AUSTRALASIAN JOURNAL OF COMBINATORICS
Volume **92(3)** (2025), Pages 266–319

Scaling limit of graph classes through split decomposition

Frédérique Bassino

Université Sorbonne Paris Nord, LIPN
CNRS UMR 7030, F-93430 Villetaneuse, France
bassino@lipn.univ-paris13.fr

Mathilde Bouvel

Université de Lorraine, CNRS, Inria
LORIA, F-54000 Nancy, France
mathilde.bouvel@loria.fr

Valentin Féray

Université de Lorraine, CNRS, IECL
F-54000 Nancy, France
valentin.feray@univ-lorraine.fr

Lucas Gerin

CMAP, École Polytechnique, CNRS
Route de Saclay, 91128 Palaiseau Cedex, France
gerin@cmap.polytechnique.fr

Adeline Pierrot

LISN, Université Paris-Saclay Bat. 650 Ada Lovelace
91405 Orsay Cedex, France
adeline.pierrot@lri.fr

Abstract

We prove that Aldous' Brownian CRT is the scaling limit, with respect to the Gromov–Prokhorov topology, of uniformly chosen random graphs in each of the three following families of graphs: distance-hereditary graphs, 2-connected distance-hereditary graphs and 3-leaf power graphs. Our approach is based on the split decomposition and on analytic combinatorics.

ISSN: 2202-3518

1 Introduction

In the present article we obtain scaling limit results for large graphs taken uniformly at random in the class of distance-hereditary graphs (DH-graphs for short) and in two interesting subclasses: 2-connected distance-hereditary graphs and 3-leaf power graphs. In all cases, the limit is the celebrated Brownian continuum random tree (Brownian CRT for short). We start by giving some background on these graph classes.

1.1 Distance-hereditary graphs and interesting subclasses

DH-graphs are the connected graphs for which the distances in any connected induced subgraph[1] are the same as in the original graph. (See an example and a counter-example in Fig. 1.) They enjoy many other characterizations, for instance by avoidance of induced subgraphs. Among other properties, they form a subclass of perfect graphs and have clique-width at most three. They have been widely studied in the algorithmic literature: in particular, it has been proved that many NP-hard problems can be solved in polynomial time for DH-graphs (see *e.g.* [15]); additionally, DH-graphs can be recognized efficiently, both in the static and dynamic framework (see [23], and references therein).

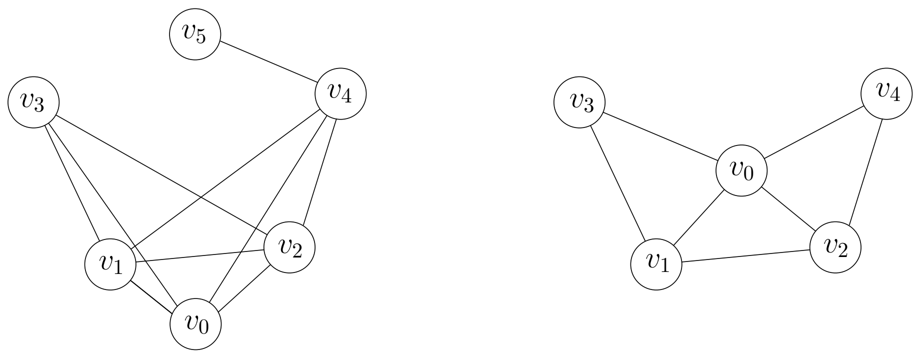

Figure 1: Left: A DH-graph with 5 vertices: for instance, $d_G(v_3, v_4) = 2$ in the original graph G, and considering an induced subgraph H containing v_3 and v_4, either we have $d_H(v_3, v_4) = 2$ (if H contains v_0, v_1 or v_2), or H is disconnected (if the vertex set of H is $\{v_3, v_4\}$ or $\{v_3, v_4, v_5\}$). Right: A graph which is not a DH-graph: $d_G(v_3, v_4) = 2$ while $d_H(v_3, v_4) = 3$ in the subgraph H induced by $\{v_1, v_2, v_3, v_4\}$ (which is the path $v_3 - v_1 - v_2 - v_4$).

To establish such algorithmic properties, a key feature of distance-hereditary graphs is that they are nicely decomposable for the so-called *split decomposition*. More recently, this split decomposition has also been used to give precise enumerative results and sampling algorithms on the class of distance-hereditary graphs and some

[1] We recall that, if W is a subset of the vertex set V of a graph G, the subgraph of G induced by W consists of the vertices in W, and of all edges of G connecting two vertices in W.

of its subclasses [14, 4]. The analysis of distance-hereditary graphs (and subclasses) via the symbolic method, as done by Chauve, Fusy and Lumbroso [14] (and reviewed in Section 3 below) is a starting point for the present paper. More precisely, in our work, we aim at illustrating the usefulness of the split decomposition (combined with symbolic and analytic combinatorics) to study large random DH-graphs.

Let us comment on the choice of graph classes considered in this article, in addition to the DH-graphs already discussed. The class of 3-leaf power graphs has been studied in [23] (respectively [14]) to illustrate the versatility of algorithmic (respectively enumerative) results obtained through the split decomposition. It is therefore natural for us to use it to illustrate as well the versatility of the probabilistic approach through the split decomposition. Since 3-leaf power graphs are defined via trees (see Theorem 7.1), their convergence to an infinite tree might seem expected. On the contrary, conditioning random DH-graphs to be 2-connected makes them less tree-like. Our result indicates that, nevertheless, at the level of scaling limits, 2-connected DH-graphs are tree-like and converge to the Brownian CRT.

Another motivation for considering 3-leaf power graphs and 2-connected DH-graphs is that, unlike unconstrained DH-graphs, they do not form what is called a *block-stable* class of graphs. Indeed, such block-stable graph classes have already been studied in the discrete probability literature [18, 19]. In particular, a scaling limit result for random graphs in such classes (under an additional subcriticality hypothesis) is provided in [32], covering the case of unconstrained DH-graphs. When a class is not block-stable, performing the decomposition often causes one to leave the original category, making it impossible to straightforwardly apply the previously established global convergence results. It is therefore important to show that our approach through split decomposition works also for classes which are not block-stable; and an obvious way to obtain a class of graphs which is not block-stable is to impose the constraint of being 2-connected.

1.2 The results

A standard question in the theories of random trees, random maps and more recently random graphs is to look for limits of random graph sequences, for various topologies. To this end, we consider graphs as discrete metric measure spaces. A metric measure space (mm-space for short) is a triple (X, d, μ), where (X, d) is a complete and separable metric space and μ a probability measure on X. A finite connected graph can be seen as a mm-space, where X is the vertex set of the graph, d the graph distance, and μ the uniform probability measure on X. In this setting scaling limits of random graphs correspond to the convergence of random mm-spaces, after renormalization of the distances.

For metric measure spaces there are two classical topologies used in the literature, the Gromov–Prohorov (GP) topology and the stronger Gromov–Hausdorff–Prohorov (GHP) topology. Our result holds with respect to the GP topology (see Section 2 for the definition). We believe that it could be extended to the GHP topology, using a criterion provided by Athreya-Löhr-Winter [3]. However, this would likely require tools and methods very different from those of the present paper, and is therefore beyond its scope.

For $n \geq 1$ denote by $\mathcal{G}_d^{(n)}$ (respectively $\mathcal{G}_{2c}^{(n)}$, respectively $\mathcal{G}_{3\ell}^{(n)}$) the set of DH-graphs (respectively 2-connected DH-graphs, respectively 3-leaf power graphs) with vertex set $[n] := \{1, \dots, n\}$ (we say that such graphs have *size* n).

We also denote by $(\mathcal{T}_\infty, d_\infty, \mu_\infty)$ the Brownian CRT equipped with its natural mass measure μ_∞. The Brownian CRT has been introduced by Aldous in [2] and is a now standard object in the discrete probability literature (for details and references, see Section 2).

Theorem 1.1 *For every family $f \in \{d, 2c, 3\ell\}$ and $n \geq 1$ let $\boldsymbol{G}_f^{(n)}$ be a uniform random graph in $\mathcal{G}_f^{(n)}$. Let μ_n be the uniform probability measure on the set of vertices $[n]$ and $d_{\boldsymbol{G}_f^{(n)}}$ be the graph distance in $\boldsymbol{G}_f^{(n)}$. Then there exists a constant $c_f > 0$ such that the following convergence holds in distribution for the Gromov–Prohorov topology:*

$$\left([n], \frac{c_f}{\sqrt{n}} d_{\boldsymbol{G}_f^{(n)}}, \mu_n\right) \overset{n\to+\infty}{\to} (\mathcal{T}_\infty, d_\infty, \mu_\infty). \tag{1}$$

Constants in Eq. (1) are explicit: namely,

$$c_d = \frac{\sqrt{2}}{\gamma_H} \approx 0.3602 \quad \text{where } \gamma_H \text{ is defined in Eq. (31) p. 292,}$$
$$c_{2c} = \frac{\sqrt{2}}{\gamma_{H,2c}} \approx 0.1885 \quad \text{where } \gamma_{H,2c} \text{ is defined in Eq. (58) p. 306,}$$
$$c_{3\ell} = \frac{\sqrt{2}}{\gamma_{E,3\ell}} \approx 0.9266 \quad \text{where } \gamma_{E,3\ell} = \sqrt{2(1+e)\log(1+e^{-1})}, \text{ see Eq. (64) p. 310.}$$

Figs. 2 to 4 show two realizations of uniform distance-hereditary graphs with a few hundred vertices, respectively in the unconstrained, 2-connected and 3-leaf power graph cases.

As mentioned above, in the case $f = d$ (*i.e.* random unconstrained DH-graphs), Theorem 1.1 is not new. Indeed, DH-graphs form a subcritical block-stable class of graphs, and it is proved in [32] that uniform random graphs in such classes converge to the Brownian CRT[2]. On the contrary, 3-leaf power graphs and 2-connected DH-graphs are not block-stable graph classes, and Theorem 1.1 is new in these cases. The stronger connectivity of 2-connected DH-graphs is reflected in the value of the renormalizing constant, which is smaller in the 2-connected case than in the unconstrained and 3-leaf power cases.

[2]In [32], the convergence is proven only for the Gromov-Hausdorff (GH) topology, which is incomparable with the GP topology we use here. We believe however that without much further effort, their argument in fact proves convergence in the stronger GHP topology, see Section C for details.

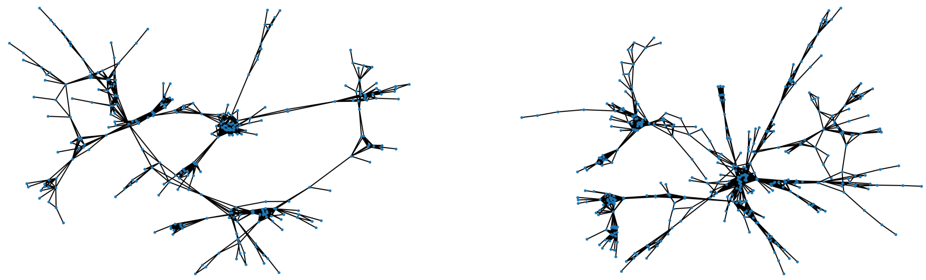

Figure 2: Two samples of uniform random DH-graphs of respective sizes $n = 290$ and $n = 388$. Both graphs were generated with a Boltzmann sampler (see [21]) using the combinatorial specification given in Eq. (5) p. 283 and plotted with the function `spring_layout` from the `python` library `networkx` (*i.e.* using a force-directed graph drawing algorithm).

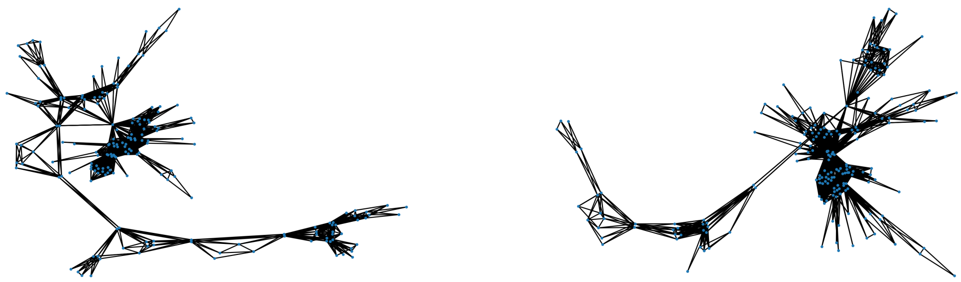

Figure 3: Two samples of uniform random 2-connected DH-graphs of respective sizes $n = 186$ and $n = 197$. Graphs were generated with the combinatorial specification given in Eq. (44) p. 304.

We can restate Theorem 1.1 in more concrete terms, which actually describe how we intend to prove Theorem 1.1. It is known (see [25] or Section 2 below) that convergence in distribution in the Gromov–Prohorov sense is equivalent to the convergence in distribution of the relative distances between k uniform vertices in the graph, for every k. For $k = 2$ Theorem 1.1 says that if $\mathbf{v}_0, \mathbf{v}_1$ are uniform i.i.d. vertices in $\boldsymbol{G}_f^{(n)}$ then

$$\frac{c_f}{\sqrt{n}} d_{\boldsymbol{G}_f^{(n)}}(\mathbf{v}_0, \mathbf{v}_1) \xrightarrow[n \to +\infty]{(d)} d_\infty(v_0, v_1) \tag{2}$$

where v_0, v_1 are independent and μ_∞-distributed in $\mathcal{T}_\infty$. It turns out that the random variable $d_\infty(v_0, v_1)$ is known to follow the Rayleigh distribution, *i.e.* has density $xe^{-x^2/2}$ on $\mathbb{R}_+$. More generally Theorem 1.1 amounts to saying that (2) holds jointly for k uniform i.i.d. vertices in $\boldsymbol{G}_f^{(n)}$. The joint limiting distribution, *i.e.* the distribution of the distances between k random points in the CRT, is given below in Theorem 2.4 (see also [2]).

We finish by discussing how our result fits in the literature on convergence of

Figure 4: Two samples of uniform random 3-leaf power graphs of respective sizes $n = 170$ and $n = 231$. Graphs were generated with the combinatorial specification given in Eq. (61) p. 309.

discrete graph models to the CRT. It is now well established that the Brownian CRT is the universal limit of many important families of random trees, see, *e.g.*, [28]. In addition, a few families of graphs which are *not* trees are also known to converge towards the CRT, although such results are less common in the literature. We can cite some models of random planar maps [1, 10, 13, 27], some models of random dissections [16], and random graphs in subcritical *block-stable* graph classes [32], as mentioned above. Our paper exhibits two new families of nontree graphs classes converging to the CRT (and an alternative proof for a third class).

1.3 Proof strategy

In this section, we present the main lines of the proof of Theorem 1.1. The starting point is a standard representation of distance-hereditary graphs as trees whose internal nodes are decorated with stars and cliques (see Theorem 3.3), called here DH-trees. Vertices of the graph correspond to leaves of the DH-tree. Moreover, it turns out that the distance between two vertices in the graph can then be read on the DH-tree, by couting the number of some particular patterns, which we call here *jumps*, along the path between the corresponding leaves (Theorem 3.4).

As mentioned above, convergence for the Gromov–Prohorov topology is equivalent to the convergence for each $k \geq 1$, of the random matrix recording the distances between k uniform random vertices in the graph. We therefore need to consider DH-trees with k marked leaves and to compute the distribution of the number of jumps between these leaves. To do that, we define a notion of enriched tree induced by k marked leaves, which records the genealogy between these leaves and the number of jumps between the marked leaves and their common ancestors. Then we prove, for each of the three models, a local limit theorem for the enriched tree induced by k uniform random leaves in the DH-tree of a uniform random graph in the family. Namely, given $(t_0, a_1, ...a_{2k-2})$, where t_0 is tree with k leaves, and $a_1, \ldots, a_{2k-2}$ are prescribed numbers of jumps (one for each edge of t_0), we find an asymptotic equivalent for the number of DH-trees in the family with k marked leaves inducing $(t_0, a_1, ...a_{2k-2})$. These local limit theorems are stated below as Theorems 5.7, 6.2

and 7.7. Such theorems imply the convergence in distribution of the induced trees with renormalized numbers of jumps, and, consequently, of the distance matrices of uniform graphs whithin each family. The limiting distance matrix can then be identified with that of the Brownian CRT, using the work of Aldous, see [2] and Theorem 2.4 here. These final arguments are the same for all three families and are presented only once, after the local limit theorem for induced trees in all DH-graphs (see in particular Theorems 5.8 and 5.10).

