for trees and forests, it remains to prove tightness of this sequence (see precisely [Ald93, Equation 25]). Here no general tightness result as in Theorem 1.1 holds and the maximal height of the tree can be much larger than E_n/σ_n. Indeed, similarly to Boltzmann maps, size-conditioned Bienaymé–Galton–Watson trees can be thought of as mixtures of this model where d_n is itself random;

then when the offspring distribution is subcritical, Kortchemski [Kor15] proved that the height of the tree grows logarithmically, whereas E_n/σ_n is of constant order, and the rescaled trees are not tight. Still, this regime, where the trees exhibit a so-called condensation phenomenon, with a unique huge degree, is extreme and we wonder wether beside such a case tightness holds; in particular, does (1.1) hold as soon as $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$?

There is a particular case that our method allows to treat, and which seems already new, which is that of regular trees: for $n, k \geqslant 1$, a (strict) *k-ary tree* with size n is a tree made of n vertices with k offspring each, and $n(k-1)+1$ leaves, and no other vertex. With the preceding notation, a k-ary tree $T_{n,k}$ with size n sampled uniformly at random has the law of T_{d_n} where $d_n(k) = n$ and $d_n(i) = 0$ for all $i \in \mathbf{N} \setminus \{k\}$, so $E_n = kn$ and $\sigma_n^2 = k(k-1)n$. For such trees, one can adapt our results to prove the analogue of Corollary 1.4, that is: for any sequence $(k_n)_{n \geqslant 1} \in \{2, 3, \dots\}^{\mathbf{N}}$, the convergence in distribution

$$(1.2) \qquad \left(V\left(T_{n,k_n}\right), \sqrt{\frac{k_n - 1}{4 k_n n}} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} \left(T_{X^0}, d_{X^0}, p_{X^0} \right)$$

holds in the Gromov–Hausdorff–Prokhorov topology. Actually, the contour or height process of the rescaled tree converges in distribution towards the standard Brownian excursion X^0. See Remark 4.7 for a further discussion.

1.3. Strategy of the proof and further discussion

The study of the random forests $T_{d_n}^{\varrho_n}$ is also an important part of the proof of our main results on maps. Indeed, the combination of the bijections from [BDFG04, JS15] relates maps in $\mathbf{M}_{d_n}^{\varrho_n}$ with forests in $\mathbf{T}_{d_n}^{\varrho_n}$ where each vertex carries an integer label. Thanks to the work of Le Gall [LG13] and Miermont [Mie13] our main theorems shall easily follow from the study of such random labelled forests and most of this paper is devoted to this. Let us mention at this point that the bijection from [BDFG04] holds also for maps which are not bipartite and it was used by Le Gall [LG13] to prove that random triangulations converge to the Brownian sphere. Recently, Addario-Berry & Albenque [ABA21] used it to prove that random p-angulations converge to the Brownian sphere for every $p \geqslant 5$ odd. However the labelled trees from [BDFG04] are more complicated to control and extending our present results to the non-bipartite case does not seem technically easy.

The graph distance on the map can be partially encoded by the *label process* which reads the labels of the forest in depth-first search order, and we prove that this process is always tight when suitably rescaled, see Theorem 2.5. This tightness relies only on the so-called *Łukasiewicz path* of the forest, which is a very simple process: up to a discrete Vervaat transform (also known as cyclic shift), the latter is a uniformly random path which makes $d_n(k)$ jumps of size $k-1$ for each $k \geqslant 0$. This is the simplest example of a random path with exchangeable increments and its asymptotic behaviour is well understood. From a technical point of view, tightness of the label process requires a precise control of the Łukasiewicz path obtained in

particular via Chernoff bounds for martingales. Going from tightness of the label process to tightness of the map as in Theorem 1.1 is then very standard in the theory.

Similarly, Theorems 1.2 and 1.3 follow easily from the convergence of the label process stated in Theorem 2.6. After tightness is proved, it only remains to consider its finite-dimensional marginals. A first step is to prove Theorem 2.4 on the convergence of the forest without the labels, in the weak sense of the sub-forests spanned by finitely many random vertices; for this we introduce a new *spinal decomposition* which describes the genealogy of such random vertices. This shall allow us to compare the so-called height process and Łukasiewicz path of the forest. We want to stress that only this weak convergence of the forest is needed, and not a strong convergence, such as the functional convergence of the contour process used in the previous works on random maps. Once we control the length of the branches of the subforests, the joint convergence of the label of these random vertices follows by a Central Limit Theorem for independent but not identically distributed random variables; here again we strongly rely on the spinal decomposition.

Finally the last section is devoted to stable Boltzmann maps; in this case the Łukasiewicz path of the associated forest is a conditioned first-passage bridge of a left-continuous random walk. We rely on previous works and familiar techniques in the study of stable Bienaymé–Galton–Watson trees to prove that the random degree distribution fits in our general framework.

This work leaves open several questions on maps (in addition to those on trees and on the non-bipartite case discussed above). Let us only briefly mention that of the asymptotic behaviour of (planar bipartite) maps in the case of large degrees, of order σ_n, that we are currently investigating. Under suitable assumptions, the Łukasiewicz path converges towards the excursion of a process with exchangeable increments which makes no negative jump. Aldous, Miermont, & Pitman [AMP04] constructed the analogue of the height process of the associated "Inhomogeneous Continuum Random Tree" which is a family of random excursions with continuous-path which extends the Brownian excursion. One can then try to define "Inhomogeneous Continuum Random Maps" by adding random labels on such trees in a similar way as in [LGM11] and to prove convergence of the discrete maps towards these objects (after extraction of a subsequence). In this regard our tightness results from Theorems 1.1 and 2.5 represent a first step towards such a convergence.

1.4. Organisation of the paper

In Section 2, we first recall the definition of labelled plane forests and their encoding by paths and we briefly discuss the bijection with planar maps; then we state the three results whose proof will occupy most of this paper: first Theorem 1.5 on the convergence of reduced forest in Section 2.3, and then Theorems 2.5 and 2.6 on the label process in Section 2.4. In Section 3 we prove the main theorems from the introduction by relying on the aforementioned results. Section 4 is devoted to the study of the random forests, we prove in particular Theorem 1.5 there. We study the label process in Section 5 where we prove Theorems 2.5 and 2.6. Finally in Section 6

we describe the model of stable Boltzmann planar maps, we state and prove our results on these models by relating them to our general setup.

Acknowledgement

I am grateful to Igor Kortchemski who spotted an error in a first draft, as well as to all the referees involved for their constructive remarks.

2. Planar maps as labelled trees

As alluded in the introduction, we study planar maps via a bijection with labelled forests. Let us first recall formally the definition of the latter and set the notation we shall need. Then we state the main contributions of this paper on scaling limits of the paths which code these forests; the proofs are differed to Sections 4 and 5.

2.1. The key bijection

A (rooted plane) tree is a planar map with a unique face. We shall interpret it as the genealogical tree of a population. First, the origin of the root edge is the ancestor of the family, denoted by $\varnothing$, then for any given vertex x, its neighbour $pr(x)$ closer to the ancestor is its parent, and all the other neighbours are its offspring; the distance of x to the root is its generation and is denoted by $|x|$. A vertex with no offspring is called a leaf, and the other ones are called internal vertices. Note that due to the embedding on the sphere, the offspring of a given internal vertex are ordered from left to right and we denote them by $x1, \ldots, xk_x$ where k_x is the offspring number of x; finally the root edge points to the left-most offspring of the ancestor. For any vertex $x \neq \varnothing$, we let $\chi_x \in \{1, \ldots, k_{pr(x)}\}$ be the only index such that $x = pr(x)\chi_x$, which is the relative position of x amongst its siblings. We shall denote by $[\![x, y]\!]$ the unique geodesic path between x and y. Finally, we shall always list the vertices of a tree as $\varnothing = x_0 < x_1 < \ldots$ in the so-called *depth-first search* order, which is constructed recursively as follows: suppose that $x_0 < x_1 < \cdots < x_i$ are constructed for some $i \geqslant 0$, then let $j \leqslant i$ denote the largest index such that x_j has an offspring which does not belong to the list $(x_0, \ldots, x_i)$ and let x_{i+1} denote the left-most of these missing offspring.

A (plane) forest is a finite ordered list of plane trees; we shall alternatively view a forest as a single tree by attaching all the roots to an extra root vertex (when focusing on its geometry), or as another connected graph by linking two consecutive roots in a chain (when focusing on labels as defined below).

For every $k \geqslant 1$, let us consider the following set of discrete bridges:

$$(2.1) \qquad B_k^{\geqslant -1} = \Big\{(b_1, \ldots, b_k) : b_1, b_2 - b_1, \ldots, b_k - b_{k-1} \in \mathbf{Z}_{\geqslant -1}, \, b_k = 0\Big\}.$$

Then a labelling of a plane forest, with say, $\varrho \geqslant 1$ trees rooted at the vertices $r_1 < \cdots < r_\varrho$ respectively, is a function ℓ from its vertices to $\mathbf{Z}$ such that:

(1) the sequence $(\ell(r_1), \ell(r_2), \ldots, S\ell(r_\varrho))$ belongs to $B_\varrho^{\geqslant -1}$,
(2) for every vertex x with $k_x \geqslant 1$ offspring, the sequence of increments $(\ell(x1) - \ell(x), \ldots, \ell(xk_x) - \ell(x))$ belongs to $B_{k_x}^{\geqslant -1}$.

We emphasise the asymmetry of the model: we have $\ell(x1) \geqslant \ell(x) - 1$ in general, but $\ell(xk_x) = \ell(x)$. Also, in the case of a single tree, the first condition reduces to the fact that its root has label 0.

Recall from the introduction that for every integer n, we consider an integer $\varrho_n \in \mathbf{N}$ and a sequence $d_n = (d_n(k))_{k \geqslant 1} \in \mathbf{Z}_+^{\mathbf{N}}$ which sums up to n, and we denote by $\mathbf{T}_{d_n}^{\varrho_n}$ the set of all plane forests with ϱ_n trees and n internal vertices, amongst which $d_n(k)$ have k offspring for every $k \geqslant 1$. Such a forest always has

$$E_n := \sum_{k \geqslant 1} k d_n(k), \qquad d_n(0) := \varrho_n + \sum_{k \geqslant 1} (k-1) d_n(k), \qquad V_n := \sum_{k \geqslant 0} d_n(k),$$

edges, leaves, and vertices respectively. We let $\mathbf{LT}_{d_n}^{\varrho_n}$ denote the set of forests in $\mathbf{T}_{d_n}^{\varrho_n}$ equipped with a labelling as above. For a single tree, we shall drop the exponent 1.

Let $\mathbf{PM}_{d_n}^{\varrho_n}$ be the set of *pointed maps* $(M_n, x_\star)$ where M_n is a map in $\mathbf{M}_{d_n}^{\varrho_n}$ and $x_\star$ is a distinguished vertex of M_n. If $(M_n, x_\star)$ is a pointed map, then the tip of the root edge is either farther (by one) to $x_\star$ than its origin, in which case the map is said to be *positive* by Marckert & Miermont [MM07], or it is closer (again by one), in which case the map is said to be *negative*. Let us immediately note that every map in $\mathbf{M}_{d_n}^{\varrho_n}$ has $\varrho_n + \sum_{k \geqslant 1} k d_n(k) = V_n$ edges in total and so $V_n - n + 1 = d_n(0) + 1$ vertices by Euler's formula. In particular a uniformly random map in $\mathbf{M}_{d_n}^{\varrho_n}$ in which we further distinguish a vertex $x_\star$ independently and uniformly at random has the uniform distribution in $\mathbf{PM}_{d_n}^{\varrho_n}$. Moreover, half of the $2\varrho_n$ edges on the boundary are "positively oriented" and half of them are "negatively oriented", so if $\mathbf{M}_{d_n}^{\varrho_n}$ is positive, we may re-root it to get a negative map. Therefore it is equivalent to work with random negative maps in $\mathbf{PM}_{d_n}^{\varrho_n}$ instead of maps in $\mathbf{M}_{d_n}^{\varrho_n}$.

Combining the bijections due to Bouttier, Di Francesco, & Guitter [BDFG04] and to Janson & Stefánsson [JS15], we obtain that the set $\mathbf{LT}_{d_n}^{\varrho_n}$ is in one-to-one correspondence with the set of negative maps in $\mathbf{PM}_{d_n}^{\varrho_n}$. Let us refer to these papers as well as to [Mar18a] for a direct construction of the bijection. In a few words the map is constructed from the forest in two steps: first as in the standard Schaeffer bijection and more generally as in [BDFG04] we link every vertex of the forest with the next one in *depth-first search order* (as opposed to the contour order on corners for the other bijections) with a smaller label, or to an extra vertex for those whose label is minimal; then we merge every internal vertex of the forest with its last offspring (recall that $\ell(xk_x) = \ell(x)$); see Figure 2.1 for an illustration. The bijection enjoys the following properties:

(1) The leaves of the forest are in one-to-one correspondence with the vertices different from the distinguished one in the map, and the label of a leaf minus the infimum over all labels, plus one, equals the graph distance between the corresponding vertex of the map and the distinguished vertex.
(2) The internal vertices of the forest are in one-to-one correspondence with the inner faces of the map, and the outdegree of the vertex is half the degree of the face.

(3) The root face of the map corresponds to the collection of roots of the forest, and the number of trees is half the perimeter of the map.

The third property only holds for negative maps, which is the reason why we restricted ourselves to this case.
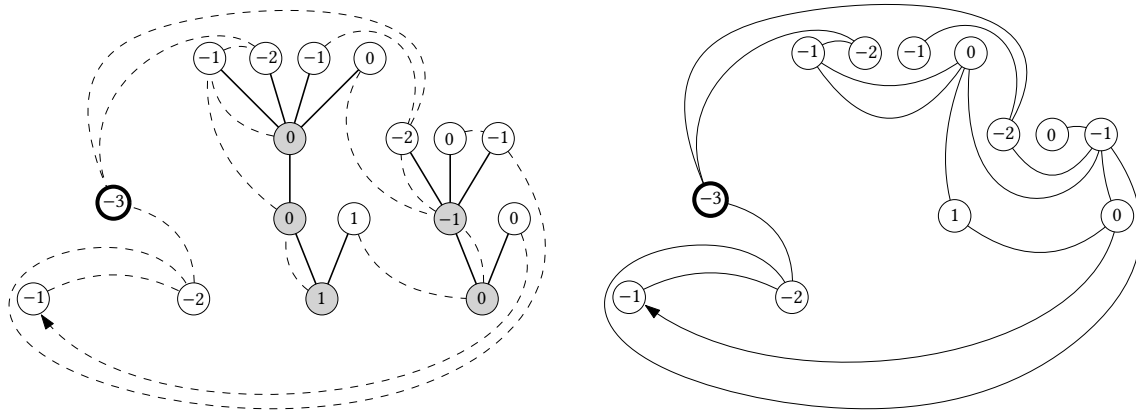

Figure 2.1. Left: a labelled forest in solid lines with an extra vertex labelled −3 and in dashed the links between each vertex and the next one with smaller label (in cyclic order). Right: the negative pointed map obtained by removing the edges of the forest and merging each internal vertex (in gray) with its last offspring.

2.2. Labelled forests and discrete paths

Fix a forest F with $\varrho \geqslant 1$ trees and $V \geqslant 1$ vertices in total, listed $x_0 < x_1 < \cdots < x_{V-1}$ in lexicographical order. It is well known that F is described by each of the following two discrete paths. First, its *Łukasiewicz path* $W = (W(j); 0 \leqslant j \leqslant V)$ is defined by $W(0) = 0$ and

$$W(j+1) = W(j) + k_{x_j} - 1 \quad (0 \leqslant j \leqslant V - 1).$$

One easily checks that $W(V) = -\varrho$ and $W(j) > -\varrho$ for every $0 \leqslant j \leqslant V - 1$. Next the *height process* $H = (H(j); 0 \leqslant j \leqslant V)$ is defined by setting $H(V) = 0$ for convenience and

$$H(j) = |x_j| \quad (0 \leqslant j \leqslant V - 1);$$

see Figure 2.2 for an illustration. Clearly the maps $F \mapsto W$ and $F \mapsto H$ are injective, and in order to recover the forest from one of those paths, one may first observe that the infimum of W up to time j, or the number of passages of H by 0 up to time j tells to which tree of the forest the vertex x_j belongs, and then the parent of this vertex is x_k, where $k = \sup\{i < j : H(i) < H(j)\}$ for the height process, and $k = \sup\{i < j : W(i) \leqslant W(j)\}$ for the Łukasiewicz path. Let us refer to e.g. Le Gall [LG05] for a thorough discussion of such encodings.

In the case of a labelled forest, we encode the labels into the *label process* given by $L(V) = 0$ for convenience and

$$L(j) = \ell(x_j) \quad (0 \leqslant j \leqslant V - 1);$$

see Figure 2.2. The labelled forest is encoded both by the pair (H, L) and by the pair (W, L).
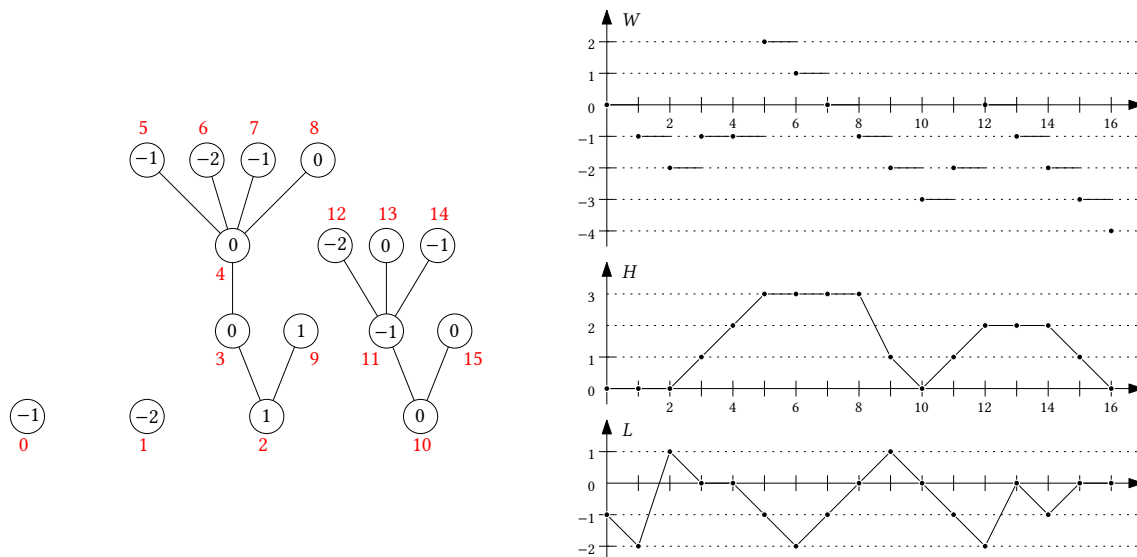

Figure 2.2. Left: A labelled forest on the left, with labels indicated on the nodes and next to them in red their lexicographical order. Right, from top to bottom: its Łukasiewicz path, its height process, and its label process.

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

Let us end with another useful construction of the Łukasiewicz path. The next lemma, whose proof is left as an exercise, gathers some deterministic results that we shall need. In order to simplify the notation, we shall often identify the vertices of a tree with their index in the lexicographic order: here if x is the i'^{th} vertex of a tree whose Łukasiewicz path is W, then we write $W(x)$ for $W(i)$.

LEMMA 2.1. — *Let T be a plane tree and W be its Łukasiewicz path. Fix a vertex $x \in T$, then*

$$W(xk_x) = W(x), \quad W(xj') = \inf_{[xj, \, xj']} W, \quad \text{and} \quad j' - j = W(xj) - W(xj')$$

for every $1 \leqslant j \leqslant j' \leqslant k_x$.

For a vertex x of a tree T, let us denote by $\mathsf{L}(x)$ and $\mathsf{R}(x)$ respectively the number of vertices y whose parent is a strict ancestor of x and which lie strictly to the left, respectively to the right, of the ancestral line $[\![\varnothing, x[\![$; then we put $\mathsf{LR}(x) = \mathsf{L}(x) + \mathsf{R}(x)$. Then a consequence of the preceding lemma is that, in a tree, we have $\mathsf{R}(x) = W(x)$. In the case of a forest, we define $\mathsf{L}(x)$ and $\mathsf{R}(x)$ (and so $\mathsf{LR}(x)$) as the same quantities in the tree containing x, and then we have more generally

$$(2.2) \qquad\qquad \mathsf{R}(x) = W(x) - \min_{0 \leqslant y \leqslant x} W(y).$$

For example, if x is the vertex number 7 in Figure 2.2, then $\mathsf{R}(x) = W(7) - \min_{0 \leqslant i \leqslant 7} W(i) = 2$, which corresponds to the vertices 8 and 9.

2.3. Geometry of random forests

The first key step to understand the asymptotic behaviour of a random map in $\mathbf{PM}_{d_n}^{\varrho_n}$ is to study the associated random labelled forest $(T_{d_n}^{\varrho_n}, \ell)$ in $\mathbf{LT}_{d_n}^{\varrho_n}$. Let us first discuss the behaviour of the forest, and more precisely of its Łukasiewicz path and height process, before stating our main results on the labels. For every integer n, we are thus given a degree sequence $d_n = (d_n(k))_{k \geqslant 0} \in \mathbf{Z}_+^{\mathbf{Z}_+}$ and $T_{d_n}^{\varrho_n}$ is sampled uniformly at random in the set $\mathbf{T}_{d_n}^{\varrho_n}$ of forests which possess $\varrho_n = \sum_{k \geqslant 0}(1 - k)d_n(k)$ trees and $V_n = \sum_{k \geqslant 0} d_n(k)$ vertices, amongst which $d_n(k)$ have k children for every $k \geqslant 0$. Recall also the notation $E_n := \sum_{k \geqslant 1} k d_n(k) = V_n - \varrho_n$ for the number of edges, $\Delta_n := \max\{k \geqslant 0 : d_n(k) > 0\}$ the largest offspring, and finally $\sigma_n^2 := \sum_{k \geqslant s0} k(k-1)d_n(k)$, so σ_n^2/V_n is the second factorial moment of a vertex chosen uniformly at random.

It is well known that the Łukasiewicz path $W_{d_n}^{\varrho_n}$ of our random forest $T_{d_n}^{\varrho_n}$ can be constructed as a cyclic shift of a bridge, also called the discrete Vervaat transform, see e.g. Pitman [Pit06, Chapter 6]. More precisely, let $B_{d_n}^{\varrho_n} = (B_{d_n}^{\varrho_n}(i))_{0 \leqslant i \leqslant V_n}$ be a discrete path sampled uniformly at random amongst all those started from $B_{d_n}^{\varrho_n}(0) = 0$ and which make exactly $d_n(k)$ jumps with value $k - 1$ for every $k \geqslant 0$, so $B_{d_n}^{\varrho_n}$ is a bridge from 0 to $B_{d_n}^{\varrho_n}(V_n) = \sum_{k \geqslant 0}(k-1)d_n(k) = -\varrho_n$. Independently, sample p_n uniformly at random in $\{0, \ldots, \varrho_n - 1\}$ and set

$$i_n = \inf\left\{i \in \{1, \ldots, V_n\} : B_{d_n}^{\varrho_n}(i) = p_n + \inf_{1 \leqslant j \leqslant V_n} B_{d_n}^{\varrho_n}(j)\right\}.$$

Then $W_{d_n}^{\varrho_n}$ has the law of the concatenation

$$\left(\left(B_{d_n}^{\varrho_n}(i_n + k) - B_{d_n}^{\varrho_n}(i_n)\right)_{0 \leqslant k \leqslant V_n - i_n}, \left(B_{d_n}^{\varrho_n}(k) + B_{d_n}^{\varrho_n}(V_n) - B_{d_n}^{\varrho_n}(i_n)\right)_{1 \leqslant k \leqslant i_n}\right).$$

Moreover, the time i_n has the uniform distribution on $\{1, \ldots, V_n\}$ and is independent of the cyclically shifted path $W_{d_n}^{\varrho_n}$. Given the excursion $W_{d_n}^{\varrho_n}$ and i_n, one recovers the bridge $B_{d_n}^{\varrho_n}$ by the same cyclic shift.

The next result states that vertices with a given outdegree are homogeneously spread in the forest in the sense that, asymptotically, the number of vertices with degree in a given set of integers, when read in depth-first search order, grows linearly. We shall apply it with $A = \{-1\}$ in order to deal with maps. For a subset $A \subset \mathbf{Z}_{\geqslant -1} = \{-1, 0, 1, 2, \ldots\}$ and a real number $r \in [0, V_n]$, let $\Lambda_A(r; W_{d_n}^{\varrho_n})$ denote the number of jumps with value in A amongst the first $\lfloor r \rfloor$ jumps of $W_{d_n}^{\varrho_n}$, and consider its inverse $\zeta_A(t; W_{d_n}^{\varrho_n}) = \inf\{r : \Lambda_A(r; W_{d_n}^{\varrho_n}) = \lfloor t \rfloor\}$ for all $t \in [0, \infty)$ the time needed to see at least $\lfloor t \rfloor$ jumps with value in A. Let us set $d_n(A + 1) = \sum_{k \in A} d_n(k + 1)$.

LEMMA 2.2. — *Fix any subset $A \subset \mathbf{Z}_{\geqslant -1}$, if $d_n(A + 1) \to \infty$, then we have the convergence in probability*

$$\left(\frac{\Lambda_A\left(V_n t; W_{d_n}^{\varrho_n}\right)}{d_n(A+1)}, \frac{\zeta_A\left(d_n(A+1)t; W_{d_n}^{\varrho_n}\right)}{V_n}\right)_{t \in [0,1]} \xrightarrow[n \to \infty]{\mathbf{P}} (t, t)_{t \in [0,1]}.$$

Proof. — Note first that since the two functions are inverse of one another and non-decreasing, the two convergences are equivalent. Second, by constructing the bridge $B_{d_n}^{\varrho_n}$ by cyclically shifting $W_{d_n}^{\varrho_n}$ at an independent uniformly random time, it is sufficient to prove the claim when the Łukasiewicz path $W_{d_n}^{\varrho_n}$ is replaced by the bridge $B_{d_n}^{\varrho_n}$.

For an integer $j \leqslant V_n$, the quantity $\Lambda_A(j; B_{d_n}^{\varrho_n})$ is the sum of *dependent* Bernoulli random variables: we start with an urn containing V_n balls in total, amongst which $d_n(A+1)$ are labelled A, we sample j balls without replacement and count the number of balls labelled A picked. If those picks were with replacement, then the first convergence would simply follow from the law of large numbers; actually when sampling without replacement, $\Lambda_A(j; B_{d_n}^{\varrho_n})$ is even more concentrated around its mean $jd_n(A+1)/V_n$ and e.g. the Chernoff bound applies in the same way it does with i.i.d. Bernoulli random variables (this is already shown in Hoeffding's seminal paper [Hoe63]); we deduce the convergence of $t \mapsto \Lambda_A(V_n t; B_{d_n}^{\varrho_n})/d_n(A+1)$ to the identity and thus also that of $t \mapsto \zeta_A(d_n(A+1)t; B_{d_n}^{\varrho_n})/V_n$. □

Let us next focus on the convergence of $W_{d_n}^{\varrho_n}$. For $\varrho \in [0, \infty)$, let us denote by $B^\varrho = (B_t^\varrho)_{t \in [0,1]}$ the standard Brownian bridge from 0 to $-\varrho$ with duration 1. Analogously one can construct X^ϱ the *first-passage Brownian bridge* from 0 to $-\varrho$ (which reduces to the standard Brownian excursion when $\varrho = 0$) by cyclically shifting B^ϱ, see [BCP03, Theorem 7]. The starting point of this work is the following result, which extends [BM14, Lemma 7] and [Lei19, Theorem 1.6].

PROPOSITION 2.3. — *Assume that $\lim_{n \to \infty} \sigma_n = \infty$ and $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ for some $\varrho \in [0, \infty]$.*

(1) *Suppose that $\varrho < \infty$. Then from every sequence of integers, one can extract a subsequence along which the sequence $\sigma_n^{-1} B_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ converges in distribution in the Skorokhod's J_1 topology.*

(2) *Furthermore this sequence converges in distribution towards B^ϱ if and only if $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. In this case, the sequence of processes $\sigma_n^{-1} W_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ converges in distribution towards X^ϱ.*

(3) *Suppose that $\varrho = \infty$, then both processes $\varrho_n^{-1} B_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ and $\varrho_n^{-1} W_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ converge in probability towards $t \mapsto -t$.*

From (2.2), by continuity, when $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$ and $\varrho < \infty$, the process

$$\left(\sigma_n^{-1} \mathsf{R}\left(\lfloor V_n t \rfloor\right)\right)_{t \in [0,1]}$$

converges in distribution towards

$$\widetilde{X^\varrho} = \left(X_t^\varrho - \inf_{s \in [0,t]} X_s^\varrho\right)_{t \in [0,1]}.$$

Proof. — Let us first start with the case $\varrho < \infty$. For every $1 \leqslant i \leqslant V_n$, let $b_i = B_{d_n}^{\varrho_n}(i) - B_{d_n}^{\varrho_n}(i-1)$. Then $\sum_{i=1}^{V_n} b_i = \sum_{k \geqslant 0} (k-1)d_n(k) = -\varrho_n$ and

$$\sum_{i=1}^{V_n} b_i^2 = \sum_{k \geqslant 0} (k-1)^2 d_n(k) = \sum_{k \geqslant 0} k(k-1)d_n(k) - \sum_{k \geqslant 0} (k-1)d_n(k) = \sigma_n^2 + \varrho_n.$$

Therefore $\sum_{i=1}^{V_n}(b_i + V_n^{-1}\varrho_n) = 0$ and $\sum_{i=1}^{V_n}(b_i + V_n^{-1}\varrho_n)^2 = \sigma_n^2 + \varrho_n - V_n^{-1}\varrho_n^2 = \sigma_n^2(1 + o(1))$. Then the sequence given by

$$x_i = \frac{b_i + V_n^{-1}\varrho_n}{\left(\sigma_n^2 + \varrho_n - V_n^{-1}\varrho_n^2\right)^{1/2}}, \quad 1 \leqslant i \leqslant V_n,$$

is called a "normalised urn" by Aldous [Ald85, Chapter 20]. For every $t \in [0, 1]$, let us define

$$S_n(t) = \sum_{i \leqslant \lfloor V_n t \rfloor} x_i = \frac{B_{d_n}^{\varrho_n}\left(\lfloor V_n t \rfloor\right)}{\sigma_n(1 + o(1))} + \frac{\varrho_n \lfloor V_n t \rfloor}{V_n \sigma_n(1 + o(1))}.$$

Then by [Ald85, Proposition 20.3], the sequence $(S_n)_{n \geqslant 1}$ is always tight in the Skorokhod's J_1 topology, this implies the first claim. Furthermore, by [Ald85, Theorem 20.7], this sequence converges in distribution towards a Brownian bridge B^0 as soon as there is no macroscopic jump, i.e. $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. Of course if there is such a large jump, a limit cannot have continuous paths. In the absence of jumps, we deduce that $\sigma_n^{-1} B_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ converges in distribution towards $(-\varrho t + B_t^0)_{t \in [0,1]}$ which has the same law as B^ϱ. Since $W_{d_n}^{\varrho_n}$ and X^ϱ are obtained by cyclically shifting $B_{d_n}^{\varrho_n}$ and B^ϱ respectively, this implies further the convergence in distribution of $(\sigma_n^{-1} W_{d_n}^{\varrho_n}(\lfloor V_n t \rfloor))_{t \in [0,1]}$ towards X^ϱ from the continuity of this operation, see Lei [Lei19, Section 4].