Our method to obtain the local limit theorems for enriched induced trees is through analytic combinatorics. Thanks to the description of DH-graphs via DH-trees, one can write down systems of equations characterizing their generating series, as well as those of the subclasses that we consider (see [14] or Theorem 3.9 and Equations (44) and (61)). We must refine these systems of equations to account for series of DH-trees with marked leaves with a given genealogy, introducing additional variables to keep track of the number of jumps along each edge of the induced tree. Writing equations for these refined multivariate series is a delicate task, requiring the introduction and analysis of several intermediate series (see Section 4). The local limit theorems are then obtained by extracting coefficients from these series; namely we use a slightly generalized version of the *Semi-large powers Theorem* (see Section A). Since this theorem is known for providing an analytic explanation for the appearance of the Rayleigh distribution, it is not surprising that we use it here.

The main steps of our proof strategy can then be summarized as follows:

- encoding DH-graphs as DH-trees;
- finding combinatorial equations for multivariate generating series of families of DH-trees with one or several marked leaves, counting both the size and numbers of jumps;
- performing asymptotic analysis of the generating functions;
- proving the convergence to the CRT via distance matrices.

Let us compare the method employed here with those used in previous works on related models. The correspondence between DH-graphs and DH-trees is a special case of the split decomposition for graphs. While we are not aware of prior use of split decomposition for probabilistic purposes, other tree decompositions have been used in the past to obtain scaling limit results: the substitution decomposition for permutations [6, 7, 8, 11], and the modular decomposition for graphs [9, 33]. Most of these papers also use analytic combinatorics to analyze the tree induced by k uniform random leaves in the decomposition tree. However, the need to track the number of jumps on each edge, and therefore to use multivariate generating series, is an important novelty of the present work. In particular, the Semi-large powers Theorem and the Rayleigh distribution do not appear in the above-cited papers.

On the other hand, in his seminal paper [2], Aldous proves the convergence of conditioned Galton–Watson trees to the Brownian CRT by establishing local limit theorems for the trees induced by k random vertices. Such local limit theorems are obtained through a mixed of combinatorial and probabilistic techniques. Such techniques seem hard to apply to our case where nodes of the trees have different types, making the combinatorics much more involved. This is why we need powerful techniques from analytic combinatorics to establish our local limit theorems.

To conclude on our proof strategy, we are not aware of other works where convergence to the Brownian CRT is proved through the same set of tools as here, and we hope that this method will prove useful in other contexts in the future.

Remark 1.2 A natural alternative strategy to prove our main result would be the following: first prove that the split decomposition tree $\boldsymbol{T}_f^{(n)}$ associated with $\boldsymbol{G}_f^{(n)}$ tends to the CRT, and then prove that $\boldsymbol{G}_f^{(n)}$ and $\boldsymbol{T}_f^{(n)}$ are close, up to some scaling factor, for the GP topology. This is in essence the approach used in [32] for subcritical block-stable classes of graphs, except that the block-decomposition tree is used instead of the split decomposition tree. There are however important difficulties to implement this strategy in our case (though they are not necessarily impossible to overcome).

First, the split decomposition trees $\boldsymbol{T}_f^{(n)}$ associated to our three models can be represented as multitype Galton-Watson trees conditioned to having a given number of leaves (as witnessed by the systems of equations (5), (44) and (61)). Convergence results to the CRT for conditioned multitype Galton-Watson trees are available in the literature (see, *e.g.*, [30]). However, such results are usually obtained for trees conditioned to having a given number of *vertices*, and in the irreducible case. Here we want to condition on the number of leaves, and, in one of our models, namely for 3-leaf power graphs, the system of equations defining the class is not irreducible; see Eq. (61). Therefore proving the convergence of the split decomposition trees to the CRT would need some work on models of random trees.

A second difficulty is that the convergence of the split decomposition trees does not imply directly the convergence of the associated graphs. For this, we would need to prove that distances in the graph are close, up to a constant factor, to that in the tree. But distances in the graph are determined by the decoration of vertices in the split decomposition (see Section 3.2). One would therefore need to understand the distribution of such decorations (*i.e.* of types in our multi-type model) on paths between marked leaves and branching points in split decomposition trees. Again, this might be feasible but certainly requires more work.

We have preferred to develop an approach via analytic combinatorics, as explained above, which is in some sense more direct and more original.

1.4 Outline of the paper

In order to simplify the presentation of the proofs we chose to focus first on the class of unconstrained DH-graphs. We explain later (in Sections 6 and 7) how to adapt the result to 2-connected DH-graphs and to 3-leaf power graphs.

- In Section 2 we state a criterion for the convergence towards the Brownian CRT with respect to the Gromov–Prohorov topology. This criterion essentially follows from [25, 29] and from Aldous' construction of the Brownian CRT [2].

- In Section 3 we give the necessary background of graph theory. We will see that there is a correspondence between DH-graphs and certain *clique-star trees*. Section 3 ends with exact and asymptotic enumerative formulas for DH-graphs. The material of this section is mainly taken from papers of Gioan–Paul [23] and Chauve–Fusy–Lumbroso [14].

- Section 4 is devoted to the combinatorial and analytic study of clique-star trees with a marked leaf. These are building blocks for the combinatorial decomposition of trees with several marked leaves done in Section 5, keeping track of distances in the graph between the corresponding vertices. The convergence of a uniform random unconstrained DH-graph to the Brownian CRT, *i.e.* the case $f = d$ in Theorem 1.1, is proved at the end of Section 5.

- In Sections 6 and 7 we extend the main result to 2-connected DH-graphs and to 3-leaf power graphs, respectively.

- In Section A we give a complete proof of a (minor) generalization of the *Semi-large powers Theorem* ([22, Theorem IX.16]), which is central in our proofs.

- Section B and Section C clarify the relation between the present work and the paper [32].

Note: Some computations in the proofs of our main results require the use of a computer algebra system. To help the reader, we provide a companion Maple worksheet, both in mw and pdf formats. These files are embedded into this pdf (alternatively you can download the source of the arXiv version to get the files).

2 Toolbox: the Gromov–Prohorov topology and the Brownian CRT

2.1 A criterion for Gromov–Prohorov convergence

Definition 2.1 *A metric measure space (called mm-space for short) is a triple (X, d, μ), where (X, d) is a complete and separable metric space and μ a probability measure on X.*

Gromov–Prohorov distance. We let $\mathbb{M}$ be the set of all mm-spaces[3], modulo the following relation: $(X, d, \mu) \sim (X', d', \mu')$ if there is an isometric embedding $\Phi : X \to X'$ such that the image measure (or push-forward) of μ by Φ, denoted by $\Phi_*(\mu)$, satisfies $\Phi_*(\mu) = \mu'$, *i.e.* $\mu(\Phi^{-1}(A')) = \mu'(A')$ for every Borel set A' of X'.

[3]To avoid Russell's paradox, throughout the section, we actually take the set of mm-spaces *whose elements are not themselves metric spaces.*

Note that Φ does not need to be surjective, so that we need to consider the transitivity and reflexivity closure of that relation. In particular one always has $(X, d, \mu) \sim (\mathrm{Supp}(\mu), d, \mu)$, where $\mathrm{Supp}(\mu)$ is the support of μ.

On the set $\mathbb{M}$, one can define a distance as follows. First we recall the notion of Prohorov distance: for Borel probability measures μ and ν on the same metric space Y, we set

$$d_P(\mu,\nu) = \inf\big\{\varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon \text{ and } \nu(A) \leq \mu(A^\varepsilon) + \varepsilon \text{ for all measurable sets } A \subseteq Y\big\},$$

where A^ε is the ε-halo of A, *i.e.* the set of all points at distance at most ε of A. This distance metrizes the weak convergence of probability measures. Then, given two mm-spaces (X, d, μ) and (X', d', μ'), we set

$$d_{\mathrm{GP}}\big((X, d, \mu), (X', d', \mu')\big) = \inf_{(Y,d_Y),\Phi,\Phi'} d_P\big(\Phi_*(\mu), \Phi'_*(\mu')\big),$$

where the infimum is taken over isometric embeddings $\Phi : X \to Y$ and $\Phi' : X' \to Y$ into a common metric space (Y, d_Y). One can prove [25, Section 5] that d_{GP} is a distance on $\mathbb{M}$ and that the resulting metric space $(\mathbb{M}, d_{\mathrm{GP}})$ is complete and separable.

Criterion of convergence. Let $\mathcal{X} = (X, d, \mu)$ be an mm-space and fix an integer $k \geq 0$. We let $x_1, \ldots, x_k$ be i.i.d. random elements of X, with law μ. We record their pairwise distances in a matrix, namely we set

$$A_k^{\mathcal{X}} = \big(d(x_i, x_j)\big)_{1\leq i,j\leq k}.$$

This is a random $k \times k$ square matrix, whose law depends on the mm-space $\mathcal{X}$ we start with.

We will also consider random mm-spaces, which we denote with boldface letters. In this case, *conditionally on* $\boldsymbol{\mathcal{X}} = (\boldsymbol{X}, \boldsymbol{d}, \boldsymbol{\mu})$, we let $x_1, \ldots, x_k$ be i.i.d. random elements of $\boldsymbol{X}$, with law $\boldsymbol{\mu}$ and we define as above $A_k^{\boldsymbol{\mathcal{X}}}$ to be their distance matrix.

We have the following characterization of convergence in distribution in $(\mathbb{M}, d_{\mathrm{GP}})$, essentially given in [25, 29].

Theorem 2.2 *Let* $\boldsymbol{\mathcal{X}}_n = (\boldsymbol{X_n}, \boldsymbol{d_n}, \boldsymbol{\mu_n})$ *for any* $n \geq 1$ *and* $\boldsymbol{\mathcal{X}} = (\boldsymbol{X}, \boldsymbol{d}, \boldsymbol{\mu})$ *be random mm-spaces. Then the following properties are equivalent:*

i) $\boldsymbol{\mathcal{X}}_n$ *converges in distribution to* $\boldsymbol{\mathcal{X}}$ *for the Gromov–Prohorov distance* d_{GP} *as* $n \to +\infty$.

ii) For any fixed $k \geq 1$, *the random distance matrix* $A_k^{\boldsymbol{\mathcal{X}}_n}$ *converges in distribution to* $A_k^{\boldsymbol{\mathcal{X}}}$ *as* n *tends to* $+\infty$.

Proof. In [25, Theorem 5], it is proved in the deterministic setting that convergence for Gromov–Prohorov distance is equivalent to the convergence of the so-called *polynomial functions*, *i.e.* of bounded continuous functions of (the entries of) distance matrices. It is then observed in [29, Corollary 2.8] that polynomial functions are *convergence-determining*, *i.e.* one has convergence in distribution of random mm-spaces if the expectations of all polynomial functions converge. On the other hand since polynomial functions are the continuous bounded functions of distance matrices, the convergence of expectations of polynomial functions is equivalent to the convergence in distribution of the distance matrices. This completes the proof. □

2.2 Distance matrix of the Brownian CRT

We now present the key properties of the Brownian Continuum Random Tree (CRT), which is the limiting object appearing in Theorem 1.1 (it is also known as the *Brownian Tree* or *Aldous' Tree*). This consists in a random metric measure space $(\mathcal{T}_\infty, d_\infty, \mu_\infty)$ that appears as a scaling limit for many sequences of random trees with n nodes, when the lengths of edges are divided by $\sqrt{n}$. (Among these families are critical Galton-Watson trees conditioned to having n nodes, when the reproduction law has finite variance.)

We first recall the usual construction of the CRT (see [28, Section 2] for more details). Starting from a normalized Brownian excursion $\mathfrak{e}$, $\mathcal{T}_\infty$ is defined as the quotient $[0,1]/\sim_{\mathfrak{e}}$ where $\sim_{\mathfrak{e}}$ is the "gluing" procedure which identifies any two points of $\mathfrak{e}$ at the same height having only higher points of $\mathfrak{e}$ between them. Through the latter construction, the mass measure μ_∞ is defined as the push-forward of the Lebesgue measure by the quotient map associated with $\sim_{\mathfrak{e}}$.

We now introduce another characterization of the CRT, which is the one we use in this paper. Informally, this states that the mutual distances of k points in $(\mathcal{T}_\infty, d_\infty, \mu_\infty)$ have the same distribution as the distances between the k leaves of a uniform random *k-proper tree* (defined below) in which edges have random length distributed according a multivariate Rayleigh distribution.

Definition 2.3 *A k-proper tree t_0 is an (unrooted) nonplane tree with $k+1$ leaves labeled $\ell_0, \ell_1, \dots, \ell_k$ and where each internal node has degree 3. The leaf ℓ_0 is considered as the root-leaf.*

Let $\mathcal{T}_k$ be the set of k-proper trees. In the sequel we let cnt and Leb be respectively the counting measure on $\mathcal{T}_k$ and the Lebesgue measure on $\mathbb{R}_+$. From [2, Lemma 21], we know that the function

$$f(t_0, x_0, \cdots, x_{2k-2}) = \left(\sum_{i=0}^{2k-2} x_i\right) \exp\left(-\frac{1}{2}\left(\sum_{i=0}^{2k-2} x_i\right)^2\right) \tag{3}$$

defines a density on $\mathcal{T}_k \times \mathbb{R}_+^{2k-1}$ with respect to the measure $\mathrm{cnt} \otimes \mathrm{Leb}^{\otimes 2k-1}$. Note that f does not depend on t_0, which means that if $(\boldsymbol{t_0}, X_0, \dots, X_{2k-2})$ is a random tuple in $\mathcal{T}_k \times \mathbb{R}_+^{2k-1}$ with that density, then $\boldsymbol{t_0}$ is uniform and independent from $(X_0, \dots, X_{2k-2})$.

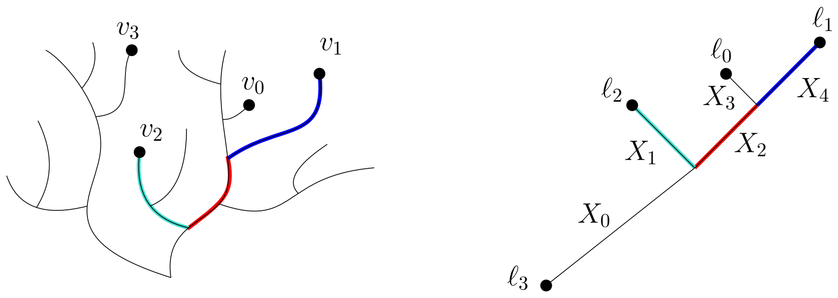

Figure 5: Illustration of Theorem 2.4. Left: an artistic view of the Brownian CRT $\mathcal{T}_\infty$, with $k+1 = 4$ points $v_0, \ldots, v_3$ independently drawn with distribution μ_∞ in $\mathcal{T}_\infty$. Right: A k-proper tree, whose edge lengths are given by distances in $\mathcal{T}_\infty$. For instance $d_\infty(v_2, v_1) = X_1 + X_2 + X_4$, as emphasized by colors.