Let us finally suppose that $\varrho_n \gg \sigma_n$. Then similarly, for every $t \in [0, 1]$, let us set

$$S_n'(t) = \frac{\sigma_n}{\varrho_n} S_n(t) = \frac{B_{d_n}^{\varrho_n}\left(\lfloor V_n t \rfloor\right)}{\varrho_n(1 + o(1))} + \frac{\varrho_n \lfloor V_n t \rfloor}{V_n \varrho_n(1 + o(1))}.$$

Since the sequence $(S_n)_{n \geqslant 1}$ is tight, then S_n' converges in probability to the null process and therefore $\varrho_n^{-1} B_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ converges in probability towards $t \mapsto -t$. The convergence of $\varrho_n^{-1} W_{d_n}^{\varrho_n}(\lfloor V_n \cdot \rfloor)$ follows again by cyclic shift.   □

Let us next turn to the height process $H_{d_n}^{\varrho_n}$ associated with the random forest $T_{d_n}^{\varrho_n}$. Let us fix a continuous function $g : [0, 1] \to \mathbf{R}$, then for every $0 \leqslant s \leqslant t \leqslant 1$, set

$$d_g(s, t) = d_g(t, s) = g(s) + g(t) - 2 \min_{r \in [s,t]} g(r).$$

One easily checks that d_g is a continuous pseudo-metric on $[0, 1]$. Consider the quotient space $T_g = [0, 1]/\{d_g = 0\}$; we let π_g be the canonical projection $[0, 1] \to T_g$, then d_g induces a metric on T_g that we still denote by d_g, and the Lebesgue measure on $[0, 1]$ induces a measure p_g on T_g. The space $T_g = (T_g, d_g, p_g)$ is a so-called compact measured real-tree, naturally rooted at $\pi_g(0)$. When $g = X^0$, the space T_{X^0} is the celebrated Brownian Continuum Random Tree of Aldous [Ald93]. More generally, for $\varrho \in (0, \infty)$, the space T_{X^ϱ} describes a forest of Brownian trees, attached by their root along an interval of length ϱ as in [BM17]. For every $t \in [0, 1]$, set $\underline{X}_t^\varrho = \min_{0 \leqslant s \leqslant t} X_s^\varrho$ and then $\widetilde{X}_t^\varrho = X_t^\varrho - \underline{X}_t^\varrho$; then the continuum tree $T_{\widetilde{X}^\varrho}$ describes the same forest of Brownian trees, but now after all the roots have been merged together.

Similarly, in the discrete setting, let us view a plane forest as a single tree by attaching all the roots to an extra root vertex. The next proposition shows that, under the assumption of no macroscopic degree, the forest $T_{d_n}^{\varrho_n}$ spanned by i.i.d. uniform random vertices, i.e. the smallest connected subset of $T_{d_n}^{\varrho_n}$ viewed as a

tree containing these vertices, converges towards the analogue for $T_{\widetilde{X}^\varrho}$ and i.d.d. points sampled from its mass measure $p_{\widetilde{X}^\varrho}$. This result implies Theorem 1.5 from the introduction.

THEOREM 2.4. — *Assume that* $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ *for some* $\varrho \in [0, \infty)$ *and that* $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. *Fix* $q \geqslant 1$ *and let* $U_1, \ldots, U_q$ *be i.i.d. uniform random variables in* $[0, 1]$ *independent of the rest and denote by* $0 = U_{(0)} < U_{(1)} < \cdots < U_{(q)}$ *their ordered statistics. Then the convergence in distribution*

$$\frac{\sigma_n}{2E_n} \left(H_{d_n}^{\varrho_n} \left(V_n U_{(i)} \right), \inf_{U_{(i-1)} \leqslant t \leqslant U_{(i)}} H_{d_n}^{\varrho_n} \left(V_n t \right) \right)_{1 \leqslant i \leqslant q}$$

$$\xrightarrow[n \to \infty]{(d)} \left(\widetilde{X}_{U_{(i)}}^\varrho, \inf_{U_{(i-1)} \leqslant t \leqslant U_{(i)}} \widetilde{X}_t^\varrho \right)_{1 \leqslant i \leqslant q}$$

holds jointly with that of $\sigma_n^{-1} W_{d_n}^{\varrho_n}(V_n \cdot)$ *towards* X^ϱ *in Proposition 2.3.*

The proof is more involved and is differed to Section 4.4. We already mentioned that the question of tightness of the process $H_{d_n}^{\varrho_n}$ (or even its maximum, as discussed in the introduction), once suitably rescaled, was not easy; at least we can say that it is simply not true in general, but we wonder wether it is the case under the assumptions of Theorem 2.4.

2.4. Random labelled forests

Let us next consider random labelled forests $(T_{d_n}^{\varrho_n}, \ell)$ in $\mathbf{LT}_{d_n}^{\varrho_n}$. The law of the latter can be constructed in the following way: first $T_{d_n}^{\varrho_n}$ is sampled uniformly at random in $\mathbf{T}_{d_n}^{\varrho_n}$ as previously, and then given this forest, we sample the label of the roots uniformly at random in $B_{\varrho_n}^{\geqslant -1}$ and, independently for every branchpoint with, say, $k \geqslant 1$ offspring, the label increments between these offspring and the branchpoint are sampled uniformly at random in $B_k^{\geqslant -1}$. Let $L_{d_n}^{\varrho_n}$ be the associated label process.

THEOREM 2.5. — *Fix any sequence* $(\varrho_n)_{n \geqslant 1}$ *and any sequence of degree sequence* $(d_n)_{n \geqslant 1}$. *Then from every increasing sequence of integers, one can extract a subsequence along which the label processes*

$$\left((\sigma_n + \varrho_n)^{-1/2} L_{d_n}^{\varrho_n}(V_n t) \right)_{t \in [0, 1]}$$

converge for the uniform topology.

In order to deduce Theorems 1.2 and 1.3 we need to identify these subsequential limits. We shall need to deal with the root vertices of $T_{d_n}^{\varrho_n}$ separately; let us therefore denote by $(b_{d_n}^{\varrho_n}(k))_{1 \leqslant k \leqslant \varrho_n}$ the labels of the roots of $T_{d_n}^{\varrho_n}$, and set $b_{d_n}^{\varrho_n}(0) = b_{d_n}^{\varrho_n}(\varrho_n) = 0$. Recall that $W_{d_n}^{\varrho_n}$ is the Łukasiewicz path of $T_{d_n}^{\varrho_n}$, for every $0 \leqslant k \leqslant V_n$, let us set $\underline{W}_{d_n}^{\varrho_n}(k) = \min_{0 \leqslant i \leqslant k} W_{d_n}^{\varrho_n}(i)$ and write

(2.3) $$L_{d_n}^{\varrho_n}(k) = \widetilde{L}_{d_n}^{\varrho_n}(k) + b_{d_n}^{\varrho_n} \left(1 - \underline{W}_{d_n}^{\varrho_n}(k) \right).$$

Then $b_{d_n}^{\varrho_n}(1 - \underline{W}_{d_n}^{\varrho_n}(k))$ gives the value of the label of the root vertex of the tree containing the k'th vertex, so $\widetilde{L}_{d_n}^{\varrho_n}(k)$ gives the value of the label of the k'^{th} vertex

minus the value of the label of the root vertex of its tree; in other words, $\widetilde{L}^{\varrho_n}_{d_n}$ is the concatenation of the label process of each tree taken individually, so where all labels have been shifted so that each root receives label 0. The point is that, conditionally given $T^{\varrho_n}_{d_n}$, the two processes $\widetilde{L}^{\varrho_n}_{d_n}(\cdot)$ and $b^{\varrho_n}_{d_n}(1 - \underline{W}^{\varrho_n}_{d_n}(\cdot))$ are independent.

The Brownian sphere and disks are described similarly by "continuum labelled trees" which we next recall, following Bettinelli & Miermont [BM17]. Fix $\varrho \in [0, \infty)$ and recall that X^ϱ denotes the standard Brownian first passage bridge from 0 to $-\varrho$, and that for every $t \in [0, 1]$, we set $\underline{X}^\varrho_t = \min_{0 \leqslant s \leqslant t} X^\varrho_s$ and then $\widetilde{X}^\varrho_t = X^\varrho_t - \underline{X}^\varrho_t$; the processes X^ϱ and $\widetilde{X}^\varrho$ each encode in a different way a Brownian forest as explained previously. For every $y \in [0, \varrho]$, let us set $\tau_y = \inf\{t \in [0, 1] : X^\varrho_t = -y\}$. We construct next another process $Z^\varrho = (Z^\varrho_t)_{t \in [0, 1]}$ on the same probability space as X^ϱ. First, conditionally on X^ϱ, let $\widetilde{Z}^\varrho$ be a centred Gaussian process with covariance

$$\mathbf{E}\left[\widetilde{Z}^\varrho_s \widetilde{Z}^\varrho_t \mid X^\varrho\right] = \min_{r \in [s, t]} \widetilde{X}^\varrho_r \quad \text{for every} \quad 0 \leqslant s \leqslant t \leqslant 1.$$

It is classical that $\widetilde{Z}^\varrho$ admits a continuous version and, without further notice, we shall work throughout this paper with this version. In the case $\varrho = 0$, we simply set $Z^0 = \widetilde{Z}^0$. If $\varrho > 0$, independently of $\widetilde{Z}^\varrho$, let $\mathbf{b}^\varrho$ be a standard Brownian bridge from 0 to 0 with duration ϱ, which has the law of $(\varrho^{1/2}\mathbf{b}(\varrho t))_{t \in [0, 1]}$ where $\mathbf{b}$ is a standard Brownian bridge on the time interval $[0, 1]$, and set

$$Z^\varrho_t = \widetilde{Z}^\varrho_t + \sqrt{3} \cdot \mathbf{b}^\varrho_{-\underline{X}^\varrho_t} \quad \text{for every} \quad 0 \leqslant t \leqslant 1.$$

This construction is the continuum analogue of the decomposition of the process $L^{\varrho_n}_{d_n}$ in (2.3). Observe that, almost surely, $Z^\varrho_s = Z^\varrho_t$ whenever $d_{X^\varrho}(s, t) = 0$ so Z^ϱ can be seen as a process indexed by T_{X^ϱ} by setting $Z^\varrho_x = Z^\varrho_t$ if $x = \pi_{X^\varrho}(t)$. We interpret Z^ϱ_x as the label of an element $x \in T_{X^\varrho}$; the pair $(T_{X^\varrho}, (Z^\varrho_x; x \in T_{X^\varrho}))$ is a continuum analogue of labelled plane forests.

For $\varrho = 0$, the process Z^0 is interpreted as a Brownian motion on the Brownian tree T_{X^0} started from 0 on the root. For $\varrho > 0$, the space T_{X^ϱ} is interpreted as a collection of Brownian trees glued at their root along the interval $[0, \varrho]$; for every $y \in [0, \varrho]$, it holds that $-\underline{X}^\varrho_{\tau_y} = y$ so each point $\pi_{X^\varrho}(\tau_y)$ "at position y" on this interval receives label $\sqrt{3}\,\mathbf{b}_y$ and then the labels on each of the trees $\pi_{X^\varrho}((\tau_{y-}, \tau_y])$ evolve like independent Brownian motions on them. The Brownian sphere and disks are constructed from these processes, see Section 3.2 below. The second main result on the label process is the following.

THEOREM 2.6. — *Suppose that* $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ *for some* $\varrho \in [0, \infty]$.

(1) *If* $\varrho = \infty$, *then the convergence in distribution*

$$\left((2\varrho_n)^{-1/2} L^{\varrho_n}_{d_n}(V_n t); t \in [0, 1]\right) \xrightarrow[n \to \infty]{(d)} (\mathbf{b}_t; t \in [0, 1])$$

holds in $C([0, 1], \mathbf{R})$.

(2) *If* $\varrho < \infty$ *and furthermore* $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$, *then the convergence in distribution*

$$\left(\left(\frac{3}{2\sigma_n}\right)^{1/2} L^{\varrho_n}_{d_n}(V_n t); t \in [0, 1]\right) \xrightarrow[n \to \infty]{(d)} (Z^\varrho_t; t \in [0, 1])$$

holds in $C([0, 1], \mathbf{R})$.

Let us point out that in Theorem 2.6 (1), when $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \infty$, we have more precisely:

$$\left((2\varrho_n)^{-1/2} \left(b_{d_n}^{\varrho_n} (\varrho_n t), \widetilde{L}_{d_n}^{\varrho_n} (V_n t) \right); t \in [0, 1] \right) \xrightarrow[n \to \infty]{(d)} \left((\mathbf{b}_t, 0); t \in [0, 1] \right),$$

which, combined with Proposition 2.3 and the decomposition (2.3), implies the claim of the theorem.

3. Convergence of random maps

The proof of Theorems 2.4, 2.5, and 2.6 on labelled forests will occupy most of this paper. Before proving them, let us first in this section deduce from them the results on random maps stated in Theorems 1.1, 1.2 and 1.3. The argument finds its root in the work of Le Gall [LG13] and has already been adapted in many contexts [ABA17, ABA21, Abr16, BJM14, BM17, LG13, Mar18a, Mar18b] so we shall be very brief and refer to the preceding references for details.

Let us first define the topology we use in these theorems. Recall that we view a planar map as a compact metric space equipped with a Borel probability measure. In words, two such spaces (X, d_X, p_X) and (Y, d_Y, p_Y) are close to each other if one can find a subset of each which carries most of the mass and which are close to be isometric. Formally, a *correspondence* between X and Y is a subset $R \subset X \times Y$ such that for every $x \in X$, there exists $y \in Y$ such that $(x, y) \in R$ and vice-versa. The *distortion* of R is defined as

$$\operatorname{dis}(R) = \sup \left\{ |d_X(x, x') - d_Y(y, y')| ; (x, y), (x', y') \in R \right\}.$$

Then the *Gromov–Hausdorff–Prokhorov distance* between these spaces is the infimum of all the values $\varepsilon > 0$ such that there exists a coupling ν between p_X and p_Y and a compact correspondence R between X and Y such that

$$\nu(R) \geqslant 1 - \varepsilon \quad \text{and} \quad \operatorname{dis}(R) \leqslant 2\varepsilon.$$

This is only a pseudo-distance, but after taking the quotient by measure-preserving isometries, one gets a genuine distance which is separable and complete, see Miermont [Mie09, Proposition 6].

Recall that starting from a uniformly random map in $\mathbf{M}_{d_n}^{\varrho_n}$, we may sample a vertex $x_\star$ independently and uniformly at random to obtain a uniformly random pointed map in $\mathbf{PM}_{d_n}^{\varrho_n}$, which we may then re-root at one of the ϱ_n possible edges on the boundary chosen uniformly at random which make this pointed map $(M_{d_n}^{\varrho_n}, x_\star)$ negative. Let $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$ be the metric measured space given by the vertices of $M_{d_n}^{\varrho_n}$ different from $x_\star$, their graph distance *in* $M_{d_n}^{\varrho_n}$ and the uniform probability measure and note that the Gromov–Hausdorff–Prokhorov distance between $M_{d_n}^{\varrho_n}$ and $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$ is bounded by one so it suffices to prove our claims for $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$. This will enable us to rely on the bijection with a labelled forest $(T_{d_n}^{\varrho_n}, \ell)$.

We shall implicitly assume that $d_n(0) \to \infty$, otherwise we may extract a subsequence along which the number of vertices in our maps is uniformly bounded, so if

one removes all the double edges (i.e. glue together both sides of a face with degree 2), which play no role in the geometry of the associated metric measured space, then only a bounded number of edges remain and tightness is clear.

3.1. Tightness of planar maps

Let us first construct the possible subsequential limits of $M_{d_n}^{\varrho_n}$, relying on a subsequential limit of the label process provided by Theorem 2.5 before proving Theorem 1.1.

Recall that in the bijection relating $(M_{d_n}^{\varrho_n}, x_\star)$ to $(T_{d_n}^{\varrho_n}, \ell)$, the vertices of the former different from $x_\star$ correspond to the leaves of the latter, and the internal vertices of $T_{d_n}^{\varrho_n}$ are identified with their last child. Therefore, for every vertex x of $T_{d_n}^{\varrho_n}$, we let $\varphi(x)$ be the vertex of $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$ in one-to-one correspondence with the rightmost leaf amongst the descendants of x in $T_{d_n}^{\varrho_n}$. Let us list the vertices of $T_{d_n}^{\varrho_n}$ as $x_0 < x_1 < \cdots < x_{V_n-1}$ in lexicographical order, set $x_{V_n} = x_{V_n-1}$, and for every $i, j \in \{0, \ldots, V_n\}$, let us set

$$d_n(i,j) = d_{\mathrm{gr}}\big(\varphi(x_i), \varphi(x_j)\big),$$

where d_{gr} is the graph distance in $M_{d_n}^{\varrho_n}$. We then extend d_n to a continuous function on $[0, V_n]^2$ by "bilinear interpolation" on each square of the form $[i, i+1] \times [j, j+1]$ as in [LG13, Section 2.5] or [LGM11, Section 7]. For every $0 \leqslant s \leqslant t \leqslant 1$, let us set

$$d_{(n)}(s,t) = (\sigma_n + \varrho_n)^{-1/2}\, d_n\,(V_n s, V_n t) \quad \text{and} \quad L_{(n)}(t) = (\sigma_n + \varrho_n)^{-1/2}\, L_{d_n}^{\varrho_n}\,(V_n t)\,.$$

For a continuous function $g : [0,1] \to \mathbf{R}$, let us set for every $0 \leqslant s \leqslant t \leqslant 1$,

$$(3.1) \qquad D_g(s,t) = D_g(t,s) = g(s) + g(t) - 2\max\left\{\min_{r \in [s,t]} g(r);\ \min_{r \in [0,s] \cup [t,1]} g(r)\right\}.$$

For $0 \leqslant i < j \leqslant V_n$, let $[i,j]$ denote the set of integers from i to j, and let $[j,i]$ denote $[j, V_n] \cup [0,i]$. Recall that we construct our map from a labelled forest, using a Schaeffer-type bijection; following the chain of edges drawn starting from two points of the forest to the next one with smaller label until they merge, one obtains the following upper bound on distances:

$$(3.2) \qquad d_n\,(V_n s, V_n t) \leqslant D_{L_{d_n}^{\varrho_n}}\,(V_n s, V_n t) + 2,$$

see Le Gall [LG07, Lemma 3.1] for a detailed proof in a different context.

According to Theorem 2.5, from every increasing sequence of integers, one can extract a subsequence along which the processes $L_{(n)}$ converge in distribution to some limit process, say L. From the previous bound, one can extract a further subsequence along which we have

$$(3.3) \quad \Big(L_{(n)}(t), D_{L_{(n)}}(s,t), d_{(n)}(s,t)\Big)_{s,t \in [0,1]} \xrightarrow[n\to\infty]{(d)} \big(L_t, D_L(s,t), d_\infty(s,t)\big)_{s,t \in [0,1]},$$

where $(d_\infty(s,t))_{s,t \in [0,1]}$ depends a priori on the subsequence and, by (3.2), satisfies $d_\infty \leqslant D_L$, see [LG07, Proposition 3.2] for a detailed proof in a similar context.

The function d_∞ is continuous on $[0,1]^2$ and is a pseudo-distance (as limit of $d_{(n)}$). Let then M_∞ be the quotient $[0,1]/\{d_\infty = 0\}$ equipped with the metric induced

by d_∞, which we still denote by d_∞. We let Π_∞ be the canonical projection from $[0, 1]$ to M_∞ which is continuous (since d_∞ is) so (M_∞, d_∞) is a compact metric space, which finally we endow with the Borel probability measure p_∞ given by the push-forward by Π_∞ of the Lebesgue measure on $[0, 1]$.

Theorem 1.1 directly follows from the following result.

PROPOSITION 3.1. — *On a subsequence along which* (3.3) *holds we have the convergence*

$$\left(V(M_{d_n}^{\varrho_n}) \setminus \{x_\star\}, (\sigma_n + \varrho_n)^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}}\right) \xrightarrow[n \to \infty]{(d)} (M_\infty, d_\infty, p_\infty)$$

for the Gromov–Hausdorff–Prokhorov topology. Furthermore, the following identity in law holds:

$$(3.4) \qquad d_\infty(U, U') \stackrel{(d)}{=} L_U - \min_{t \in [0, 1]} L_t,$$

where U, U' are i.i.d. uniform random variables on $[0, 1]$ and independent of everything else.

Proof. — Let us implicitly restrict ourselves to a subsequence along which (3.3) holds. Appealing to Skorokhod's representation theorem, let us assume furthermore that it holds almost surely. Recall that for a vertex x of $T_{d_n}^{\varrho_n}$, we denote by $\varphi(x)$ the vertex of $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$ in one-to-one correspondence with the leaf at the extremity of the right-most ancestral line starting from x in $T_{d_n}^{\varrho_n}$. Then the sequence $(\varphi(x_i))_{0 \leqslant i \leqslant V_n}$ lists *with redundancies* the vertices of $M_{d_n}^{\varrho_n}$ different from $x_\star$. For every $1 \leqslant i \leqslant d_n(0)$, let us denote by $\lambda(i) \in \{0, \ldots, V_n - 1\}$ the index such that $x_{\lambda(i)}$ is the i'th leaf of $T_{d_n}^{\varrho_n}$, and extend λ linearly between integer times. The function λ corresponds to $\zeta_{\{-1\}}(\,\cdot\,; W_{d_n}^{\varrho_n})$ with the notation of Lemma 2.2. According to this lemma, we have

$$(3.5) \qquad \left(V_n^{-1} \lambda\left(\lfloor d_n(0)t \rfloor\right); t \in [0, 1]\right) \xrightarrow[n \to \infty]{\mathbf{P}} (t; t \in [0, 1]),$$

Observe that the sequence $(\varphi(x_{\lambda(i)}))_{1 \leqslant i \leqslant d_n(0)}$ now lists *without redundancies* the vertices of $M_{d_n}^{\varrho_n}$ different from $x_\star$. The set

$$R_n = \left\{\left(\varphi\left(x_{\lambda(\lfloor d_n(0)t \rfloor)}\right), \Pi_\infty(t)\right); t \in [0, 1]\right\}.$$

is a correspondence between $M_{d_n}^{\varrho_n} \setminus \{x_\star\}$ and M_∞. Let further ν be the coupling between p_{unif} and p_∞ given by

$$\int_{\left(M_{d_n}^{\varrho_n} \setminus \{x_\star\}\right) \times M_\infty} f(x, z) \mathrm{d}\nu(x, z) = \int_0^1 f\left(\varphi\left(x_{\lambda(\lfloor d_n(0)t \rfloor)}\right), \Pi_\infty(t)\right) \mathrm{d}t,$$

for every test function f. Then ν is supported by R_n by construction. Finally, the distortion of R_n is given by

$$\sup_{s, t \in [0, 1]} \left| d_{(n)}\left(\frac{\lambda\left(\lfloor d_n(0)s \rfloor\right)}{V_n}, \frac{\lambda\left(\lfloor d_n(0)t \rfloor\right)}{V_n}\right) - d_\infty(s, t) \right|,$$

which, appealing to (3.5), tends to 0 whenever the convergence (3.3) holds, which concludes the proof of the convergence of the maps.

Let us next consider the identity (3.4) which we shall need shortly in the proof of Theorem 1.2. If U, U' are i.i.d. uniform random variables on $[0, 1]$ and independent

of everything else, then $X_n = \varphi(x_{\lambda(\lfloor d_n(0)U \rfloor)})$ and $Y_n = \varphi(x_{\lambda(\lfloor d_n(0)U' \rfloor)})$ are uniform random vertices of $V(M_{d_n}^{\varrho_n}) \setminus \{x_\star\}$, they can therefore be coupled with two independent uniform random vertices X_n' and Y_n' of $M_{d_n}^{\varrho_n}$ in such a way that the conditional probability given $M_{d_n}^{\varrho_n}$ that $(X_n, Y_n) \neq (X_n', Y_n')$ converges to 0; we implicitly assume in the sequel that $(X_n, Y_n) = (X_n', Y_n')$. Since $x_\star$ is also a uniform random vertex of $M_{d_n}^{\varrho_n}$, we obtain that

$$d_{\mathrm{gr}}(X_n, Y_n) \overset{(d)}{=} d_{\mathrm{gr}}(x_\star, Y_n).$$

By definition we have $d_{\mathrm{gr}}(X_n, Y_n) = d_n(\lambda(\lfloor d_n(0)U \rfloor), \lambda(\lfloor d_n(0)U' \rfloor))$ and by construction of the labels on $T_{d_n}^{\varrho_n}$, we have

$$d_{\mathrm{gr}}(x_\star, Y_n) = L_{d_n}^{\varrho_n}(\lambda(\lfloor d_n(0)U' \rfloor)) - \min_{0 \leqslant j \leqslant V_n} L_{d_n}^{\varrho_n}(j) + 1.$$

Letting $n \to \infty$ along the same subsequence as in (3.3) and appealing to (3.5), we obtain (3.4). □

3.2. Convergence towards a Brownian disk

Let us next turn to the proof of Theorem 1.2; in view of the proof of Theorem 1.1 it suffices to identify the subsequential limit d_∞. Fix $\varrho \in [0, \infty)$ and let us recall the distance D^ϱ of the Brownian disk, following Le Gall [LG07] and Bettinelli & Miermont [BM17] to which we refer for details. Recall the process Z^ϱ defined in Section 2.4 and let us construct the function D_{Z^ϱ} on $[0, 1]^2$ as in (3.1); first, let us view D_{Z^ϱ} as a function on the forest T_{X^ϱ} by setting

$$D_{Z^\varrho}(x, y) = \inf\left\{ D_{Z^\varrho}(s, t); s, t \in [0, 1], x = \pi_{X^\varrho}(s) \ \text{and} \ y = \pi_{X^\varrho}(t) \right\},$$

for every $x, y \in T_{X^\varrho}$, where we recall the notation π_{X^ϱ} for the canonical projection. This function is not a pseudo-distance on T_{X^ϱ} but one can construct one as follows: for every $x, y \in T_{X^\varrho}$, set

$$D^\varrho(x, y) = \inf\left\{ \sum_{i=1}^{k} D_{Z^\varrho}(a_{i-1}, a_i); k \geqslant 1, (x = a_0, a_1, \ldots, a_{k-1}, a_k = y) \in T_{X^\varrho} \right\}.$$

The function D^ϱ is in fact the largest pseudo-distance on T_{X^ϱ} which is bounded by D_{Z^ϱ}. Indeed if D is another pseudo-distance with $D \leqslant D_{Z^\varrho}$, then for every $x, y \in T_{X^\varrho}$, for every $k \geqslant 1$ and every chain $x = a_0, a_1, \ldots, a_{k-1}, a_k = y$ in T_{X^ϱ}, by the triangle inequality, $D(x, y) \leqslant \sum_{i=1}^{k} D(a_{i-1}, a_i) \leqslant \sum_{i=1}^{k} D_{Z^\varrho}(a_{i-1}, a_i)$ and so $D(x, y) \leqslant D^\varrho(x, y)$. Furthermore, it can be seen as a pseudo-distance on $[0, 1]$ by setting $D^\varrho(s, t) = D^\varrho(\pi_{X^\varrho}(s), \pi_{X^\varrho}(t))$ for every $s, t \in [0, 1]$. Then for all $s, t \in [0, 1]$ such that $d_{X^\varrho}(s, t) = 0$ we have $\pi_{X^\varrho}(s) = \pi_{X^\varrho}(t)$ and so $D^\varrho(\pi_{X^\varrho}(s), \pi_{X^\varrho}(t)) = 0$. We deduce from the previous maximality property that D^ϱ is the largest pseudo-distance D on $[0, 1]$ satisfying the following two properties:

$$D \leqslant D_{Z^\varrho} \quad \text{and} \quad d_{X^\varrho}(s, t) = 0 \quad \text{implies} \quad D(s, t) = 0.$$

The Brownian disk is then given by the quotient $[0, 1]/\{D^\varrho = 0\}$, endowed with the metric D^ϱ and the push-forward p^ϱ of the Lebesgue measure on $[0, 1]$.

Let us now prove Theorem 1.2.

Proof of Theorem 1.2. — Let us first assume that $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$ and that $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ with $\varrho \in [0, \infty)$ and let us prove that our maps converge to the Brownian disk. Let us set for all $t \in [0, 1]$,

$$W_{[n]}(t) = \frac{1}{\sigma_n} W_{d_n}^{\varrho_n}\left(\lfloor V_n t \rfloor\right) \qquad \text{and} \qquad L_{[n]}(t) = \left(\frac{3}{2\sigma_n}\right)^{1/2} L_{d_n}^{\varrho_n}(V_n t),$$

and for all $s, t \in [0, 1]$

$$d_{[n]}(s, t) = \left(\frac{3}{2\sigma_n}\right)^{1/2} d_n\left(V_n s, V_n t\right).$$

Then we infer from Proposition 2.3, Theorem 2.6 (2), and the preceding subsection that from every increasing sequence of integers, one can extract a subsequence along which we have

$$\left(W_{[n]}(t), L_{[n]}(t), D_{L_{[n]}}(s, t), d_{[n]}(s, t)\right)_{s, t \in [0, 1]}$$
$$\xrightarrow[n \to \infty]{(d)} \left(X^\varrho, Z_t^\varrho, D_{Z^\varrho}(s, t), d_\infty(s, t)\right)_{s, t \in [0, 1]},$$

where d_∞ depends a priori on the subsequence. Again by Skorokhod's representation theorem, let us assume that this holds almost surely. It remains to prove that $d_\infty = D^\varrho$ almost surely. Our argument is adapted from the work of Bettinelli & Miermont [BM17, Lemma 32].

First recall from the previous proof that d_∞ is a pseudo-distance on $[0, 1]$ which satisfies $d_\infty \leqslant D_{Z^\varrho}$ almost surely. Let $0 \leqslant s < t \leqslant 1$ be such that $d_{X^\varrho}(s, t) = 0$, i.e. such that $X_s^\varrho = X_t^\varrho = \min_{r \in [s, t]} X_r^\varrho$. Suppose first that $X_r^\varrho > X_t^\varrho$ for all $r \in (s, t)$, then by the preceding convergence one can find two sequences of integers $(i_n)_{n \geqslant 1}$ and $(j_n)_{n \geqslant 1}$ such that $i_n/V_n \to s$ and $j_n/V_n \to t$ and for all n large enough $W_{d_n}^{\varrho_n}(i_n) = W_{d_n}^{\varrho_n}(j_n) \leqslant W_{d_n}^{\varrho_n}(k)$ for every $i_n + 1 \leqslant k \leqslant j_n - 1$. In terms of the forest $T_{d_n}^{\varrho_n}$, this means that the vertex x_{i_n} is an ancestor of x_{j_n} and that furthermore the latter lies on the right-most ancestral line amongst the descendants of x_{i_n} in $T_{d_n}^{\varrho_n}$, so in particular $\varphi(x_{i_n}) = \varphi(x_{j_n})$, hence $d_n(i_n, j_n) = 0$. Letting $n \to \infty$ along the same subsequence as above, we conclude that $d_\infty(s, t) = 0$. If there exists $r \in (s, t)$ with $X_s^\varrho = X_t^\varrho = X_r^\varrho$, then it is unique and we conclude similarly that $d_\infty(s, r) = d_\infty(r, t) = 0$. From the maximality property of D^ϱ, we deduce the bound

$$d_\infty \leqslant D^\varrho \quad \text{almost surely.}$$

Next let U, U' be i.i.d. uniform random variables on $[0, 1]$ and independent of everything else. Recall the identity (3.4) which reads here $d_\infty(U, U') = Z_U^\varrho - \min Z^\varrho$ in distribution. The key point is that, according to Le Gall [LG13, Corollary 7.3] for $\varrho = 0$ and Bettinelli & Miermont [BM17, Lemma 17 & Corollary 21] for $\varrho > 0$, the right-hand side is also distributed as $D^\varrho(U, U')$. We conclude that the identity

$$d_\infty(U, U') = D^\varrho(U, U')$$

holds in distribution and thus almost surely by the previous bound. The identity $d_\infty = D^\varrho$ follows by a density argument.