Lemma 2.4 (See Fig. 5 for notation.) *For every $k \geq 2$ and every k-proper tree t_0 in $\mathcal{T}_k$, we fix a labeling of its edges $e_0, \ldots, e_{2k-2}$.*

The distribution of $(\mathcal{T}_\infty, d_\infty, \mu_\infty)$ is characterized by the property that for every $k \geq 2$, if we take $v_0, \ldots, v_k$ uniform and independent in $\mathcal{T}_\infty$ with distribution μ_∞, then

$$\Big(d_\infty(v_i, v_j)\Big)_{0 \leq i,j \leq k} \stackrel{(d)}{=} \Bigg(\sum_{r:\ e_r \in \mathcal{P}_{i,j}^{\boldsymbol{t}_0}} X_r\Bigg)_{0 \leq i,j \leq k}, \tag{4}$$

where in the RHS, the random tuple $(\boldsymbol{t}_0, X_0, \ldots, X_{2k-2})$ has density given by (3) and the sum runs over edges e_r on the path $\mathcal{P}_{i,j}^{\boldsymbol{t}_0}$ joining leaves ℓ_i and ℓ_j in $\boldsymbol{t}_0$.

Aldous proved that this property properly defines a random metric space [2, Lemma 21] which coincides with the previous construction of the CRT (as gluing of a Brownian excursion) [2, Cor.22].

3 Combinatorial analysis of distance-hereditary trees

In this section, we first recall the encoding of distance-hereditary graphs by clique-star trees (which is a special case of the encoding of general graphs by split decomposition trees). This is done in Section 3.1 and largely follows [23, Sections 2.1-2.2] (itself inspired by [17]). We then explain how distances in a DH-graph can be recovered from the associated clique-star tree (Section 3.2). We could not find this result in the literature, though this might be known to experts. The last two sections provide a combinatorial and analytic study of the generating series of DH-graphs (or rather of the associated trees); this mainly follows the work of Chauve–Fusy–Lumbroso [14]. This whole section can be seen as combinatorial preliminaries for the proof of the convergence of unconstrained DH-graphs to the Brownian CRT (case $f = d$ in Theorem 1.1).

Throughout the paper, all trees are labeled on their leaves.

3.1 Clique-star trees

Definition 3.1 *A graph-decorated*[4] *tree is a (nonplane unrooted) tree τ in which every internal node v of degree k is decorated with a graph Γ_v with k vertices; moreover, for each v, we fix a bijection ρ_v from the tree-edges incident to v to the vertices of Γ_v.*

See Fig. 6 left for an example of graph-decorated tree.

We fix some terminology and conventions. To avoid confusion between decoration graphs Γ_v and other graphs, we use the term *decoration* for Γ_v and *marker vertices* for its vertices. An edge e of τ between two nodes v and v' is sometimes seen as connecting the marker vertices $q = \rho_v(e)$ to $q' = \rho_{v'}(e)$. In particular in graphical representations, we draw an edge e of the tree between nodes v and v' from $q = \rho_v(e)$ to $q' = \rho_{v'}(e)$. When we refer to the bijection ρ_v, we say that an edge e incident to v is *attached* to the corresponding marker vertex (say, x) of Γ_v. When e is incident to v and to a leaf ℓ, we make a small abuse of notation by saying that ℓ is attached to x.

Let τ be a graph-decorated tree and ℓ, ℓ' be leaves of τ. We consider the (unique) path p from ℓ to ℓ' in τ. For any node v on this path, we denote $e_{in}(v)$ (respectively $e_{out}(v)$) the edge of p entering (respectively leaving) v. Then ℓ' is said to be accessible from ℓ (or equivalently ℓ accessible from ℓ') if, for every node v on p, the pair $\{\rho_v(e_{in}(v)), \rho_v(e_{out}(v))\}$ is an edge of the decoration Γ_v. With this notion in hand, we can associate to τ a graph $\mathrm{Gr}(\tau)$, whose vertex set is the leaf set of τ, and where $\{\ell, \ell'\}$ is an edge in $\mathrm{Gr}(\tau)$ if and only if ℓ is accessible from ℓ' in τ. Since τ is labeled on its leaves, the vertices of $\mathrm{Gr}(\tau)$ are naturally labeled as well. This construction is illustrated on Fig. 6.

In the sequel, we only consider graph-decorated trees τ where all decorations Γ_v are either cliques (*i.e.*, complete graphs) or stars (*i.e.*, graphs where one vertex, called the center, is connected to all others, which form an independent set among themselves) – following [14], we speak of *clique-star trees*. It is known (see [23, Section 3.1]) that the graphs which can be obtained as $\mathrm{Gr}(\tau)$ where τ is a clique-star tree, are precisely the distance-hereditary graphs (DH-graphs). By convention the graph with a single vertex and the connected graph with two vertices are DH-graphs.

We note that a DH-graph G can possibly be obtained as $\mathrm{Gr}(\tau)$ for several clique-star trees τ. Uniqueness can nevertheless be ensured by adding extra conditions on τ.

Definition 3.2 *A clique-star tree τ is called* reduced *if it satisfies the following conditions:*

i) every internal node v has degree at least 3*;*

ii) no edge of τ connects two internal nodes both decorated with cliques;

[4]In [23], the term *graph-labeled tree* is used; we prefer here to speak of *graph-decorated tree* to avoid confusion with labeling in the sense of labeled combinatorial classes [22], a notion that we will use throughout the article.

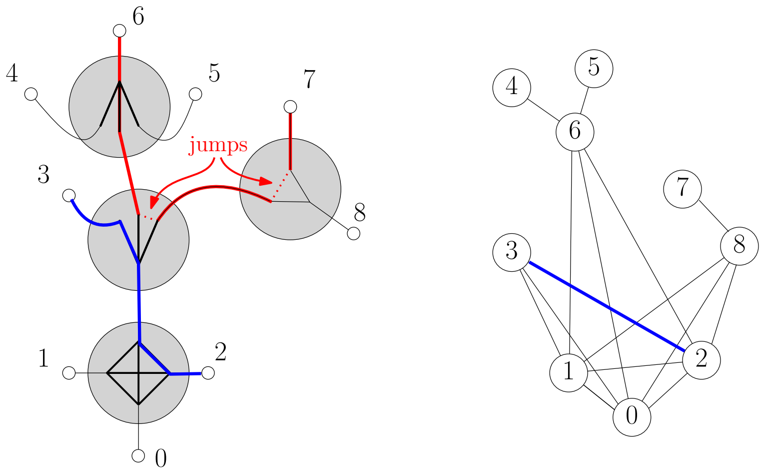

Figure 6: Left: A clique-star tree τ with $n = 9$ leaves drawn with its 4 decorations. Right: The corresponding graph $\mathrm{Gr}(\tau)$. To illustrate the construction of $\mathrm{Gr}(\tau)$, we have highlighted two pairs of leaves and the paths between them. Following the blue path, we say that 3 is accessible from 2 in τ; accordingly $\{2, 3\}$ is an edge in $\mathrm{Gr}(\tau)$. On the opposite, 6 is not accessible from 7 in τ; accordingly $\{6, 7\}$ is not an edge in $\mathrm{Gr}(\tau)$. (Jumps are defined in Section 3.2.)

iii) no edge of τ connects marker vertices q and q' where q is the center of a star Γ_v and q' a leaf of another star $\Gamma_{v'}$.

Then uniqueness follows directly from [23, Theorem 2.9] (which considers all graphs, not only DH-graphs). Namely, the following holds.

Proposition 3.3 *For every labeled DH-graph G of size at least 3, there exists a unique reduced clique-star tree τ such that $G = \mathrm{Gr}(\tau)$.*

3.2 Distances in DH-graphs through their clique-star trees

Let τ be a clique-star tree and $G = \mathrm{Gr}(\tau)$ be the corresponding graph (which is a DH-graph as we have seen). We denote by d_G the graph distance in G. In this section, we explain how d_G can be read on the tree τ. We recall that the leaves of τ are identified with the vertices of G.

For a path p in τ, the *jumps* of p are defined as follows. When p goes through a node v, it enters and exits through edges $e_{in}(v)$ and $e_{out}(v)$ (both incident to v). If $\{\rho_v(e_{in}(v)), \rho_v(e_{out}(v))\}$ is not an edge in Γ_v, we say that v is a jump of p. (In particular, and unless otherwised specified, the starting and ending points of p are not jumps of p.) Now, for two leaves ℓ and ℓ' of τ, letting p be the unique path from ℓ to ℓ' in τ, the number of jumps of p is denoted by $\mathrm{jp}(\tau, \ell, \ell')$.

Lemma 3.4 *Let τ be a clique-star tree with corresponding DH-graph $G = \mathrm{Gr}(\tau)$, and let ℓ, ℓ' be leaves of τ. Then we have $d_G(\ell, \ell') = \mathrm{jp}(\tau, \ell, \ell') + 1$.*

Example 3.5 Consider the clique-star tree τ of Fig. 6, and its leaves 6 and 7. The path from 6 to 7 (in red on the picture) has exactly two jumps. Accordingly, the distance between vertices 6 and 7 in the associated DH-graph (also drawn on Fig. 6) is 3.

Remark 3.6 According to Lemma 3.4, ℓ is accessible from ℓ' in τ (*i.e.* $\mathrm{jp}(\tau, \ell, \ell') = 0$) if and only if $\{\ell, \ell'\}$ is an edge of G (*i.e.* $d_G(\ell, \ell') = 1$). In other words, the lemma superseeds and generalizes the definition of the edge set of $\mathrm{Gr}(\tau)$.

Proof. We proceed by induction. If τ has a single internal node, then G is isomorphic to the decoration of that node (hence, either a clique or a star), and the statement holds trivially.

Let τ have $k > 1$ internal nodes and assume that the statement holds for all clique-star trees with fewer internal nodes.

Consider a node v of τ, all of whose neighbors but one are leaves (such a node always exists). Denote by $d \geq 3$ the degree of v, by $\ell_1, \ldots, \ell_{d-1}$ the leaves adjacent to v, and by u the internal node of τ adjacent to v. We also denote by Γ_v the decoration of v, and by x the marker vertex of Γ_v corresponding to the edge (v, u). We let $\tau^\star$ be the clique-star tree obtained by replacing v and $\ell_1, \ldots, \ell_{d-1}$ by a single leaf $\ell^\star$ (adjacent to u), and denote by $G^\star = \mathrm{Gr}(\tau^\star)$ the associated graph. All these notations are sumarized on Fig. 7 for the reader's convenience.

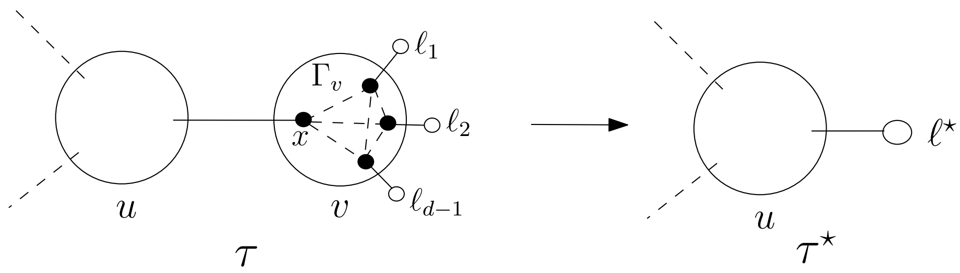

Figure 7: Illustration of the proof of Theorem 3.4.

As we shall see, G can be obtained by performing some local modifications on $G^\star$, which depend on Γ_v and x. First note that leaves of τ and $\tau^\star$ different from $\ell^\star$, ℓ_1, $\ldots$, ℓ_{d-1} are the same and are therefore vertices in both G and $G^\star$; we call them *old* vertices, while we call $\ell_1, \ldots, \ell_{d-1}$ *new* vertices. By construction, adjacency relations between old vertices are identical in G and $G^\star$. So, knowing $G^\star$, to know G entirely, we just have to describe the adjacency relations among new vertices, and between the new vertices and the old ones. To this end, we distinguish several cases.

- If Γ_v is a clique, then the definition of the construction Gr implies that G is obtained from $G^\star$ by replacing $\ell^\star$ with $d-1$ vertices $\ell_1, \ldots, \ell_{d-1}$, which form a clique of size $d-1$, and such that the old neighbors of each ℓ_i are the neighbors of $\ell^\star$ in $G^\star$.

- If Γ_v is a star with x the center of the star, then similarly G is obtained from $G^\star$ by replacing $\ell^\star$ with $d-1$ vertices $\ell_1, \ldots, \ell_{d-1}$, which form an independent set of size $d-1$, and such that the old neighbors of each ℓ_i are the neighbors of $\ell^\star$ in $G^\star$.
- Finally, assume that Γ_v is a star and x is not the center of the star. Let ℓ_j be the leaf of τ attached to the center of Γ_v. Here, G is obtained from $G^\star$ by keeping the vertex $\ell^\star$ (with its adjacent edges) but renaming it ℓ_j, and adding $d-2$ vertices $\ell_1, \ldots, \ell_{j-1}, \ell_{j+1}, \ldots, \ell_{d-1}$, which form an independent set of size $d-2$, and all connected only to ℓ_j.

In particular, G always contains at least one vertex with exactly the same old neighbors as $\ell^\star$ in $G^\star$; call such vertices *copies* of $\ell^\star$. Moreover, new vertices of G which are not copies of $\ell^\star$ are pendant vertices incident to a copy of $\ell^\star$.

With this remark, it becomes clear that distances between old vertices are the same in G and $G^\star$. Moreover, the path between any two old leaves ℓ and ℓ' in τ also matches the path between ℓ and ℓ' in $\tau^\star$, so that we have $d_G(\ell, \ell') = d_{G^\star}(\ell, \ell') = \mathrm{jp}(\tau^\star, \ell, \ell') + 1 = \mathrm{jp}(\tau, \ell, \ell') + 1$ as claimed. When ℓ and ℓ' are both new vertices, their distance $d_G(\ell, \ell')$ is either 1 or 2, depending on whether the corresponding marker vertices in Γ_v are connected or not. Thus, in this case also, we have $d_G(\ell, \ell') = \mathrm{jp}(\tau, \ell, \ell') + 1$. The interesting case is when ℓ is a new vertex and ℓ' an old vertex. Again, we proceed by case analysis. Denote by p the path from ℓ to ℓ' in τ and by $p^\star$ the path from $\ell^\star$ to ℓ' in $\tau^\star$. The path p is obtained from $p^\star$ by replacing the first edge $(\ell^\star, u)$ by the two edges $(\ell, v), (v, u)$. (Recall that u is the only nonleaf node of τ adjacent to v, corresponding to the marker vertex x of Γ_v.)

- Assume first that ℓ is a copy of $\ell^\star$. Note that this happens when Γ_v is a clique, or when Γ_v is a star with ℓ attached to the center of Γ_v, or when Γ_v is a star with x the center of the star. Since ℓ is a copy of $\ell^\star$, of course $d_G(\ell, \ell') = d_{G^\star}(\ell^\star, \ell')$. On the other hand, in all cases, the marker vertices of Γ_v attached to ℓ and u are adjacent. Therefore, we have $\mathrm{jp}(\tau^\star, \ell^\star, \ell') = \mathrm{jp}(\tau, \ell, \ell')$, and it follows that $d_G(\ell, \ell') = \mathrm{jp}(\tau, \ell, \ell') + 1$.
- The last case to consider is when Γ_v is a star with x an extremity of the star, and ℓ attached to another extremity of the star. In this case, p has one more jump than $p^\star$, since the marker vertices to which x and ℓ are attached are not adjacent in Γ_v. On the other hand, the only neighbor of ℓ in G is the leaf of τ attached to the center of Γ_v, previously denoted ℓ_j. Since ℓ_j is a copy of $\ell^\star$, we have $d_G(\ell, \ell') = 1 + d_G(\ell_j, \ell') = 1 + d_{G^\star}(\ell^\star, \ell')$, which gives $d_G(\ell, \ell') = \mathrm{jp}(\tau, \ell, \ell') + 1$ as desired. □

3.3 Clique-star trees as a labeled combinatorial class

In Section 3.1, we have seen that DH-graphs are in bijection with reduced clique-star trees. We recall that the latter are nonplane unrooted trees. To use the symbolic method and tools of analytic combinatorics, it is more convenient to deal with rooted trees. Starting from a DH-graph with vertex set $\{0, 1, \dots, n\}$, we consider the reduced clique-star tree associated with it by Theorem 3.3 and see the leaf with label 0 as the root.

Definition 3.7 *A* distance-hereditary tree *(DH-tree for short) of size $n \geq 2$ is a reduced clique-star tree with $n+1$ leaves labeled from 0 to n, where the leaf 0 is seen as the root, therefore called the* root-leaf.

By construction, DH-trees of size n are in bijection with DH-graphs with vertex set $\{0, 1, \dots, n\}$. Most of the time, we forget the root-leaf and think at the tree as rooted in the internal node to which the root-leaf is attached; this node is referred to as root-node below. The root-leaf is represented by the symbol $\perp$ in pictures (see for instance Fig. 8).

Having broken the symmetry when selecting a root, a node v decorated with a star can be of two types.

- Either the path from v to the root-leaf exits v through an edge attached to *an extremity* of the star Γ_v. In this case, we say that v is of type $\mathcal{S}_{\mathrm{X}}$. Note that one of the children of v is attached to the center of the star. We see this child as distinguished.

- Or the path from v to the root-leaf exits v through the edge attached to *the center* of the star. In this case, we say that v is of type $\mathcal{S}_{\mathrm{C}}$. Note that all children of v are attached to extremities of the star so that there is no distinguished child in this case.

A node decorated with a clique is of type $\mathcal{K}$.

With this in mind, and recalling the conditions of Theorem 3.2, one can describe DH-trees directly as follows.

Proposition 3.8 *A DH-tree is a nonplane tree T rooted at a node such that*

i) T has n leaves labeled $1, \dots, n$;
ii) internal nodes of T (including the root) carry decorations, called types*, taken from the set $\{\mathcal{K}, \mathcal{S}_{\mathrm{C}}, \mathcal{S}_{\mathrm{X}}\}$;*
iii) every node of type $\mathcal{K}$ has at least 2 children, none of which can be of type $\mathcal{K}$;
iv) every node of type $\mathcal{S}_{\mathrm{C}}$ has at least 2 children, none of which can be of type $\mathcal{S}_{\mathrm{C}}$;
v) every node of type $\mathcal{S}_{\mathrm{X}}$ has at least 2 children, one of which is distinguished; the distinguished child cannot be of type $\mathcal{S}_{\mathrm{X}}$, while other are forbidden to be of type $\mathcal{S}_{\mathrm{C}}$.

Fig. 8 shows an example of DH-tree.

We will now translate this description into the framework of labeled combinatorial classes (see [22] for an introduction). We recall that $+$ is used for the disjoint union of combinatorial classes; $\mathcal{A} \times \mathcal{B}$ is the set of pairs (a, b) where a is in $\mathcal{A}$ and b in $\mathcal{B}$ (with

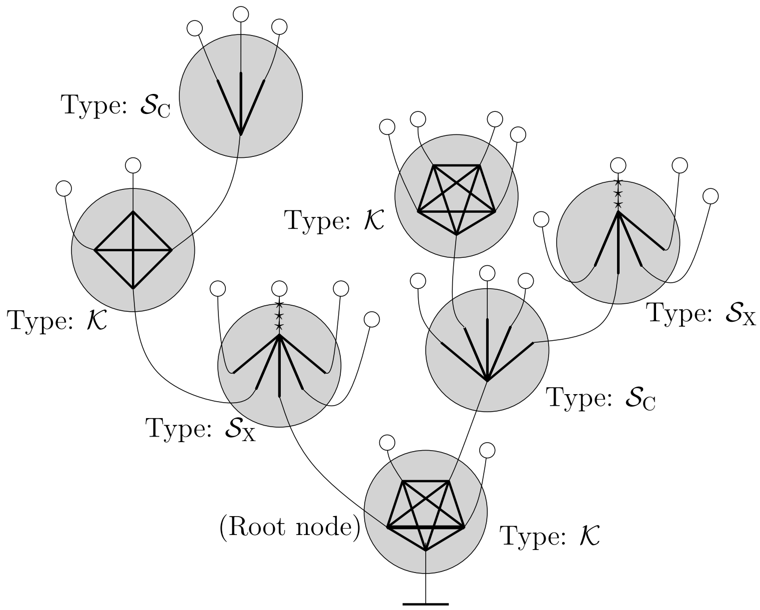

Figure 8: A DH-tree of size $n = 22$, omitting the labels of the leaves for readability. For both nodes of type $\mathcal{S}_\mathrm{X}$ one indicates with three $\star$'s the edge going to its distinguished child (which is, because of Item v, not of type $\mathcal{S}_\mathrm{X}$.)

the convention that the label sets of a and b are disjoint; we refer to [22] for details on how to deal with labelings in combinatorial classes). Also, if $\mathcal{C}$ is a combinatorial class with no element of size 0, then $\mathrm{Set}(\mathcal{C})$ is the class of (unordered) sets of elements of $\mathcal{C}$. An index on Set indicates restrictions on the number of elements in the set.

We say that a DH-tree is of type t if its root-node is of type t. We let $\mathcal{D}_\mathcal{K}$ (respectively $\mathcal{D}_{\mathcal{S}_\mathrm{C}}$, $\mathcal{D}_{\mathcal{S}_\mathrm{X}}$) be the (labeled) combinatorial class of DH-trees of type $\mathcal{K}$ (respectively $\mathcal{S}_\mathrm{C}$, $\mathcal{S}_\mathrm{X}$). As usual, we use the symbol $\mathcal{Z}$ to represent the trivial tree reduced to one vertex (which is a leaf).

Proposition 3.9 (Chauve-Fusy-Lumbroso[14]) [5] *The combinatorial classes $\mathcal{D}_\mathcal{K}, \mathcal{D}_{\mathcal{S}_\mathrm{C}}, \mathcal{D}_{\mathcal{S}_\mathrm{X}}$ have the following specification:*

$$\begin{cases} \mathcal{D}_\mathcal{K} = & \mathrm{Set}_{\geq 2}(\mathcal{Z} + \mathcal{D}_{\mathcal{S}_\mathrm{C}} + \mathcal{D}_{\mathcal{S}_\mathrm{X}}); \\ \mathcal{D}_{\mathcal{S}_\mathrm{C}} = & \mathrm{Set}_{\geq 2}(\mathcal{Z} + \mathcal{D}_\mathcal{K} + \mathcal{D}_{\mathcal{S}_\mathrm{X}}); \\ \mathcal{D}_{\mathcal{S}_\mathrm{X}} = & (\mathcal{Z} + \mathcal{D}_\mathcal{K} + \mathcal{D}_{\mathcal{S}_\mathrm{C}}) \times \mathrm{Set}_{\geq 1}(\mathcal{Z} + \mathcal{D}_\mathcal{K} + \mathcal{D}_{\mathcal{S}_\mathrm{X}}). \end{cases} \tag{5}$$

The class $\mathcal{D}$ of all DH-trees is simply the disjoint union of the three classes above, i.e.

$$\mathcal{D} = \mathcal{D}_\mathcal{K} + \mathcal{D}_{\mathcal{S}_\mathrm{C}} + \mathcal{D}_{\mathcal{S}_\mathrm{X}}.$$

[5] The equation given for $\mathcal{D}_{\mathcal{S}_\mathrm{X}}$ in [14, Theorem 3] is different from the one given here. The one given here can however be found in the proof of [14, Theorem 3].

3.4 Singularity analysis of the specification

We associate to each combinatorial class of DH-trees a generating function $D = D(z)$, $D_{\mathcal{K}} = D_{\mathcal{K}}(z)$, $D_{\mathcal{S}_{\mathrm{C}}} = D_{\mathcal{S}_{\mathrm{C}}}(z)$ and $D_{\mathcal{S}_{\mathrm{X}}} = D_{\mathcal{S}_{\mathrm{X}}}(z)$:

$$D = \sum_{T\in\mathcal{D}} \frac{z^{|T|}}{|T|!}\ , \quad D_{\mathcal{K}} = \sum_{T\in\mathcal{D}_{\mathcal{K}}} \frac{z^{|T|}}{|T|!}\ , \quad D_{\mathcal{S}_{\mathrm{C}}} = \sum_{T\in\mathcal{D}_{\mathcal{S}_{\mathrm{C}}}} \frac{z^{|T|}}{|T|!}\ , \quad D_{\mathcal{S}_{\mathrm{X}}} = \sum_{T\in\mathcal{D}_{\mathcal{S}_{\mathrm{X}}}} \frac{z^{|T|}}{|T|!}\ .$$

By loose estimates on the number of DH-trees, it is easy to see that each of the above series has a positive radius of convergence. A key step in the proof of our main theorem will be given by the singularity analysis of the above series. A similar analysis is provided in [14] in the *unlabeled* case; we here give all the details of the labeled case.

We first note that using Eqs. (5) and an immediate induction on $i \geq 0$, we have $[z^i]D_{\mathcal{K}} = [z^i]D_{\mathcal{S}_{\mathrm{C}}}$ for all $i \geq 0$, *i.e.* $D_{\mathcal{K}} = D_{\mathcal{S}_{\mathrm{C}}}$ as formal power series. We will therefore drop $D_{\mathcal{S}_{\mathrm{C}}}$ and use only $D_{\mathcal{K}}$. Using the standard translation of equations from combinatorial families to exponential generating functions [22, Theorem II.1], Eqs. (5) yield:

$$\begin{cases} D_{\mathcal{K}} = & \exp_{\geq 2}(z + D_{\mathcal{K}} + D_{\mathcal{S}_{\mathrm{X}}}); \\ D_{\mathcal{S}_{\mathrm{X}}} = & (z + 2D_{\mathcal{K}}) \exp_{\geq 1}(z + D_{\mathcal{K}} + D_{\mathcal{S}_{\mathrm{X}}})\ , \end{cases} \tag{6}$$

where $\exp_{\geq r}(y) = \sum_{\ell \geq r} y^\ell/\ell!$.

The system (6) satisfies the assumptions of the Drmota–Lalley–Woods Theorem (see [8, Theorem A.6][6]). It follows that the series $D_{\mathcal{K}}, D_{\mathcal{S}_{\mathrm{X}}}$ have the same radius of convergence ρ and both have a square-root singularity at ρ. Moreover they are Δ-*analytic*, meaning that they are defined and analytic on some set of the form

$$\{z \in \mathbb{C}, |z| < R_1 \text{ and } |\operatorname{Arg}(z-\rho)| > \theta\},$$

for some $R_1 > \rho$ and $\theta > 0$, where Arg is the principal determination of the logarithm. The notion of Δ-analyticity is standard in analytic combinatorics, see [22, Chapter VI]. Let us introduce an auxiliary series

$$F(z) := \exp_{\geq 1}(z + D_{\mathcal{K}} + D_{\mathcal{S}_{\mathrm{X}}}). \tag{7}$$

Lemma 3.10 *We have*

$$\begin{cases} D_{\mathcal{K}} = & \frac{1}{2}\left(\frac{F}{1+F} - z\right); \\ D_{\mathcal{S}_{\mathrm{X}}} = & \frac{F^2}{1+F}. \end{cases} \tag{8}$$

Proof. Using F, we can rewrite the system (6) as

$$\begin{cases} D_{\mathcal{K}} = & F - (z + D_{\mathcal{K}} + D_{\mathcal{S}_{\mathrm{X}}}); \\ D_{\mathcal{S}_{\mathrm{X}}} = & (z + 2D_{\mathcal{K}})\, F. \end{cases}$$

[6] More classical references for variants of this theorem are [22, Section VII.6] and [20, Section 2.2.5], but the first one assumes that we have a polynomial system, while the second one has a different well-posedness condition, which is not satisfied here (and uses extra parameters which are not needed here).

We solve this linear system for $D_{\mathcal{K}}$ and $D_{\mathcal{S}_X}$, seeing F as a parameter. This gives the formulas of the lemma. □

Proposition 3.11 *The series F is Δ-analytic at ρ and admits the following singular expansion around ρ:*

$$F(z) = F(\rho) - \gamma_F \sqrt{1 - z/\rho} + O\big(1 - z/\rho\big), \tag{9}$$

where

- $F(\rho) = \frac{\sqrt{3}-1}{2}$ *is the unique positive root of* $2F(\rho)^2 + 2F(\rho) - 1 = 0$*;*
- $\gamma_F = \frac{\sqrt{2}(1+\sqrt{3})^2}{4\sqrt{3+2\sqrt{3}}}\sqrt{\rho}$.

The expression for γ_F is computed in the companion Maple worksheet. This also holds for other constants $\gamma_H, \gamma_{H,2c}, \gamma_{E,3\ell}$ arising later.

Throughout the paper, when a series S has a square-root singularity, we denote by γ_S the coefficient of the square-root term in the singular expansion of S near its radius of convergence, with the same sign convention as above. Also, for a variable x and a (multivariate) function $A(x, \dots)$, we denote by A_x the partial derivative of A with respect to x.

Proof. By Eq. (7) the series F is Δ-analytic at ρ and has a square-root singularity at ρ, therefore the expansion of F around ρ is given by Eq. (9) for some $F(\rho), \gamma_F$ which are to be determined. In addition, since F is a series in z with nonnegative coefficients, the transfer theorem ensures that $\gamma_F > 0$.

Thanks to Eqs. (8) one can eliminate $D_{\mathcal{K}}$ and $D_{\mathcal{S}_X}$ in Eq. (7). We obtain that F is the solution of the equation $F = G(z, F)$, where

$$G(z, w) = \exp_{\geq 1}\left[z + \frac{1}{2}\left(\frac{w}{1+w} - z\right) + \frac{w^2}{1+w}\right]. \tag{10}$$

Plugging Eq. (9) into $F = G(z, F)$ and comparing the expansions of both sides show that necessarily

$$F(\rho) = G(\rho, F(\rho)) \quad \text{and} \quad G_w(\rho, F(\rho)) = 1. \tag{11}$$

These equations are usually referred to as the characteristic system [22, Section VII.4]; if this system has a solution within the domain of convergence of G, then ρ is the radius of convergence of F, and F has a square-root singularity in ρ, see the Singular Implicit Function Lemma [22, Lemma VII.3].

In our case, observing that

$$G_w(z, w) = (1 - \tfrac{1}{2(1+w)^2})(1 + G(z, w)),$$

the characteristic system yields the following equation

$$2F(\rho)^2 + 2F(\rho) - 1 = 0, \tag{12}$$

whose only positive solution is $\frac{\sqrt{3}-1}{2}$.

Using that $F(\rho) = \frac{\sqrt{3}-1}{2}$, we can solve for ρ the first of Eqs. (11), giving an explicit expression for ρ and the numerical estimate $\rho \approx 0.1597$ (see Maple worksheet). Furthermore, using the Singular Implicit Function Lemma [22, Lemma VII.3], the constant γ_F is given by

$$\gamma_F = \sqrt{\frac{2\rho G_z(\rho, F(\rho))}{G_{ww}(\rho, F(\rho))}}.$$

We note that $G_z(z, w) = \frac{1}{2}(1 + G(z, w))$, so that $2G_z(\rho, F(\rho)) = 1 + F(\rho)$. Thus we have

$$\gamma_F = \sqrt{\frac{1 + F(\rho)}{G_{ww}(\rho, F(\rho))}}\,\sqrt{\rho} = \frac{(1+\sqrt{3})^2}{2\sqrt{6+4\sqrt{3}}}\sqrt{\rho},$$

where the last equality is justified in the companion Maple worksheet. □

Remark 3.12 Since F is the solution of the implicit equation $F = G(z, F)$, it is tempting to use the smooth implicit-function schema [22, Theorem VII.3] to find its dominant singularity and asymptotic expansion. We can however not proceed like this since the expansion of G contains negative coefficients, contradicting [22, Hypothesis ($\boldsymbol{I_2}$) p. 468]. This explains the indirect path used here. In short, the system (6) has the advantage of having nonnegative coefficients: it is used to prove without effort that all series have square-root singularities. On the other hand, F is defined by a single equation, giving simpler computations to determine explicitly the coefficients in its singular expansion.