We end this proof by arguing that our assumption on the largest degree is necessary. Indeed, extracting a subsequence if necessary, let us assume that $\sigma_n^{-1} \varrho_n$ and $\sigma_n^{-1} \Delta_n$

converge respectively to $\varrho \geqslant 0$ and $\delta > 0$ and that $(V(M_{d_n}^{\varrho_n}), \sigma_n^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}})$ converges in distribution to some random space (M, D, p) and let us prove that the latter does not have the topology of M^ϱ, which is that of the sphere if $\varrho = 0$ [LGP08, Mie08] or the disk if $\varrho > 0$ [Bet15]. Let us label all the vertices by their graph distance to the distinguished vertex $x_\star$. Let Φ_n be an inner face with degree $2\Delta_n \sim 2\delta\sigma_n$, then the labels of its vertices read in clockwise order form a (shifted) bridge with ± 1 steps which, when rescaled by a factor of order $\Delta_n^{-1/2}$ converges towards a Brownian bridge. Let x_n^-, x_n^+ be two vertices of Φ_n such that their respective labels are the minimum and the maximum over all labels on Φ_n. Then when $\varepsilon > 0$ is small, with high probability we have that $\ell(x_n^+) - \ell(x_n^-) > 6\varepsilon\sigma_n^{1/2}$.

Our argument is depicted on Figure 3.1; let us describe it. Let us read the vertices on the face Φ_n from x_n^- to x_n^+ in clockwise order, there is a vertex which is the last one with label smaller than $\ell(x_n^-) + \varepsilon\sigma_n^{1/2}$ and another one which is the first one with label larger than $\ell(x_n^+) - \varepsilon\sigma_n^{1/2}$; let us call "blue vertices" all the vertices visited between these two. Let us similarly call "green vertices" the vertices defined similarly when going from x_n^- to x_n^+ in counter-clockwise order. A vertex may be simultaneously green and blue, but in this case, and more generally if there exists a pair of blue and green vertices at graph distance $o(\sigma_n^{1/2})$ in the whole map, then this creates at the limit a pinch-point separating the map into two parts each with diameter larger than $\varepsilon\sigma_n^{1/2}$.
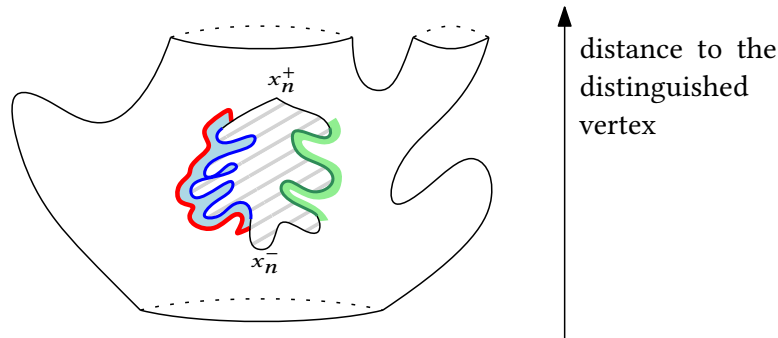
Figure 3.1. A portion of the pointed map represented as a surface (the so-called "cactus" representation), seen from the distinguished vertex. If the grey face is macroscopic and the light blue and light green regions are disjoint, then the red simple path separates the map into two macroscopic parts.

We assume henceforth that the distance between green and blue vertices is larger than $4\eta\sigma_n^{1/2}$ for some $\eta > 0$; the light blue and light green regions in Figure 3.1 represent the hull of the set of vertices at distance smaller than $\eta\sigma_n^{1/2}$ from the blue and green vertices respectively. Consider the simple red path obtained by taking the boundary of the light blue region: its extremities lie at macroscopic distance and it separates two parts of the map with macroscopic diameter. A sphere cannot be separated by a simple curve with distinct extremities so this yields our claim when $\varrho = 0$. This is possible on a disk, if the path touches the boundary twice. Therefore, in order to avoid a contradiction, the red path must contain (at least) two points at macroscopic distance from each other, and both at microscopic distance from the

boundary of the map. If none of these points is at microscopic distance from an extremity of the path, then the preceding argument still applies, so both extremities, which belong to Φ_n must lie at microscopic distance from the boundary, but then this creates pinch-points at the limit. $\qquad\square$

3.3. Convergence to the Brownian tree

Let us finally prove Theorem 1.3 relying on Theorem 2.6 (1). The idea is that the greatest distance in the map to the boundary is small compared to the scaling so only the boundary remains in the limit, and furthermore this boundary, whose distances are related to a discrete bridge, converges to the Brownian tree, which is encoded by the Brownian bridge.

Proof of Theorem 1.3. — Let us denote by $\mathbf{b}$ and X^0 respectively the standard Brownian bridge and the standard Brownian excursion. Recall the construction of the Brownian tree $(T_{X^0}, d_{X^0}, p_{X^0})$ and let us define a random pseudo-distance $D_{\mathbf{b}}$ as in (3.1); let further $(T_{\mathbf{b}}, D_{\mathbf{b}}, p_{\mathbf{b}})$ be the space constructed as M_∞ where d_∞ is replaced by $D_{\mathbf{b}}$, so T_{X^0} and $T_{\mathbf{b}}$ are obtained by taking the quotient of $[0,1]$ by $\{d_{X^0} = 0\}$ and $\{D_{\mathbf{b}} = 0\}$ respectively. Comparing the definition of $D_{\mathbf{b}}$ and d_{X^0}, since $\mathbf{b}$ and X^0 are related by the Vervaat transform, it is easy to prove that the spaces $(T_{X^0}, d_{X^0}, p_{X^0})$ and $(T_{\mathbf{b}}, D_{\mathbf{b}}, p_{\mathbf{b}})$ are isometric so it is equivalent to prove the convergence in distribution

$$\left(V\left(M_{d_n}^{\varrho_n}\right) \setminus \{x_\star\}, (2\varrho_n)^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}}\right) \xrightarrow[n \to \infty]{(d)} (T_{\mathbf{b}}, D_{\mathbf{b}}, p_{\mathbf{b}})$$

in the Gromov–Hausdorff–Prokhorov topology. From the previous proofs it actually suffices to prove the convergence in distribution for the uniform topology

$$(3.6) \qquad \left((2\varrho_n)^{-1/2} d_n\left(V_n s, V_n t\right)\right)_{s,t \in [0,1]} \xrightarrow[n \to \infty]{(d)} (D_{\mathbf{b}}(s,t))_{s,t \in [0,1]}.$$

Here we may adapt the argument of the proof of [Bet15, Theorem 5] to which we refer for details. First, recall the upper bound (3.2):

$$d_n(i,j) \leqslant L_{d_n}^{\varrho_n}(i) + L_{d_n}^{\varrho_n}(j) + 2 - 2\max\left\{\min_{k \in [i,j]} L_{d_n}^{\varrho_n}(k); \min_{k \in [j,i]} L_{d_n}^{\varrho_n}(k)\right\},$$

where for $0 \leqslant i < j \leqslant V_n$, the interval $[i,j]$ denotes the set of integers from i to j, and $[j,i]$ denotes $[j, V_n] \cup [0,i]$. We have a very similar lower bound, see the proof of [CLGM13, Corollary 4.4]:

$$(3.7) \qquad d_n(i,j) \geqslant L_{d_n}^{\varrho_n}(i) + L_{d_n}^{\varrho_n}(j) - 2\max\left\{\min_{k \in [\![i,j]\!]} L_{d_n}^{\varrho_n}(k); \min_{k \in [\![j,i]\!]} L_{d_n}^{\varrho_n}(k)\right\},$$

where the intervals are defined as follows. Let us link the roots of two consecutive trees in the forest, as well as the first and last one in order to create a cycle. Then $[\![i,j]\!]$ denotes the set of all indices k such that x_k lies in the geodesic path between x_i and x_j in this graph; in other words, $k \geqslant 1$ belongs to $[\![i,j]\!]$ if either x_k is an ancestor of x_i or of x_j (and it is an ancestor of both if and only if it is their last common one), or if it is the root of a tree which lies between x_i and x_j in the original

forest. We define $[\![j, i]\!]$ similarly as the set of all indices $k \geqslant 1$ such that either x_k is an ancestor of x_i or of x_j, or it is the root of a tree which *does not* lie between x_i and x_j.

Let us suppose that $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \infty$; according to Theorem 2.6 (1) the convergence in distribution

$$\left((2\varrho_n)^{-1/2} L_{d_n}^{\varrho_n} (V_n t) ; t \in [0, 1] \right) \xrightarrow[n \to \infty]{(d)} (\mathbf{b}_t ; t \in [0, 1])$$

holds in $C([0, 1], \mathbf{R})$. Then the right-hand side of (3.2) with $i = V_n s$ and $j = V_n t$ converges in distribution to $D_\mathbf{b}(s, t)$ once rescaled by a factor $(2\varrho_n)^{-1/2}$. The same holds for the right-hand side of (3.7); indeed, the root vertices visited in $[i, j]$ and in $[\![i, j]\!]$ are the same, so the only difference lies in the non-root vertices, but their label differ by that of the root of their tree by at most $\max \widetilde{L}_{d_n}^{\varrho_n} - \min \widetilde{L}_{d_n}^{\varrho_n} = o(\varrho_n^{1/2})$ in probability by the proof of Theorem 2.6 (1). This proves the convergence (3.6) and hence our claim. □

4. Geometry of random forests with a prescribed degree sequence

As we already mentioned, most of this paper is devoted to the study of labelled forests. In this section we focus on the random forest $T_{d_n}^{\varrho_n}$, the labels are studied in Section 5. We first state in Section 4.1 a technical spinal decomposition whose proof is postponed to the appendix, which approximates the offspring distribution of the ancestors of random vertices by random picks, without replacement, from the *size-biased* offspring distribution $\sum_{k \geqslant 1} \frac{k d_n(k)}{E_n} \delta_k$. Then in Section 4.2 we state and prove some exponential concentration for the random bridge $B_{d_n}^{\varrho_n}$ and the excursion $W_{d_n}^{\varrho_n}$ which are combined with the spinal decomposition in Section 4.3 to prove exponential bounds for the height of a random vertex. Finally we prove Theorem 2.4 in Section 4.4 on the convergence of the reduced trees to reduced Brownian trees in the case of no macroscopic degrees, relying again on the spinal decomposition.

Throughout this section we are given for every $n \geqslant 1$ a degree sequence $d_n = (d_n(k))_{k \geqslant 0} \in \mathbf{Z}_+^{\mathbf{Z}_+}$ and $T_{d_n}^{\varrho_n}$ is sampled uniformly at random in the set $\mathbf{T}_{d_n}^{\varrho_n}$ of forests which possess $\varrho_n = \sum_{k \geqslant 0} (1 - k) d_n(k)$ trees and $V_n = \sum_{k \geqslant 0} d_n(k)$ vertices, amongst which $d_n(k)$ have k children for every $k \geqslant 0$; it therefore has $E_n = \sum_{k \geqslant 1} k d_n(k)$ edges and we recall the notation $\Delta_n = \max\{k \geqslant 0 : d_n(k) > 0\}$ for the largest offspring and $\sigma_n^2 = \sum_{k \geqslant 1} k(k - 1) d_n(k)$.

4.1. A spinal decomposition

We describe in this section the ancestral lines of i.d.d. random vertices of the random forest $T_{d_n}^{\varrho_n}$. A related result on trees was first obtained by Broutin & Marckert [BM14] for a single random vertex and it was extended in [Mar18b] to several vertices. The present one is different, we shall compare them after the statement.

Suppose that an urn contains initially $kd_n(k)$ balls labelled k for every $k \geqslant 1$, so E_n balls in total; let us pick balls repeatedly one after the other *without* replacement; for every $1 \leqslant i \leqslant E_n$, we denote the label of the i'th ball by $\xi_{d_n}(i)$. Conditionally on $(\xi_{d_n}(i))_{1 \leqslant i \leqslant E_n}$, let us sample independent random variables $(\chi_{d_n}(i))_{1 \leqslant i \leqslant E_n}$ such that each $\chi_{d_n}(i)$ is uniformly distributed in $\{1, \ldots, \xi_{d_n}(i)\}$. We shall use the fact that the $\xi_{d_n}(i)$'s are identically distributed, with

$$\mathbf{E}\left[\xi_{d_n}(i) - 1\right] = \sum_{k \geqslant 1} (k-1) \frac{kd_n(k)}{E_n} = \frac{\sigma_n^2}{E_n},$$

(4.1)

$$\mathrm{Var}\left(\xi_{d_n}(i) - 1\right) \leqslant \mathbf{E}\left[(\xi_{d_n}(1) - 1) \cdot \max_i (\xi_{d_n}(i) - 1)\right] \leqslant \Delta_n \frac{\sigma_n^2}{E_n}.$$

Let us fix a plane forest F. Recall that we denote by χ_x the relative position of a vertex x amongst its siblings; for every $0 \leqslant i \leqslant |x|$, let us also denote by $a_i(x)$ the ancestor of x at height i, so $a_{|x|}(x) = x$ and $a_0(x)$ is the root of the tree containing x. Define next for every vertex x the *content* of the branch $[\![a_0(x), x[\![$ as

$$(4.2) \qquad \mathsf{Cont}(x) = \left(\left(k_{a_{i-1}(x)}, \chi_{a_i(x)}\right); 1 \leqslant i \leqslant |x|\right).$$

In words, $\mathsf{Cont}(x)$ lists the number of children of each strict ancestor of x and the position of the next ancestor amongst these children. More generally, let $x_1, \ldots, x_q$ be q vertices of F and let us consider the forest F *reduced* to its root and these vertices: $F(x_1, \ldots, x_q)$ contains only the vertices $x_1, \ldots, x_q$ and their ancestors, and it naturally inherits a plane forest structure from F. Now let $\mathsf{Cont}(x_1, \ldots, x_q)$ be defined as the collection of pairs $(k_{pr(y)}, \chi_y)$ where these quantities are those *in the original forest F*, and y ranges in lexicographical order over the set of vertices of the reduced forest whose parent has only one offspring in this reduced forest. We refer to Figure 4.1 for an example. Let us stress that the branchpoints of the reduced forest are excluded from $\mathsf{Cont}(x_1, \ldots, x_q)$.
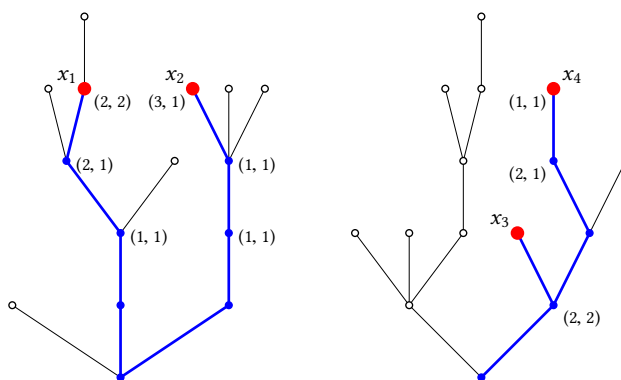

Figure 4.1. *A forest F consisting of two trees with four distinguished red vertices, in blue is the reduced forest $F(x_1, \ldots, x_4)$; the pairs $(k_{pr(y)}, \chi_y)$ which form the associated content $\mathsf{Cont}(x_1, \ldots, x_4)$ are indicated next to their node y, the children of the blue branchpoints are excluded.*

LEMMA 4.1. — *Fix $n \in \mathbf{N}$ and a degree sequence $d_n = (d_n(k))_{k \geqslant 1}$. Let $q \geqslant 1$ and sample $x_{n,1}, \ldots, x_{n,q}$ independently uniformly at random in $T_{d_n}^{\varrho_n}$. Let $(k_i, j_i)_{i=1}^{h}$ be positive integers such that $1 \leqslant j_i \leqslant k_i$ for each i. If $q \geqslant 2$, assume that $h, q \leqslant V_n/4$. Then for every integers $b \geqslant 0$ and $c \geqslant 1$ with $b + c \leqslant q$, the probability that $\mathsf{Cont}(x_{n,1}, \ldots, x_{n,q}) = (k_i, j_i)_{i=1}^{h}$, and that the reduced forest $T_{d_n}^{\varrho_n}(x_{n,1}, \ldots, x_{n,q})$ possesses c trees, q leaves, and b branchpoints is bounded above by*

$$q^2 2^{q-1} \left(\frac{\sigma_n}{V_n}\right)^{q+b} \left(\frac{\varrho_n}{\sigma_n}\right)^{c-1} \frac{\varrho_n + (q-1)\Delta_n + \sum\limits_{i=1}^{h}(k_i - 1)}{\sigma_n}$$

times

$$\mathbf{P}\left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right).$$

The proof consists in straightforward but long combinatorial calculations derived from the decomposition of the forest in the spirit of [BM14] and is given in Appendix A. Let us make two comments. First in the case of a single random vertex, we have $b = 0$ and $c = 1$ so the upper bound reads simply

$$(4.3) \qquad \frac{\varrho_n + \sum\limits_{i=1}^{h}(k_i - 1)}{V_n} \cdot \mathbf{P}\left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right).$$

Second, the reduced forest $T_{d_n}^{\varrho_n}(x_{n,1}, \ldots, x_{n,q})$ possesses q leaves when no $x_{n,i}$ is an ancestor of another; we shall see below that the height of a random vertex is of order E_n/σ_n, so this occurs with high probability as soon as $\sigma_n \to \infty$. Also, if the reduced forest has b branchpoints, then $q + b$ denotes the number of branches once we remove these branchpoints; these branches typically have length of order E_n/σ_n in the original forest, so the factor $(\sigma_n/V_n)^{q+b}$ is important. The other factors will be typically bounded in our applications so the upper bound will read

$$\mathrm{Cst} \cdot \left(\frac{\sigma_n}{V_n}\right)^{q+b} \cdot \mathbf{P}\left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right).$$

In words, roughly speaking, along distinguished paths (removing the branchpoints), up to a multiplicative factor, the individuals reproduce according to the *size-biased law* $(kd_n(k)/E_n)_{k \geqslant 1}$, and conditionally on the offspring, the paths continue via offspring chosen uniformly at random. Note that these size-biased picks are not independent, since we are sampling without replacement. In the case of a single tree $\varrho_n = 1$, an analogous result when sampling *with* replacement was obtained by Broutin & Marckert [BM14] for a single random vertex and it was extended in [Mar18b] to several vertices. The significant difference is that when comparing to sampling with replacement, an extra factor of order e^{h^2/E_n} appears in the upper bound, and one cannot remove it. This was not an issue in [BM14, Mar18b] which focus on the "finite-variance regime", when σ_n^2 is of order E_n since then the bound (4.5) below provided by Addario-Berry [AB12] ensures that h^2 is at most of order E_n with high probability, so the exponential factor does not explode. We mentioned already that

the height of uniform random vertices is of order E_n/σ_n so it does not explode anyway, but we shall prove this fact by relying on this lemma. Moreover, we feel that sampling without replacement is more natural in this model.

4.2. Concentration for bridges and excursions

Let us prove two technical bounds which we shall need in the next subsection, as well as for the proof of Theorem 2.5 on tightness of the label process in Section 5. Recall the notation $B_{d_n}^{\varrho_n} = (B_{d_n}^{\varrho_n}(i))_{0 \leqslant i \leqslant V_n}$ for a bridge from $B_{d_n}^{\varrho_n}(0) = 0$ to $B_{d_n}^{\varrho_n}(V_n) = -\rho_n$ sampled uniformly at random amongst all those which make exactly $d_n(k)$ jumps with value $k-1$ for every $k \geqslant 0$. Recall the construction of the Łukasiewicz path $W_{d_n}^{\varrho_n}$ of our random forest $T_{d_n}^{\varrho_n}$ as a cyclic shift of $B_{d_n}^{\varrho_n}$. In Proposition 2.3 we argued that then $B_{d_n}^{\varrho_n}$ and so $W_{d_n}^{\varrho_n}$ scale like $\sigma_n + \varrho_n$. The next proposition provides a precise Hölder-continuity like bound for the infimum process which is tailored for our applications.

PROPOSITION 4.2. — *Fix a sequence $(d_n)_{n \geqslant 1}$ and $\varepsilon \in (0,1)$, then there exists $C > 0$ such that, for every $n \geqslant 1$, with probability at least $1 - \varepsilon$, it holds*

$$W_{d_n}^{\varrho_n}\left(\lfloor V_n s \rfloor\right) - \min_{s \leqslant r \leqslant t} W_{d_n}^{\varrho_n}\left(\lfloor V_n r \rfloor\right) \leqslant C \cdot (\sigma_n + \varrho_n) \cdot |t - s|^{(1-\varepsilon)/2}$$

uniformly for $0 \leqslant s < t \leqslant 1$.

The proof relies on the following exponential tail bound for the random bridge $B_{d_n}^{\varrho_n}$, adapted from [AB12, Section 3] in the case $\varrho_n = 1$.

LEMMA 4.3. — *For every $\alpha \in (0,1)$ and every $z \geqslant 0$, we have*

$$\mathbf{P}\left(\min_{i \leqslant \alpha V_n} B_{d_n}^{\varrho_n}(i) + \alpha \varrho_n \leqslant -z\right) \leqslant \exp\left(-\frac{(1-\alpha)^3 z^2}{2\alpha(\sigma_n^2 + \varrho_n) + z}\right).$$

Proof. — It is classical that the "remaining sequence" (remaining space divided by the remaining time)

$$M_i = \frac{-\varrho_n - B_{d_n}^{\varrho_n}(i)}{V_n - i} \qquad (0 \leqslant i \leqslant V_n - 1)$$

is a martingale for the natural filtration $(F_i)_{0 \leqslant i \leqslant V_n - 1}$, started from $M_0 = -\varrho_n/V_n$. Indeed, for each time $i \geqslant 1$, set $\Delta B_{d_n}^{\varrho_n}(i) = B_{d_n}^{\varrho_n}(i) - B_{d_n}^{\varrho_n}(i-1)$ and let d_n^i be the remaining degree sequence given by $d_n^i(k) = d_n(k) - \#\{j \leqslant i : \Delta B_{d_n}^{\varrho_n}(j) = k-1\}$ for every $k \geqslant 0$; note that d_n^i is random and F_i-measurable. Fix $0 \leqslant i \leqslant V_n - 2$; then we have

$$\mathbf{P}\left(\Delta B_{d_n}^{\varrho_n}(i+1) = k-1 \mid F_i\right) = \frac{d_n^i(k)}{V_n - i}, \quad k \geqslant 0,$$

from which we obtain

$$\mathbf{E}\left[\Delta B_{d_n}^{\varrho_n}(i+1) \mid F_i\right] = \sum_{k \geqslant 0}(k-1)\frac{d_n^i(k)}{V_n - i} = -\frac{\varrho_n + B_{d_n}^{\varrho_n}(i)}{V_n - i},$$

which shows the martingale property. Furthermore,

$$\mathbf{E}\left[\Delta B_{d_n}^{\varrho_n}(i+1)^2 \mid F_i\right] = \sum_{k \geqslant 0}(k-1)^2 \frac{d_n^i(k)}{V_n - i} \leqslant \frac{\sigma_n^2 + \varrho_n}{V_n - i}.$$

Since $B_{d_n}^{\varrho_n}(i+1) \geqslant B_{d_n}^{\varrho_n}(i) - 1$, then for every $0 \leqslant i \leqslant V_n - 2$,

$$M_{i+1} - M_i \leqslant \frac{V_n - i - B_{d_n}^{\varrho_n}(i) - \varrho_n}{(V_n - (i+1))(V_n - i)} \leqslant \frac{E_n}{(V_n - (i+1))^2}$$

and

$$\text{Var}\left(M_{i+1} \mid F_i\right) \leqslant \frac{1}{(V_n - (i+1))^2}\mathbf{E}\left[\Delta B_{d_n}^{\varrho_n}(i+1)^2 \mid F_i\right] \leqslant \frac{\sigma_n^2 + \varrho_n}{(V_n - (i+1))^3}.$$

Then the martingale Chernoff-type bound, see e.g. McDiarmid [McD98, Theorem 3.15 and the remark at the end of Section 3.5], there about the maximum, shows that for every $z \geqslant 0$ and every $1 \leqslant k \leqslant V_n - 1$,

$$\mathbf{P}\left(\max_{i \leqslant k} M_i - M_0 \geqslant z\right) \leqslant \exp\left(-\frac{z^2}{2\frac{k(\sigma_n^2 + \varrho_n)}{(V_n - k)^3} + \frac{2}{3}z\frac{E_n}{(V_n - k)^2}}\right)$$

$$\leqslant \exp\left(-\frac{(V_n - k)^3 z^2}{2k\left(\sigma_n^2 + \varrho_n\right) + E_n(V_n - k)z}\right).$$

Observe that for every $i \leqslant k$,

$$-(M_i - M_0) = \frac{B_{d_n}^{\varrho_n}(i) + \varrho_n}{V_n - i} - \frac{\varrho_n}{V_n} = \frac{B_{d_n}^{\varrho_n}(i) + i\varrho_n/V_n}{V_n - i} \leqslant \frac{B_{d_n}^{\varrho_n}(i) + k\varrho_n/V_n}{V_n - i},$$

we infer that for every $z \geqslant 0$ and every $1 \leqslant k \leqslant V_n - 1$,

$$\mathbf{P}\left(\min_{i \leqslant k}\left(B_{d_n}^{\varrho_n}(i) + \frac{k\varrho_n}{V_n}\right) \leqslant -z\right) \leqslant \mathbf{P}\left(\min_{i \leqslant k}\frac{B_{d_n}^{\varrho_n}(i) + i\varrho_n/V_n}{V_n - i} \leqslant -\frac{z}{V_n}\right)$$

$$\leqslant \exp\left(-\frac{(V_n - k)^3 z^2/V_n^2}{2k\left(\sigma_n^2 + \varrho_n\right) + E_n\left(V_n - k\right)z/V_n}\right).$$

The claim follows by taking $k = \alpha V_n$ with $\alpha \in (0,1)$. $\square$

We may now prove Proposition 4.2.

Proof of Proposition 4.2. — Let us first focus on the bridge $B_{d_n}^{\varrho_n}$ and observe that, with no jump smaller than -1, we have $\min_{i \leqslant \alpha V_n} B_{d_n}^{\varrho_n}(i) + \alpha\varrho_n \geqslant -\alpha V_n + \alpha\varrho_n = -\alpha E_n$. Therefore, in Lemma 4.3, the probability on the left-hand side is zero as soon as $z > \alpha E_n$ so for every $\alpha \in (0, 1/2]$ and every $z \geqslant 0$,

$$\mathbf{P}\left(\min_{i \leqslant \alpha V_n} B_{d_n}^{\varrho_n}(i) + \alpha\varrho_n \leqslant -z\right) \leqslant \exp\left(-\frac{(1-\alpha)^3 z^2}{\alpha\left(2\left(\sigma_n^2 + \varrho_n\right) + E_n\right)}\right)$$

$$\leqslant \exp\left(-\frac{z^2}{16\alpha\left(\sigma_n^2 + \varrho_n + E_n\right)}\right).$$

By exchangeability, for any integers $i < j$ in $[0, V_n]$ with $|j - i| \leqslant V_n/2$, we have

$$\mathbf{P}\left(B_{d_n}^{\varrho_n}(i) - \min_{i \leqslant k \leqslant j} B_{d_n}^{\varrho_n}(k) - \varrho_n \frac{|j - i|}{V_n} > \sqrt{(\sigma_n^2 + \varrho_n + E_n)\frac{|j - i|}{V_n}x}\right) \leqslant \exp\left(-\frac{x^2}{16}\right).$$

By integrating this tail bound applied to $x^{1/p}$, we obtain that for every $0 \leqslant s < t \leqslant 1$ with $|t - s| \leqslant 1/2$,

$$\mathbf{E}\left[\left(\frac{B_{d_n}^{\varrho_n}(\lfloor V_n s \rfloor) - \min_{s \leqslant r \leqslant t} B_{d_n}^{\varrho_n}(\lfloor V_n r \rfloor) - \varrho_n|t - s|}{(\sigma_n^2 + \varrho_n + E_n)^{1/2}|t - s|^{1/2}}\right)^p\right] \leqslant \int_0^\infty \exp\left(-\frac{x^{2/p}}{16}\right) \mathrm{d}x.$$

Let $c(p)$ denote the right hand side. Since $a^p + b^p \leqslant (a + b)^p \leqslant 2^{p-1}(a^p + b^p)$ for every $a, b \geqslant 0$, we conclude that

$$\mathbf{E}\left[\left(B_{d_n}^{\varrho_n}(\lfloor V_n s \rfloor) - \min_{s \leqslant r \leqslant t} B_{d_n}^{\varrho_n}(\lfloor V_n r \rfloor)\right)^p\right]$$
$$\leqslant 2^{p-1}\left(\mathbf{E}\left[\left(B_{d_n}^{\varrho_n}(\lfloor V_n s \rfloor) - \min_{s \leqslant r \leqslant t} B_{d_n}^{\varrho_n}(\lfloor V_n r \rfloor) - \varrho_n|t - s|\right)^p\right] + (\varrho_n|t - s|)^p\right)$$
$$\leqslant 2^{p-1}\left(c(p)\left(\sigma_n^2 + \varrho_n + E_n\right)^{p/2}|t - s|^{p/2} + \varrho_n^p|t - s|^{p/2}\right)$$
$$\leqslant C(p)\left(\left(\sigma_n^2 + E_n\right)^{1/2} + \varrho_n\right)^p|t - s|^{p/2},$$

for some $C(p) > 0$. Upon changing the constant $C(p)$, the restriction $|t - s| \leqslant 1/2$ can be lifted by the triangle inequality.

Note that the scaling $(\sigma_n^2 + E_n)^{1/2} + \varrho_n$ that appears here is larger than the claimed one $\sigma_n + \varrho_n$. However, in the case $d_n(1) = 0$, it holds that $\sigma_n^2 = \sum_{k \geqslant 2} k(k-1)d_n(k) \geqslant \sum_{k \geqslant 2} kd_n(k) = E_n$ so both scalings are of the same order. In this case, for some other constant $C(p) > 0$, for every $s < t$, it holds

$$\mathbf{E}\left[\left(B_{d_n}^{\varrho_n}(\lfloor V_n s \rfloor) - \min_{s \leqslant r \leqslant t} B_{d_n}^{\varrho_n}(\lfloor V_n r \rfloor)\right)^p\right] \leqslant C(p)\left(\sigma_n + \varrho_n\right)^p|t - s|^{p/2}.$$

In the general case, let us replace the degree sequence d_n by $d_n'(k) = d_n(k)\mathbf{1}_{\{k \neq 1\}}$ and couple the two forests with degree sequence d_n and d_n', by removing all the vertices with outdegree 1 in $T_{d_n}^{\varrho_n}$. Let us denote by $V_n' = V_n - d_n(1)$ and $E_n' = E_n - d_n(1)$ the number of vertices and edges respectively of the new forest $T_{d_n}^{\varrho_n\prime}$; note that both σ_n and ϱ_n are unchanged. Note also that the increments which are removed from $B_{d_n}^{\varrho_n}$ in this operation are all null. Let $s < t$ and assume for notational convenience that $V_n s$ and $V_n t$ are both integers. In the forest $T_{d_n}^{\varrho_n\prime}$, these two instants correspond to, say, $V_n's_n$ and $V_n't_n$ respectively, and $V_n'|t_n - s_n|$ has the binomial distribution with parameters $V_n|t - s|$ and V_n'/V_n and is independent of the path $B_{d_n}^{\varrho_n}$ to which we have removed the null increments. Then the moment of order $p/2$ of $V_n'|t_n - s_n|$ is bounded above by some constant times $(V_n'|t - s|)^{p/2}$, so we conclude from the

preceding case that

$$\mathbf{E}\left[\left(B_{d_n}^{\varrho_n}\left(\lfloor V_n s\rfloor\right)-\min_{s\leqslant r\leqslant t}B_{d_n}^{\varrho_n}\left(\lfloor V_n r\rfloor\right)\right)^p\right]$$

$$=\mathbf{E}\left[\mathbf{E}\left[\left(B_{d_n}^{\varrho_n\prime}\left(\lfloor V_n' s_n\rfloor\right)-\min_{s_n\leqslant r\leqslant t_n}B_{d_n}^{\varrho_n\prime}\left(\lfloor V_n' r\rfloor\right)\right)^p\;\Big|\;s_n,t_n\right]\right]$$

$$\leqslant C(p)(\sigma_n+\varrho_n)^p\mathbf{E}\left[|t_n-s_n|^{p/2}\right]$$

$$\leqslant C'(p)(\sigma_n+\varrho_n)^p|t-s|^{p/2}.$$

The proof of the standard Kolmogorov criterion then shows that the claim of the proposition holds when we replace $W_{d_n}^{\varrho_n}$ by $B_{d_n}^{\varrho_n}$.