In the sequel we also need the asymptotic expansion of $D_{\mathcal{K}}$, $D_{\mathcal{S}_\mathrm{X}}$ and its derivative. Recall from Theorem 3.10 that $D_{\mathcal{S}_\mathrm{X}} = \frac{F^2}{1+F}$. Usually, $D_{\mathcal{S}_\mathrm{X}}$ is a function of z, but it will be useful here to see it also as a function of F; we will denote by $\frac{\partial D_{\mathcal{S}_\mathrm{X}}}{\partial F}$ its derivative with respect to F and $D'_{\mathcal{S}_\mathrm{X}}(z)$ the derivative with respect to z. Using Theorem 3.11 and basic calculus, we get that $D_{\mathcal{S}_\mathrm{X}}$ is Δ-analytic and that

$$D_{\mathcal{S}_\mathrm{X}}(z) = D_{\mathcal{S}_\mathrm{X}}(\rho) - \gamma_X \sqrt{1 - z/\rho} + O\big(1 - z/\rho\big), \tag{13}$$

where:

- $D_{\mathcal{S}_\mathrm{X}}(\rho) = \frac{F(\rho)^2}{1+F(\rho)} = \frac{2-\sqrt{3}}{1+\sqrt{3}}$;
- $\gamma_X = \frac{\partial D_{\mathcal{S}_\mathrm{X}}}{\partial F}(F(\rho))\,\gamma_F = \left(1 - \frac{1}{(1+F(\rho))^2}\right)\gamma_F = \frac{\sqrt{6}}{2\sqrt{3+2\sqrt{3}}}\sqrt{\rho}$.

By singular differentiation [22, Theorem VI.8], we also have that $D'_{\mathcal{S}_\mathrm{X}}$ is Δ-analytic and that

$$D'_{\mathcal{S}_\mathrm{X}}(z) = \frac{\gamma_X}{2\rho}(1 - z/\rho)^{-1/2} + O\big(1\big). \tag{14}$$

Similarly we obtain that $D_{\mathcal{K}}$ is Δ-analytic and that

$$D_{\mathcal{K}}(z) = \frac{F(\rho)}{2(1+F(\rho))} - \frac{\rho}{2} - \gamma_K \sqrt{1 - z/\rho} + O\big(1 - z/\rho\big), \qquad (15)$$

where $\gamma_K = \frac{\partial D_{\mathcal{K}}}{\partial F}(F(\rho))\, \gamma_F = \frac{1}{2(1+F(\rho))^2}\, \gamma_F = \frac{1}{\sqrt{6+4\sqrt{3}}}\sqrt{\rho}$.

Summing, we have for D the following expansion:

$$D(z) = D(\rho) - \gamma_D \sqrt{1 - z/\rho} + O\big(1 - z/\rho\big), \qquad (16)$$

where $\gamma_D = \frac{2+\sqrt{3}}{\sqrt{6+4\sqrt{3}}}\sqrt{\rho}$.

4 DH-trees with a marked leaf

In this section, we introduce and analyze combinatorial classes of DH-trees with a marked leaf and certain conditions. This is a first step in the proof of Theorem 1.1 for unconstrained DH-graphs ($f = d$). Indeed, the classes studied here are building blocks in the decomposition of trees with several marked leaves, which we will consider in the next section in order to study distance matrices of uniform random DH-graphs.

4.1 A combinatorial system of equations for DH-trees with a marked leaf

Definition 4.1 *Let T be a DH-tree and v a vertex of T which is neither its root-leaf nor its root-node. Let p be the parent of v in T. Informally, the* cotype *of v is the type that p would have if v were the root-leaf or root-node. More precisely,*

- *if p is of type $\mathcal{K}$, then v is of cotype $\mathcal{K}$;*
- *if p is of type $\mathcal{S}_{\mathrm{C}}$, then v is of cotype $\mathcal{S}_{\mathrm{X}}$;*
- *if p is of type $\mathcal{S}_{\mathrm{X}}$ and v is the distinguished child of p, then v is of cotype $\mathcal{S}_{\mathrm{C}}$;*
- *if p is of type $\mathcal{S}_{\mathrm{X}}$ and v is not the distinguished child of p, then v is of cotype $\mathcal{S}_{\mathrm{X}}$.*

In most cases, v will be a leaf of T, but we will need occasionally to consider the cotype of internal vertices as well. The reader is invited to look at the example of Fig. 9.

Definition 4.2 *Let $a, b \in \{\mathcal{K}, \mathcal{S}_{\mathrm{C}}, \mathcal{S}_{\mathrm{X}}\}$. We define $\mathcal{D}_a^b$ as the set of DH-trees of type a with one marked leaf of cotype b.*

We further set, for $a, b \in \{\mathcal{K}, \mathcal{S}_{\mathrm{C}}, \mathcal{S}_{\mathrm{X}}\}$,

$$\mathcal{D}_a^{\bullet} = \mathcal{D}_a^{\mathcal{K}} + \mathcal{D}_a^{\mathcal{S}_{\mathrm{C}}} + \mathcal{D}_a^{\mathcal{S}_{\mathrm{X}}};$$
$$\mathcal{D}_{\bullet}^{b} = \mathcal{D}_{\mathcal{K}}^{b} + \mathcal{D}_{\mathcal{S}_{\mathrm{C}}}^{b} + \mathcal{D}_{\mathcal{S}_{\mathrm{X}}}^{b}.$$

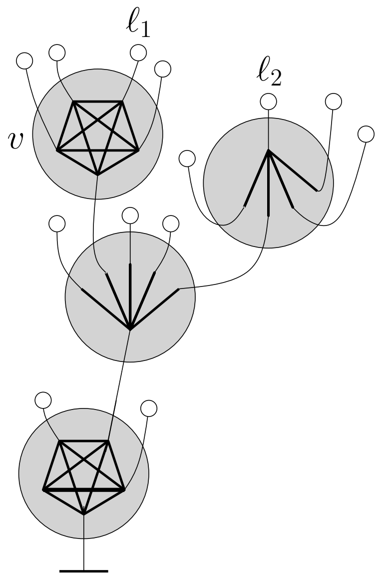

Figure 9: In this DH-tree, the leaf ℓ_1 has cotype $\mathcal{K}$, the leaf ℓ_2 has cotype $\mathcal{S}_\text{C}$ and the node v has cotype $\mathcal{S}_\text{X}$: its parent has type $\mathcal{S}_\text{C}$ but would have type $\mathcal{S}_\text{X}$ if v were the root-node (meaning that the root-leaf would be one of the four children of v).

In other words, a bullet as index (respectively exponent) denotes an unconstrained type of the root-node (respectively cotype of the marked leaf).

We now introduce the following statistics. Let (T, ℓ) be a DH-tree with one marked leaf ℓ. We denote by $\text{jp}(T, \ell)$ the number of jumps on the path from the marked leaf ℓ to the root-leaf in T (in particular, the root-node might be a jump in this path, see Fig. 10).

We consider (exponential) bivariate generating series of families of DH-trees with one marked leaf with respect to the size (variable z) and to the number of jumps (variable u). Namely, for $a, b \in \{\bullet, \mathcal{K}, \mathcal{S}_\text{C}, \mathcal{S}_\text{X}\}$,

$$D_a^b(z,u) = \sum_{(T,\ell)\in\mathcal{D}_a^b} \frac{z^{|(T,\ell)|}}{|(T,\ell)|!} u^{\text{jp}(T,\ell)}.$$

We take the convention that for a DH-tree with a marked leaf (T, ℓ), its size $|(T, \ell)|$ is the number of unmarked non-root leaves of (T, ℓ) (in other words, the root-leaf and the marked leaf are not counted).

Proposition 4.3 *The bivariate series $D_a^b(z,u)$ for $a, b \in \{\mathcal{K}, \mathcal{S}_\text{C}, \mathcal{S}_\text{X}\}$ are solutions* *of the following systems of equations:*

$$\begin{cases} D_{\mathcal{K}}^{\mathcal{K}} = (1 + D_{\mathcal{S}_\mathrm{C}}^{\mathcal{K}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{K}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}} + z); \\ D_{\mathcal{S}_\mathrm{X}}^{\mathcal{K}} = (D_{\mathcal{S}_\mathrm{C}}^{\mathcal{K}} + D_{\mathcal{K}}^{\mathcal{K}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z) \\ \qquad\qquad +u \cdot (D_{\mathcal{K}}^{\mathcal{K}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{K}}) (D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z) \exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z); \\ D_{\mathcal{S}_\mathrm{C}}^{\mathcal{K}} = (D_{\mathcal{K}}^{\mathcal{K}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{K}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z); \end{cases} \tag{17}$$

$$\begin{cases} D_{\mathcal{K}}^{\mathcal{S}_\mathrm{X}} = (D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{X}} + D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{X}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}} + z); \\ D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{X}} = (D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}}^{\mathcal{S}_\mathrm{X}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z) \\ \qquad\qquad +u \cdot (1 + D_{\mathcal{K}}^{\mathcal{S}_\mathrm{X}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{X}}) (D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z) \exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z); \\ D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{X}} = (1 + D_{\mathcal{K}}^{\mathcal{S}_\mathrm{X}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{X}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z); \end{cases} \tag{18}$$

$$\begin{cases} D_{\mathcal{K}}^{\mathcal{S}_\mathrm{C}} = (D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{C}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}} + z); \\ D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{C}} = (1 + D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}}^{\mathcal{S}_\mathrm{C}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z) \\ \qquad\qquad +u \cdot (D_{\mathcal{K}}^{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{C}}) (D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z) \exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z); \\ D_{\mathcal{S}_\mathrm{C}}^{\mathcal{S}_\mathrm{C}} = (D_{\mathcal{K}}^{\mathcal{S}_\mathrm{C}} + D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{C}}) \exp_{\geq 1}(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z). \end{cases} \tag{19}$$

Proof. We prove in details the case of $D_{\mathcal{S}_\mathrm{X}}^{\mathcal{S}_\mathrm{C}}$ (second equation in the system (19)). The eight other equations are proved in a similar way.

Hence we consider a DH-tree T of type $\mathcal{S}_\mathrm{X}$ with a marked leaf of cotype $\mathcal{S}_\mathrm{C}$. We can decompose T as a root-node v, to which several subtrees are attached. (Notations are summarized in Fig. 10.) The subtrees attached to v are:

- The subtree T' containing the marked leaf. In order to keep track of variable u we need to consider two cases.
  - First, T' may be attached to the center of Γ_v (Case A of Fig. 10). In this case, there is no jump in v. Also, T' (if not reduced to a leaf) is of type $\mathcal{K}$ or $\mathcal{S}_\mathrm{C}$. Note that T' may also be reduced to a leaf (hence, the marked leaf), since a leaf attached to the center of Γ_v has indeed cotype $\mathcal{S}_\mathrm{C}$.
  - Otherwise, T' is attached to an extremity of Γ_v (Case B of Fig. 10). In this case, there is a jump in v. Here, T' can be of type $\mathcal{K}$ or $\mathcal{S}_\mathrm{X}$, and T' cannot be reduced to a leaf since a leaf attached to an extremity of Γ_v would have cotype $\mathcal{S}_\mathrm{X}$.
- Attached to every (other) extremity of v one has a tree of type $\mathcal{K}$ or $\mathcal{S}_\mathrm{X}$ or a leaf.
- Attached to the center of v (if this is not where T' is attached, *i.e.* in Case B in Fig. 10) there is a tree of type $\mathcal{K}$ or $\mathcal{S}_\mathrm{C}$ or a leaf.

We now translate this decomposition on generating functions. According to the case analysis above, T' is counted by:

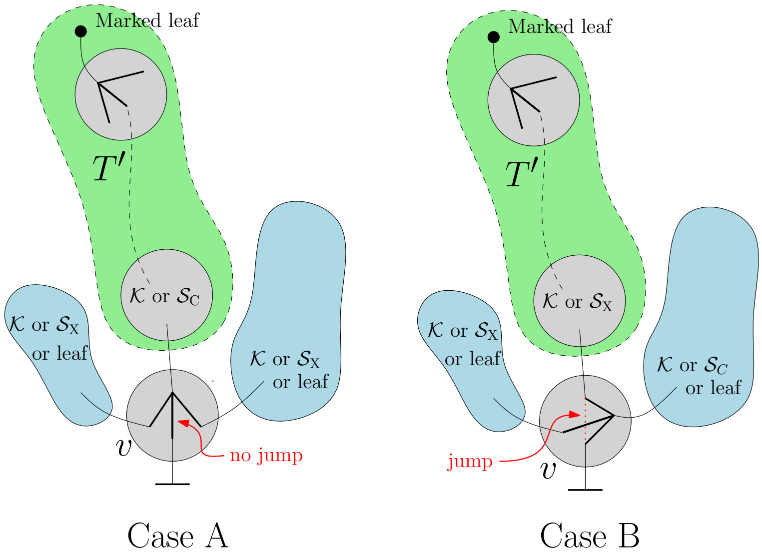

Figure 10: Decomposition of a tree in $\mathcal{D}_{\mathcal{S}_{\mathrm{X}}}^{\mathcal{S}_{\mathrm{C}}}$.

- in Case A: $1 + D_{\mathcal{S}_{\mathrm{C}}}^{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}}^{\mathcal{S}_{\mathrm{C}}}$ (since a single marked leaf, counted by 1, is allowed for T');
- in Case B: $D_{\mathcal{K}}^{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{S}_{\mathrm{X}}}^{\mathcal{S}_{\mathrm{C}}}$.

The remaining trees attached to v are counted by:

- in Case A: $\exp_{\geq 1}(D_{\mathcal{S}_{\mathrm{X}}} + D_{\mathcal{K}} + z)$ (because they form a nonempty unordered sequence of trees in $\mathcal{D}_{\mathcal{S}_{\mathrm{X}}} + \mathcal{D}_{\mathcal{K}} + \mathcal{Z}$)
- in Case B: $(D_{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}} + z)\exp(D_{\mathcal{S}_{\mathrm{X}}} + D_{\mathcal{K}} + z)$ (with a distinguished tree attached to the center which is either in $\mathcal{D}_{\mathcal{S}_{\mathrm{C}}}$ or in $\mathcal{D}_{\mathcal{K}}$ or a leaf, and other trees which form an unordered sequence of trees in $\mathcal{D}_{\mathcal{S}_{\mathrm{X}}} + \mathcal{D}_{\mathcal{K}} + \mathcal{Z}$).