We finally want to transfer this bound to $W_{d_n}^{\varrho_n}$ by cyclic shift; some care is needed here. Indeed, recall that we may couple $W_{d_n}^{\varrho_n}$ and $B_{d_n}^{\varrho_n}$ in such a way that $B_{d_n}^{\varrho_n}$ is obtained by cyclically shifting $W_{d_n}^{\varrho_n}$ at a uniform random time $V_n - i_n$ independent of $W_{d_n}^{\varrho_n}$. Fix again $s < t$ and assume that $i = V_n s$ and $j = V_n t$ are integers. There are two cases: either $V_n - i_n$ falls (strictly) between i and j, or it does not. If it does not, then the path of $W_{d_n}^{\varrho_n}$ between i and j is moved without change in $B_{d_n}^{\varrho_n}$ so the claim follows from the preceding bound applied to the image of i and j after the cyclic shift operation, which are still at distance $|j - i|$ from each other. Assume henceforth that $V_n - i_n$ does fall between i and j, then the part of $W_{d_n}^{\varrho_n}$ between i and $V_n - i_n$ is moved to the part of $B_{d_n}^{\varrho_n}$ between $i + i_n$ and V_n, and the part of $W_{d_n}^{\varrho_n}$ between $V_n - i_n$ and j is moved to the part of $B_{d_n}^{\varrho_n}$ between 0 and $j + i_n - V_n$, so

$$W_{d_n}^{\varrho_n}(i)-\min_{i\leqslant k\leqslant j}W_{d_n}^{\varrho_n}(k)$$

$$\leqslant\left(W_{d_n}^{\varrho_n}(i)-\min_{i\leqslant k\leqslant V_n-i_n}W_{d_n}^{\varrho_n}(k)\right)+\left(W_{d_n}^{\varrho_n}(V_n-i_n)-\min_{V_n-i_n\leqslant k\leqslant j}W_{d_n}^{\varrho_n}(k)\right)$$

$$=\left(B_{d_n}^{\varrho_n}(i+i_n)-\min_{i+i_n\leqslant k\leqslant V_n}B_{d_n}^{\varrho_n}(k)\right)+\left(B_{d_n}^{\varrho_n}(0)-\min_{0\leqslant k\leqslant j+i_n-V_n}W_{d_n}^{\varrho_n}(k)\right).$$

Using again that for every $a, b \geqslant 0$, we have $a^p + b^p \leqslant (a+b)^p \leqslant 2^{p-1}(a^p + b^p)$, and since $|j - i| = |V_n - (i + i_n)| + |j + i_n - V_n|$, we conclude from the bound on $B_{d_n}^{\varrho_n}$. $\qquad\square$

4.3. Exponential tails for the height and width

Recall from Section 2.2 that for a vertex x in a tree, we denote by $\mathsf{L}(x)$, resp. $\mathsf{R}(x)$, the number of vertices whose parent is a strict ancestor of x and which lie strictly before, resp. strictly after, x in depth-first search order, so $\mathsf{LR}(x) = \mathsf{L}(x) + \mathsf{R}(x)$ denotes the total number of vertices different from x branching off the path $[\![\varnothing, x[\![$. In the case of a forest, we consider these quantities in the tree containing x. Recall finally the identity (2.2): if W is the Łukasiewicz path of the forest, then $\mathsf{R}(x) = W(x) - \min_{0\leqslant y\leqslant x}W(y)$.

Let us note that, as in Lemma 4.3, our argument is not optimal since our bounds get worse when the number of trees ϱ_n grows, but in most cases, this factor is negligible.

PROPOSITION 4.4. — *Let x_n be a uniformly random vertex of $T_{d_n}^{\varrho_n}$, then*

$$\mathbf{P}\left(\mathsf{LR}(x_n) \geqslant z\left(\sigma_n^2 + \varrho_n\right)^{1/2}\right) \leqslant 4\exp\left(-\frac{z}{288}\right)$$

uniformly in $z \geqslant 1/2$ and $n \in \mathbf{N}$.

Proof. — Observe that the "mirror forest" obtained from $T_{d_n}^{\varrho_n}$ by flipping the order of the children of every vertex has the same law as $T_{d_n}^{\varrho_n}$ so the random variables $\mathsf{L}(x_n)$ and $\mathsf{R}(x_n)$ have the same law; of course they are not independent. Still, since $\mathsf{LR}(x_n) = \mathsf{L}(x_n) + \mathsf{R}(x_n)$, it suffices to consider the tail of $\mathsf{R}(x_n)$ with $z/2$ in place of z and to use a union bound. Recall also that $W_{d_n}^{\varrho_n}$ can be obtained by cyclically shifting the bridge $B_{d_n}^{\varrho_n}$ at the random time i_n which is uniformly distributed in $\{1, \ldots, V_n\}$ and is independent of $W_{d_n}^{\varrho_n}$. In this coupling, we have that

$$W_{d_n}^{\varrho_n}(V_n - i_n) - \min_{0 \leqslant j \leqslant V_n - i_n} W_{d_n}^{\varrho_n}(j) = B_{d_n}^{\varrho_n}(V_n) - \min_{0 \leqslant j \leqslant V_n} B_{d_n}^{\varrho_n}(j)$$
$$= -\varrho_n - \min_{0 \leqslant j \leqslant V_n} B_{d_n}^{\varrho_n}(j).$$

Note that $V_n - i_n$ is uniformly distributed in $\{0, \ldots, V_n - 1\}$ and is independent of $W_{d_n}^{\varrho_n}$, so the vertex visited at this time has the same law as x_n and therefore $\mathsf{R}(x_n)$ has the same law as $-\varrho_n - \min_{0 \leqslant j \leqslant V_n} B_{d_n}^{\varrho_n}(j)$. By considering the two cases where the minimum of $B_{d_n}^{\varrho_n}$ is achieved on the first half or on the second half, we see that this is bounded above by

$$\left(-\frac{\varrho_n}{2} - \min_{0 \leqslant j \leqslant V_n/2} B_{d_n}^{\varrho_n}(j)\right) + \left(-\frac{\varrho_n}{2} - \min_{V_n/2 \leqslant j \leqslant V_n} B_{d_n}^{\varrho_n}(j) + B_{d_n}^{\varrho_n}\left(\frac{V_n}{2}\right)\right),$$

and the two terms on the right have the same law. Hence for every $z \geqslant 1/2$ we have after two union bounds

$$\mathbf{P}\left(\mathsf{LR}(x_n) \geqslant z\left(\sigma_n^2 + \varrho_n\right)^{1/2}\right) \leqslant 4 \cdot \mathbf{P}\left(\min_{0 \leqslant j \leqslant V_n/2} B_{d_n}^{\varrho_n}(j) + \frac{\varrho_n}{2} \leqslant -\frac{z\left(\sigma_n^2 + \varrho_n\right)^{1/2}}{4}\right).$$

Finally, by Lemma 4.3,

$$\mathbf{P}\left(\min_{i \leqslant V_n/2} B_{d_n}^{\varrho_n}(i) + \frac{\varrho_n}{2} \leqslant -\frac{z\left(\sigma_n^2 + \varrho_n\right)^{1/2}}{4}\right) \leqslant \exp\left(-\frac{(1/2)^3\left(\frac{z\left(\sigma_n^2+\varrho_n\right)^{1/2}}{4}\right)^2}{\left(\sigma_n^2 + \varrho_n\right) + \frac{z\left(\sigma_n^2+\varrho_n\right)^{1/2}}{4}}\right)$$

$$\leqslant \exp\left(-\frac{2^{-5}\left(\sigma_n^2 + \varrho_n\right)^{1/2} z^2}{4\left(\sigma_n^2 + \varrho_n\right)^{1/2} + z}\right).$$

Since both $z \geqslant 1/2$ and $\left(\sigma_n^2 + \varrho_n\right)^{1/2} \geqslant 1$, then $4(\sigma_n^2 + \varrho_n)^{1/2} \leqslant 8(\sigma_n^2 + \varrho_n)^{1/2}z$ and $z \leqslant (\sigma_n^2 + \varrho_n)^{1/2}z$ so the denominator in the last exponential is bounded above by $9(\sigma_n^2 + \varrho_n)^{1/2}z$, which yields our claim. □

Let us next focus on the case of a single tree $\varrho_n = 1$. Let us denote by $\mathrm{wid}(T_{d_n})$ the maximum over all $i \geqslant 1$ of the number of vertices of T_{d_n} at distance i from the root, called the *width* of T_{d_n}, and by $\mathrm{ht}(T_{d_n})$ the greatest $i \geqslant 1$ such that there exists at least one vertex of T_{d_n} at distance i from the root, called the *height* of T_{d_n}. The

preceding result gives a universal tail bound on the first quantity. Indeed, with a similar reasoning as in the preceding proof, one can check that

$$(4.4) \qquad \mathbf{P}\left(\min_{0 \leqslant j \leqslant V_n/2} B_{d_n}(j) + \frac{1}{2} \leqslant -z\sigma_n\right) \leqslant \exp\left(-\frac{z}{48}\right),$$

for all $z \geqslant 1$. Then following [AB12], by replacing Equation 2 there by this bound, this yields the existence of two universal constants $c_1, c_2 > 0$ such that for every $z \geqslant 1$,

$$\mathbf{P}\left(\mathrm{wid}(T_{d_n}) \geqslant z\sigma_n\right) \leqslant c_1 e^{-c_2 z}.$$

By observing that $\mathrm{wid}(T_{d_n}) \times \mathrm{ht}(T_{d_n}) \geqslant E_n$, we also get

$$\mathbf{P}\left(\mathrm{ht}(T_{d_n}) \leqslant E_n/(z\sigma_n)\right) \leqslant c_1 e^{-c_2 z}.$$

Following [AB12] again we may obtain the following upper bound for the height of the tree T_{d_n}:

$$(4.5) \qquad \mathbf{P}\left(\mathrm{ht}\left(T_{d_n}\right) \geqslant z\left(\sigma_n^2 + E_n\right)^{1/2}\right) \leqslant c_1 e^{-c_2 z},$$

for some other constants $c_1, c_2 > 0$. However this scaling may be much larger than E_n/σ_n outside the "finite-variance" regime, when σ_n is of order $E_n^{1/2}$. We already mentioned in Section 1.2 that an upper-bound at the scaling E_n/σ_n in full generality is not possible; nonetheless, the scaling E_n/σ_n is correct for typical vertices.

PROPOSITION 4.5. — *There exists two universal constants $c_1, c_2 > 0$ such that the following holds: if x_n is a uniformly random vertex of T_{d_n}, then its height $|x_n|$ satisfies*

$$\mathbf{P}\left(|x_n| \geqslant zE_n/\sigma_n\right) \leqslant c_1 e^{-c_2 z}$$

uniformly for $z \geqslant 1$ and $n \in \mathbf{N}$.

Our argument is the following: if x_n is at height at least zE_n/σ_n and if its ancestors reproduce according to the size-biased law as the ξ_{d_n}'s in the preceding section, then, according to (4.1), the number of vertices $\mathsf{LR}(x_n)$ branching off its ancestral line is in average at least $(zE_n/\sigma_n) \times (\sigma_n^2/E_n) = z\sigma_n$, and we know from Proposition 4.4 that this occurs with a sub-exponential probability.

Proof. — According to Proposition 4.4, it is sufficient to bound

$$\mathbf{P}\left(|x_n| \geqslant z\frac{E_n}{\sigma_n} \text{ and } \mathsf{LR}\left(x_n\right) \leqslant \frac{z}{2}\sigma_n\right) = \sum_{h \geqslant zE_n/\sigma_n} \mathbf{P}\left(|x_n| = h \text{ and } \mathsf{LR}(x_n) \leqslant \frac{z}{2}\sigma_n\right).$$

Recall the definition of $\mathsf{Cont}(x_n) = ((k_{a_{i-1}(x_n)}, \chi_{a_i(x_n)}))_{1 \leqslant i \leqslant |x_n|}$ from (4.2) and note that we have $\mathsf{LR}(x_n) = \sum_{1 \leqslant i \leqslant |x_n|}(k_{a_{i-1}(x_n)} - 1)$. Let us fix $h \geqslant zE_n/\sigma_n$. We deduce

from Lemma 4.1 with a single random vertex that

$$\mathbf{P}\left(|x_n| = h \text{ and } \mathsf{LR}(x_n) \leqslant \frac{z}{2}\sigma_n\right)$$

$$= \sum_{(k_i, j_i)_{i=1}^h} \mathbf{P}\left(\mathsf{Cont}\,(x_n) = (k_i, j_i)_{i=1}^h\right) \mathbf{1}_{\left\{\sum_{i=1}^h (k_i-1) \leqslant \frac{z}{2}\sigma_n\right\}}$$

$$\leqslant \sum_{\substack{(k_i, j_i)_{i=1}^h \\ \sum_{i=1}^h (k_i-1) \leqslant \frac{z}{2}\sigma_n}} \frac{1 + \sum_{i=1}^h (k_i-1)}{V_n} \cdot \mathbf{P}\left(\bigcap_{i=1}^h \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right)$$

$$\leqslant \frac{\frac{z}{2}\sigma_n + 1}{E_n} \mathbf{P}\left(\sum_{i=1}^h (\xi_{d_n}(i) - 1) \leqslant \frac{z}{2}\sigma_n\right).$$

Let us write $X_n(i) = \xi_{d_n}(i) - 1$ in order to simplify the notation. Recall from (4.1) that these random variables have mean σ_n^2/E_n so for $h \geqslant zE_n/\sigma_n$, the probability in the last line is bounded above by

$$\mathbf{P}\left(\sum_{i=1}^h \left(X_n(i) - \mathbf{E}[X_n(i)]\right) \leqslant -\frac{h\sigma_n^2}{2E_n}\right).$$

The $X_n(i)$'s come from sampling balls without replacement; as we already mentioned, by [Ald85, Proposition 20.6], their sum satisfies any concentration inequality based on controlling the Laplace transform the similar sum when sampling with replacement does. In particular, we may apply [McD98, Theorem 2.7], and get

$$\mathbf{P}\left(\sum_{i=1}^h \left(X_n(i) - \mathbf{E}\left[X_n(i)\right]\right) \leqslant -\frac{h\sigma_n^2}{2E_n}\right) \leqslant \exp\left(-\frac{\left(\frac{h\sigma_n^2}{2E_n}\right)^2}{2h\frac{\Delta_n\sigma_n^2}{E_n} + \frac{2}{3}\frac{h\sigma_n^2}{2E_n}\frac{\sigma_n^2}{E_n}}\right)$$

$$= \exp\left(-\frac{h\sigma_n^2}{8E_n\Delta_n + \frac{4}{3}\sigma_n^2}\right).$$

Observe that $\sigma_n^2 \leqslant E_n\Delta_n$ so $8E_n\Delta_n + \frac{4}{3}\sigma_n^2 \leqslant 10E_n\Delta_n$, hence

$$\mathbf{P}\left(|x_n| \geqslant z\frac{E_n}{\sigma_n} \text{ and } \mathsf{LR}(x_n) \leqslant \frac{z}{2}\sigma_n\right) \leqslant \frac{z\sigma_n + 2}{2E_n} \sum_{h \geqslant z E_n/\sigma_n} \exp\left(-\frac{h\sigma_n^2}{10E_n\Delta_n}\right)$$

$$\leqslant \frac{z\sigma_n + 2}{2E_n} \frac{\exp\left(-\frac{z\sigma_n}{10\Delta_n}\right)}{1 - \exp\left(-\frac{\sigma_n^2}{10E_n\Delta_n}\right)}.$$

We next appeal to the following two bounds: first $(1 - \mathrm{e}^{-t}) \geqslant t(1 - t/2) \geqslant 19t/20$ for every $0 \leqslant t \leqslant 1/10$, second $t\mathrm{e}^{-t} \leqslant \mathrm{e}^{-t/2}$ for all $t \geqslant 0$. We thus have

$$\mathbf{P}\left(|x_n| \geqslant z \text{ and } \mathsf{LR}(x_n) \leqslant \frac{z}{2}\sigma_n\right) \leqslant \frac{z\sigma_n + 2}{2E_n} \frac{10\Delta_n}{z\sigma_n} \exp\left(-\frac{z\sigma_n}{20\Delta_n}\right) \frac{200E_n\Delta_n}{19\sigma_n^2}$$

$$\leqslant \frac{1000\Delta_n^2}{19\sigma_n^2}\left(1 + \frac{2}{z\sigma_n}\right) \exp\left(-\frac{z\sigma_n}{20\Delta_n}\right).$$

Recall that we assume that $\Delta_n \geqslant 2$, which implies $\Delta_n^2 \leqslant 2\sigma_n^2$ and also $\sigma_n \geqslant 1$. We thus obtain for every $z \geqslant 1$,

$$\mathbf{P}\left(|x_n| \geqslant z \text{ and } \mathsf{LR}(x_n) \leqslant \frac{z}{2}\sigma_n\right) \leqslant \frac{2000 \times 3}{19} \exp\left(-\frac{z}{20\sqrt{2}}\right).$$

Jointly with the exponential bound from Proposition 4.4, this completes the proof of Proposition 4.5. □

Remark 4.6. — In the more general case of forests, the same argument applies and the only difference lies in (4.3) where in the numerator of the ratio in front of the probability, the term "+1" is more generally "+ϱ_n". Then in the last display of the preceding proof, the term "3" is replaced by "$1 + \varrho_n/\sigma_n$". If this ratio is bounded, as in Theorem 2.4, then the result of Proposition 4.5 stills holds, except that c_1 now depends on the sequences $(\varrho_n)_{n \geqslant 1}$ and $(d_n)_{n \geqslant 1}$, but it can be chosen independently of n and $z \geqslant 1$.

Remark 4.7. — Recall the particular case of uniformly random k_n-ary trees discussed in the introduction, in which every internal vertex has exactly k_n offspring, so $E_n = nk_n$ and $\sigma_n^2 = nk_n(k_n-1)$. In this case, Proposition 4.5 extends to the maximal height of the tree, without appealing to Lemma 4.1 since in this case, if a vertex x lies at generation zE_n/σ_n, say, then $\mathsf{LR}(x) = z(k_n-1)E_n/\sigma_n = z\sigma_n$. Then, by (4.4) and the Vervaat transform, the probability that there exists such a vertex is bounded by $c_1 e^{-c_2 z}$, where $c_1, c_2 > 0$ are universal. With this bound at hand, it is not difficult then to extend the convergence of the reduced trees provided by Theorem 1.5 to the convergence (1.2) by proving [Ald93, Equation 25]. Precisely, one can show that for every $\varepsilon > 0$, one can fix q large enough so that with high probability as $n \to \infty$, each tree in the forest consisting of the complement in $T_{d_n}^{\varrho_n}$ of the ancestors of the q random vertices has height smaller $\varepsilon E_n/\sigma_n$. We hope to obtain a more general result in the future and therefore refrain to provide the details here in this case.

4.4. Convergence of reduced forests

Let us close this section with the proof of Theorem 2.4. Recall that for $\varrho \in [0, \infty)$, we denote by X^ϱ the first-passage Brownian bridge from 0 to $-\varrho$ with duration 1 and that we denote by $\widetilde{X}_t^\varrho = X_t^\varrho - \min_{0 \leqslant s \leqslant t} X_s^\varrho$ for every $t \in [0, 1]$. Let us assume that $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ and that $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. We have shown with Proposition 2.3 the convergence

$$\left(\frac{1}{\sigma_n} W_{d_n}^{\varrho_n}(V_n t); t \in [0, 1]\right) \xrightarrow[n \to \infty]{(d)} (X^\varrho(t); t \in [0, 1]).$$

We aim at showing that, jointly with this convergence, if for $q \geqslant 1$, we sample $U_1, \ldots, U_q$ i.i.d. uniform random variables in $[0, 1]$ independently of the rest, and if

we denote by $0 = U_{(0)} < U_{(1)} < \cdots < U_{(q)}$ their ordered statistics, then we have

$$
\frac{\sigma_n}{2E_n} \left(H_{d_n}^{\varrho_n} \left(V_n U_{(i)} \right), \inf_{U_{(i-1)} \leqslant t \leqslant U_{(i)}} H_{d_n}^{\varrho_n} (V_n t) \right)_{1 \leqslant i \leqslant q}
$$

$$
\xrightarrow[n \to \infty]{(d)} \left(\widetilde{X}_{U_{(i)}}^{\varrho}, \inf_{U_{(i-1)} \leqslant t \leqslant U_{(i)}} \widetilde{X}_t^{\varrho} \right)_{1 \leqslant i \leqslant q}.
$$

The proof is inspired from the work of Broutin & Marckert [BM14] which itself finds its root in the work of Marckert & Mokkadem [MM03]; the ground idea is to compare the process $H_{d_n}^{\varrho_n}$ which describes the height of the vertices with $W_{d_n}^{\varrho_n}$ which counts the number of vertices branching off to the right of their ancestral lines. As opposed to these works, these two processes have different scaling here so one has to be more careful.

Proof of Theorem 2.4. — Let U have the uniform distribution on $[0, 1]$ independently of the forest and let x_n be the $\lfloor V_n U \rfloor'^{\text{th}}$ vertex of $T_{d_n}^{\varrho_n}$ in lexicographical order, so it has the uniform distribution in $T_{d_n}^{\varrho_n}$. Then $H_{d_n}^{\varrho_n}(\lfloor V_n U \rfloor) = |x_n|$ denotes its generation in its tree, whereas $\mathsf{R}(x_n)$ as defined in (2.2) is the number of individuals branching off strictly to the right of its ancestral line in its tree. According to Proposition 2.3 and (2.2), the process $\sigma_n^{-1} \mathsf{R}(\lfloor V_n \cdot \rfloor)$ converges in distribution towards $\widetilde{X}^{\varrho}$; we claim that

$$
(4.6) \qquad \left| \frac{1}{\sigma_n} \mathsf{R}(x_n) - \frac{\sigma_n}{2E_n} |x_n| \right| \xrightarrow[n \to \infty]{\mathbf{P}} 0.
$$

Recall the notation $\mathsf{LR}(x_n) = \mathsf{L}(x_n) + \mathsf{R}(x_n)$ for the total number of individuals branching off the ancestral line of x_n and recall that, by symmetry, $\mathsf{L}(x_n)$ and $\mathsf{R}(X_n)$ have the same law (but they are not independent in general). Fix $\varepsilon, \eta > 0$. Let $K > 0$ be such that

$$
\mathbf{P}\Big(\mathsf{LR}(x_n) \leqslant K \sigma_n \quad \text{and} \quad |x_n| \leqslant K E_n / \sigma_n \Big) \geqslant 1 - \eta
$$

for every n large enough. This is ensured e.g. by Proposition 4.4 and Remark 4.6. Then the probability that $|\frac{1}{\sigma_n} \mathsf{R}(x_n) - \frac{\sigma_n}{2E_n} |x_n|| > \varepsilon$ is bounded above by

$$
\eta + \mathbf{P} \left(\left| \mathsf{R}(x_n) - \frac{\sigma_n^2}{2E_n} |x_n| \right| > \varepsilon \sigma_n, \mathsf{LR}(x_n) \leqslant K \sigma_n, \text{ and } |x_n| \leqslant K E_n / \sigma_n \right).
$$

We next proceed similarly to the previous proof, appealing to the spinal decomposition obtained in Lemma 4.1 with $q = 1$. Recall the notation from this lemma; consider the event that $|x_n| = h$ and that for all $0 \leqslant i < h$, the ancestor of x_n at generation i has k_i offspring and the j_i'th one is the ancestor of x_n at generation $i + 1$; according to (4.3) its probability is bounded by

$$
\frac{\varrho_n + \sum\limits_{i=1}^{h} (k_i - 1)}{V_n} \cdot \mathbf{P} \left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\} \right).
$$

By decomposing according to the height of x_n and taking the worst case, we see that the probability that $|\frac{1}{\sigma_n}\mathsf{R}(x_n) - \frac{\sigma_n}{2E_n}|x_n|| > \varepsilon$ is bounded above by

$$\eta + \frac{KE_n}{\sigma_n}\frac{\varrho_n + K\sigma_n}{V_n} \sup_{h \leqslant K\,E_n/\sigma_n} \mathbf{P}\left(\left|\sum_{i \leqslant h}\left(\xi_{d_n}(i) - \chi_{d_n}(i)\right) - \frac{\sigma_n^2}{2E_n}h\right| > \varepsilon\sigma_n\right).$$

From our assumption, $\frac{KE_n}{\sigma_n}\frac{\varrho_n + K\sigma_n}{V_n}$ converges to $K(\varrho + K)$. From the triangle inequality, the last probability is bounded above by the sum of

$$\mathbf{P}\left(\left|\sum_{i \leqslant h}\left((\xi_{d_n}(i) - \chi_{d_n}(i)) - \frac{\xi_{d_n}(i) - 1}{2}\right)\right| > \frac{\varepsilon\sigma_n}{2}\right)$$

and

$$\mathbf{P}\left(\left|\sum_{i \leqslant h}\frac{\xi_{d_n}(i) - 1}{2} - \frac{\sigma_n^2}{2E_n}h\right| > \frac{\varepsilon\sigma_n}{2}\right).$$

Recall that the $\xi_{d_n}(i)$'s are identically distributed, with mean and variance given in (4.1). Furthermore, as discussed in the end of Section 4.1 these random variables are obtained by successive picks without replacement in an urn, and therefore are negatively correlated. In particular, the variance of their sum is bounded by the sum of their variances, see e.g. [Ald85, Proposition 20.6]. The Markov inequality then yields for every $h \leqslant KE_n/\sigma_n$

$$\mathbf{P}\left(\left|\sum_{i \leqslant h}\frac{\xi_{d_n}(i) - 1}{2} - \frac{\sigma_n^2}{2E_n}h\right| > \frac{\varepsilon\sigma_n}{2}\right) \leqslant \frac{h}{\varepsilon^2\sigma_n^2}\mathrm{Var}\left(\xi_{d_n}(1) - 1\right) \leqslant \frac{K\Delta_n}{\varepsilon^2\sigma_n},$$

which converges to 0. Moreover, conditionally on the $\xi_{d_n}(i)$'s, the random variables $\xi_{d_n}(i) - \chi_{d_n}(i)$ are independent and uniformly distributed on $\{0, \ldots, \xi_{d_n}(i) - 1\}$, with mean $(\xi_{d_n}(i) - 1)/2$ and variance $(\xi_{d_n}(i)^2 - 1)/12$. Similarly, the Markov inequality applied conditionally on the $\xi_{d_n}(i)$'s yields for every $h \leqslant KE_n/\sigma_n$:

$$\mathbf{P}\left(\left|\sum_{i \leqslant h}\left((\xi_{d_n}(i) - \chi_{d_n}(i)) - \frac{\xi_{d_n}(i) - 1}{2}\right)\right| > \frac{\varepsilon\sigma_n}{2}\right) \leqslant \frac{4}{\varepsilon^2\sigma_n^2}\cdot\mathbf{E}\left[\sum_{i \leqslant h}\frac{\xi_{d_n}(i)^2 - 1}{12}\right]$$

$$\leqslant \frac{\Delta_n + 1}{3\varepsilon^2\sigma_n^2}\cdot h\cdot\mathbf{E}\left[\xi_{d_n}(1) - 1\right]$$

$$\leqslant K\frac{\Delta_n + 1}{3\varepsilon^2\sigma_n},$$

which also converges to 0. This completes the proof of (4.6), which, combined with Proposition 2.3 and (2.2), yields the convergence

$$\frac{\sigma_n}{2E_n}\left(H_{d_n}^{\varrho_n}\left(V_nU_{(i)}\right)\right)_{0 \leqslant i \leqslant q} \xrightarrow[n \to \infty]{(d)} \left(\widetilde{X}_{U_{(i)}}^{\varrho}\right)_{0 \leqslant i \leqslant q}$$

for any $q \geqslant 1$ fixed.

In order to obtain the full statement of the proposition, we need to prove the following: Assume that the forest $T_{d_n}^{\varrho_n}$ reduced to the ancestors of q i.d.d. vertices has q leaves (this occurs with high probability since the height of such a vertex

is at most of order $E_n/\sigma_n = o(E_n)$ as we have seen) and a random number, say, $b \in \{1, \ldots, q-1\}$ of branchpoints; then remove these b branchpoints from the reduced forest to obtain a collection of $q + b$ single branches. Then we claim that uniformly for $1 \leqslant i \leqslant q + b$, the length of the i'th branch multiplied by $\frac{\sigma_n}{2E_n}$ is close to σ_n^{-1} times the number of vertices branching off strictly to the right of this path in the original forest. Since this number is encoded by the Łukasiewicz path, the proposition then follows from Proposition 2.3. Such a comparison follows similarly as in the case $q = 1$ above from the spinal decomposition of Lemma 4.1: now we have to consider not only the length h of a single branch, but those $h_1, \ldots, h_{q+b}$ of all the branches, which is compensated by the factor $(\sigma_n/V_n)^{q+b}$ in this lemma, the other terms before the probability are bounded. We leave the details to the reader. $\qquad\square$

5. On the label process

The aim of this section is to prove Theorems 2.5 and 2.6 on the label process $L_{d_n}^{\varrho_n}$ of a labelled forest $(T_{d_n}^{\varrho_n}, \ell)$ sampled uniformly at random in $\mathbf{LT}_{d_n}^{\varrho_n}$. Let us first prove Theorem 2.5 which asserts that the sequence

$$\left((\sigma_n + \varrho_n)^{-1/2} L_{d_n}^{\varrho_n} (V_n t) ; t \in [0, 1] \right)_{n \geqslant 1}$$

is tight in the space $C([0, 1], \mathbf{R})$. We shall rely on Proposition 4.2 and adapt the argument from [Mar18b]. Then in Section 5.2 we prove that in the case $\varrho_n \gg \sigma_n$ of a large number of trees, this process converges in distribution towards a Brownian bridge as stated in Theorem 2.6 (1) by identifying the limit of the finite-dimensional marginals. On the other hand in the case of no macroscopic degree, where both $\varrho_n \sim \varrho\sigma_n$ and $\Delta_n \ll \sigma_n$, we deduce Theorem 2.6 (2) from the convergence of the finite-dimensional *random* marginals which is itself proved in Section 5.3, appealing to Theorem 2.4. Throughout this section we shall make an extensive use of Lemma 4.1.

5.1. Tightness of the label process

Theorem 2.5 extends [Mar18b, Proposition 7] which is restricted to the case of a single tree and in a "finite-variance regime", when σ_n^2 is of order E_n. Many arguments generalise here so shall only briefly recall them and focus on the main difference. First, we shall need a technical result which resembles [Mar18b, Corollary 3], but a slight adaptation is needed here.

Recall the notation $\chi_z \in \{1, \ldots, k_{pr(z)}\}$ for the relative position of a vertex $z \in T_{d_n}^{\varrho_n}$ amongst its siblings as well as the interval notation $\rrbracket x, y \rrbracket$ for the simple path going from the vertex x (excluded) to the vertex y.

LEMMA 5.1. — *Consider the following event which we denote by $\mathcal{E}_n$: for every pair of vertices $\widehat{x}$ and x such that $\widehat{x}$ is an ancestor of x and $\#\{z \in \rrbracket\widehat{x}, x\rrbracket : k_{pr(z)} \geqslant 2\} > 2 \cdot 16^2 \ln(4V_n)$, it holds that*

$$\frac{\#\left\{ z \in \rrbracket\widehat{x}, x\rrbracket : \chi_z = 1 \text{ and } k_{pr(z)} \geqslant 2 \right\}}{\#\left\{ z \in \rrbracket\widehat{x}, x\rrbracket : k_{pr(z)} \geqslant 2 \right\}} \leqslant \frac{3}{4}.$$

Then $\mathbf{P}(\mathcal{E}_n^c) \leqslant 2(V_n - d_n(1))^{-2}$.

In words, if the number of vertices with degree different from 1 tends to infinity, then with high probability, there is no branch along which, amongst the individuals which are not single child, as soon as this quantity is at least logarithmic in V_n, the proportion of individuals which are the left-most (or right-most by symmetry) child of their parent is larger than $3/4$. This will be needed to use a symmetry argument in the proof of Theorem 2.5; note that we exclude the individuals with only one child since the label increment in this case is null.

Proof. — Let us first explain the values appearing in the lemma. Since the statement completely ignores the individuals with only one child, then we may replace the degree sequence d_n by $d_n'(k) = d_n(k)\mathbf{1}_{\{k \neq 1\}}$ and couple the two labelled forests with degree sequence d_n and d_n', by removing all the vertices with outdegree 1 in $T_{d_n}^{\varrho_n}$. Let us denote by $V_n' = V_n - d_n(1)$, $E_n' = E_n - d_n(1)$, and $d_n'(0) = d_n(0)$ the number of vertices, edges, and leaves respectively of the new forest $T_{d_n}^{\varrho_n \prime}$. Let us set

$$g_n = \frac{d_n'(0)}{8V_n'}, \qquad c_n = 2g_n + \frac{V_n' - d_n'(0)}{E_n'}, \qquad l_n = g_n^{-2} \ln\left(g_n^{-1}\left(V_n'\right)^2\right).$$

Then $g_n \leqslant 1/8$ and since each inner vertex has degree at least 2, then $V_n' - d_n'(0) \leqslant E_n'/2$, and thus $c_n \leqslant 3/4$. On the other hand we also have $g_n \geqslant 1/16$ so $l_n \leqslant 16^2 \ln(16(V_n')^2) \leqslant 2 \cdot 16^2 \ln(4V_n)$. We shall prove that the probability to find a pair $(\widehat{x}, x)$ such that $\widehat{x}$ is an ancestor of x and $\#\{z \in [\![\widehat{x}, x]\!] : k_{pr(z)} \geqslant 2\} > l_n$ and

$$\frac{\#\left\{z \in [\![\widehat{x}, x]\!] : \chi_z = 1 \text{ and } k_{pr(z)} \geqslant 2\right\}}{\#\left\{z \in [\![\widehat{x}, x]\!] : k_{pr(z)} \geqslant 2\right\}} > c_n$$

is smaller than or equal to $2(V_n')^{-2}$. Let $\mathcal{E}_n' \supset \mathcal{E}_n^c$ denote this event, this implies our claim.

For a vertex x in the forest $T_{d_n}^{\varrho_n \prime}$ and $1 \leqslant i \leqslant |x|$, let us denote by $\alpha_i(x)$ the ancestor of x at height $|x| - i + 1$, so $\alpha_1(x)$ is x itself, $\alpha_2(x)$ is its parent, etc. We may then rewrite the event $\mathcal{E}_n'$ as

$$\mathcal{E}_n' = \bigcup_{x \in T_{d_n}^{\varrho_n \prime}} \bigcup_{l_n \leqslant l \leqslant |x|} \left\{\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{\alpha_i(x)}=1\right\}} > c_n l\right\}.$$

Let x_n be a vertex sampled uniformly at random in $T_{d_n}^{\varrho_n \prime}$; a union bound yields

$$\mathbf{P}\left(\mathcal{E}_n'\right) \leqslant V_n' \cdot \mathbf{P}\left(\bigcup_{l=l_n}^{|x_n|} \left\{\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{\alpha_i(x_n)}=1\right\}} > c_n l\right\}\right).$$

Appealing to Lemma 4.1, since the ratio $(\varrho_n + \sum_{i=1}^{h}(k_i - 1))/V_n'$ in (4.3) is bounded by 1, then the previous right-hand side is bounded above by

$$V_n' \sum_{h \leqslant V_n'} \sum_{l=l_n}^{h} \mathbf{P}\left(\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{d_n'}(i)=1\right\}} > c_n l\right).$$

Recall that $c_n = 2g_n + \frac{V'_n - d'_n(0)}{E'_n}$, then a union bound yields

$$\mathbf{P}\left(\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{d'_n}(i)=1\right\}} > c_n l\right)$$

$$= \mathbf{P}\left(\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{d'_n}(i)=1\right\}} - l\frac{V'_n - d'_n(0)}{E'_n} > 2g_n l\right)$$

$$\leqslant \mathbf{P}\left(\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{d'_n}(i)=1\right\}} - \sum_{i=1}^{l} \frac{1}{\xi_{d'_n}(i)} > g_n l\right) + \mathbf{P}\left(\sum_{i=1}^{l} \frac{1}{\xi_{d'_n}(i)} - l\frac{V'_n - d'_n(0)}{E'_n} > g_n l\right).$$

Note that each $\xi_{d'_n}(i)^{-1}$ takes values in $[0, 1]$ and has mean

$$\mathbf{E}\left[\xi_{d'_n}(i)^{-1}\right] = \sum_{k \geqslant 1} k^{-1} \frac{k d'_n(k)}{E'_n} = \frac{V'_n - d'_n(0)}{E'_n}.$$

If these variables were obtained by sampling *with* replacement, then we could apply a well known concentration result, see e.g. [McD98, Theorem 2.3] to obtain

$$\mathbf{P}\left(\sum_{i=1}^{l} \xi_{d'_n}(i)^{-1} - l\frac{V'_n - d'_n(0)}{E'_n} > g_n l\right) \leqslant \exp\left(-2g_n^2 l\right).$$

As already noted, this remains true here since this concentration is obtained by controlling the Laplace transform of $\sum_{i=1}^{l} \xi_{d'_n}(i)^{-1}$ and the expectation of any convex function of this sum is bounded above by the corresponding expectation when sampling with replacement, see e.g. [Ald85, Proposition 20.6]. Further, conditionally on the $\xi_{d'_n}(i)$'s, the variables $\mathbf{1}_{\{\chi_{d'_n}(i)=1\}}$ are independent and Bernoulli distributed, with parameter $\xi_{d'_n}(i)^{-1}$ respectively, so we have similarly

$$\mathbf{P}\left(\sum_{i=1}^{l} \mathbf{1}_{\left\{\chi_{d'_n}(i)=1\right\}} - \sum_{i=1}^{l} \xi_{d'_n}(i)^{-1} > g_n l\right) \leqslant \exp\left(-2g_n^2 l\right).$$

Since $g_n \in (0, 1/2)$, then it holds that $1 - \exp(-2g_n^2) \geqslant g_n^2$, so we obtain

$$\mathbf{P}\left(\mathcal{E}'_n\right) \leqslant V_n \sum_{h \leqslant V'_n} \sum_{l=l_n}^{h} 2 \exp\left(-2g_n^2 l\right)$$

$$\leqslant 2 \left(V'_n\right)^2 \frac{\exp\left(-2g_n^2 l_n\right)}{1 - \exp\left(-2g_n^2\right)}$$

$$\leqslant 2 \left(V'_n\right)^2 g_n^{-2} \exp\left(-2g_n^2 l_n\right).$$

Finally, we have chosen $l_n = g_n^{-2} \ln(g_n^{-1}(V'_n)^2)$, so $\mathbf{P}(\mathcal{E}'_n) \leqslant 2(V'_n)^{-2}$. $\qquad\square$

We may now prove the tightness of the label process, relying on this result and Proposition 4.2.

Proof of Theorem 2.5. — Suppose first that the number $V_n - d_n(1)$ of vertices with outdegree different from 1 does not tend to infinity; upon extracting a subsequence, assume that it is uniformly bounded. Then the maximal degree Δ_n, and so the maximal label increment along an edge, is bounded (since the number of leaves is), and there are a bounded number of edges along which the label increment is not

zero. In this case, if furthermore the number ϱ_n of trees is also bounded, then it is clear that the label process is tight, and the scaling factor is of constant order. On the other hand, if $V_n - d_n(1)$ is bounded but ϱ_n tends to infinity, then the largest gap between labels in the same tree is uniformly bounded, so it tends to 0 after the scaling, only the labels of the roots of the trees can give a non zero contribution; the latter precisely converge after scaling to a Brownian bridge as discussed more precisely in the proof of Theorem 2.6 (1) below.

We henceforth assume that $V_n - d_n(1) \to \infty$ as $n \to \infty$. Fix $\varepsilon > 0$ arbitrarily small, let C be as in Proposition 4.2 and let $\mathcal{E}_n$ be the intersection of the event in this proposition and that of Lemma 5.1, whose probability is therefore at least $1 - 2\varepsilon$ for every n large enough. We claim that for every $q > 4$, for every $\beta \in (0, q/4 - 1)$, there exists a constant $C > 0$, which depends on q, β and possibly $(d_n)_{n \geqslant 0}$, such that for every n large enough, for every $0 \leqslant s \leqslant t \leqslant 1$, it holds that

$$(5.1) \qquad \mathbf{E}\left[\left|L_{d_n}^{\varrho_n}(V_n s) - L_{d_n}^{\varrho_n}(V_n t)\right|^q \,\middle|\, \mathcal{E}_n\right] \leqslant C \cdot (\sigma_n + \varrho_n)^{q/2} \cdot |t - s|^{1+\beta}.$$

The standard Kolmogorov criterion then implies that for every $\gamma \in (0, 1/4)$, for another constant $K > 0$,

$$\lim_{K \to \infty} \limsup_{n \to \infty} \mathbf{P}\left(\sup_{s \neq t} \frac{\left|L_{d_n}^{\varrho_n}(V_n s) - L_{d_n}^{\varrho_n}(V_n t)\right|}{(\sigma_n + \varrho_n)^{1/2} \cdot |t - s|^\gamma} > K \,\middle|\, \mathcal{E}_n\right) = 0.$$

Then the same holds for the unconditioned probability and this implies the tightness as claimed.

Exactly as in the proof of Proposition 4.2, it suffices to prove (5.1) for the "one-reduced" forest $T_{d_n}^{\varrho_n\prime}$ with degree sequence $d_n'(k) = d_n(k)\mathbf{1}_{\{k \neq 1\}}$ obtained by removing all the vertices with outdegree 1 in $T_{d_n}^{\varrho_n}$. Indeed, as there, these vertices induce a null label increment and they are located uniformly at random in the forest; therefore, with the notation of the proof of Proposition 4.2, the suitable rescaled time intervals $[s, t]$ and $[s_n, t_n]$ corresponding to the same vertices respectively in $T_{d_n}^{\varrho_n}$ and in $T_{d_n}^{\varrho_n\prime}$ are of the same order in the sense that $\mathbf{E}[|t_n - s_n|^{1+\beta}] \leqslant c \cdot |t - s|^{1+\beta}$ for some constant c independent of s and t; this allows to deduce the bound (5.1) from $T_{d_n}^{\varrho_n\prime}$ to $T_{d_n}^{\varrho_n}$ (with another constant C). Note that the event $\mathcal{E}_n$ is not affected by the reduction.

Let us summarise: we assume henceforth that the degree sequence is such that $d_n(1) = 0$ and $V_n \to \infty$ as $n \to \infty$, and our aim is to prove (5.1). The rest of the proof follows closely that of [Mar18b, Proposition 7]. We may, and do, suppose that $V_n s$ and $V_n t$ are integers. Let us view $T_{d_n}^{\varrho_n}$ as a tree by attaching all the roots to an extra root vertex, let x and y be the vertices visited at time $V_n s$ and $V_n t$ respectively, and let $\widehat{x}$ and $\widehat{y}$ be the children of their last common ancestor which are ancestor of x and y respectively. Note that if x and y belong to different trees of $T_{d_n}^{\varrho_n}$, then $\widehat{x}$ and $\widehat{y}$ are the roots of these two trees and then the term $\chi_{\hat{y}} - \chi_{\hat{x}}$ below counts the number of other roots strictly between them, plus one. Let us decompose the label increment $L_{d_n}^{\varrho_n}(V_n s) - L_{d_n}^{\varrho_n}(V_n t) = \ell(x) - \ell(y)$ as follows:

$$\ell(x) - \ell(y) = \sum_{z \in \rrbracket \hat{x}, x \rrbracket} \left(\ell(z) - \ell(pr(z))\right) + (\ell(\widehat{y}) - \ell(\widehat{y})) + \sum_{z \in \rrbracket \hat{y}, y \rrbracket} \left(\ell(z) - \ell(pr(z))\right).$$

Then it was argued in [Mar18b], see Equation 22 and the next few lines there, that $\mathbf{E}[|\ell(x) - \ell(y)|^q \mid T_{d_n}^{\varrho_n}]$ is bounded above by some constant times

$$\left(\sum_{z \in \rrbracket \hat{x}, x \rrbracket} \left(k_{pr(z)} - \chi_z \right) + (\chi_{\hat{y}} - \chi_{\hat{x}}) \right)^{q/2} + \left(\sum_{z \in \rrbracket \hat{y}, y \rrbracket} \chi_z \right)^{q/2}.$$

Furthermore, from Lemma 2.1 and then Proposition 4.2, one gets that there exists $C > 0$ such that, for every $n \geqslant 1$, with probability at least $1 - \varepsilon$, it holds

$$\sum_{z \in \rrbracket \hat{x}, x \rrbracket} \left(k_{pr(z)} - \chi_z \right) + (\chi_{\hat{y}} - \chi_{\hat{x}}) = W_{d_n}^{\varrho_n}(V_n s) - \inf_{r \in [s,t]} W_{d_n}^{\varrho_n}(V_n r)$$

$$\leqslant C \cdot (\sigma_n + \varrho_n) \cdot |t - s|^{(1-\varepsilon)/2}.$$

uniformly for $0 \leqslant s < t \leqslant 1$. By choosing ε small enough, we deduce that

$$\left(\sum_{z \in \rrbracket \hat{x}, x \rrbracket} \left(k_{pr(z)} - \chi_z \right) + (\chi_{\hat{y}} - \chi_{\hat{x}}) \right)^{q/2} \leqslant C^{q/2} \cdot (\sigma_n + \varrho_n)^{q/2} \cdot |t - s|^{1+\beta},$$

uniformly for $0 \leqslant s < t \leqslant 1$ with probability $1 - \varepsilon$.

We next want a similar bound for the moments of the other term $\sum_{z \in \rrbracket \hat{y}, y \rrbracket} \chi_z$. We would like to proceed symmetrically, using the 'mirror forest' obtained by flipping the order of the children of every vertex, but there is a difference with the first term: the quantity $\sum_{z \in \rrbracket \hat{x}, x \rrbracket} (k_{pr(z)} - \chi_z)$ counts the individuals which lie strictly to the right of the branch $\rrbracket \hat{x}, x \rrbracket$ in the sense that their parent belongs to $\rrbracket \hat{x}, x \llbracket$ but not themselves, and they lie after this branch in the depth-first search order; on the other hand, the term $\sum_{z \in \rrbracket \hat{y}, y \rrbracket} \chi_z$ counts similarly the number, say $\mathsf{L}(\rrbracket \hat{y}, y \rrbracket)$ of individuals which lie strictly to the left of the branch $\rrbracket \hat{y}, y \rrbracket$ *plus* the length $\#\rrbracket \hat{y}, y \rrbracket$ of this branch. In other words, whilst individuals which are the right-most child of their parent are not counted in the first term (since $\ell(z k_z) = \ell(z)$), those which are the left-most child of their parent *are* counted in the second term (since $\ell(z1) \neq \ell(z)$ in general); Lemma 5.1 allows us to control this number as follows. Recall that we assume $d_n(1) = 0$. Under the event $\mathcal{E}_n$, if the branch $\rrbracket \hat{y}, y \rrbracket$ has length greater than $l_n = 2 \cdot 16^2 \ln(4V_n)$, then the proportion of individuals which are the first child of their parent is at most $3/4$; all other vertices (a proportion at least $1/4$) contribute to $\mathsf{L}(\rrbracket \hat{y}, y \rrbracket)$ so there are at most $\mathsf{L}(\rrbracket \hat{y}, y \rrbracket)$ of them and therefore on the event $\mathcal{E}_n$, it holds that

$$\#\rrbracket \hat{y}, y \rrbracket + \mathsf{L}(\rrbracket \hat{y}, y \rrbracket) \leqslant 4\mathsf{L}(\rrbracket \hat{y}, y \rrbracket) \mathbf{1}_{\{\#\rrbracket \hat{y}, y \rrbracket \geqslant l_n\}} + \left(l_n + \mathsf{L}(\rrbracket \hat{y}, y \rrbracket) \right) \mathbf{1}_{\{\#\rrbracket \hat{y}, y \rrbracket < l_n\}}$$

$$\leqslant l_n + 5\mathsf{L}(\rrbracket \hat{y}, y \rrbracket).$$

Now, using the mirror forest, the previous bound on the left branch yields

$$\mathbf{E}\left[\mathsf{L}(\rrbracket \hat{y}, y \rrbracket)^{q/2} \right] \leqslant C (\sigma_n + \varrho_n)^{q/2} |t - s|^{1+\beta}$$

for some $C > 0$ and so, for some other constant $C > 0$,

$$\mathbf{E}\left[\left(\sum_{z \in \llbracket \widehat{y}, y \rrbracket} \chi_z\right)^{q/2} \mathbf{1}_{\mathcal{E}_n}\right] = \mathbf{E}\left[\left(\#\rrbracket\widehat{y}, y\rrbracket + \mathsf{L}\left(\rrbracket\widehat{y}, y\rrbracket\right)\right)^{q/2}\mathbf{1}_{\mathcal{E}_n}\right]$$

$$\leqslant C\left(l_n^{q/2} + (\sigma_n + \varrho_n)^{q/2}|t - s|^{1+\beta}\right).$$

Recall that we assume $d_n(1) = 0$, so $\sigma_n^2 + \varrho_n \geqslant E_n + \varrho_n = V_n$ so $V_n^{-1/2}(\sigma_n + \varrho_n)$ is bounded away from 0. Finally, recall also that we have assumed $V_n s$ and $V_n t$ to be integers so $|t - s| \geqslant V_n^{-1}$, and thus, for $\beta < q/4 - 1$, we have $(\sigma_n + \varrho_n)^{q/2}|t - s|^{1+\beta} \geqslant (\sigma_n + \varrho_n)^{q/2}V_n^{-(1+\beta)}$ which is bounded below by some constant times a positive power of V_n, so, for n large enough (but independently of s and t), it is larger than l_n which is of order $\ln V_n$. It follows that for another constant $C > 0$,

$$\mathbf{E}\left[\left(\sum_{z \in \llbracket \widehat{y}, y \rrbracket} \chi_z\right)^{q/2} \mathbf{1}_{\mathcal{E}_n}\right] \leqslant C \cdot (\sigma_n + \varrho_n)^{q/2} \cdot |t - s|^{1+\beta},$$

which completes the proof of (5.1). $\square$

5.2. Brownian labelled forests

Recall from Section 2.4, in particular (2.3), that the label process $L_{d_n}^{\varrho_n}$ can be written as

$$L_{d_n}^{\varrho_n}(k) = \widetilde{L}_{d_n}^{\varrho_n}(k) + b_{d_n}^{\varrho_n}\left(1 - \underline{W}_{d_n}^{\varrho_n}(k)\right), \qquad (0 \leqslant k \leqslant V_n)$$

where, conditionally given $T_{d_n}^{\varrho_n}$, the two processes $\widetilde{L}_{d_n}^{\varrho_n}(\cdot)$ and $b_{d_n}^{\varrho_n}(1 - \underline{W}_{d_n}^{\varrho_n}(\cdot))$ are independent. On the one hand $\widetilde{L}_{d_n}^{\varrho_n}$ is the concatenation of the label process of each tree taken individually, so where all labels have been shifted so that each root has label 0. On the other hand $b_{d_n}^{\varrho_n}(1 - \underline{W}_{d_n}^{\varrho_n}(k))$ gives the value of the label of the root vertex of the tree containing the k'th vertex, with $\underline{W}_{d_n}^{\varrho_n}(k) = \min_{0 \leqslant i \leqslant k} W_{d_n}^{\varrho_n}(i)$ and $(b_{d_n}^{\varrho_n}(i))_{1 \leqslant i \leqslant \varrho_n}$ which is independent of $T_{d_n}^{\varrho_n}$ and uniformly distributed in $B_{\varrho_n}^{\geqslant -1}$.

We claim that if $\lim_{n \to \infty} \sigma_n^{-1}\varrho_n = \infty$, then the convergence in distribution

$$\left((2\varrho_n)^{-1/2}\left(b_{d_n}^{\varrho_n}(\varrho_n t), \widetilde{L}_{d_n}^{\varrho_n}(V_n t)\right); t \in [0, 1]\right) \xrightarrow[n \to \infty]{(d)} \left((\mathbf{b}_t, 0); t \in [0, 1]\right),$$

holds in $C([0, 1], \mathbf{R}^2)$, where $\mathbf{b}$ is an independent standard Brownian bridge from 0 to 0 with duration 1. This yields Theorem 2.6 (1) as noted just after the statement of this theorem.

Proof of Theorem 2.6 (1). — First, consider the random walk S with step distribution $\mathbf{P}(S_1 = i) = 2^{-(i+2)}$ for all $i \geqslant -1$, which is centred and with variance 2. One can check that for any $k \geqslant 1$, the law of the bridge obtained by conditioning S to satisfy $S_k = 0$ has the uniform distribution in $B_k^{\geqslant -1}$. Therefore the process $b_{d_n}^{\varrho_n}$ has the law of such a random walk bridge with length ϱ_n and a conditional version of Donsker's invariance principle, see e.g. [Bet10, Lemma 10] for a detailed proof, shows

that $(2\varrho_n)^{-1/2} b^{\varrho_n}_{d_n}(\varrho_n \cdot)$ converges in distribution towards $\mathbf{b}$. Recall Proposition 2.3, we conclude that the convergence in distribution

$$\left((2\varrho_n)^{-1/2} b^{\varrho_n}_{d_n}\left(1 - \underline{W}^{\varrho_n}_{d_n}(\lfloor V_n t \rfloor)\right); t \in [0,1]\right) \xrightarrow[n \to \infty]{(d)} (\mathbf{b}_t; t \in [0,1])$$

holds in $C([0,1], \mathbf{R})$.

We next claim that $\varrho_n^{-1/2} \widetilde{L}^{\varrho_n}_{d_n}$ converges in probability to the null process, which yields our result by the decomposition (2.3). By this decomposition and the preceding convergence, tightness follows from Theorem 2.5 so it remains to prove the convergence of the (one-dimensional) marginals. Fix $k \geqslant 1$, and $X_k = (X_{k,1}, \ldots, X_{k,k})$ a uniform random bridge in $B^{\geqslant -1}_k$, and set $X_{k,0} = 0$; using the representation as a bridge of the random walk S, one easily infers that for every $q \geqslant 2$, there exists $C(q) > 0$ such that and every $i, j \in \{0, \ldots, k\}$,

$$(5.2) \qquad \mathbf{E}\left[|X_{k,i} - X_{k,j}|^q\right] \leqslant C(q) \cdot |i - j|^{q/2},$$

see e.g. Le Gall & Miermont [LGM11, Lemma 1]. Fix $t \in [0,1]$; building on this bound, as in the preceding proof, we obtain

$$\mathbf{E}\left[\left|\widetilde{L}^{\varrho_n}_{d_n}(\lfloor V_n t \rfloor)\right|^q \,\bigg|\, T^{\varrho_n}_{d_n}\right] \leqslant C(q) \cdot \left|W^{\varrho_n}_{d_n}(\lfloor V_n t \rfloor) - \underline{W}^{\varrho_n}_{d_n}\lfloor V_n t \rfloor\right|^{q/2}.$$

See e.g. [Mar18b, Equation 22] and the next few lines there, with $u = 0$ and $v = \lfloor V_n t \rfloor$; only the case of a single tree is considered there but the argument extends readily. By Proposition 2.3, the right-hand side divided by $\varrho_n^{q/2}$ converges in probability to 0 which yields the convergence in probability to 0 of $\varrho_n^{-1/2} \widetilde{L}^{\varrho_n}_{d_n}(\lfloor V_n t \rfloor)$ and the proof is complete. □

Let us next turn to the second part of Theorem 2.6. Recall from Section 2.4 the process Z^ϱ which describes Brownian labels on a forest of Brownian trees coded by the first-passage bridge X^ϱ; it takes the form

$$Z^\varrho_t = \widetilde{Z}^\varrho_t + \sqrt{3} \cdot \mathbf{b}^\varrho_{-\underline{X}^\varrho_t} \quad \text{for every} \quad 0 \leqslant t \leqslant 1,$$

where $\widetilde{Z}^\varrho$ is defined conditionally on $\widetilde{X}^\varrho = X^\varrho - \underline{X}^\varrho$, where $\underline{X}$ is the running infimum of X^ϱ, as a centred Gaussian process with covariance $\mathbf{E}[\widetilde{Z}^\varrho_s \widetilde{Z}^\varrho_t \mid \widetilde{X}^\varrho] = \min_{r \in [s,t]} \widetilde{X}^\varrho_r$ for every $0 \leqslant s \leqslant t \leqslant 1$, and where $\mathbf{b}^\varrho$ is an independent standard Brownian bridge from 0 to 0 with duration ϱ.

Then Theorem 2.6(2) claims that if $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ with $\varrho \in [0, \infty)$ and $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$ then the convergence in distribution

$$\left(\left(\frac{3}{2\sigma_n}\right)^{1/2} L^{\varrho_n}_{d_n}(V_n t); t \in [0,1]\right) \xrightarrow[n \to \infty]{(d)} (Z^\varrho_t; t \in [0,1])$$

holds in $C([0,1], \mathbf{R})$. Since we already proved tightness of this sequence, it only remains to characterise the subsequential limits.

PROPOSITION 5.2. — *Suppose that* $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho$ *for some* $\varrho \in [0, \infty)$ *and* $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. *For every* $n, q \geqslant 1$, *sample* (T_{d_n}, ℓ) *uniformly at random in* $\mathbf{LT}^{\varrho_n}_{d_n}$ *and, independently, sample* $U_1, \ldots, U_q$ *uniformly at random in* $[0,1]$. *Sample*

$\widetilde{Z}^\varrho$ *independently of* $(U_1, \ldots, U_q)$, *then we have*

$$(5.3) \qquad \left(\frac{3}{2\sigma_n}\right)^{1/2} \left(\widetilde{L}^{\varrho_n}_{d_n}(V_n U_1), \ldots, \widetilde{L}^{\varrho_n}_{d_n}(V_n U_q)\right) \xrightarrow[n \to \infty]{(d)} \left(\widetilde{Z}^\varrho_{U_1}, \ldots, \widetilde{Z}^\varrho_{U_q}\right)$$

jointly with the convergence of $W^{\varrho_n}_{d_n}$ *and* $H^{\varrho_n}_{d_n}$ *in Proposition 2.3 and Theorem 2.4.*

Let us differ the proof of the proposition to the next subsection and finish now the proof of Theorem 2.6.

Proof of Theorem 2.6(2). — Suppose that $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho \in [0, \infty)$ and $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n = 0$. Then the sequence $(\sigma_n^{-1} W^{\varrho_n}_{d_n}(\lfloor V_n t \rfloor))_{t \in [0,1]}$ converges in distribution towards X^ϱ by Proposition 2.3 and, as in the previous proof, the sequence $((2\varrho_n)^{-1/2} b^{\varrho_n}_{d_n}(\varrho_n t))_{n \geqslant 1}$ converges in distribution towards $\mathbf{b}$ whenever $\varrho_n \to \infty$. Since the two are independent, we deduce that, when $\varrho > 0$,

$$\left(\left(\frac{3}{2\sigma_n}\right)^{1/2} b^{\varrho_n}_{d_n}\left(1 - \underline{W}^{\varrho_n}_{d_n}(V_n t)\right)\right)_{t \in [0,1]} \xrightarrow[n \to \infty]{(d)} \left((3\varrho)^{1/2} \mathbf{b}_{-\varrho^{-1} \underline{X}^\varrho_t}\right)_{t \in [0,1]}$$

in $C([0,1], \mathbf{R})$ and the limit has the same law as $(\sqrt{3} \mathbf{b}^\varrho_{-\underline{X}^\varrho_t})_{t \in [0,1]}$; when $\varrho = 0$, the limit is instead the null process.

Next, still as in the previous proof, by Theorem 2.5, the sequence $(\sigma_n^{-1/2} \widetilde{L}^{\varrho_n}_{d_n})_{n \geqslant 1}$ is tight; then this equicontinuity combined with the uniform continuity of the process $\widetilde{Z}^\varrho$ allows to approximate the marginals evaluated at deterministic times by sampling sufficiently many i.i.d. uniform random times. Hence Proposition 5.2 yields actually a convergence in $C([0,1], \mathbf{R})$. Combined with Proposition 2.3 we obtain the joint convergence

$$\left(\frac{1}{\sigma_n} W^{\varrho_n}_{d_n}(V_n t), \left(\frac{3}{2\sigma_n}\right)^{1/2} \widetilde{L}^{\varrho_n}_{d_n}(V_n t)\right)_{t \in [0,1]} \xrightarrow[n \to \infty]{(d)} \left(X^\varrho_t, \widetilde{Z}^\varrho_t\right)_{t \in [0,1]}$$

in $C([0,1], \mathbf{R}^2)$. Recall the decomposition (2.3) of $L^{\varrho_n}_{d_n}$; we deduce from the conditional independence of $b^{\varrho_n}_{d_n}$ and $\widetilde{L}^{\varrho_n}_{d_n}$, that

$$\left(\left(\frac{3}{2\sigma_n}\right)^{1/2} L^{\varrho_n}_{d_n}(V_n t)\right)_{t \in [0,1]} \xrightarrow[n \to \infty]{(d)} \left(\widetilde{Z}^\varrho_t + 3^{1/2} \mathbf{b}^\varrho_{-\underline{X}^\varrho_t}\right)_{t \in [0,1]},$$

and the right-hand side is the definition of the process Z^ϱ. $\qquad \square$

5.3. Marginals of the label process

Let us now prove Proposition 5.2 which is the last remaining ingredient to Theorem 2.6 and therefore to Theorem 1.2. We start with the case $q = 1$ and comment then on the general case.