Finally, a factor u appears in Case B to take into account the jump in v. Hence

$$D_{\mathcal{S}_{\mathrm{X}}}^{\mathcal{S}_{\mathrm{C}}} = (1 + D_{\mathcal{S}_{\mathrm{C}}}^{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}}^{\mathcal{S}_{\mathrm{C}}}) \exp_{\geq 1}(D_{\mathcal{S}_{\mathrm{X}}} + D_{\mathcal{K}} + z) \\ + u \cdot (D_{\mathcal{K}}^{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{S}_{\mathrm{X}}}^{\mathcal{S}_{\mathrm{C}}})(D_{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}} + z) \exp(D_{\mathcal{S}_{\mathrm{X}}} + D_{\mathcal{K}} + z).$$

□

4.2 Resolution of the system

Recall (see Section 3.4) that

$$D_{\mathcal{S}_{\mathrm{C}}} = D_{\mathcal{K}}, \quad \exp_{\geq 1}(D_{\mathcal{K}} + D_{\mathcal{S}_{\mathrm{X}}} + z) = F \text{ and } D_{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}} + z = 2D_{\mathcal{K}} + z = \tfrac{F}{1+F}.$$

This implies $(D_{\mathcal{S}_{\mathrm{C}}} + D_{\mathcal{K}} + z)\exp(D_{\mathcal{S}_{\mathrm{X}}} + D_{\mathcal{K}} + z) = F$, allowing simplification of the systems (17) to (19) as follows:

$$\begin{cases} D^{\mathcal{K}}_{\mathcal{K}} = (1 + D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{C}}} + D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{X}}})\, F; \\ D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{X}}} = (D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{C}}} + D^{\mathcal{K}}_{\mathcal{K}})\, F + u\, F \cdot (D^{\mathcal{K}}_{\mathcal{K}} + D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{X}}}); \\ D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{C}}} = (D^{\mathcal{K}}_{\mathcal{K}} + D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{X}}})\, F; \end{cases} \tag{20}$$

$$\begin{cases} D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{K}} = (D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{X}}} + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{C}}})\, F; \\ D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{X}}} = (D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{C}}} + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{K}})\, F + u\, F \cdot (1 + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{K}} + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{X}}}); \\ D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{C}}} = (1 + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{K}} + D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{X}}})\, F; \end{cases} \tag{21}$$

$$\begin{cases} D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{K}} = (D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{C}}} + D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{X}}})\, F; \\ D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{X}}} = (1 + D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{C}}} + D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{K}})\, F + u\, F \cdot (D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{K}} + D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{X}}}); \\ D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{C}}} = (D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{K}} + D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{X}}})\, F. \end{cases} \tag{22}$$

Solving the system[7] gives the following formulas (put under a suitable form for the subsequent asymptotic analysis):

$$D^{\mathcal{K}}_{\mathcal{K}} = \frac{F}{F+1} + \frac{F^2}{(1+F)(1-2F) - Fu}; \tag{23}$$

$$D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{X}}} = \frac{-1}{F+1} + \frac{(1-F)^2}{(1+F)(1-2F) - Fu}; \tag{24}$$

$$D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{C}}} = \frac{F^2}{(1+F)(1-2F) - Fu}; \tag{25}$$

$$D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{K}} = D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{X}}} = \frac{-F}{F+1} + \frac{F\,(1-F)}{(1+F)(1-2F) - Fu}; \tag{26}$$

$$D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{K}} = D^{\mathcal{K}}_{\mathcal{S}_{\mathrm{C}}} = \frac{F^2}{(1+F)(1-2F) - Fu}; \tag{27}$$

$$D^{\mathcal{S}_{\mathrm{C}}}_{\mathcal{S}_{\mathrm{X}}} = D^{\mathcal{S}_{\mathrm{X}}}_{\mathcal{S}_{\mathrm{C}}} = \frac{F(1-F)}{(1+F)(1-2F) - Fu}. \tag{28}$$

Remark 4.4 Symmetries $D^b_a = D^a_b$ in above equations can easily be explained combinatorially. Indeed, we can see a DH-tree with root-leaf r and a marked leaf ℓ as a DH-tree rooted in ℓ where r is a marked leaf; doing so, the type of the (old) root becomes the cotype of r and the cotype of ℓ becomes the type of the (new) root.

Recalling that F depends only on z (not on u), in each case, the series can be written under the form

$$D^b_a = Q^b_a(z) + \frac{M^b_a(z)}{1 - u\, H^b_a(z)}, \tag{29}$$

[7] see Maple worksheet.

where Q_a^b, M_a^b and H_a^b are rational functions in F. For example, looking at Eq. (23), we have

$$Q_{\mathcal{K}}^{\mathcal{K}} = \frac{F}{F+1}, \ M_{\mathcal{K}}^{\mathcal{K}} = \frac{F^2}{(1+F)(1-2F)} \text{ and } H_{\mathcal{K}}^{\mathcal{K}} = \frac{F}{(1+F)(1-2F)}.$$

Similar formulas are easily written for other $a, b \in \{\mathcal{K}, \mathcal{S}_{\mathrm{C}}, \mathcal{S}_{\mathrm{X}}\}$, looking at Eqs. (24) to (28). Such expressions are keys for applying the semi-large powers theorem (see [22, Theorem IX.16], or Appendix A) as we will see in Section 5.

Interestingly (and we shall later use these remarks), $H_a^b = H = \frac{F}{(1+F)(1-2F)}$ is the same for all $a, b \in \{\mathcal{K}, \mathcal{S}_{\mathrm{C}}, \mathcal{S}_{\mathrm{X}}\}$ and M_a^b factorizes as $M_a^b = \Lambda_a \Lambda_b$, where

$$\Lambda_{\mathcal{K}} = \Lambda_{\mathcal{S}_{\mathrm{C}}} = \frac{F}{\sqrt{(1+F)(1-2F)}}; \quad \Lambda_{\mathcal{S}_{\mathrm{X}}} = \frac{1-F}{\sqrt{(1+F)(1-2F)}}.$$

Recall from Section 3.4 (see in particular Eq. (12)), that the radius of convergence ρ of F satisfies $1 - 2F(\rho) = 2F(\rho)^2$. Consequently, $(1 + F(z))(1 - 2F(z)) \notin (-\infty, 0)$ for z close to ρ. Hence, the definitions of $\Lambda_{\mathcal{K}}, \Lambda_{\mathcal{S}_{\mathrm{C}}}$ and $\Lambda_{\mathcal{S}_{\mathrm{X}}}$ make sense near ρ; we shall only use them in this domain.

Moreover all these formulas immediately extend to the case where a or b or both is/are equal to $\bullet$ (unconstrained type of the root-node or cotype of the marked leaf), with the natural convention that

$$\begin{aligned} H_\bullet^b &= H_a^\bullet = H; \\ M_\bullet^b &= M_{\mathcal{K}}^b + M_{\mathcal{S}_{\mathrm{C}}}^b + M_{\mathcal{S}_{\mathrm{X}}}^b, \\ \Lambda_\bullet &= \Lambda_{\mathcal{K}} + \Lambda_{\mathcal{S}_{\mathrm{C}}} + \Lambda_{\mathcal{S}_{\mathrm{X}}}; \end{aligned}$$

and conventions similar to the second line for $M_a^\bullet$, $Q_a^\bullet$ and $Q_\bullet^b$.

Since F has nonnegative coefficients and $2F(\rho) = \sqrt{3} - 1 < 1$, the denominators of $Q_a^b(z)$, $M_a^b(z)$, $\Lambda_a^b(z)$ and $H(z)$ are positive for z in $[0, \rho]$ and thus, the series Q_a^b, M_a^b, $\Lambda_a^b(z)$ and H all have the same radius of convergence ρ as F, and a square-root singularity in ρ, inherited from that of F. Later (in the proof of Proposition 5.7) the function H will play a particular role in the asymptotic analysis so let us now compute its expansion at ρ:

$$H(z) = H(\rho) - \gamma_H \sqrt{1 - z/\rho} + O(1 - z/\rho),$$

where

$$H(\rho) = \frac{F(\rho)}{(1+F(\rho))(1-2F(\rho))} = 1, \tag{30}$$

$$\begin{aligned} \gamma_H = \frac{\partial H}{\partial F}(F(\rho))\,\gamma_F &= \left(\frac{1}{3(1+F(\rho))^2} + \frac{2}{3(1-2F(\rho))^2} \right) \gamma_F \\ &= \frac{2 \cdot (3 - \sqrt{3})}{\sqrt{6 + 4\sqrt{3}}\,(2 - \sqrt{3})^2} \sqrt{\rho} \end{aligned} \tag{31}$$

whose numerical estimate is $\gamma_H \approx 3.9258$ (see Maple worksheet).

5 k-point distances and induced subtrees

The goal of this section is to obtain the joint convergence in distribution of distances between marked leaves in a uniform DH-tree (see Theorem 5.8 below). This allows us to complete the proof of Theorem 1.1 in the case of unconstrained DH-graphs ($f = d$).

5.1 Marked leaves and induced subtrees

In this section, we consider DH-trees with k marked leaves $(\ell_1, \dots, \ell_k)$ with the following convention.

Definition 5.1 *A DH-tree of size n with k marked leaves is a nonplane tree T such that*

- *T is rooted at a leaf labeled 0;*
- *T has n nonmarked non-root leaves labeled $1, \dots, n$;*
- *additionally, T has k ordered non-root leaves carrying marks $(\ell_1, \dots, \ell_k)$;*
- *items ii) to v) p. 282 are satisfied.*

Equivalently, it is a DH-tree of size $n + k$ where leaves with labels $n + 1, \dots, n + k$ are seen as marked and get marks $\ell_1, \dots, \ell_k$, respectively. These marked leaves are not counted in the size. With this convention, the exponential generating series of DH-trees with k marked leaves is $D^{(k)}(z)$ (the k-th derivative of D). Theorem 3.3 is immediately rephrased as follows.

Proposition 5.2 *Labeled DH-graphs of size $n+k+1$ are in bijection with DH-trees of size n with k marked leaves (when $n + k + 1 \geq 3$).*

To simplify notation, we write $\boldsymbol{\ell} = (\ell_1, \dots, \ell_k)$.

We recall the definition of induced subtrees.

Definition 5.3 *Let T be a DH-tree with k marked leaves $\boldsymbol{\ell}$. We call* essential vertices *of T (with respect to the marked leaves $\boldsymbol{\ell}$) its root-leaf, its k marked leaves $\boldsymbol{\ell}$ and their pairwise first common ancestors. Then, the subtree of T induced by $\boldsymbol{\ell}$ is obtained as follows:*

- *its vertices are the essential vertices of T;*
- *its genealogy (ancestor/descendant relation) is inherited from that of T.*

Fig. 11 illustrates this definition. We remark that the subtree of T induced by k marked leaves $\boldsymbol{\ell}$ is naturally rooted at the vertex corresponding to the root-leaf of T. This vertex is always of degree 1, and will be called root-leaf of the induced subtree.

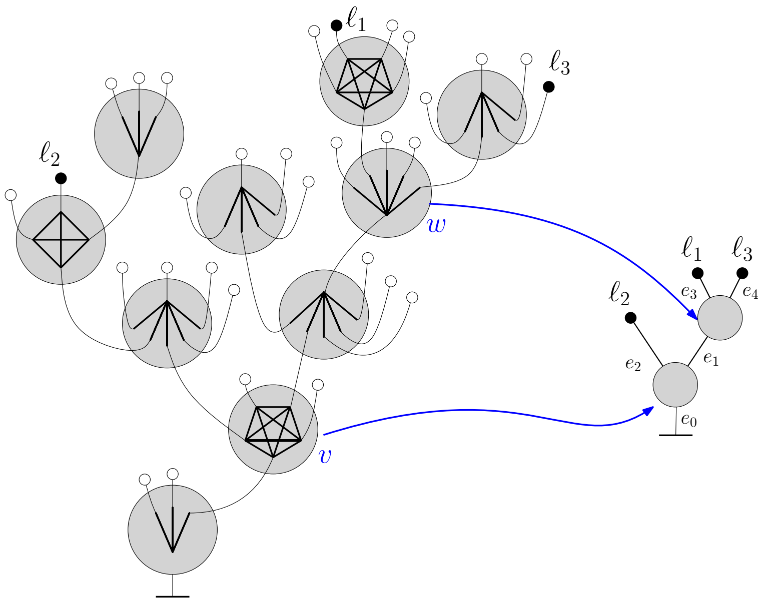

Figure 11: Left: A DH-tree T of size $n = 28$ with $k = 3$ marked leaves. The nodes v and w are the first common ancestors of ℓ_1, ℓ_2 and ℓ_3. Right: The subtree t_0 of T induced by (ℓ_1, ℓ_2, ℓ_3). We have highlighted the correspondence between first common ancestors in T and internal vertices of the induced subtree t_0. The enriched subtree of $(T, (\ell_1, \ell_2, \ell_3))$ for the ordering of the edges shown on the right is $(t_0, 0, 0, 1, 0, 1)$.

We now enrich the notion of induced subtrees to record the number of jumps along some paths of T. Consider a DH-tree T with k marked leaves $\boldsymbol{\ell}$. Let t_0 be the associated induced subtree. Each edge e in t_0 corresponds to a path between two consecutive essential vertices of T. We define $\mathrm{jp}_e(T; \boldsymbol{\ell})$ as the number of jumps of the path corresponding to e, with the convention that essential vertices are not counted as jumps (but note that the root-node of T can be a jump). We call *enriched induced subtree* of $(T, \boldsymbol{\ell})$ the induced subtree t_0, with the quantities $\mathrm{jp}_e(T; \boldsymbol{\ell})$ attached to its edges. It will be convenient to fix for each tree t with k leaves an enumeration $(e_0, e_1, \dots)$ of its edges such that e_0 is the edge adjacent to the root-leaf (for instance a breath-first traversal of the tree with an arbitrary planar embedding). Then the enriched induced subtree of $(T, \boldsymbol{\ell})$ can be written as a tuple $(t_0, a_0, \dots, a_m)$, where t_0 is the induced subtree of $(T, \boldsymbol{\ell})$ and $a_i = \mathrm{jp}_{e_i}(T; \boldsymbol{\ell})$. In the following we denote $r(T, \boldsymbol{\ell}) := (t_0, a_0, \dots, a_m)$.

We recall from Theorem 3.4 that the distances in a DH-graph are closely related to the jumps in the associated DH-trees. Hence understanding the enriched subtree induced by uniform random leaves in a uniform random DH-tree is a crucial step to understand the asymptotic behaviour of the distance matrices of DH-graphs. We now study these enriched subtrees via combinatorial decompositions and analytic combinatorics.

5.2 Combinatorial decomposition

Recall that a k-proper tree is an (unrooted) nonplane tree with $k+1$ leaves where each internal node has degree 3 (with one leaf considered as the root-leaf and the other leaves denoted $\{\ell_1, \dots, \ell_k\}$). It is easily observed that a k-proper tree has $k+1$ leaves, $k-1$ internal vertices and $2k-1$ edges. It is also a standard fact (see, *e.g.*, [2]) that the cardinality of the set of k-proper trees is exactly $(2k-3)!!$, where $(2k-3)!!$ is the product of all odd positive integers less than or equal to $2k-3$. Indeed, a k-proper tree can be obtained in a unique way from a $(k-1)$-proper tree by selecting one of its $2k-3$ edges and grafting in the middle a new edge with a leaf ℓ_k at its extremity.

Let us fix a k-proper tree t_0. We consider the following class of marked DH-trees.

Definition 5.4 *We let $\mathcal{D}_{t_0}$ be the labeled combinatorial class of DH-trees $(T, \boldsymbol{\ell})$ with k marked leaves such that:*

i) the subtree of T induced by $\boldsymbol{\ell}$ is t_0;

ii) no two essential vertices of T are neighbors of each other.

Item ii) is a technical condition to have a nicer combinatorial decomposition in Eq. (33) below.

Recall that we have fixed an enumeration $(e_0, e_1, \dots, e_{2k-2})$ of the $(2k-1)$ edges of our k-proper tree t_0, in which e_0 is the edge attached to the root-leaf of t_0. Consider the following multivariate generating series for $\mathcal{D}_{t_0}$:

$$D_{t_0}(z, u_0, \dots, u_{2k-2}) = \sum_{(T;\boldsymbol{\ell}) \in \mathcal{D}_{t_0}} \frac{z^{|(T,\boldsymbol{\ell})|}}{|(T,\boldsymbol{\ell})|!} u_0^{\mathrm{jp}_{e_0}(T,\boldsymbol{\ell})} \cdots u_{2k-2}^{\mathrm{jp}_{e_{2k-2}}(T,\boldsymbol{\ell})},$$

where $|(T, \boldsymbol{\ell})|$ is the number of non-root unmarked leaves of $(T, \boldsymbol{\ell})$.

In order to compute the series D_{t_0} we introduce the following new classes of DH-trees. For $a, b, c \in \{\bullet, \mathcal{K}, \mathcal{S}_{\mathrm{X}}, \mathcal{S}_{\mathrm{C}}\}$, let $\mathcal{J}_a^{b\,c}$ be the set of DH-trees T with two (ordered) marked leaves such that

- the two marked leaves are children of the root-node;
- if T_1 is a DH-tree of type b, one can glue T_1 on the first marked leaf of T (merging the marked leaf and the root-node of T_1) without violating the adjacency restrictions defining DH-trees (conditions iii) to v) p. 282);
- the same condition holds with gluing a DH-tree of type c on the second marked leaf;
- additionally, if T_0 is a DH-tree with a marked leaf of cotype a, one can glue T on the marked leaf of T_0 without violating the adjacency restrictions defining DH-trees.