We shall need the following technical estimate. Let x_n denote a vertex in $T^{\varrho_n}_{d_n}$, with height $|x_n|$ and for every $1 \leqslant p \leqslant |x_n|$, let $k_p(x_n)$ denote the number of offspring of the ancestor of x_n at height $p - 1$ and let $j_p(x_n)$ the index of the offspring of this ancestor which is also an ancestor of x_n. Conditionally on $T^{\varrho_n}_{d_n}$ and x_n, let $(B_p(x_n); 1 \leqslant p \leqslant x_n)$ be independent random variables, where each $B_p(x_n)$ has the

law of a uniformly random bridge in $B^{\geqslant -1}_{k_p(x_n)}$ evaluated at time $j_p(x_n)$. Note that they are centred, let us denote by $\sigma^2_p(x_n)$ the variance of $B_p(x_n)$, which is measurable with respect to $T^{\varrho_n}_{d_n}$ and x_n.

LEMMA 5.3. — *Assume that* $\lim_{n \to \infty} \sigma_n^{-1} \varrho_n = \varrho \in [0, \infty)$ *and* $\lim_{n \to \infty} \sigma_n^{-1} \Delta_n$ $= 0$. *For every* $n \geqslant 1$, *sample* $(T^{\varrho_n}_{d_n}, \ell)$ *uniformly at random in* $\mathbf{LT}^{\varrho_n}_{d_n}$ *and independently sample a vertex* x_n *uniformly at random. Then*

$$\frac{3E_n}{\sigma_n^2 |x_n|} \sum_{p=1}^{|x_n|} \sigma^2_p(x_n) \xrightarrow[n \to \infty]{\mathbf{P}} 1.$$

Proof. — Let us first prove the claim on the event $\mathcal{E}_n(K)$ defined as follows with $K > 0$ large but fixed:

(5.4) $$\mathcal{E}_n(K) = \left\{ \frac{\sigma_n}{E_n} |x_n| \in \left[K^{-1}, K \right] \text{ and } \mathsf{LR}(x_n) \leqslant K \sigma_n \right\}.$$

Indeed by combining Proposition 4.4, Remark 4.6, as well as Theorem 2.4 for the lower bound on $|x_n|$, we see that $\lim_{K \to \infty} \liminf_{n \to \infty} \mathbf{P}(\mathcal{E}_n(K)) = 1$. Let $\sigma^2(k, j)$ denotes the variance of a uniformly random bridge in $B^{\geqslant -1}_k$ evaluated at time j, which is known explicitly, see e.g. Marckert & Miermont [MM07, page 1664[(1)]]: we have

$$\sigma^2(k, j) = \frac{2j(k - j)}{k + 1}, \qquad \text{so} \qquad \sum_{j=1}^{k} \sigma^2(k, j) = \frac{k(k - 1)}{3}.$$

Fix $\delta > 0$, then by (4.3),

$$\mathbf{P}\left(\left\{ \left| \frac{3E_n}{\sigma_n^2 |x_n|} \sum_{p=1}^{|x_n|} \sigma^2(x_n) - 1 \right| > \delta \right\} \cap \mathbf{1}_{\mathcal{E}_n(K)} \right)$$

$$\leqslant K(\varrho + K + o(1)) \sup_{K^{-1} E_n/\sigma_n \leqslant h \leqslant K E_n/\sigma_n} \mathbf{P}\left(\left| \frac{3E_n}{\sigma_n^2 h} \sum_{p=1}^{h} \sigma^2_p(\xi_{d_n}(p), \chi_{d_n}(p)) - 1 \right| > \delta \right)$$

We calculate the first two moments: to ease notation, let (ξ, ξ', χ, χ') have the law of the quadruple $(\xi_{d_n}(1), \xi_{d_n}(2), \chi_{d_n}(1), \chi_{d_n}(2))$. Since χ has the uniform distribution on $\{1, \ldots, \xi\}$, then

$$\mathbf{E}\left[\sum_{p=1}^{h} \sigma^2(\xi_{d_n}(p), \chi_{d_n}(p)) \right] = h \cdot \mathbf{E}\left[\xi^{-1} \sum_{j=1}^{\xi} \sigma^2(\xi, j) \right] = h \cdot \mathbf{E}\left[\frac{\xi - 1}{3} \right] = \frac{\sigma_n^2 h}{3E_n}.$$

It only remains to prove that the variance is small; we have similarly

$$\mathbf{E}\left[\left(\frac{3E_n}{\sigma_n^2 h} \sum_{p=1}^{h} \sigma^2(\xi_{d_n}(p), \chi_{d_n}(p)) \right)^2 \right]$$

$$= \left(\frac{3E_n}{\sigma_n^2 h} \right)^2 \left(h \mathbf{E}\left[\sigma^2(\xi, \chi)^2 \right] + h(h - 1)\mathbf{E}\left[\sigma^2(\xi, \chi)\sigma^2(\xi', \chi') \right] \right).$$

[(1)] Note that they consider uniform random bridges in $B^{\geqslant -1}_{k+1}$!

Since the χ's are independent conditionally on the ξ's, we have

$$\mathbf{E}\left[\sigma^2(\xi,\chi)\sigma^2(\xi',\chi')\right] = \mathbf{E}\left[\frac{(\xi-1)(\xi'-1)}{9}\right] = \left(\frac{\sigma_n^2}{3E_n}\right)^2 (1+o(1)),$$

where we used that two samples without replacement decorrelate as the number of possible picks tends to infinity. We conclude that

$$\mathbf{E}\left[\frac{3E_n}{\sigma_n^2 h}\sum_{p=1}^{h}\sigma^2\left(\xi_{d_n}(p),\chi_{d_n}(p)\right)\right] = 1,$$

and, since $\sigma^2(\xi,\chi) \leqslant \xi \leqslant \Delta_n$, then this random variable has variance

$$\left(\frac{3E_n}{\sigma_n^2 h}\right)^2\left(h\mathbf{E}\left[\sigma^2(\xi,\chi)^2\right] + h(h-1)\mathbf{E}\left[\sigma^2(\xi,\chi)\sigma^2(\xi',\chi')\right]\right) - 1$$

$$\leqslant \left(\frac{3E_n}{\sigma_n^2 h}\right)^2\left(h\Delta_n\frac{\sigma_n^2}{3E_n} + h(h-1)\left(\frac{\sigma_n^2}{3E_n}\right)^2(1+o(1))\right) - 1$$

$$\leqslant \frac{3E_n}{\sigma_n^2 h}\Delta_n + o(1),$$

which converges to 0 uniformly in $h \in \left[K^{-1}E_n/\sigma_n, KE_n/\sigma_n\right]$ since we assume that $\Delta_n = o(\sigma_n)$. This proves that $\frac{3E_n}{\sigma_n^2|x_n|}\sum_{p=1}^{|x_n|}\sigma_p^2(x_n)$ converges in probability to 1 on the event $\mathcal{E}_n(K)$ and therefore also unconditionally by letting further K tend to infinity. $\square$

We may now prove Proposition 5.2.

Proof of Proposition 5.2. — Let us start with the case $q = 1$. Let U have the uniform distribution on $[0,1]$ and let x_n be the uniformly random vertex visited at the time $\lfloor V_n U\rfloor$ in lexicographical order, with label $\ell(x_n) = L_{d_n}^{\varrho_n}(\lfloor V_n U\rfloor)$ and height $|x_n| = H_{d_n}^{\varrho_n}(\lfloor V_n U\rfloor)$. Let us write

$$\left(\frac{3}{2\sigma_n}\right)^{1/2}\ell(x_n) = \left(\frac{\sigma_n}{2E_n}|x_n|\right)^{1/2}\left(\frac{3E_n}{\sigma_n^2|x_n|}\right)^{1/2}\ell(x_n).$$

Recall from Theorem 2.4 that $\frac{\sigma_n}{2E_n}|x_n|$ converges in distribution towards $\widetilde{X}_U^{\varrho}$, it is therefore equivalent to show that, jointly with this convergence, we have

$$(5.5)\qquad\qquad \left(\frac{3E_n}{\sigma_n^2|x_n|}\right)^{1/2}\ell(x_n) \underset{n\to\infty}{\Longrightarrow} N(0,1),$$

where $N(0,1)$ denotes the standard Gaussian distribution and "$\Rightarrow$" is a slight abuse of notation to refer to the weak convergence of the law of the random variable. The idea is to decompose $\ell(x_n)$ as the sum of the label increments $(B_p(x_n); 1 \leqslant p \leqslant |x_n|)$ between two consecutive ancestors, as defined before the preceding lemma. To ease notation, let us set $\Sigma_n^2 = \sum_{p=1}^{|x_n|}\sigma_p^2(x_n)$, then the Central Limit Theorem for triangular arrays shows that the conditional law of $\Sigma_n^{-1}\ell(x_n)$ converges to the standard Gaussian distribution as soon as for every $\varepsilon > 0$,

$$\Sigma_n^{-2} \sum_{p=1}^{|x_n|} \mathbf{E}\left[B_p(x_n)^2 \mathbf{1}_{\{|B_p(x_n)| > \varepsilon\Sigma_n\}} \;\Big|\; T_{d_n}^{\varrho_n}, x_n \right] \xrightarrow[n \to \infty]{} 0. \tag{5.6}$$

Our claim (5.5) then follows from (5.6) and Lemma 5.3. As in the proof of this lemma we shall implicitly work on the intersection of the event $\mathcal{E}_n(K)$ defined in (5.4) and $\{|\Sigma_n^2 \frac{3E_n}{\sigma_n^2 |x_n|} - 1| \leqslant K^{-1}\}$ with $K > 0$ fixed, whose probability tends to 1 when first $n \to \infty$ and then $K \to \infty$. Note that this allows to replace equivalently Σ_n^2 by σ_n in (5.6). As proved in [LGM11, Lemma 1], for every $r \geqslant 1$, there exists a constant $K_{(r)} > 0$ such that for every $k \geqslant j \geqslant 1$, the $(2r)$'th moment of a uniformly random bridge in $B_k^{\geqslant -1}$ evaluated at time j is bounded above by $K_{(r)}|k-j|^r$. Fix $\varepsilon > 0$, then, by first using the Cauchy–Schwarz and the Markov inequalities,

$$\mathbf{E}\left[B_p(x_n)^2 \mathbf{1}_{\{|B_p(x_n)| > \varepsilon\Sigma_n\}} \;\Big|\; T_{d_n}^{\varrho_n}, x_n \right]^2 \leqslant \mathbf{E}\left[B_p(x_n)^4 \;\Big|\; T_{d_n}^{\varrho_n}, x_n \right] \frac{\mathbf{E}\left[|B_p(x_n)|^2 \;\Big|\; T_{d_n}^{\varrho_n}, x_n \right]}{\varepsilon^2 \Sigma_n^2}$$

$$\leqslant K_{(2)} |k_p(x_n) - j_p(x_n)|^2 \frac{K_{(1)} |k_p(x_n) - j_p(x_n)|}{\varepsilon^2 \Sigma_n^2}$$

$$\leqslant \frac{K'}{\varepsilon^2} \frac{|k_p(x_n) - j_p(x_n)|^3}{\sigma_n},$$

where $K' > 0$ depends on K. Therefore the left hand side in (5.6) is upper bounded by some constant that depends on K and ε times

$$\frac{1}{\sigma_n} \sum_{p=1}^{|x_n|} \sqrt{\frac{|k_p(x_n) - j_p(x_n)|^3}{\sigma_n}} \leqslant \sqrt{\frac{\Delta_n}{\sigma_n}} \frac{1}{\sigma_n} \sum_{p=1}^{|x_n|} |k_p(x_n) - j_p(x_n)|,$$

where the inequality is obtained by upper bounding $|k_p(x_n) - j_p(x_n)|$ by Δ_n. Observe that the last sum precisely equals the value of the Łukasiewicz path at time $\lfloor V_n U \rfloor$, which is of order σ_n with high probability. Since $\Delta_n/\sigma_n \to 0$, then we conclude that indeed (5.6) holds on an event which is measurable with respect to $T_{d_n}^{\varrho_n}$ and x_n and whose probability is arbitrarily close to 1. Combined with Lemma 5.3 this concludes the proof of (5.5) and thus of the proposition for $q = 1$.

We next sketch the proof of the claim for $q = 2$, the general case $q \geqslant 3$ is exactly the same and only requires more notation. We shall need the following result, which claims that in the case of no macroscopic degree, there is no macroscopic label increment in the sense that the maximal difference along an edge is small:

$$\sigma_n^{-1/2} \max_{x \in T_{d_n}^{\varrho_n}} \left| \max_{1 \leqslant j \leqslant k_x} \ell(xj) - \min_{1 \leqslant j \leqslant k_x} \ell(xj) \right| \xrightarrow[n \to \infty]{\mathbf{P}} 0. \tag{5.7}$$

This extends [Mar18b, Proposition 2] to which we refer for a proof, just replace $N_{\mathbf{n}}$ there by σ_n^2. Let x_n and y_n be independent uniform random vertices of $T_{d_n}^{\varrho_n}$ and let us consider the three branches consisting of: their common ancestors on one side, the ancestors of x_n only on another side, and finally those of y_n only. According to Theorem 2.4, their lengths jointly converge in distribution, when rescaled by a factor $\sigma_n/(2E_n)$. By (5.7), the label increments between the last common ancestor of x_n and y_n and its offspring which are ancestors of x_n and y_n respectively are

small compared to our scaling, and the total label increment on each of the three remaining branches are independent; it only remains to prove that the law of these increments, multiplied by $(3E_n/\sigma_n^2)^{1/2}$ and divided by the square-root of their length converges to the standard Gaussian law, as in (5.5). This can be obtained in the very same way as in the case $q = 1$, appealing to Lemma 4.1 to compare the content of each branch with the sequence $(\xi_{d_n}(p), \chi_{d_n}(p))_p$. As in the proof of Theorem 2.4, the fact that we now have three branches is compensated by the factor $(\sigma_n/V_n)^3$ in Lemma 4.1, we leave the details to the reader. $\square$

6. Stable Boltzmann maps

In this last section, we study size-conditioned stable Boltzmann random maps by relying on our main results. In Sections 6.1 and 6.2, we present the precise setup and the assumptions we shall make on such laws, by relying on the Łukasiewicz path of the associated labelled forest. Then in Section 6.3 we state and prove the main result of this section, Theorem 6.2, which shows that the random degree distribution satisfies the main assumptions from the introduction, according to the value of the stability index. We then easily derive the behaviour of the maps in Sections 6.4 and 6.5, see Theorems 6.3 and 6.5.

Throughout this section, we shall divide by real numbers which depend on an integer n, and consider conditional probabilities with respect to events which depend on n; we shall therefore always implicitly restrict ourselves to those values of n for which such quantities are well defined and statements such as "as $n \to \infty$" should be understood along the appropriate subsequence of integers.

6.1. On stable domains of attraction

Throughout this section, we work with a sequence $(X_i)_{i \geqslant 1}$ of i.i.d. copies of a random variable X with distribution, say ν, supported by (a subset of) $\mathbf{Z}_{\geqslant -1}$ with $\nu(-1) \neq 0$, with finite first moment and $\mathbf{E}[X_1] = 0$. For every integer $n \geqslant 1$, we let $S_n = X_1 + \cdots + X_n$, and set $S_0 = 0$. For every integers $n \geqslant 1$ and $k \geqslant -1$, let $K_k(n) = \#\{1 \leqslant i \leqslant n : X_i = k\}$ be the number of jumps of size k up to time n; for a subset $A \subset \mathbf{Z}_{\geqslant -1}$, set then $K_A(n) = \sum_{k \in A} K_k(n)$. Finally, for $\varrho \geqslant 1$, let $\zeta(S, \varrho) = \inf\{i \geqslant 1 : S_i = -\varrho\}$, and simply write $\zeta(S)$ for $\zeta(S, 1)$.

We recall that a measurable function $l : [0, \infty) \to [0, \infty)$ is said to be *slowly varying* (at infinity) if for every $c > 0$, we have $\lim_{x \to \infty} l(cx)/l(x) = 1$. For $\alpha \in [1, 2]$, we say that ν satisfies (H_α) if the tail distribution can be written as

$$\mathbf{P}(X \geqslant n) = n^{-\alpha} L_1(n),$$

where L_1 is slowly varying. We also include the case where X has finite variance in (H_2). Finally, when $\alpha = 1$, we say that ν satisfies $(\mathrm{H}_1^{\mathrm{loc}})$ when the mass function can be written as

$$\mathbf{P}(X = n) = n^{-2} L_1(n),$$

where L_1 is slowly varying. The assumption (H_α) corresponds to the *domain of attraction* of a stable law with index α, and $(\mathrm{H}_1^{\mathrm{loc}})$ is more restrictive than (H_1).

When $\alpha \in (1,2]$, it is well known that there exists a sequence $(a_n)_{n \geqslant 1}$ such that $(n^{-1/\alpha}a_n)_{n \geqslant 1}$ is slowly varying and $a_n^{-1}(X_1 + \cdots + X_n)$ converges in distribution to $X^{(\alpha)}$ with Laplace transform given by $\mathbf{E}[\exp(-tX^{(\alpha)})] = \exp(t^\alpha)$ for $t > 0$. Note that $X^{(2)}$ has the Gaussian distribution with variance 2; as a matter of fact, if X has finite variance σ^2, then we may take $a_n = (n\sigma^2/2)^{1/2}$. Moreover, there exists another slowly varying function L such that for every $n \geqslant 1$, we have $\mathrm{Var}(X\mathbf{1}_{\{X \leqslant n\}}) = n^{2-\alpha}L(n)$. This function is related to L_1 by

$$(6.1) \qquad \lim_{n \to \infty} \frac{L_1(n)}{L(n)} = \lim_{n \to \infty} \frac{n^2 \mathbf{P}(X \geqslant n)}{\mathrm{Var}\left(X\mathbf{1}_{\{X \leqslant n\}}\right)} = \frac{2-\alpha}{\alpha},$$

see Feller [Fel71, Chapter XVII, Equation 5.16]. Finally, according to [Kor17, Equation 7], we have

$$(6.2) \qquad \lim_{n \to \infty} \frac{nL(a_n)}{a_n^\alpha} = \frac{1}{(2-\alpha)\Gamma(-\alpha)},$$

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

In the case $\alpha = 1$, when (H_1) is the domain of attraction of a Cauchy distribution, in addition to the sequence $(a_n)_{n \geqslant 1}$, there exists another sequence $(b_n)_{n \geqslant 1}$ such that both $(n^{-1}a_n)_{n \geqslant 1}$ and $(n^{-1}b_n)_{n \geqslant 1}$ are slowly varying, with $b_n \to -\infty$ and $b_n/a_n \to -\infty$, and now $a_n^{-1}(X_1 + \cdots + X_n - b_n)$ converges in distribution to $X^{(1)}$ with Laplace transform given by $\mathbf{E}[\exp(-tX^{(1)})] = \exp(t \ln t)$ for $t > 0$. An example to have in mind when dealing with this rather unusual regime is given by Kortchemski & Richier [KR19]: take $\nu(n) \sim \frac{c}{n^2 \ln(n)^2}$, then $a_n \sim \frac{cn}{\ln(n)^2}$ and $b_n \sim -\frac{cn}{\ln(n)}$.

6.2. Local limit theorems and conditioned paths

Let us now describe our model of random paths, these assumptions and notations shall be used throughout this section. We shall assume that ν satisfies either (H_α) for some $\alpha \in [1,2]$, or $(\mathrm{H}_1^{\mathrm{loc}})$, and then the sequences $(a_n)_{n \geqslant 1}$ and $(b_n)_{n \geqslant 1}$ will always be as in the preceding subsection. Fix $n \geqslant 1$, $\varrho_n \geqslant 1$, and a set $A \subset \mathbf{Z}_{\geqslant -1}$ with $\nu(A) > 0$, then:

- Let us denote by W^{ϱ_n} the path S stopped when first hitting $-\varrho_n$; if $\varrho_n = 1$, simply write W.
- If ν satisfies (H_α) for some $\alpha \in (1,2]$, then we let $W_{n,A}^{\varrho_n}$ be the path W^{ϱ_n} conditioned to have made n jumps with value in the set A.
- If ν satisfies $(\mathrm{H}_1^{\mathrm{loc}})$, then we only consider $A = \mathbf{Z}_{\geqslant -1}$ and we define $W_n^{\varrho_n} = W_{n,\mathbf{Z}_{\geqslant -1}}^{\varrho_n}$ similarly so it reduces to conditioning the walk S to first hit $-\varrho_n$ at time n and stopping it there.
- If ν satisfies (H_1), then again $A = \mathbf{Z}_{\geqslant -1}$ but now $\varrho_n = 1$, and we let $W_{\geqslant n}$ be the walk S conditioned to stay nonnegative at least up to time n and stopped when first hitting -1 (at a random time).

We recalled how to construct the path $W_n^{\varrho_n}$ by cyclically shifting a bridge, i.e. the walk S up to time n, conditioned to satisfy $S_n = -\varrho_n$. This construction can be generalised to any subset A, by cyclically shifting the walk S up to its n'th jump with value in A, and conditioned to be at $-\varrho_n$ at this moment, see Kortchemski [Kor12, Lemma 6.4] for a detailed proof in the case $\varrho_n = 1$ and $A = \{-1\}$; it extends here: replacing "$= -1$" by "$\in A$" does not change anything, and when $\varrho_n \geqslant 2$, there are ϱ_n cyclically shifted bridges which are first-passage bridges, but this factor ϱ_n cancels since the cycle lemma is used twice.

The study of such bridges relies strongly on local limit theorems, which are used to compare them to unconditioned random walks. By [Kor12, Equation 50] the following holds: suppose that ν satisfies (H_α) with $\alpha \in (1, 2]$, fix $A \subset \mathbf{Z}_{\geqslant -1}$ such that $\nu(A) > 0$, and, if ν has infinite variance, suppose that either A or $\mathbf{Z}_{\geqslant -1} \setminus A$ is finite. Then for any sequence $(\varrho_n)_{n \geqslant 1}$ such that $\limsup_{n \to \infty} a_n^{-1} \varrho_n < \infty$, we have

$$(6.3) \qquad \lim_{n \to \infty} \left| n \cdot \mathbf{P} \left(K_A \left(\zeta \left(S, \varrho_n \right) \right) = n \right) - \frac{\nu(A)^{1/\alpha} \varrho_n}{a_n} \cdot p_1 \left(-\frac{\nu(A)^{1/\alpha} \varrho_n}{a_n} \right) \right| = 0,$$

where p_1 is the density of $X^{(\alpha)}$. We shall mostly be interested in the cases $A = \mathbf{Z}_{\geqslant -1}$, $A = \{-1\}$, and $A = \mathbf{Z}_{\geqslant 0}$. Nevertheless, Thévenin [Thé20] proposed a different method, relying on a multivariate local limit theorem, in order to lift the restriction that either A or its complement should be finite; Proposition 6.5 there corresponds to (6.3) in the case $\rho_n = 1$ on which the latter work focuses, but the general case can be adapted similarly.

In the Cauchy regime $\alpha = 1$, such a local limit theorem is not known for general sets A, which is the reason why we restrict ourselves to the case $A = \mathbf{Z}_{\geqslant -1}$. Then $K_{\mathbf{Z}_{\geqslant -1}}(k) = k$ so what we consider in (6.3) is simply the first hitting time of $-\varrho_n$. After cyclic shift, the probability that the latter equals n is given by $n^{-1} \varrho_n \mathbf{P}(S_n = -\varrho_n)$. When ν satisfies (H_1^{loc}), we read from the recent work of Berger [Ber19, Theorem 2.4] that, when $\varrho_n = O(a_n) = o(|b_n|)$, we have

$$(6.4) \qquad \mathbf{P} \left(\zeta \left(S, \varrho_n \right) = n \right) \underset{n \to \infty}{\sim} \varrho_n \cdot \frac{L_1 \left(|b_n| - \varrho_n \right)}{\left(|b_n| - \varrho_n \right)^2},$$

see [Ber19, Equation 2.10] with $p = \alpha = 1$ and $x = |b_n| - \varrho_n$. When ν only satisfies (H_1), Kortchemski & Richier [KR19, Proposition 12] proved that, in the case $\varrho_n = 1$,

$$(6.5) \qquad \mathbf{P} \left(\zeta(S) \geqslant n \right) \underset{n \to \infty}{\sim} \Lambda(n) \cdot \frac{L_1(|b_n|)}{|b_n|},$$

where Λ is some other slowly varying function unimportant here.

6.3. On the empirical jump distribution

We study in this subsection the random paths $W_{n,A}^{\varrho_n}$ and $W_{\geqslant n}$ described above. Let $\zeta(W^{\varrho_n})$ denote the number of steps of the path W^{ϱ_n}, and for every subset $B \subset \mathbf{Z}_{\geqslant -1}$, let

$$J_B \left(W^{\varrho_n} \right) = \# \left\{ 1 \leqslant i \leqslant \zeta \left(W^{\varrho_n} \right) : W^{\varrho_n}(i) - W^{\varrho_n}(i-1) \in B \right\},$$

and define similar quantities for $W_{n,A}^{\varrho_n}$ and $W_{\geqslant n}$. Our first result is a strong law of large numbers.

LEMMA 6.1. — *Let B any subset of $\mathbf{Z}_{\geqslant -1}$. Assume that either ν satisfies (H_α) for some $\alpha \in (1,2]$ and then take $A \subset \mathbf{Z}_{\geqslant -1}$ arbitrary (with $\nu(A) > 0$), or that ν satisfies $(\mathrm{H}_1^{\mathrm{loc}})$ and then take $A = \mathbf{Z}_{\geqslant -1}$. Finally assume that $\limsup_{n \to \infty} a_n^{-1} \varrho_n < \infty$. Then we have*

$$n^{-1} J_B \left(W_{n,A}^{\varrho_n} \right) \xrightarrow[n \to \infty]{a.s.} \frac{\nu(B)}{\nu(A)}.$$

If ν satisfies (H_1), then

$$\zeta \left(W_{\geqslant n} \right)^{-1} J_B \left(W_{\geqslant n} \right) \xrightarrow[n \to \infty]{a.s.} \nu(B).$$

In particular, when $\alpha > 1$, recalling that $n^{-1/\alpha} a_n$ is slowly varying, we have for $B = \mathbf{Z}_{\geqslant -1}$:

$$n^{-1} \zeta \left(W_{n,A}^{\varrho_n} \right) \xrightarrow[n \to \infty]{a.s.} \nu(A)^{-1} \quad \text{and so} \quad a_n^{-1} a_{\zeta \left(W_{n,A}^{\varrho_n} \right)} \xrightarrow[n \to \infty]{a.s.} \nu(A)^{-1/\alpha}.$$

When ν satisfies $(\mathrm{H}_1^{\mathrm{loc}})$, we simply have $\zeta(W_n^{\varrho_n}) = n$; finally, when ν satisfies (H_1), we obtain appealing to [KR19, Theorem 30] that

$$|b_n|^{-1} \left| b_{\zeta \left(W_n^{\varrho_n} \right)} \right| \xrightarrow[n \to \infty]{(d)} I,$$

where I has the law $\mathbf{P}(I \geqslant x) = x^{-1}$ for all $x \geqslant 1$. Therefore in any case the natural scaling factor for our conditioned paths, which should involve their total length ζ, may be replaced by a factor depending only on n. The more general statement, for an arbitrary set B, shall be used later when dealing with random maps.

Proof. — Let us detail the case $\alpha \in (1,2]$. Fix $\delta > 0$ small, we claim that there exist $c, C > 0$ such that for every n, we have

$$(6.6) \qquad \mathbf{P} \left(\left| \frac{n}{J_B \left(W_{n,A}^{\varrho_n} \right)} - \frac{\nu(A)}{\nu(B)} \right| > \frac{\delta}{n^{1/4}} \right) \leqslant C \mathrm{e}^{-cn^{1/2}}.$$

Since the right-hand side is summable, this indeed shows that $n/J_B(W_{n,A}^{\varrho_n})$ converges to $\nu(A)/\nu(B)$ almost surely. Let us write this conditional probability as

$$\frac{1}{\mathbf{P} \left(J_A \left(W^{\varrho_n} \right) = n \right)} \cdot \mathbf{P} \left(\left| \frac{J_A \left(W^{\varrho_n} \right)}{J_B \left(W^{\varrho_n} \right)} - \frac{\nu(A)}{\nu(B)} \right| > \frac{\delta}{n^{1/4}} \text{ and } J_A \left(W^{\varrho_n} \right) = n \right).$$

Recall the local limit estimate in (6.3); it is known that p_1 is continuous and positive, so in particular bounded away from 0 and ∞ on any compact interval. Hence $\varrho_n^{-1} n a_n \mathbf{P}(J_A(W^{\varrho_n}) = n)$ is bounded away from 0 and ∞ since $\varrho_n = O(a_n)$. Moreover, $\varrho_n \geqslant 1$ and $a_n = O(n^{3/2})$ and finally $\zeta(W^{\varrho_n}) \geqslant J_A(W^{\varrho_n})$, so the probability that $\left| \frac{n}{J_B(W_{n,A}^{\varrho_n})} - \frac{\nu(A)}{\nu(B)} \right| > \frac{\delta}{n^{1/4}}$ is bounded above by some constant times

$$n^{5/2} \mathbf{P} \left(\left| \frac{J_A \left(W^{\varrho_n} \right)}{J_B \left(W^{\varrho_n} \right)} - \frac{\nu(A)}{\nu(B)} \right| > \frac{\delta}{n^{1/4}} \text{ and } \zeta \left(W^{\varrho_n} \right) \geqslant n \right),$$

and it remains to bound this last probability. Straightforward calculations show that if n is large enough (so e.g. $\delta n^{-1/4} < \nu(B)/2$), then the event

$$\left\{ \left| \frac{J_A(W^{\varrho_n})}{\zeta(W^{\varrho_n})} - \nu(A) \right| \leqslant \frac{\delta}{n^{1/4}} \right\} \cap \left\{ \left| \frac{J_B(W^{\varrho_n})}{\zeta(W^{\varrho_n})} - \nu(B) \right| \leqslant \frac{\delta}{n^{1/4}} \right\}$$

is included in

$$\left\{ \left| \frac{J_A(W^{\varrho_n})}{J_B(W^{\varrho_n})} - \frac{\nu(A)}{\nu(B)} \right| \leqslant \frac{\delta'}{n^{1/4}} \right\},$$

for some explicit δ' which depends on δ, $\nu(A)$ and $\nu(B)$. We may write

$$\mathbf{P} \left(\left| \frac{J_A(W^{\varrho_n})}{\zeta(W^{\varrho_n})} - \nu(A) \right| > \frac{\delta}{n^{1/4}} \text{ and } \zeta(W^{\varrho_n}) \geqslant n \right)$$

$$\leqslant \sum_{N \geqslant n} \mathbf{P} \left(\left| \frac{J_A(W^{\varrho_n})}{N} - \nu(A) \right| > \frac{\delta}{N^{1/4}} \text{ and } \zeta(W^{\varrho_n}) = N \right)$$

$$\leqslant \sum_{N \geqslant n} \mathbf{P} \left(\left| \frac{\#\{1 \leqslant i \leqslant N : X_i \in A\}}{N} - \nu(A) \right| > \frac{\delta}{N^{1/4}} \right)$$

$$\leqslant \sum_{N \geqslant n} 2 \exp\left(-2\delta^2 N^{1/2} \right),$$

where the last bound follows from the Chernoff bound for binomial distributions. The last sum is bounded by some constant times $n^{1/2} \exp(-2\delta^2 n^{1/2})$; the same holds with the set B, so we obtain after a union bound,

$$\mathbf{P} \left(\left| \frac{n}{J_B(W_{n,A}^{\varrho_n})} - \frac{\nu(A)}{\nu(B)} \right| > \frac{\delta}{n^{1/4}} \right) \leqslant K n^3 \mathrm{e}^{-2\delta^2 n^{1/2}},$$

for some $K > 0$ and the proof in the case $\alpha \in (1, 2]$ is complete.

The argument is very similar in the case $\alpha = 1$, appealing to (6.4) and (6.5), we leave the details to the reader. $\square$

In the next Theorem 6.2, we prove that the empirical jump distribution of our conditioned paths fits in our general framework. For any path $P = (P_i)_{i \geqslant 0}$, we let $\Delta(P) = \max\{i \geqslant 1 : P_i - P_{i-1}\}$ be the largest jump, and we let $\Delta'(P)$ be the second largest jump.

THEOREM 6.2. — *Assume that ν satisfies (H_α) for some $\alpha \in [1, 2]$ or $(\mathrm{H}_1^{\mathrm{loc}})$ and fix $A \subset \mathbf{Z}_{\geqslant -1}$ with $\nu(A) > 0$ arbitrary if $\alpha > 1$, but take $A = \mathbf{Z}_{\geqslant -1}$ otherwise. Let $(\varrho_n)_{n \geqslant 1}$ be such that $\lim_{n \to \infty} a_n^{-1} \varrho_n = \varrho \nu(A)^{-1/\alpha}$ for some $\varrho \in [0, \infty)$.*

(1) *If ν satisfies (H_2), then*

$$a_n^{-2} \sum_{k \geqslant 1} k(k+1) J_k\left(W_{n,A}^{\varrho_n} \right) \xrightarrow[n \to \infty]{\mathbf{P}} \frac{2}{\nu(A)} \qquad \text{and} \qquad a_n^{-1} \Delta\left(W_{n,A}^{\varrho_n} \right) \xrightarrow[n \to \infty]{\mathbf{P}} 0.$$

(2) *If ν satisfies (H_α) with $\alpha \in (1, 2)$, then $a_n^{-2} \sum_{k \geqslant 1} k(k+1) J_k(W_{n,A}^{\varrho_n})$ converges in distribution to a random variable Y_α^ϱ whose law does not depend on ν.*

(3) *If ν satisfies $(\mathrm{H}_1^{\mathrm{loc}})$, recall that we take $A = \mathbf{Z}_{\geqslant -1}$, then*

$$|b_n|^{-1}\Delta\left(W_n^{\varrho_n}\right) \xrightarrow[n\to\infty]{\mathbf{P}} 1 \quad and \quad |b_n|^{-2} \sum_{k \leqslant \Delta'\left(W_n^{\varrho_n}\right)} k(k+1) J_k{}^{\prime}\left(W_n^{\varrho_n}\right) \xrightarrow[n\to\infty]{\mathbf{P}} 0.$$

(4) *Similarly, if ν satisfies (H_1) and $\varrho_n = 1$, let I be distributed as $P(I \geqslant x) = x^{-1}$ for every $x \geqslant 1$, then*

$$|b_n|^{-1}\Delta\left(W_{\geqslant n}\right) \xrightarrow[n\to\infty]{(d)} I \quad and \quad |b_n|^{-2} \sum_{k \leqslant \Delta'\left(W_{\geqslant n}\right)} k(k+1) J_k\left(W_{\geqslant n}\right) \xrightarrow[n\to\infty]{\mathbf{P}} 0.$$

In the case $\varrho_n = 1$, when moreover ν has finite variance, Theorem 6.2 (1) was first obtained by Broutin & Marckert [BM14] for $A = \mathbf{Z}_{\geqslant -1}$ and generalised to any A in [Mar18b]. The last two statements when $\alpha = 1$ shall follow easily from [KR19]. A key idea to prove the first two statements, when $\alpha > 1$, as in [BM14, Mar18b], is first to observe that the claims are invariant under cyclic shift, so we may consider a random walk bridge instead of a first-passage bridge, and then to compare the law of these bridges with that of the unconditioned random walk S for which the claims are easy to prove. Precisely, when ν satisfies (H_α) for some $\alpha \in (1,2]$ and $A \subset \mathbf{Z}_{\geqslant -1}$, let us denote for every real $t \geqslant 1$ by

$$J_{A,t}^-(S) = \inf\left\{k \geqslant 1 : J_A\left((S_i)_{i \leqslant k}\right) = \lfloor t \rfloor\right\}$$

the instant at which the walk S makes its $\lfloor t \rfloor'^{\mathrm{th}}$ step in A. Then, as we discussed in the preceding subsection, the path $W_{n,A}^{\varrho_n}$ has the law of the Vervaat transform of the path S up to time $J_{A,n}^-(S)$ conditioned to be at $-\varrho_n$ at this time. Further, from the Markov property applied to S at time $J_{A,n/2}^-(S)$ and the local limit theorem, one can see that there exists a constant $C > 0$ such that for every $n \geqslant 1$ and every event $E_{A,n/2}(S)$ which is measurable with respect to the $J_{A,n/2}^-(S)$ first steps of the path S, we have that

$$(6.7) \qquad \limsup_{n\to\infty} \mathbf{P}\left(E_{A,n/2}(S) \mid S_{J_{A,n}^-(S)} = -\varrho_n\right) \leqslant C \limsup_{n\to\infty} \mathbf{P}\left(E_{A,n/2}(S)\right),$$

see e.g. Kortchemski [Kor12, Lemmas 6.10 and 6.11 and Equation 44], there.

We may now prove Theorem 6.2. Let us start with the Gaussian regime $\alpha = 2$.

Proof of Theorem 6.2(1). — According to the preceding discussion, it suffices to prove the convergences

$$(6.8) \qquad a_n^{-2} \sum_{k \geqslant 1} k(k+1) J_k\left((S_i)_{i \leqslant J_{A,n}^-(S)}\right) \xrightarrow[n\to\infty]{\mathbf{P}} \frac{2}{\nu(A)},$$

$$a_n^{-1}\Delta\left((S_i)_{i \leqslant J_{A,n}^-(S)}\right) \xrightarrow[n\to\infty]{\mathbf{P}} 0,$$

under the conditional probability

$$\mathbf{P}\left(\cdot \,\Big|\, S_{J_{A,n}^-(S)} = -\varrho_n\right).$$

Moreover, if we cut this bridge at time $J_{A,n/2}^-(S)$, then the time- and space-reversal of the second part has the same law as the first one. Therefore it suffices to prove

that (6.8) holds when n is replaced by $n/2$. By (6.7), it suffices finally to prove that (6.8) holds for the unconditioned random walk.

Recall the two slowly varying functions L and L_1, which are respectively given by $L_1(x) = x^2 \mathbf{P}(X \geqslant x)$ and $L(x) = \mathrm{Var}(X \mathbf{1}_{\{|X| \leqslant x\}})$ for every $x > 0$; recall from (6.1) and (6.2) that L_1/L converges to 0 and $na_n^{-2}L(a_n)$ converges to 2. Then for every $\varepsilon > 0$, it holds that

$$\mathbf{P}\left(a_n^{-1} \max_{1 \leqslant i \leqslant n} X_i \geqslant \varepsilon\right) \leqslant n\mathbf{P}(X \geqslant \varepsilon a_n) = n(\varepsilon a_n)^{-2} L_1(\varepsilon a_n) \xrightarrow[n \to \infty]{} 0,$$

where we used the fact that L_1 is slowly varying, so $L_1(\varepsilon a_n) \sim L_1(a_n) = o(L(a_n))$. Concerning the first convergence, we aim at showing that $a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i(X_i + 1)$ converges in probability to 2, which is equivalent to the fact that $a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i^2$ converges in probability to 2 since $n^{-1} \sum_{1 \leqslant i \leqslant n} X_i$ converges in probability to 0 by the law of large numbers, and $n = O(a_n^2)$. Let us fix $\varepsilon > 0$, then we have that

$$\mathbf{E}\left[a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i^2 \mathbf{1}_{\{|X_i| \leqslant \varepsilon a_n\}}\right] = na_n^{-2} \mathbf{E}\left[X^2 \mathbf{1}_{\{|X| \leqslant \varepsilon a_n\}}\right] = na_n^{-2} L(a_n)(1 + o(1)),$$

which converges to 2 and, similarly,

$$\mathrm{Var}\left(a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i^2 \mathbf{1}_{\{|X_i| \leqslant \varepsilon a_n\}}\right)$$
$$= na_n^{-4} \mathrm{Var}\left(X^2 \mathbf{1}_{\{|X| \leqslant \varepsilon a_n\}}\right) \leqslant \varepsilon^2 na_n^{-2} \mathbf{E}\left[X^2 \mathbf{1}_{\{|X| \leqslant \varepsilon a_n\}}\right]$$

which converges to $2\varepsilon^2$. We have shown that with high probability, we have $|X_i| \leqslant \varepsilon a_n$ for every $i \leqslant n$, so we conclude that indeed $a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i^2$ converges in probability to 2. We have thus shown that

$$a_n^{-2} \sum_{k \geqslant 1} k(k+1) J_k\left((S_i)_{i \leqslant n}\right) \xrightarrow[n \to \infty]{\mathbf{P}} 2 \quad \text{and} \quad a_n^{-1} \Delta\left((S_i)_{i \leqslant n}\right) \xrightarrow[n \to \infty]{\mathbf{P}} 0,$$

for the unconditioned random walk, and so, since $a_{cn} \sim c^{1/2} a_n$,

$$a_n^{-2} \sum_{k \geqslant 1} k(k+1) J_k\left((S_i)_{i \leqslant n/\nu(A)}\right) \xrightarrow[n \to \infty]{\mathbf{P}} \frac{2}{\nu(A)} \quad \text{and} \quad a_n^{-1} \Delta\left((S_i)_{i \leqslant n/\nu(A)}\right) \xrightarrow[n \to \infty]{\mathbf{P}} 0.$$

Since $n^{-1} J_{A,n}^-(S)$ converges almost surely to $\nu(A)^{-1}$ by Lemma 6.1, we obtain (6.8) for the unconditioned random walk and the proof is complete. $\qquad \square$

We next consider the regime $1 < \alpha < 2$.

Proof of Theorem 6.2(2). — The claim for the unconditioned random walk S is easy: let X be the α-stable Lévy process whose law at time 1 is $X^{(\alpha)}$, then by a classical result on random walks, the convergence in distribution of $a_n^{-1} S_n$ towards $X^{(\alpha)}$ is equivalent to that of $(a_n^{-1} S_{\lfloor nt \rfloor})_{t \geqslant 0}$ towards $(X_t)_{t \geqslant 0}$ in the Skorokhod's J_1 topology. Let $\Delta X_s = X_s - X_{s-} \geqslant 0$, then the sum $S_t = \sum_{s \leqslant t} (\Delta X_s)^2$ is well defined; in fact it is well known that the process $(S_t)_{t \geqslant 0}$ is an $\alpha/2$-stable subordinator, whose law can be easily derived from that of X. Furthermore it is easily seen to be the limit in distribution of $(a_n^{-2} \sum_{i \leqslant \lfloor nt \rfloor} X_i(X_i + 1))_{t \geqslant 0}$; note that no centring is needed here since $\alpha/2 < 1$. A simple consequence of the fact that S is a pure jump process is that the non-increasing rearrangement of the vector $(a_n^{-2} X_i(X_i + 1))_{1 \leqslant i \leqslant n}$ converges in

distribution in the ℓ^1 topology towards the decreasing rearrangement of the nonzero jumps of $(S_t)_{t \in [0,1]}$.

Let us return to our random bridges; one may adapt the arguments from [Kor12] when $\varrho_n = 1$ to obtain the convergence in distribution under

$$\mathbf{P}\left(\cdot \,\middle|\, S_{J_{A,n}^-(S)} = -\varrho_n \right)$$

of the paths

$$\left(\nu(A)^{1/\alpha} a_n^{-1} \sum_{1 \leqslant i \leqslant \lfloor J_{A,n}^-(S)t \rfloor} X_i \right)_{t \in [0,1]}$$

towards the bridge $(X_t^\varrho)_{t \in [0,1]}$ which is informally the process $(X_t)_{t \in [0,1]}$ conditioned to be at $-\varrho$ at time 1. This implies the convergence of the N largest values amongst $(X_i(X_i + 1))_{1 \leqslant i \leqslant n}$ towards the N largest values amongst $((\Delta X_t^\varrho)^2)_{0 \leqslant t \leqslant 1}$ for every N; by (6.7) and the preceding paragraph, if N is chosen large enough, the sum of all the other jumps is small so we obtain under the law

$$\mathbf{P}\left(\cdot \,\middle|\, S_{J_{A,n}^-(S)} = -\varrho_n \right), \nu(A)^{2/\alpha} a_n^{-2} \sum_{1 \leqslant i \leqslant n} X_i \left(X_i + 1 \right) \xrightarrow[n \to \infty]{(d)} \sum_{t \in [0,1]} (\Delta X_t^\varrho)^2 .$$

We conclude as in the preceding proof from the space-time reversal invariance. □

We finally consider the Cauchy regime. Let us start with the local conditioning. Recall that here we assume that $A = \mathbf{Z}_{\geqslant -1}$, and $W_n^{\varrho_n}$ is simply the walk S conditioned to first hit $-\varrho_n$ at time n.

Proof of Theorem 6.2(3). — In the case $\varrho_n = 1$, Kortchemski & Richier [KR19, Theorem 3] found the joint limit of the N largest jumps of W_n^1 for any N fixed: the largest one is equivalent to $|b_n|$, whereas the others are of order $a_n = o(|b_n|)$ which implies our first claim. This can be generalised to any $\varrho_n = O(a_n) = o(b_n)$; indeed, the key is Proposition 20 there which still applies when one takes the $X_i^{(n)}$'s to be the jumps of the path S (denoted by W there!) conditioned to $S_n = -\varrho_n$ instead of $S_n = -1$: the only feature of the case $\varrho_n = 1$ which is used is that $\mathbf{P}(S_n = -\varrho_n) \sim n\varrho_n \cdot \mathbf{P}(X = |b_n|)$, which remains valid as soon as $\varrho_n = O(a_n) = o(b_n)$ by (6.4).

By this extension of [KR19, Proposition 20] the following holds as $n \to \infty$: let $U_n = \inf\{i \leqslant n : X_i = \max_{j \leqslant n} X_j\}$ be the first time at which the path S up to time n makes its largest jump, then the law under $\mathbf{P}(\cdot \mid S_n = -\varrho_n)$ of the vector $(X_1, \ldots, X_{U_n-1}, X_{U_n+1}, \ldots, X_n)$ is close in total variation to $n-1$ i.d.d. copies of X. From this, the proof of Theorem 21 from [KR19] can be extended to show that if we first run the walk S up to time $n-1$ and then send it to $-\varrho_n$, and then we construct $Z_n^{\varrho_n}$ as the Vervaat transform of this path $(S_0, \ldots, S_{n-1}, -\varrho_n)$, then

$$(6.9) \qquad d_{TV}\left((W_n^{\varrho_n}(i))_{0 \leqslant i \leqslant n}, (Z_n^{\varrho_n}(i))_{0 \leqslant i \leqslant n} \right) \xrightarrow[n \to \infty]{} 0.$$

Finally, from these results, the proof of [KR19, Theorem 3] remains unchanged and our first claim follows. Concerning the second claim, it is equivalent to showing that $|b_n|^{-2} \sum_{i \leqslant n} X_i^2 \mathbf{1}_{\{i \neq U_n\}}$ converges to 0 under $\mathbf{P}(\cdot \mid S_n = -\varrho_n)$. As in the case

$1 < \alpha < 2$, we have that $a_n^{-2}(X_1^2 + \cdots + X_n^2)$ converges in distribution towards some $1/2$-stable random variable. Our claim then follows by (6.9) since $a_n = o(|b_n|)$. □

It only remains to consider the tail conditioning. Recall that here $A = \mathbf{Z}_{\geqslant -1}$ and $\varrho_n = 1$.

Proof of Theorem 6.2(4). — Compared to the previous proof, we do not need any adaption here and the first claim now directly follows from [KR19, Theorem 6]. The second claim is more subtle. First, we have a similar approximation to (6.9) given by [KR19, Theorem 27]; what replaces $Z_n^{\varrho_n}$ there is the process $\vec{Z}^{(n)}$ defined as follows: first I_n is the last weak ladder time of $(S_i)_{i \leqslant n}$, then conditionally on $\{I_n = j\}$, the path $\vec{Z}^{(n)}$ consists in three independent parts:

(1) First $(\vec{Z}_i^{(n)})_{i < j}$ has the law of $(S_i)_{i < j}$ conditioned to satisfy $\min_{i \leqslant j} S_i \geqslant 0$.
(2) Then we make a big jump $\vec{Z}_j^{(n)} - \vec{Z}_{j-1}^{(n)}$, sampled from $\mathbf{P}(\,\cdot\, \mid X \geqslant |b_n|)$.
(3) Finally $(\vec{Z}_{j+i}^{(n)} - \vec{Z}_j^{(n)})_{i \geqslant 0}$ continues as the unconditioned random walk S.

The big jump will be excluded in our sum and we have seen previously that the sum of N copies of X^2 is of order a_N^2 as $N \to \infty$. Therefore, if we consider only the jumps of $W_{\geqslant n}$ after its big jump, then there are less than $\zeta(W_{\geqslant n})$ of them, and $|b_n|^{-1}\zeta(W_{\geqslant n})$ converges in distribution to I as $n \to \infty$ by [KR19, Proposition 3.1], so the sum of the square of these jumps is at most of order $a_{|b_n|}^2$. We recall that a_n is defined by $n\mathbf{P}(X \geqslant a_n) \to 1$; since $\mathbf{P}(X \geqslant x) = x^{-1}L_1(x)$ and that $\mathbf{E}[|X|] < \infty$, then $L_1(x) \to 0$, therefore $a_n = o(n)$, from which we deduce that the sum of the square of the jumps of the last part is at most of order $a_{|b_n|}^2 = o(|b_n|^2)$. Finally, let us consider the jumps of the first part, up to time $I_n - 1$. As in the preceding proof, the sum of the square of the first $I_n - 1$ jumps of $\vec{Z}^{(n)}$ grows like $a_{I_n}^2 = o(|b_n|^2)$, and precisely this path converges after scaling to the meander given by a $1/2$-stable process conditioned to be nonnegative at least up to time 1. □

6.4. Boltzmann planar maps

Let us briefly recall the definition of Boltzmann distributions on planar maps, as introduced by Marckert & Miermont [MM07]. Let us also refer to [BM17] for more details. For every $\varrho \geqslant 1$, let $\mathbf{M}^\varrho$ be the set of all finite rooted bipartite maps M with perimeter 2ϱ, and let $\mathbf{PM}^\varrho$ be the set of all such rooted and pointed maps $(M, x_\star)$. Recall that $\mathbf{M}^1$ and $\mathbf{PM}^1$ can be seen as the set of maps and pointed maps without boundary by gluing the two boundary edges, we then drop the exponent 1. Let us fix a sequence of nonnegative real numbers $\mathbf{q} = (q_i; i \geqslant 0)$ which, in order to avoid trivialities, satisfies $q_i > 0$ for at least one $i \geqslant 2$; we define a measure $\mathbf{w}^\varrho$ on $\mathbf{M}^\varrho$ by setting

$$\mathbf{w}^\varrho(M) = \prod_{f \text{ inner face}} q_{\deg(f)/2}, \qquad M \in \mathbf{M}^\varrho,$$

where $\deg(f)$ is the degree of the face f. We set $\mathbf{W}^\varrho = \mathbf{w}^\varrho(\mathbf{M}^\varrho)$. We define similarly a measure $\mathbf{w}_\star^\varrho$ on $\mathbf{PM}^\varrho$ by $\mathbf{w}_\star^\varrho((M, x_\star)) = \mathbf{w}^\varrho(M)$ for every $(M, x_\star) \in \mathbf{PM}^\varrho$ and we put $\mathbf{W}_\star^\varrho = \mathbf{w}_\star^\varrho(\mathbf{PM}^\varrho)$. We say that $\mathbf{q}$ is *admissible* when $\mathbf{W}_\star = \mathbf{W}_\star^1$ is finite; this

seems stronger than requiring $\mathbf{W}^1$ to be finite, but it is not, see [BCM19]; moreover, this implies that $\mathbf{W}^\varrho_\star$ (and so $\mathbf{W}^\varrho$) is finite for any $\varrho \geqslant 1$. If $\mathbf{q}$ is admissible, we set

$$\mathbf{P}^\varrho(\cdot) = \frac{1}{\mathbf{W}^\varrho}\mathbf{w}^\varrho(\cdot) \qquad \text{and} \qquad \mathbf{P}^{\varrho,\star}(\cdot) = \frac{1}{\mathbf{W}^\varrho_\star}\mathbf{w}^\varrho_\star(\cdot).$$

For every integers $n, \varrho \geqslant 1$, let $\mathbf{M}^\varrho_{E=n}$, $\mathbf{M}^\varrho_{V=n}$ and $\mathbf{M}^\varrho_{F=n}$ be the subsets of $\mathbf{M}$ of maps with respectively n edges, $n+1$ vertices, and n inner faces. More generally, for every $A \subset \mathbf{N}$, let $\mathbf{M}_{(F,A)=n}$ be the subset of $\mathbf{M}$ of maps with n inner faces whose degree belongs to $2A$ (and possibly other faces, but with a degree in $2\mathbf{N} \setminus 2A$). For every $S \in \{E, V, F\} \cup \bigcup_{A \subset \mathbf{N}}\{(F, A)\}$ and every $n \geqslant 2$, we define

$$\mathbf{P}^\varrho_{S=n}(M) = \mathbf{P}^\varrho\left(M | M \in \mathbf{M}^\varrho_{S=n}\right), \qquad M \in \mathbf{M}^\varrho_{S=n},$$

the law of a rooted Boltzmann map with perimeter 2ϱ conditioned to have 'size' n. We also let $\mathbf{P}_{E \geqslant n}$ be the law of a rooted Boltzmann map without boundary conditioned to have at least n edges. We define similarly the laws $\mathbf{P}^{\varrho,\star}_{S=n}$ on such pointed maps in $\mathbf{PM}^\varrho_{S=n}$ and $\mathbf{P}^\star_{E \geqslant n}$ on $\bigcup_{k \geqslant n} \mathbf{PM}_{E=k}$.

According to [MM07] (see [Mar18b] for a presentation closer to the present work), the admissibility of a sequence can be checked analytically: $\mathbf{q}$ is admissible if and only if the power series $g_\mathbf{q} : x \mapsto 1 + \sum_{k \geqslant 1}\binom{2k-1}{k-1}q_k x^k$ (which is convex and strictly increasing) possesses a fixed point $Z_\mathbf{q}$ such that $g'_\mathbf{q}(Z_\mathbf{q}) \leqslant 1$; in this case it holds that $Z_\mathbf{q} = (\mathbf{W}_\star + 1)/2$. When $\mathbf{q}$ is admissible, the sequence

$$\mu_\mathbf{q}(k) = (Z_\mathbf{q})^{k-1}\binom{2k-1}{k-1}q_k \quad (k \geqslant 0),$$

where $\mu_\mathbf{q}(0)$ is understood as $(Z_\mathbf{q})^{-1}$, thus defines a probability measure with mean smaller than or equal to one, and we say that $\mathbf{q}$ is *critical* when $\mu_\mathbf{q}$ has mean exactly one.

Sample a random labelled forest (T^ϱ, ℓ) as follows: first T^ϱ has the law of a Bienaymé–Galton–Watson forest with ϱ trees and offspring distribution $\mu_\mathbf{q}$; this means that its Łukasiewicz path has the law of the first-passage bridge W^ϱ of the preceding subsections associated with the distribution $\nu_\mathbf{q}(\cdot) = \mu_\mathbf{q}(\cdot + 1)$. Then, conditionally on T^ϱ, the random labelling ℓ is obtained by sampling the label of the roots uniformly at random in $B^{\geqslant -1}_\varrho$ as defined in (2.1) and, independently, for every branchpoint with outdegree $k \geqslant 1$, the label increments between itself and its offspring are sampled uniformly at random in $B^{\geqslant -1}_k$.

Then we first construct a negative pointed map from this forest as discussed in Section 2.1, and then we re-root it at one of the 2ϱ edges along the boundary (keeping the external face to the right) chosen uniformly at random, then this new pointed map has the law $\mathbf{P}^{\varrho,\star}$. This can be shown in the case $\varrho = 1$ by adapting the arguments from [MM07] which rely on the coding of [BDFG04], and the generalisation to $\varrho \geqslant 2$ can be obtained similarly, with a straightforward analogue of [BM17, Lemma 10] for the re-rooting along the boundary.

Moreover, the pointed map has the law $\mathbf{P}^{\varrho,\star}_{S=n}$ when the forest has the law of such a Bienaymé–Galton–Watson forest conditioned to have n vertices with outdegree in the set $B_S \subset \mathbf{Z}_+$ given by

$$B_E = \mathbf{Z}_+, \qquad B_V = \{0\}, \qquad B_F = \mathbf{N} \qquad \text{and} \qquad B_{(F,A)} = A.$$

We may therefore rely on the preceding sections to obtain information about (pointed) Boltzmann maps.

We let $\nu_{\mathbf{q}}(k) = \mu_{\mathbf{q}}(k+1)$ for every $k \geqslant -1$ which is centred if and only if $\mathbf{q}$ is critical. Let the sequences $(a_n)_{n \geqslant 1}$ and $(b_n)_{n \geqslant 1}$ be as in Section 6.2: when $\alpha > 1$, the sequence $(n^{-1/\alpha} a_n)_{n \geqslant 1}$ is slowly varying and $a_n^{-1}(X_1 + \cdots + X_n)$ converges in distribution to an α-stable random variable, where the X_i's are i.i.d. with law $\nu_{\mathbf{q}}$; when $\alpha = 1$, both $(n^{-1} a_n)_{n \geqslant 1}$ and $(n^{-1} b_n)_{n \geqslant 1}$ are slowly varying, with $b_n \to -\infty$ and $b_n/a_n \to -\infty$, and now $a_n^{-1}(X_1 + \cdots + X_n - b_n)$ converges in distribution.

THEOREM 6.3. — *Assume that $\mathbf{q}$ is an admissible and critical sequence such that the law $\nu_{\mathbf{q}}$ satisfies (H_α) for some $\alpha \in [1, 2]$ or $(\mathrm{H}_1^{\mathrm{loc}})$. Let $(\varrho_n)_{n \geqslant 1}$ be such that $\limsup_{n \to \infty} a_n^{-1} \varrho_n < \infty$.*

(1) *Fix $S \in \{E, V, F\} \cup \bigcup_{A \subset \mathbf{N}} \{(F, A)\}$. Suppose that $\nu_{\mathbf{q}}$ satisfies (H_α) with $\alpha \in (1, 2]$. For every $n \geqslant 1$, let $M_{S=n}^{\varrho_n}$ have the law $\mathbf{P}_{S=n}^{\varrho_n}$. Then from every increasing sequence of integers, one can extract a subsequence along which the spaces*

$$\left(V\left(M_{S=n}^{\varrho_n}\right), a_n^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right)_{n \geqslant 1}$$

converge in distribution in the Gromov–Hausdorff–Prokhorov topology to a limit with a nonzero diameter.

(2) *Suppose furthermore that $\alpha = 2$; let $c_{\mathbf{q}, S} = (\mu_{\mathbf{q}}(B_S)/2)^{1/2}$ and suppose that there exists $\varrho \in [0, \infty)$ such that $\lim_{n \to \infty} a_n^{-1} \varrho_n = \varrho/c_{\mathbf{q}, S}$, then the convergence in distribution*

$$\left(V\left(M_{S=n}^{\varrho_n}\right) \left(\frac{3 c_{\mathbf{q}, S}}{2 a_n}\right)^{1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} (M^\varrho, D^\varrho, p^\varrho)$$

holds in the Gromov–Hausdorff–Prokhorov topology, where M^0 is the Brownian sphere and M^ϱ is the Brownian disk with perimeter ϱ if $\varrho > 0$.

(3) *Suppose that $\nu_{\mathbf{q}}$ satisfies $(\mathrm{H}_1^{\mathrm{loc}})$. For every $n \geqslant 1$, let $M_{E=n}^{\varrho_n}$ have the law $\mathbf{P}_{E=n}^{\varrho_n}$. Then the convergence in distribution*

$$\left(V\left(M_{E=n}^{\varrho_n}\right), |2 b_n|^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} (T_{X^0}, d_{X^0}, p_{X^0})$$

holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion.

(4) *Suppose that $\nu_{\mathbf{q}}$ satisfies (H_1). For every $n \geqslant 1$, let $M_{E \geqslant n}$ have the law $\mathbf{P}_{E \geqslant n}$. Then the convergence in distribution*

$$\left(V\left(M_{E \geqslant n}\right), |2 b_n|^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} (T_{I^{1/2} X^0}, d_{I^{1/2} X^0}, p_{I^{1/2} X^0})$$

holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion, and I is independently sampled from the law $\mathbf{P}(I \geqslant x) = x^{-1}$ for all $x \geqslant 1$.

Finally, the same results hold when the maps are obtained by forgetting the distinguished vertex in the pointed versions of the laws.

For maps without boundary, the case where $\nu_{\mathbf{q}}$ satisfies (H_α) with $\alpha \in (1, 2)$ was first introduced by Le Gall & Miermont [LGM11] in a more restricted case, where $\nu_{\mathbf{q}}(n) \sim \mathrm{Cst} \cdot n^{-1-\alpha}$, and extended to (H_α) in [Mar18a]. More is known for this model than just the above statement; in particular the height and label process of the associated labelled tree converge in distribution. This could be extended to maps with a boundary in a similar way. Le Gall & Miermont [LGM11] also proved that the subsequential limits all have Hausdorff dimension 2α and showing that these limits all have the same topology and further the same distribution (at α fixed) is the object of current investigation.

Proof. — The statements for random pointed maps follow easily from Theorem 6.2 and the results from the introduction. Indeed the statement (1) for $\alpha \in (1, 2)$ follows from Theorem 1.1 and Theorem 6.2 (2): According to the latter, the factor $\sigma_n^{1/2}$ in Theorem 1.1 for the random degree sequence of $M_{S=n}^{\varrho_n}$ is of order $a_n^{1/2}$; since furthermore conditionally on these degrees the map has the uniform distribution on the set of all possible maps, we may then apply Theorem 1.1 conditionally on the degrees, with $a_n^{1/2}$ instead of $\sigma_n^{1/2}$, and then average with respect to the degrees. The statement (2) for $\alpha = 2$ follows similarly from Theorem 1.2 and Theorem 6.2 (1); here the scaling constant $(3/(2\sigma_n))^{1/2}$ is given by $((9\mu_{\mathbf{q}}(B_S))/(8a_n^2))^{1/4}$. The statements (3) and (4) finally follow from Theorem 1.3 and Theorems 6.2 (3) and 6.2 (4) respectively.

In the next proposition, we compare pointed and non-pointed maps, which allows to transfer these results to non-pointed maps. $\square$

Let $\phi : \mathbf{PM} \to \mathbf{M}$ be the projection $\phi((M, x_\star)) = M$ which "forgets the distinguished vertex". We stress that, unless in the case $S = V$ for which there is no bias, the law $\mathbf{P}_{S=n}^{\varrho}$ and the push-forward $\phi_* \mathbf{P}_{S=n}^{\varrho, \star}$ differ at n fixed. Nonetheless, this bias disappears at the limit as shown in several works [Abr16, BJM14, BM17, Mar18a, Mar18b].

PROPOSITION 6.4. — *Under the assumption of Theorem 6.3* (1), (3), *and* (4) *respectively, each corresponding total variation distance*

$$\left\| \mathbf{P}_{S=n}^{\varrho} - \phi_* \mathbf{P}_{S=n}^{\varrho, \star} \right\|_{TV}, \qquad \left\| \mathbf{P}_{E=n}^{\varrho} - \phi_* \mathbf{P}_{E=n}^{\varrho, \star} \right\|_{TV}, \quad \text{and} \quad \left\| \mathbf{P}_{E\geqslant n} - \phi_* \mathbf{P}_{E\geqslant n}^{\star} \right\|_{TV},$$

converges to 0 as $n \to \infty$.