The generating function of $\mathcal{J}_a^{b\,c}$ will be denoted by $J_a^{b\,c}(z)$.

Lemma 5.5 *For $a, b, c \in \{\bullet, \mathcal{K}, \mathcal{S}_\mathrm{X}, \mathcal{S}_\mathrm{C}\}$, it holds that*

$$\begin{aligned} J_a^{b\,c}(z) &= (\mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C)\exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z) + \mathbb{1}_{\mathcal{K}\notin\{a,b,c\}}\exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{S}_\mathrm{C}} + z) \\ &\quad + \mathbb{1}_{\mathcal{S}_\mathrm{C}\notin\{a,b,c\}}(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z)\exp(D_{\mathcal{K}} + D_{\mathcal{S}_\mathrm{X}} + z) \\ &= (\mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C + \mathbb{1}_{\mathcal{K}\notin\{a,b,c\}})(1+F) + \mathbb{1}_{\mathcal{S}_\mathrm{C}\notin\{a,b,c\}}F, \end{aligned} \tag{32}$$

where

$$A = \{(a,b,c) \mid a \neq \mathcal{S}_\mathrm{X}, b \neq \mathcal{S}_\mathrm{C}, c \neq \mathcal{S}_\mathrm{C}\},$$
$$B = \{(a,b,c) \mid b \neq \mathcal{S}_\mathrm{X}, a \neq \mathcal{S}_\mathrm{C}, c \neq \mathcal{S}_\mathrm{C}\},$$
$$C = \{(a,b,c) \mid c \neq \mathcal{S}_\mathrm{X}, a \neq \mathcal{S}_\mathrm{C}, b \neq \mathcal{S}_\mathrm{C}\}.$$

Proof. We consider different cases depending on the type of the root-node. The trees of $\mathcal{J}_a^{b\,c}$ having a root-node of type $\mathcal{K}$ are counted by $\mathbb{1}_{\mathcal{K}\notin\{a,b,c\}}\exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{S}_\mathrm{C}} + z)$. The ones having a root-node of type $\mathcal{S}_\mathrm{C}$ are counted by $\mathbb{1}_A \exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z)$. Finally the ones having a root-node of type $\mathcal{S}_\mathrm{X}$ are counted by $(\mathbb{1}_B + \mathbb{1}_C)\exp(D_{\mathcal{S}_\mathrm{X}} + D_{\mathcal{K}} + z) + \mathbb{1}_{\mathcal{S}_\mathrm{C}\notin\{a,b,c\}}(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z)\exp(D_{\mathcal{K}} + D_{\mathcal{S}_\mathrm{X}} + z)$ since the center of the star may be connected to the first marked leaf, to the second marked leaf or to neither of them.

To conclude the proof, we use that $D_{\mathcal{K}} = D_{\mathcal{S}_\mathrm{C}}$, and Eqs. (7) and (8). ☐

For $0 \leq i \leq 2k-2$ let v_i (respectively w_i) be the vertex incident to e_i in t_0 closest to (respectively farthest from) the root-leaf of t_0. In particular, some v_i's are equal to each other, v_0 is the root-leaf and some w_i are leaves (see Fig. 12, right).

If w_i is not a leaf, let s_i (respectively g_i) be the smallest (respectively greatest) index of the edges from w_i to its (two) children.

Proposition 5.6 *We have*

$$D_{t_0}(z, u_0, \ldots, u_{2k-2}) = \sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{C}} \prod_{i=0}^{2k-2} D_{tp_i}^{ct_i}(z, u_i) \prod_{\substack{i \\ w_i \text{ is an internal node}}} J_{ct_i}^{tp_{s_i}\,tp_{g_i}}(z) \tag{33}$$

where

$$\mathfrak{C} = \{(\mathbf{tp},\mathbf{ct}) = (tp_i, ct_i)_{0\leq i\leq 2k-2} \mid tp_i, ct_i \in \{\bullet, \mathcal{K}, \mathcal{S}_\mathrm{X}, \mathcal{S}_\mathrm{C}\} \text{ and } tp_i = \bullet \text{ iff } i = 0 \text{ and } ct_i = \bullet \text{ iff } w_i \text{ is a leaf}\}.$$

Proof. We shall build a size-preserving bijection from $\mathcal{D}_{t_0}$ to the disjoint union

$$\biguplus_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{C}} \prod_{i=0}^{2k-2} \mathcal{D}_{tp_i}^{ct_i} \prod \mathcal{J}_{ct_i}^{tp_{s_i}\,tp_{g_i}}.$$

Let $(T, \boldsymbol{\ell}) \in \mathcal{D}_{t_0}$. Then T is a DH-tree. Let $\bar{v}_i$ (respectively $\bar{w}_i$) be the essential vertex of T corresponding to v_i (respectively w_i). We set $tp_0 = \bullet$, and $ct_i = \bullet$ when

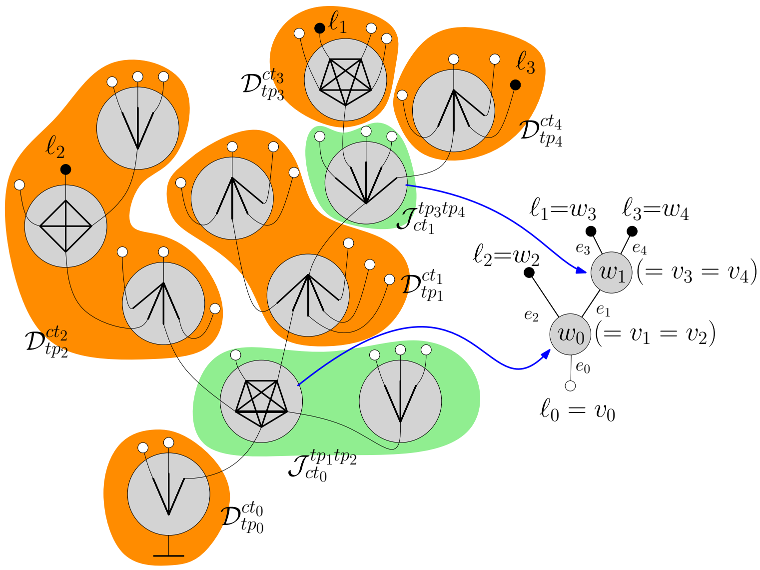

Figure 12: Decomposition of a DH-tree in $\mathcal{D}_{t_0}$. The groups of nodes indicated by orange and green areas correspond to the pieces defined in the proof of Theorem 5.6.

w_i is a leaf of t_0. Otherwise, we denote by ct_i the cotype of $\bar{w}_i$, and by tp_i the type of the child of $\bar{v}_i$ which is the root of the subtree containing $\bar{w}_i$ (this type is well-defined thanks to item ii) of Theorem 5.4). We then have $(tp_i, ct_i)_{0 \le i \le 2k-2} \in \mathfrak{E}$. For example with $(T, \boldsymbol{\ell})$ given in Fig. 12,

$$tp_0 = \bullet,\ tp_1 = tp_2 = tp_4 = \mathcal{S}_{\mathrm{X}},\ tp_3 = \mathcal{K},\ ct_0 = \mathcal{S}_{\mathrm{X}},\ ct_1 = \mathcal{S}_{\mathrm{C}},\ ct_2 = ct_3 = ct_4 = \bullet.$$

We decompose $(T, \boldsymbol{\ell})$ as follows. For each i such that w_i is an internal node of t_0, we cut the parent edge from $\bar{w}_i$, as well as the two edges incident to $\bar{w}_i$ which are the start of a path going to a marked leaf (since t_0 is a k-proper tree, there are always exactly two such edges). This operation turns $(T, \boldsymbol{\ell})$ into a disjoint union of trees, which we call pieces. Each edge that is cut is replaced by a marked leaf (in the piece closer to the root of T) and a root-leaf (in the piece further away from the root of T). Then the piece containing $\bar{w}_i$ belongs to $\mathcal{J}_{ct_i}^{tp_{s_i} tp_{g_i}}$ (for every internal node w_i). Moreover, the pieces containing none of the $\bar{w}_i$ are in bijection with the edges of t_0, and the piece corresponding to e_i belongs to $\mathcal{D}_{tp_i}^{ct_i}$.

By decomposing $(T, \boldsymbol{\ell})$ we have indeed obtained an element of

$$\biguplus_{(\mathbf{tp}, \mathbf{ct}) \in \mathfrak{E}} \prod_{i=0}^{2k-2} \mathcal{D}_{tp_i}^{ct_i} \prod \mathcal{J}_{ct_i}^{tp_{s_i} tp_{g_i}}.$$

Conversely, let $(tp_i, ct_i)_{0 \le i \le 2k-2} \in \mathfrak{E}$ and take a tuple

$$(T_{e_i})_{0 \le i \le 2k-2} \times (T_{w_i})_{w_i \text{ internal node}} \in \prod_{i=0}^{2k-2} \mathcal{D}_{tp_i}^{ct_i} \prod \mathcal{J}_{ct_i}^{tp_{s_i} tp_{g_i}}.$$

From these trees, we build a tree uniquely as follows. For every internal node w_i of t_0, we glue the root-leaf of T_{w_i} to the marked leaf of T_{e_j}, where e_j is the edge from w_i to its parent (when gluing, the two edges from the root-leaf and from the marked leaf become one edge, and the leaves disappear). Moreover, we glue the first marked leaf of T_{w_i} to the root-leaf of $T_{e_{s_i}}$ and we glue the second marked leaf of T_{w_i} to the root-leaf of $T_{e_{g_i}}$ (recall that s_i (respectively g_i) is the smallest (respectively greatest) index of the edges from w_i to its children).

Since t_0 is a k-proper tree, once these gluings are done, we obtain only one tree T, with one root-leaf (the one of T_{e_0}) and k marked leaves (those of T_{e_i} where e_i is incident to a leaf of t_0) which are in one-to-one correspondence with the leaves of t_0. By construction (recalling also the definition of $\mathcal{J}_a^{b\,c}$), T is a DH-tree, whose k marked leaves induce t_0, and which satisfies item ii) of Theorem 5.4 (since elements of $\mathcal{D}_a^b$ are DH-trees thus have one or more internal node(s)). All together, we have $T \in \mathcal{D}_{t_0}$.

Finally, we have a size-preserving bijection, since the size of T is the sum of the sizes of the T_{e_i} and of the T_{w_i}. Indeed, for $\mathcal{D}_{t_0}$, $\mathcal{D}_a^b$ and $\mathcal{J}_a^{b\,c}$, the root-leaf and marked leaves are not counted in the size, and the leaves which have disappeared when gluing are all marked leaves or root-leaves. □

5.3 Asymptotic analysis

Recall the notation $r(T, \boldsymbol{\ell})$ from the end of Section 5.1, denoting the enriched induced subtree of $(T, \boldsymbol{\ell})$.

Proposition 5.7 *Let $(\boldsymbol{T}_n, \boldsymbol{\ell})$ be a uniform random DH-tree of size n with k marked leaves (not counted in the size). Fix a k-proper tree t_0 and real numbers $x_0, \dots, x_{2k-2} > 0$. We set $a_i = \lfloor x_i \sqrt{n} \rfloor$. Then*

$$\mathbb{P}\big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \dots, a_{2k-2})\big] \sim \frac{\gamma_H^{2k}}{2^k \sqrt{n}^{2k-1}} s \exp\left(\frac{-\gamma_H^2 s^2}{4}\right), \tag{34}$$

where $s = \sum_i x_i$. Moreover, this estimate is uniform for $(x_0, \dots, x_{2k-2})$ in any compact subset of $(0, +\infty)^{2k-1}$.

Proof. We first note that, for n large enough,

$$r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \dots, a_{2k-2})$$

implies that $(\boldsymbol{T}_n, \boldsymbol{\ell})$ is in D_{t_0}. Indeed, item i) of Theorem 5.4 comes from the definition of t_0; item ii) follows from the fact that for every $i > 0$, we have $a_i > 0$ (for n large enough): so, there must be some jumps between each pair of essential vertices, and thus they cannot be neighbors.

Therefore, writing $p(n) := \mathbb{P}\big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \dots, a_{2k-2})\big]$, we have, for n large enough,

$$p(n) = \frac{[z^n u_0^{a_0} \dots u_{2k-2}^{a_{2k-2}}]\ D_{t_0}(z, u_0, \dots, u_{2k-2})}{[z^n] D^{(k)}(z)}. \tag{35}$$

We first analyze the denominator. From Eq. (16) and singular differentiation, we have

$$D^{(k)}(z) = \frac{1\cdot 3\cdots(2k-3)}{2^k}\gamma_D \rho^{-k}(1-z/\rho)^{1/2-k} + \mathcal{O}\big((1-z/\rho)^{-k}\big).$$

Applying the transfer theorem then yields

$$[z^n]D^{(k)}(z) \sim \frac{1\cdot 3\cdots(2k-3)}{2^k}\gamma_D \rho^{-k-n}\frac{n^{k-3/2}}{\Gamma(k-1/2)}.$$

Since $\Gamma(k-1/2) = (k-3/2)(k-5/2)\cdots 1/2\cdot\Gamma(1/2) = \frac{1\cdot 3\cdots(2k-3)}{2^{k-1}}\sqrt{\pi}$, we get

$$[z^n]D^{(k)}(z) \sim \frac{\gamma_D}{2\sqrt{\pi}}\rho^{-k-n}n^{k-3/2}. \tag{36}$$

Consider now the numerator of Eq. (35), i.e. the coefficient of $z^n u_0^{a_0}\dots u_{2k-2}^{a_{2k-2}}$ in the series $D_{t_0}(z,u_0,\dots,u_{2k-2})$. The exponents a_i are all positive since we assumed $x_0,\dots,x_{2k-2}>0$. We start from Eq. (33) and use that from Eq. (29) all D_a^b are of the form

$$D_a^b = Q_a^b(z) + \frac{M_a^b(z)}{1-u\,H(z)}.$$

The first term $Q_a^b(z)$ is independent of u. Therefore, when expanding the product $\prod_{i=0}^{2k-2} D_{tp_i}^{ct_i}(z,u_i)$ in Eq. (33), we can forget the terms $Q_{tp_i}^{ct_i}(z)$ without changing the coefficient of $z^n u_0^{a_0}\dots u_{2k-2}^{a_{2k-2}}$. Also, clearly, $[u_i^{a_i}]\frac{1}{1-u_i\,H(z)} = H(z)^{a_i}$. We therefore get

$$[z^n u_0^{a_0}\dots u_{2k-2}^{a_{2k-2}}]\; D_{t_0}(z,u_0,\dots,u_{2k-2}) = \sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{C}} [z^n]\,\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(z)H(z)^{a_0+\dots+a_{2k-2}}, \tag{37}$$

where

$$\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(z) = \prod_{i=0}^{2k-2} M_{tp_i}^{ct_i}(z) \prod_{\substack{i\\ w_i \text{ is an internal node}}} J_{ct_i}^{tp_{s_i}tp_{g_i}}(z). \tag{38}$$

We apply[8] the Semi-large powers Theorem (see Theorem A.1 p.313) with $s = \sum_{i=0}^{2k-2} x_i$. Using that $H(\rho)=1$ (see Eq. (30)), we have

$$[z^n]\,\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(z)H(z)^{a_0+\dots+a_{2k-2}} \sim \frac{s\gamma_H}{2}\exp\left(\frac{-s^2\,\gamma_H^2}{4}\right)\frac{1}{n\rho^n\sqrt{\pi}}\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho),$$

where we recall that γ_H is given by Eq. (31). Therefore we have

$$[z^n u_0^{a_0}\dots u_{2k-2}^{a_{2k-2}}]\; D_{t_0}(z,u_0,\dots,u_{2k-2}) \sim \frac{s\gamma_H}{2}\exp\left(\frac{-s^2\,\gamma_H^2}{4}\right)\frac{1}{n\rho^n\sqrt{\pi}}\sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{C}}\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho). \tag{39}$$

[8] Of course, when $(x_0,\dots,x_{2k-2})$ spans a compact subset of $(0,+\infty)^{2k-1}$, then s spans a compact subset of $(0,+\infty)$.