Proof. — This result generalises [Mar18b, Proposition 12] whose proof extends readily here: for notational convenience, suppose that we are in the first case of the theorem. Let $\Lambda(T_{n,B_S}^{\varrho_n})$ denote the number of leaves of the forest $T_{n,B_S}^{\varrho_n}$, which equals the number of vertices minus one in the associated map. Then the above total variation distance is bounded above by

$$\mathbf{E}\left[\left\| \frac{1}{\Lambda\left(T_{n,B_S}^{\varrho_n}\right) + 1} \frac{1}{\mathbf{E}\left[\frac{1}{\Lambda\left(T_{n,B_S}^{\varrho_n}\right)+1}\right]} - 1 \right\|\right].$$

see e.g. the proof of [Mar18b, Proposition 12]. It thus suffices to prove that this expectation tends to 0, which is the content of [Mar18b, Lemma 8]. The proof of this lemma is fairly general once we have the exponential concentration of the proportion

of leaves we obtained in the proof of Lemma 6.1: take the sets A and B in (6.6) to be respectively B_S and $\{-1\}$ here. We refer to [Mar18b] for details.   □

6.5. Subcritical maps

Let us end this paper with a condensation result similar to Theorem 6.3 (3) and (4) for *subcritical* maps, that is, when the weight sequence $\mathbf{q}$ is so that the law $\mu_{\mathbf{q}}$ has mean

$$m_{\mathbf{q}} = \sum_{k \geqslant 0} k \mu_{\mathbf{q}}(k) < 1.$$

In order to directly use the results available in the literature, we shall only work with maps without boundary, so $\varrho_n = 1$. The first convergence below, for pointed maps, is the main result of Janson & Stefánsson [JS15].

THEOREM 6.5. — *Assume that $\mathbf{q}$ is an admissible sequence such that $m_{\mathbf{q}} < 1$.*

(1) *Suppose that there exists a slowly varying function L and an index $\beta > 1$ such that we have*

$$\mu_{\mathbf{q}}(k) = k^{-(1+\beta)} L(k), \qquad k \geqslant 1.$$

   For every $n \geqslant 1$, let $M_{E=n}$ have the law $\mathbf{P}_{E=n}$. Then the convergence in distribution

$$\left(V\left(M_{E=n}\right), (2n)^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} \left(T_{\gamma^{1/2} X^0}, d_{\gamma^{1/2} X^0}, p_{\gamma^{1/2} X^0} \right)$$

   holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion and $\gamma = 1 - m_{\mathbf{q}}$.

(2) *Under the weaker assumption that there exists a slowly varying function L and an index $\beta > 1$ such that*

$$\mu_{\mathbf{q}}([k, \infty)) = k^{-\beta} L(k), \qquad k \geqslant 1,$$

   if now $M_{E \geqslant n}$ has the law $\mathbf{P}_{E \geqslant n}$, then the convergence in distribution

$$\left(V\left(M_{E \geqslant n}\right), (2n)^{-1/2} d_{\mathrm{gr}}, p_{\mathrm{unif}} \right) \xrightarrow[n \to \infty]{(d)} \left(T_{I^{1/2} X^0}, d_{I^{1/2} X^0}, p_{I^{1/2} X^0} \right)$$

   holds in the Gromov–Hausdorff–Prokhorov topology, where X^0 is the standard Brownian excursion, and I is independently sampled from the law $\mathbf{P}(I \geqslant x) = \left(\frac{1 - m_{\mathbf{q}}}{x} \right)^{\beta}$ for all $x \geqslant 1 - m_{\mathbf{q}}$.

Finally, the same results hold when the maps are obtained by forgetting the distinguished vertex in the pointed versions of the laws.

Proof. — The proof goes exactly as that of Theorem 6.3 in the case $\alpha = 1$. First, the analogue of Theorem 6.2 (3) is provided by Kortchemski [Kor15]: according to Theorem 1 there, with the notation of Theorem 6.2 here, we have in the first case,

$$n^{-1} \left(\Delta(W_n), \Delta'(W_n) \right) \xrightarrow[n \to \infty]{\mathbf{P}} ; (1 - m_{\mathbf{q}}, 0).$$

This implies that $n^{-2} \sum_{k \leqslant \Delta'(W_n)} k(k+1) J_k(W_n) \leqslant n^{-1}(1 + \Delta'(W_n))^2$ converges in probability to 0. The analogue of Theorem 6.2 (4) in the second case is provided by Kortchemski & Richier [KR20]: the convergence in distribution

$$n^{-1} \left(\Delta(W_n), \Delta'(W_n) \right) \xrightarrow[n \to \infty]{(d)} (I, 0)$$

follows from Proposition 9 there, and again this implies that $n^{-2} \sum_{k \leqslant \Delta'(W_n)} k(k+1)$ $J_k(W_n)$ converges in probability to 0. Theorem 1.3 allows us to conclude in the case of pointed maps.

In order to consider non-pointed maps, we need an analogous result to Proposition 6.4, which holds as soon as the proportion of leaves in the tree coded by W_n or $W_{\geqslant n}$ satisfies an exponential concentration as in (6.6). The proof is easily adapted: we only need the analogue of Equations (6.5) and (6.4) with $\varrho_n = 1$ on the hitting time of -1; the former, in the case of the tail conditioning, is given by [KR20, Equation 9] and references therein, and the latter, in the case of the local conditioning, is given by [Kor15, Equation 14]. □

Appendix A. Proof of the spinal decomposition

In this appendix we prove Lemma 4.1 on the geometry of random forests spanned by $q \geqslant 1$ random vertices. We start with the simpler case $q = 1$, the case $q \geqslant 2$ is very similar but requires more notation.

A.1. Proof of the one-point decomposition

Let us first consider the simpler case of a single random vertex. In this case, we have $b = 0$ and $c = 1$ so the upper bound reads simply

$$\frac{\varrho_n + \sum_{i=1}^{h} (k_i - 1)}{V_n} \cdot \mathbf{P} \left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\} \right).$$

Proof of Lemma 4.1 for $q = 1$. — The proof relies on the fact that there is a one-to-one correspondence between a plane forest with ϱ trees and a distinguished vertex x on the one hand, and on the other hand the triplet given by the knowledge of which of the ϱ trees contains x, the vector $\mathsf{Cont}(x)$, and the plane forest obtained by removing all the strict ancestors of x; note that this forest contains $\varrho + \sum_{0 \leqslant i < |x|} (k_{a_i(x)} - 1)$ trees. Recall that for any sequence $\theta = (\theta_\ell)_{\ell \geqslant 0}$ of nonnegative integers with finite sum $|\theta|$, the number of plane forests having exactly θ_ℓ vertices with ℓ offspring for every $\ell \geqslant 0$ is given by

$$\#\mathbf{F}(\theta) = \frac{r}{|\theta|} \binom{|\theta|}{(\theta_\ell)_{\ell \geqslant 0}},$$

where $r = \sum_{\ell \geqslant 0} (1 - \ell) \theta_\ell$ is the number of roots. Fix $(k_i, j_i)_{i=1}^{h}$ a sequence of positive integers such that $1 \leqslant j_i \leqslant k_i$ for each i. For every $\ell \geqslant 1$, let $m_\ell = \#\{1 \leqslant i \leqslant h :$

$k_i = \ell\}$ and assume that $m_\ell \leqslant d_n(\ell)$; let also $m_0 = 0$ and $\mathbf{m} = (m_\ell)_{\ell \geqslant 0}$. By the preceding bijection, if x_n is a vertex chosen uniformly at random, then we have

$$\mathbf{P}\left(\mathsf{Cont}(x_n) = (k_i, j_i)_{i=1}^h\right) = \frac{\varrho_n \# \mathbf{F}(d_n - \mathbf{m})}{V_n \# \mathbf{F}(d_n)}$$

$$= \frac{\varrho_n \dfrac{\varrho_n + \sum\limits_{i=1}^h (k_i - 1)}{V_n - h} \binom{V_n - h}{(d_n(\ell) - m_\ell)_{\ell \geqslant 0}}}{V_n \dfrac{\varrho_n}{V_n} \binom{V_n}{(d_n(\ell))_{\ell \geqslant 0}}}$$

$$= \frac{\varrho_n + \sum\limits_{i=1}^h (k_i - 1)}{V_n - h} \frac{(V_n - h)!}{V_n!} \prod_{\ell \geqslant 1} \frac{d_n(\ell)!}{(d_n(\ell) - m_\ell)!}.$$

Let us set

$$P\left((k_i, j_i)_{i \leqslant h}\right) = \mathbf{P}\left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right)$$

$$= \frac{\prod_{\ell \geqslant 1} \ell^{-m_\ell} (\ell d_n(\ell)) (\ell d_n(\ell) - 1) \cdots (\ell d_n(\ell) - m_\ell + 1)}{E_n (E_n - 1) \cdots (E_n - h + 1)}$$

$$= \frac{(E_n - h)!}{E_n!} \prod_{\ell \geqslant 1} \ell^{-m_\ell} \frac{(\ell d_n(\ell))!}{(\ell d_n(\ell) - m_\ell)!}.$$

Then the probability that $\mathsf{Cont}(u_n) = (k_i, j_i)_{i=1}^h$ equals $P((k_i, j_i)_{i \leqslant h})$ times

$$\frac{\varrho_n + \sum\limits_{i=1}^h (k_i - 1)}{V_n - h} \frac{(V_n - h)!}{V_n!} \frac{E_n!}{(E_n - h)!} \left(\prod_{\ell \geqslant 1} \ell^{m_\ell} \frac{d_n(\ell)! (\ell d_n(\ell) - m_\ell)!}{(d_n(\ell) - m_\ell)! (\ell d_n(\ell))!}\right).$$

Observe that $V_n \geqslant E_n + 1$, so

$$\frac{(V_n - h)!}{V_n!} \frac{E_n!}{(E_n - h)!} = \frac{V_n - h}{V_n} \prod_{i=0}^{h-1} \frac{E_n - i}{V_n - 1 - i} \leqslant \frac{V_n - h}{V_n}.$$

Moreover, we claim that $\ell^{m_\ell} \frac{d_n(\ell)! (\ell d_n(\ell) - m_\ell)!}{(d_n(\ell) - m_\ell)! (\ell d_n(\ell))!} \leqslant 1$ for every $\ell \geqslant 1$. Indeed, it equals 1 when $\ell = 1$; we suppose next that $\ell \geqslant 2$, so in particular $m_\ell \leqslant d_n(\ell) \leqslant \ell d_n(\ell)/2$. It is simple to check that for every $x \in [0, 1/2]$, we have $(1 - x)^{-1} \leqslant 2^{2x}$, and for every $x \in [0, 1]$, we have $1 - x \leqslant 2^{-x}$. It follows that for every $\ell \geqslant 2$ such that $d_n(\ell) \neq 0$,

$$\ell^{m_\ell} \frac{d_n(\ell)! (\ell d_n(\ell) - m_\ell)!}{(d_n(\ell) - m_\ell)! (\ell d_n(\ell))!} = \frac{d_n(\ell)!}{(d_n(\ell) - m_\ell)! d_n(\ell)^{m_\ell}} \frac{(\ell d_n(\ell) - m_\ell)! (\ell d_n(\ell))^{m_\ell}}{(\ell d_n(\ell))!}$$

$$= \prod_{i=0}^{m_\ell - 1} \frac{d_n(\ell) - i}{d_n(\ell)} \frac{\ell d_n(\ell)}{\ell d_n(\ell) - i}$$

$$\leqslant \prod_{i=0}^{m_\ell - 1} \exp\left(\ln 2 \left(-\frac{i}{d_n(\ell)} + \frac{2i}{\ell d_n(\ell)}\right)\right),$$

which is indeed bounded by 1 for $\ell \geqslant 2$. We have thus shown that

$$\mathbf{P}\left(\mathsf{Cont}(u_n) = (k_i, j_i)_{i=1}^h\right) \leqslant \frac{\varrho_n + \sum\limits_{i=1}^h (k_i - 1)}{V_n} P\left((k_i, j_i)_{i \leqslant h}\right),$$

which corresponds to our claim in the case $q = 1$. $\qquad\square$

A.2. Proof of the multi-point decomposition

We next turn to the proof in the general case; the difference is that one has to deal with the branchpoints of the reduced forest, which, we recall, are not considered in the vector of content.

Proof of Lemma 4.1 for $q \geqslant 2$. — Sample $x_{n,1}, \ldots, x_{n,q}$ i.i.d. uniformly random vertices in $T_{d_n}^{\varrho_n}$ and assume that none is an ancestor of another. As in the case $q = 1$, there is a one-to-one correspondence between the plane forest $T_{d_n}^{\varrho_n}$ and $x_{n,1}, \ldots, x_{n,q}$ on the one hand, and on the other hand the plane forest obtained by removing their strict ancestors, the vector $\mathsf{Cont}(x_{n,1}, \ldots, x_{n,q})$, and the following data: first the number c of trees containing at least one of the $x_{n,i}$'s and the knowledge of which ones, second for each of the b branchpoints: their total number r of children in $T_{d_n}^{\varrho_n}$, their number d of children which are ancestors of at least one $x_{n,j}$, and the relative positions z_i's of these children.

Therefore, fix $(k_i, j_i)_{i=1}^h$, also $c \in \{1, \ldots, q\}$, $b \in \{0, \ldots, q-1\}$, and for every $p \in \{1, \ldots, b\}$, fix $r(p) \geqslant d(p) \geqslant 2$ and integers $1 \leqslant z_{p,1} < \cdots < z_{p,d(p)} \leqslant r(p)$, and let us consider the following event: first $\mathsf{Cont}(x_{n,1}, \ldots, x_{n,q}) = (k_i, j_i)_{i=1}^h$, second, in the forest spanned by $x_{n,1}, \ldots, x_{n,q}$, we have c trees, b branchpoints, for every $p \leqslant b$, the p'th branchpoint in lexicographical order has $r(p)$ children in total in the original forest $T_{d_n}^{\varrho_n}$, amongst which $d(p)$ belong to the reduced forest, and the relative positions of the latter are given by the $z_{p,i}$'s. If we set $m_0 = 0$ and for every $\ell \geqslant 1$, we let $m_\ell = \#\{1 \leqslant i \leqslant h : k_i = \ell\}$ and $\overline{m}_\ell = m_\ell + \sum_{p=1}^b \mathbf{1}_{\{r(p)=\ell\}}$, then on this event, the complement of the reduced forest is a forest with degree sequence $(d_n(\ell) - \overline{m}_\ell)_{\ell \geqslant 0}$; note that it contains $R = \varrho_n - c + q + \sum_{p=1}^b (r(p) - d(p)) + \sum_{i=1}^h (k_i - 1)$ trees and $V_n - h - b$ vertices. Therefore the probability of our event is given by

$$\frac{\binom{\varrho_n}{c} \# \mathbf{F}(d_n - \overline{\mathbf{m}})}{V_n^q \# \mathbf{F}(d_n)} = \frac{\binom{\varrho_n}{c} \frac{R}{V_n - h - b} \binom{V_n - h - b}{(d_n(\ell) - \overline{m}_\ell)_{\ell \geqslant 0}}}{V_n^q \frac{\varrho_n}{V_n} \binom{V_n}{(d_n(\ell))_{\ell \geqslant 0}}}$$

$$= \frac{\binom{\varrho_n}{c} R}{\varrho_n V_n^{q-1} (V_n - h - b)} \frac{(V_n - h - b)!}{V_n!} \prod_{\ell \geqslant 1} \frac{(d_n(\ell))!}{(d_n(\ell) - \overline{m}_\ell)!}.$$

Now as previously,

$$\frac{(V_n - h - b)!}{V_n!} \leqslant \frac{(V_n - h - b)!}{(V_n - h)!} \frac{V_n - h}{V_n} \frac{(E_n - h)!}{E_n!} = \frac{(V_n - h - b)!}{V_n (V_n - h - 1)!} \frac{(E_n - h)!}{E_n!}.$$

Next,

$$\prod_{\ell \geqslant 1} \frac{(d_n(\ell))!}{(d_n(\ell) - \overline{m}_\ell)!} = \prod_{\ell \geqslant 1} \frac{(d_n(\ell))!}{(d_n(\ell) - m_\ell)!} \prod_{\ell \geqslant 1} \frac{(d_n(\ell) - m_\ell)!}{\left(d_n(\ell) - m_\ell - \sum_{p=1}^{b} \mathbf{1}_{\{r(p) = \ell\}}\right)!}$$

$$\leqslant \prod_{\ell \geqslant 1} \frac{(d_n(\ell))!}{(d_n(\ell) - m_\ell)!} \prod_{p=1}^{b} d_n(r(p)).$$

Finally, we have seen that

$$\mathbf{P}\left(\bigcap_{i \leqslant h} \left\{ (\xi_{d_n}(i), \chi_{d_n}(i)) = (k_i, j_i) \right\}\right) = \frac{(E_n - h)!}{E_n!} \prod_{\ell \geqslant 1} \ell^{-m_\ell} \frac{(\ell d_n(\ell))!}{(\ell d_n(\ell) - m_\ell)!},$$

which we shall denote by $P((k_i, j_i)_{i \leqslant h})$. We infer that

$$\frac{\binom{\varrho_n}{c} \# \mathbf{F}(d_n - \overline{\mathbf{m}})}{V_n^q \# \mathbf{F}(d_n)}$$

$$\leqslant \frac{\binom{\varrho_n}{c} R}{\varrho_n V_n^{q-1} (V_n - h - b)} \frac{(V_n - h - b)!}{V_n (V_n - h - 1)!} \frac{(E_n - h)!}{E_n!} \prod_{\ell \geqslant 1} \frac{(d_n(\ell))!}{(d_n(\ell) - m_\ell)!} \prod_{p=1}^{b} d_n(r(p))$$

$$\leqslant \frac{\binom{\varrho_n}{c} R}{\varrho_n V_n^q} \frac{(V_n - h - b - 1)!}{(V_n - h - 1)!} \underbrace{\prod_{\ell \geqslant 1} \ell^{m_\ell} \frac{(\ell d_n(\ell) - m_\ell)! (d_n(\ell))!}{(\ell d_n(\ell))! (d_n(\ell) - m_\ell)!}}_{\leqslant 1} \prod_{p=1}^{b} d_n(r(p)) \cdot P\left((k_i, j_i)_{i \leqslant h}\right),$$

where the fact that the first product is bounded by 1 was shown at the very end of the proof of the case $q = 1$.

Since $c \geqslant 1$, $b \leqslant q - 1$, and $2 \leqslant d(p) \leqslant r(p) \leqslant \Delta_n$, we have

$$-c + q + \sum_{p=1}^{b} (r(p) - d(p)) \leqslant -1 + q + b(\Delta_n - 1) \leqslant (q - 1)\Delta_n,$$

hence

$$R = \varrho_n - c + q + \sum_{p=1}^{b} (r(p) - d(p)) + \sum_{i=1}^{h} (k_i - 1) \leqslant \varrho_n + (q - 1)\Delta_n + \sum_{i=1}^{h} (k_i - 1).$$

Moreover, from our assumption that $h, q \leqslant V_n/4$,

$$\frac{(V_n - h - b - 1)!}{(V_n - h - 1)!} \leqslant \left(\frac{1}{V_n - h - b}\right)^b \leqslant \left(\frac{2}{V_n}\right)^b,$$

and finally, since $\binom{\varrho_n}{c} \leqslant \varrho_n^c$, we obtain

$$\frac{\binom{\varrho_n}{c} \# \mathbf{F}(d_n - \overline{\mathbf{m}})}{V_n^q \# \mathbf{F}(d_n)} \leqslant 2^b \varrho_n^c \frac{\varrho_n + (q - 1)\Delta_n + \sum_{i=1}^{h}(k_i - 1)}{\varrho_n V_n^{q+b}} P\left((k_i, j_i)_{i \leqslant h}\right) \cdot \prod_{p=1}^{b} d_n(r(p)).$$

On the left is the probability that $\mathsf{Cont}(x_{n,1}, \ldots, x_{n,q}) = (k_i, j_i)_{i=1}^{h}$ and that the reduced forest has c trees, q leaves, and b branchpoints and that for every $p \leqslant b$, the p'th branchpoint has $r(p)$ children, amongst which $d(p) \geqslant 2$ belong to the reduced

forest, with relative positions given by $1 \leqslant z_{p,1} < \cdots < z_{p,d(p)} \leqslant r(p)$. We now want to sum over all possible z's, all vectors d's and r's; note that the quantities on the right before the product sign do not depend on them so we only focus on the term $\prod_{p=1}^{b} d_n(r(p))$. First note that the sum of this product over all the possible z's is bounded by

$$\prod_{p=1}^{b} \binom{r(p)}{d(p)} d_n(r(p)) \leqslant \prod_{p=1}^{b} \Delta_n^{d(p)-2} r(p)(r(p)-1) d_n(r(p))$$

$$\leqslant \Delta_n^{q-c-b} \prod_{p=1}^{b} r(p)(r(p)-1) d_n(r(p)),$$

where for the last inequality, we note that, since the $d(p)$'s are the number of children of the branchpoints in the reduced forest, which contains c trees and q leaves, then we have $q = c + \sum_{p=1}^{b}(d(p)-1) = c + b + \sum_{p=1}^{b}(d(p)-2)$. Note also the very crude bound: there are less than $bq \leqslant q^2$ such vectors $(d(1), \ldots, d(b))$. We finally want to sum the last bound times q^2 over all the vectors $(r(1), \ldots, r(b))$; we have

$$\sum_{r(1),\ldots,r(b) \leqslant \Delta_n} \prod_{p=1}^{b} r(p)(r(p)-1) d_n(r(p)) = \left(\sum_{r \leqslant \Delta_n} r(r-1) d_n(r) \right)^b = \sigma_n^{2b}.$$

Since $\Delta_n \geqslant 2$, then $\Delta_n^2 \leqslant 2\Delta_n(\Delta_n - 1) \leqslant 2\sigma_n^2$, so finally, since $q - c - b \leqslant 0$, then $2^{(q-c-b)/2} \leqslant 1$ and so

$$\sum_{r(1),\ldots,r(b) \leqslant \Delta_n} q^2 \prod_{p=1}^{b} \binom{r(p)}{d(p)} d_n(r(p)) \leqslant q^2 \Delta_n^{q-c-b} \sigma_n^{2b} \leqslant q^2 \sigma_n^{q-c+b}.$$

We conclude that the probability that $\mathsf{Cont}(x_{n,1}, \ldots, x_{n,q}) = (k_i, j_i)_{i=1}^{h}$ and the reduced forest has q leaves, c trees, and b branchpoints is bounded by

$$2^b \varrho_n^c \frac{\varrho_n + (q-1)\Delta_n + \sum\limits_{i=1}^{h}(k_i - 1)}{\varrho_n V_n^{q+b}} P\left((k_i, j_i)_{i \leqslant h} \right) q^2 \sigma_n^{q-c+b}$$

$$= q^2 2^b \left(\frac{\sigma_n}{V_n} \right)^{q+b} \left(\frac{\varrho_n}{\sigma_n} \right)^{c-1} \frac{\varrho_n + (q-1)\Delta_n + \sum_{i=1}^{h}(k_i - 1)}{\sigma_n} P\left((k_i, j_i)_{i \leqslant h} \right),$$

and the proof is complete since $b \leqslant q - 1$. $\qquad\square$

BIBLIOGRAPHY

[AB12]   Louigi Addario-Berry, *Tail bounds for the height and width of a random tree with a given degree sequence*, Random Struct. Algorithms **41** (2012), no. 2, 253–261. ↑344, 345, 350

[ABA17]   Louigi Addario-Berry and Marie Albenque, *The scaling limit of random simple triangulations and random simple quadrangulations*, Ann. Probab. **45** (2017), no. 5, 2767–2825. ↑335

[ABA21]   _____, *Convergence of non-bipartite maps via symmetrization of labeled trees*, Ann. Henri Lebesgue **4** (2021), 653–683. ↑324, 335

[Abr16]   Céline Abraham, *Rescaled bipartite planar maps converge to the Brownian map*, Ann. Inst. Henri Poincaré, Probab. Stat. **52** (2016), no. 2, 575–595. ↑335, 377

[Ald85]   David Aldous, *Exchangeability and related topics*, École d'été de probabilités de Saint-Flour, XIII—1983, Lecture Notes in Mathematics, vol. 1117, Springer, 1985, pp. 1–198. ↑332, 351, 354, 357

[Ald93]   _____, *The continuum random tree. III*, Ann. Probab. **21** (1993), no. 1, 248–289. ↑321, 323, 332, 352

[AMP04]   David Aldous, Grégory Miermont, and Jim Pitman, *The exploration process of inhomogeneous continuum random trees, and an extension of Jeulin's local time identity*, Probab. Theory Relat. Fields **129** (2004), no. 2, 182–218. ↑325

[BCM19]   Olivier Bernardi, Nicolas Curien, and Grégory Miermont, *A Boltzmann approach to percolation on random triangulations*, Can. J. Math. **71** (2019), no. 1, 1–43. ↑375

[BCP03]   Jean Bertoin, Loïc Chaumont, and Jim Pitman, *Path transformations of first passage bridges*, Electron. Commun. Probab. **8** (2003), 155–166. ↑331

[BDFG04]   Jérémie Bouttier, Philippe Di Francesco, and Emmanuel Guitter, *Planar maps as labeled mobiles*, Electron. J. Comb. **11** (2004), no. 1, article no. 69. ↑324, 327, 375

[Ber19]   Quentin Berger, *Notes on random walks in the Cauchy domain of attraction*, Probab. Theory Relat. Fields **175** (2019), no. 1-2, 1–44. ↑368

[Bet10]   Jérémie Bettinelli, *Scaling limits for random quadrangulations of positive genus*, Electron. J. Probab. **15** (2010), article no. 52. ↑360

[Bet15]   _____, *Scaling limit of random planar quadrangulations with a boundary*, Ann. Inst. Henri Poincaré, Probab. Stat. **51** (2015), no. 2, 432–477. ↑319, 321, 340, 341

[BJM14]   Jérémie Bettinelli, Emmanuel Jacob, and Grégory Miermont, *The scaling limit of uniform random plane maps, via the Ambjørn–Budd bijection*, Electron. J. Probab. **19** (2014), article no. 74. ↑335, 377

[BM14]   Nicolas Broutin and Jean-François Marckert, *Asymptotics of trees with a prescribed degree sequence and applications*, Random Struct. Algorithms **44** (2014), no. 3, 290–316. ↑322, 331, 342, 344, 353, 371

[BM17]   Jérémie Bettinelli and Grégory Miermont, *Compact Brownian surfaces I: Brownian disks*, Probab. Theory Relat. Fields **167** (2017), no. 3-4, 555–614. ↑319, 320, 321, 322, 332, 334, 335, 338, 339, 374, 375, 377

[CLGM13]   Nicolas Curien, Jean-François Le Gall, and Grégory Miermont, *The Brownian cactus I. Scaling limits of discrete cactuses*, Ann. Inst. Henri Poincaré, Probab. Stat. **49** (2013), no. 2, 340–373. ↑341

[Fel71]   William Feller, *An introduction to probability theory and its applications. Vol. II.*, second ed., Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1971. ↑367

[Hoe63]   Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Am. Stat. Assoc. **58** (1963), 13–30. ↑331

[JS15]   Svante Janson and Sigurður Örn Stefánsson, *Scaling limits of random planar maps with a unique large face*, Ann. Probab. **43** (2015), no. 3, 1045–1081. ↑322, 324, 327, 378

[Kor12]   Igor Kortchemski, *Invariance principles for Galton–Watson trees conditioned on the number of leaves*, Stochastic Processes Appl. **122** (2012), no. 9, 3126–3172. ↑368, 371, 373

[Kor15]   _____, *Limit theorems for conditioned non-generic Galton–Watson trees*, Ann. Inst. Henri Poincaré, Probab. Stat. **51** (2015), no. 2, 489–511. ↑324, 378, 379

[Kor17]   _____, *Sub-exponential tail bounds for conditioned stable Bienaymé–Galton–Watson trees*, Probab. Theory Relat. Fields **168** (2017), no. 1-2, 1–40. ↑367

[KR19]   Igor Kortchemski and Loïc Richier, *Condensation in critical Cauchy Bienaymé–Galton–Watson trees*, Ann. Appl. Probab. **29** (2019), no. 3, 1837–1877. ↑322, 367, 368, 369, 371, 373, 374

[KR20]   _____, *The boundary of random planar maps via looptrees*, Ann. Fac. Sci. Toulouse, Math. **29** (2020), no. 2, 391–430. ↑379

[Lei19]   Tao Lei, *Scaling limit of random forests with prescribed degree sequences*, Bernoulli **25** (2019), no. 4A, 2409–2438. ↑322, 331, 332

[LG05]   Jean-François Le Gall, *Random trees and applications*, Probab. Surv. **2** (2005), 245–311. ↑328

[LG07]   _____, *The topological structure of scaling limits of large planar maps*, Invent. Math. **169** (2007), no. 3, 621–670. ↑319, 336, 338

[LG13]   _____, *Uniqueness and universality of the Brownian map*, Ann. Probab. **41** (2013), no. 4, 2880–2960. ↑318, 319, 320, 321, 324, 335, 336, 339

[LGM11]   Jean-François Le Gall and Grégory Miermont, *Scaling limits of random planar maps with large faces*, Ann. Probab. **39** (2011), no. 1, 1–69. ↑322, 325, 336, 361, 365, 377

[LGP08]   Jean-François Le Gall and Frédéric Paulin, *Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere*, Geom. Funct. Anal. **18** (2008), no. 3, 893–918. ↑319, 340

[Mar18a]   Cyril Marzouk, *On scaling limits of planar maps with stable face-degrees*, ALEA, Lat. Am. J. Probab. Math. Stat. **15** (2018), no. 2, 1089–1123. ↑322, 327, 335, 377

[Mar18b]   _____, *Scaling limits of random bipartite planar maps with a prescribed degree sequence*, Random Struct. Algorithms **53** (2018), no. 3, 448–503. ↑318, 320, 335, 342, 344, 355, 358, 359, 361, 365, 371, 375, 377, 378

[McD98]   Colin McDiarmid, *Concentration*, Probabilistic methods for algorithmic discrete mathematics, Algorithms and Combinatorics, vol. 16, Springer, 1998, pp. 195–248. ↑346, 351, 357

[Mie08]   Grégory Miermont, *On the sphericity of scaling limits of random planar quadrangulations*, Electron. Commun. Probab. **13** (2008), 248–257. ↑319, 340

[Mie09]   _____, *Tessellations of random maps of arbitrary genus*, Ann. Sci. Éc. Norm. Supér. **42** (2009), no. 5, 725–781. ↑335

[Mie13]   _____, *The Brownian map is the scaling limit of uniform random plane quadrangulations*, Acta Math. **210** (2013), no. 2, 319–401. ↑318, 319, 320, 324

[MM03]   Jean-François Marckert and Abdelkader Mokkadem, *The depth first processes of Galton–Watson trees converge to the same Brownian excursion*, Ann. Probab. **31** (2003), no. 3, 1655–1678. ↑353

[MM07]   Jean-François Marckert and Grégory Miermont, *Invariance principles for random bipartite planar maps*, Ann. Probab. **35** (2007), no. 5, 1642–1705. ↑327, 363, 374, 375

[Pit06]   Jim Pitman, *Combinatorial stochastic processes. école d'été de Probabilités de Saint-Flour XXXII – 2002.*, Lecture Notes in Mathematics, vol. 1875, Springer, 2006. ↑330

[Thé20]   Paul Thévenin, *Vertices with fixed outdegrees in large Galton–Watson trees*, Electron. J. Probab. **25** (2020), article no. 64. ↑368

This journal is a member of Centre Mersenne.

Cyril MARZOUK
École polytechnique,
CMAP / UMR 7641 CNRS,
Route de Saclay,
91128 Palaiseau Cedex (France)
cyril.marzouk@polytechnique.edu