To make notation lighter, we set $\kappa := \sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{E}} \widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho)$, which we will evaluate later. From Eqs. (35), (36) and (39), we have

$$p(n) \sim \frac{\frac{s\gamma_H}{2}\exp\left(\frac{-s^2\,\gamma_H^2}{4}\right)\frac{1}{n\rho^n\sqrt{\pi}}\kappa}{\frac{\gamma_D}{2\sqrt{\pi}}\rho^{-k-n}n^{k-3/2}} = n^{-k+1/2}\rho^k\frac{\gamma_H\kappa}{\gamma_D}\, s\,\exp\left(\frac{-s^2\,\gamma_H^2}{4}\right).$$

To conclude the proof of Theorem 5.7, it remains to check that

$$\rho^k\frac{\gamma_H\kappa}{\gamma_D} = \frac{\gamma_H^{2k}}{2^k}. \tag{40}$$

To this end we simplify the quantity $\kappa = \sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{E}} \widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho)$. Since $M_a^b = \Lambda_a\Lambda_b$, we have

$$\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho) = \prod_{i=0}^{2k-2}\Lambda_{tp_i}(\rho)\,\Lambda_{ct_i}(\rho)\prod_{\substack{i\\ w_i \text{ is an internal node}}} J_{ct_i}^{tp_{s_i}\,tp_{g_i}}(\rho).$$

The first product runs over edges of t_0. We can rearrange its terms according to vertices. Namely, we get a term $\Lambda_\bullet(\rho)$ for the root-leaf of t_0 and one for each leaf of t_0 (the type of the root-leaf and the cotypes of the leaves are $\bullet$; see the definition of $\mathfrak{E}$ in Theorem 5.6). Additionally, for each i such that w_i is an internal vertex, we get a factor $\Lambda_{ct_i}(\rho)$ from the parent edge e_i of w_i, and two factors $\Lambda_{tp_{s_i}}(\rho)$ and $\Lambda_{tp_{g_i}}(\rho)$ from the children edges e_{s_i} and e_{g_i} of w_i. The above display therefore rewrites as

$$\widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho) = \Lambda_\bullet(\rho)^{k+1}\prod_{\substack{i\\ w_i \text{ is an internal node}}}\Lambda_{ct_i}(\rho)\Lambda_{tp_{s_i}}(\rho)\Lambda_{tp_{g_i}}(\rho)J_{ct_i}^{tp_{s_i}\,tp_{g_i}}(\rho).$$

We now want to sum this quantity over $(\mathbf{tp},\mathbf{ct})$ in $\mathfrak{E}$. Note that choosing an element of $\mathfrak{E}$ consists in choosing ct_i, tp_{s_i} and tp_{g_i} for each internal vertex w_i. The sum $\kappa = \sum_{(\mathbf{tp},\mathbf{ct})\in\mathfrak{E}} \widetilde{M}_{(\mathbf{tp},\mathbf{ct})}(\rho)$ therefore factorizes over internal vertices of t_0 (there are $k-1$ of them) and we get

$$\kappa = \Lambda_\bullet(\rho)^{k+1}\left(\sum_{ct,tp_s,tp_g\in\{\mathcal{K},\mathcal{S}_\mathrm{X},\mathcal{S}_\mathrm{C}\}^3}\Lambda_{ct}(\rho)\Lambda_{tp_s}(\rho)\Lambda_{tp_g}(\rho)J_{ct}^{tp_s\,tp_g}(\rho)\right)^{k-1}.$$

We can write $\kappa = \mu\nu^k$, with

$$\mu = \Lambda_\bullet(\rho)\left(\sum_{ct,tp_s,tp_g\in\{\mathcal{K},\mathcal{S}_\mathrm{X},\mathcal{S}_\mathrm{C}\}^3}\Lambda_{ct}(\rho)\Lambda_{tp_s}(\rho)\Lambda_{tp_g}(\rho)J_{ct}^{tp_s\,tp_g}(\rho)\right)^{-1}; \tag{41}$$

$$\nu = \Lambda_\bullet(\rho)\left(\sum_{ct,tp_s,tp_g\in\{\mathcal{K},\mathcal{S}_\mathrm{X},\mathcal{S}_\mathrm{C}\}^3}\Lambda_{ct}(\rho)\Lambda_{tp_s}(\rho)\Lambda_{tp_g}(\rho)J_{ct}^{tp_s\,tp_g}(\rho)\right). \tag{42}$$

Then Eq. (40) holds for any $k \geq 1$ if

$$\gamma_H^2 = 2\rho\,\nu \quad \text{ and } \quad \gamma_H\mu = \gamma_D,$$

which we verify using Maple, from the definitions of the Λ_α and Theorem 5.5 for the $J_a^{b,c}$ (observing that $J_a^{b,c} = J_a^{c,b} = J_b^{a,c}$ for all a, b, c). □

Theorem 5.7 is a kind of local limit theorem for $r(\boldsymbol{T}_n, \boldsymbol{\ell})$. It is rather standard that such statements imply convergence in distribution, as we will see in our next result. Recall that $\mathcal{T}_k$ denotes the set of k-proper trees and that cnt is the counting measure on $\mathcal{T}_k$.

Corollary 5.8 *Recall that $(\boldsymbol{T}_n, \boldsymbol{\ell})$ denotes a uniform random DH-tree of size n with k marked leaves (not counted in the size). We set*

$$(\boldsymbol{t}_0^n, A_0, \dots, A_{2k-2}) = r(\boldsymbol{T}_n, \boldsymbol{\ell}).$$

Then

$$\left(\boldsymbol{t}_0^n, \frac{A_0}{\sqrt{n}}, \dots, \frac{A_{2k-2}}{\sqrt{n}}\right) \xrightarrow{(d)} (\boldsymbol{t}_0, X_0, \dots, X_{2k-2})$$

where $(\boldsymbol{t}_0, X_0, \dots, X_{2k-2})$ has density

$$\frac{\gamma_H^{2k}}{2^k} s \exp(-\gamma_H^2 s^2/4), \qquad \textit{with } s := x_0 + \dots + x_{2k-2}$$

with respect to the measure $\mathrm{cnt} \otimes \mathrm{Leb}^{\otimes 2k-1}$ *on* $\mathcal{T}_k \times (\mathbb{R}^+)^{2k-1}$.

Proof. The proof is an adaptation to the random variable $r(\boldsymbol{T}_n, \boldsymbol{\ell})$ which lives in $\mathcal{T}_k \times \mathbb{Z}^{2k-1}$ of the so-called "discrete Scheffé's lemm" for $\mathbb{Z}^d$-valued random variables (see the wikipedia page in French on Scheffé's lemma, or [34, Corollary A.3] for a published reference). We let $U_0, \dots, U_{2k-2}$ be independent uniform random variables in $[0, 1]$, and consider the variable

$$\left(\boldsymbol{t}_0^n, \frac{A_0 + U_0}{\sqrt{n}}, \dots, \frac{A_{2k-2} + U_{2k-2}}{\sqrt{n}}\right).$$

This variable has density

$$\sqrt{n}^{2k-1} P\big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, \lfloor\sqrt{n}x_0\rfloor, \dots, \lfloor\sqrt{n}x_{2k-2}\rfloor)\big]$$

with respect to the measure $\mathrm{cnt} \times \mathrm{Leb}^{\otimes 2k-1}$, where cnt is the counting measure on $\mathcal{T}_k$ and Leb the Lebesgue measure on $\mathbb{R}^+$. By Eq. (34), this density converges pointwise to

$$\frac{\gamma_H^{2k}}{2^k} s \exp(-\gamma_H^2 s^2/4), \qquad \text{with } s := x_0 + \dots + x_{2k-2}.$$

We note that the latter function defines a density on $\mathcal{T}_k \times (\mathbb{R}^+)^{2k-1}$ (make the change of variables $y_i = x_i \gamma_H/2$ in (3), which has been proven to be a density by Aldous [2, Lemma 21]).

By Sheffé's lemma, the convergence of density functions implies the convergence in distribution of the corresponding random variables. Since the terms $U_i/\sqrt{n}$ are asymptotically negligible, this proves the lemma. □

5.4 Gromov–Prohorov convergence of DH-graphs

Let $\boldsymbol{G}^{(n)}$ be the uniform DH-graph of size n. We want to deduce from Theorem 5.8 the convergence in distribution of the marginals of the distance matrix of $\boldsymbol{G}^{(n)}$. To do this recall that Theorem 3.4 allows us to estimate distances in $\boldsymbol{G}^{(n)}$ in terms of jumps in the associated DH-tree.

We first reformulate Theorem 3.4 with the vocabulary of induced subtrees. For $k \geq 2$ let G be a DH-graph of size $n+k+1$ (whose vertex set is therefore $\{1,\dots,n+k+1\}$). Let v_0 be the vertex of G with label $n+k+1$ and, for $1 \leq i \leq k$, let v_i be the vertex of G with label $n+i$. Denote by $(T, \boldsymbol{\ell})$ the DH-tree associated to G in the following way: the tree T, whose root-leaf ℓ_0 corresponds to v_0, has size n and k marked leaves $\ell_1,\dots,\ell_k$ respectively corresponding to vertices $v_1,\dots,v_k$. We denote by $(t_0, \alpha_0, \dots, \alpha_{2k-2})$ the enriched induced subtree $r(T, \boldsymbol{\ell})$ (defined at the end of Section 5.1).

Lemma 5.9 *For $0 \leq i, j \leq k$, let $\mathcal{P}^{t_0}_{i,j}$ be the path joining leaves ℓ_i and ℓ_j in t_0. Then*

$$1 \leq d_G(v_i, v_j) - \sum_{r:\ e_r \in \mathcal{P}^{t_0}_{i,j}} \alpha_r \leq k,$$

where $e_0, \dots, e_{2k-2}$ is the enumeration of edges of t_0.

Proof. Theorem 3.4 states that

$$d_G(v_i, v_j) = \sum_{r:\ e_r \in \mathcal{P}^{t_0}_{i,j}} \alpha_r + \#\{\text{ jumps in essential vertices}\} + 1.$$

The lemma follows observing that there are at most $k-1$ essential vertices. □

Proposition 5.10 *Let $(\boldsymbol{G}^{(m)})_{m\geq 3}$ be a sequence of uniform random labeled DH-graphs of size m. Let $k \geq 1$ and $V_0, V_1, \dots, V_k$ be uniform i.i.d. vertices in $\boldsymbol{G}^{(m)}$. Then we have the joint convergence in distribution:*

$$\left(\frac{1}{\sqrt{m}} \frac{\sqrt{2}}{\gamma_H} d_{\boldsymbol{G}^{(m)}}(V_i, V_j) \right)_{0 \leq i,j \leq k} \xrightarrow[m\to+\infty]{(d)} \left(d_\infty(v_i, v_j) \right)_{0 \leq i,j \leq k} \tag{43}$$

where the right-hand side denotes the marginals of distances in the Brownian CRT defined by Eq. (4)*.*

Proof. We fix $k \geq 1$. We first observe that with probability $1 - \mathcal{O}(k^2/m)$ we have that k i.i.d. uniform vertices in $\boldsymbol{G}^{(m)}$ are distinct. Therefore we can prove (43) where $(V_0, V_1, \dots, V_k)$ is a uniform $k+1$-tuple of distinct vertices.

Since the distribution of $\boldsymbol{G}^{(m)}$ is invariant by relabeling of vertices, we have that

$$\left(d_{\boldsymbol{G}^{(m)}}(V_i, V_j) \right)_{0 \leq i,j \leq k} \stackrel{(d)}{=} \left(d_{\boldsymbol{H}_{m-k-1}}(W_i, W_j) \right)_{0 \leq i,j \leq k}$$

where $\boldsymbol{H}_{m-k-1}$ is a uniform DH-graph of size $m-k-1$ with $k+1$ marked vertices $W_0, \dots, W_k$ not counted in the size.

Using Theorem 5.9 with $G = \boldsymbol{H}_{m-k-1}$ yields

$$\left(d_{\boldsymbol{G}^{(m)}}(V_i, V_j)\right)_{0\le i,j\le k} \overset{(d)}{=} \left(\sum_{r:\ e_r \in \mathcal{P}^{\boldsymbol{t}_0^m}_{i,j}} A_r + \mathcal{O}(1)\right)_{0\le i,j\le k}.$$

We finally use the convergence obtained in Theorem 5.8 (put $n = m-k-1$) and the criterion of Theorem 2.4 . □

From Theorem 2.2, Theorem 5.10 implies the convergence of uniform DH-graphs of size n towards the Brownian CRT with respect to the Gromov–Prohorov topology. Thus this concludes the proof of Theorem 1.1 in the case $f = d$.

6 The case of 2-connected DH-graphs

The goal of this section is to prove the convergence of a uniform random 2-connected DH-graph to the Brownian CRT, i.e. the case $f = 2c$ in Theorem 1.1. We start by giving a characterization of 2-connected DH-graphs through the associated (reduced) clique-star tree. The proof of the case $f = 2c$ in Theorem 1.1 then follows essentially the same steps as that of the case $f = d$ (unconstrained DH-graphs). We shall indicate all necessary modifications.

6.1 Combinatorial characterization

Recall that a vertex v in a connected graph G is called a *cut-vertex* if removing v (and edges incident to v) disconnects G. A connected graph G without cut-vertices is said to be 2-*connected*. Cut-vertices in DH-graphs, and hence 2-connected DH-graphs, are easily characterized through the associated reduced clique-star tree.

Lemma 6.1 *Let G be a DH-graph and let τ be a clique-star tree such that $G = \mathrm{Gr}(\tau)$. A vertex ℓ in G is a cut-vertex if and only if the associated leaf in τ is connected to the center of a star. Consequently, a distance-hereditary graph $G = \mathrm{Gr}(\tau)$ is 2-connected if and only if no leaf of τ is connected to the center of a star.*

Proof. We abusively call also ℓ the leaf of τ corresponding to the vertex ℓ of G, and v the unique vertex of τ adjacent to ℓ. We also denote by Γ_v the decoration of v, and by x the marker vertex of Γ_v corresponding to the edge (v, ℓ). Finally we denote $G\backslash\ell$ the graph obtained by removing ℓ (and its incident edges) from G.

By construction $G\backslash\ell = \mathrm{Gr}(\tau\backslash\ell)$, where $\tau\backslash\ell$ is the decorated tree obtained from τ by erasing the leaf ℓ and replacing in v the decoration Γ_v by $\Gamma_v\backslash x$ (note that $\tau\backslash\ell$ might not be a clique-star tree). By [23, Lemma 2.3], $G\backslash\ell$ is connected if and only if all decorations of $\tau\backslash\ell$ are connected. The only potentially non-connected decoration is $\Gamma_v\backslash x$ and it is disconnected precisely when Γ_v is a star, and x its center. This proves the characterization of cut-vertices given in the lemma. The characterization of 2-connected graphs follows immediately. □

By abuse of terminology, we say that a clique-star tree, or a DH-tree, is 2-connected if the associated DH-graph is 2-connected , *i.e.* if it does not contain a leaf (including the root-leaf in the case of DH-trees) linked to the center of a star. Specializing the bijection between DH-graphs and DH-trees to 2-connected objects, the above lemma allows to easily adapt the system of equations (5) to this setting:

$$\begin{cases} \overline{\mathcal{D}}_{\mathcal{K}} = & \mathrm{Set}_{\geq 2}(\mathcal{Z} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}}); \\ \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}} = & \mathrm{Set}_{\geq 2}(\mathcal{Z} + \overline{\mathcal{D}}_{\mathcal{K}} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}}); \\ \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}} = & (\overline{\mathcal{D}}_{\mathcal{K}} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}}) \times \mathrm{Set}_{\geq 1}(\mathcal{Z} + \overline{\mathcal{D}}_{\mathcal{K}} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}}); \\ \overline{\mathcal{D}} = & \overline{\mathcal{D}}_{\mathcal{K}} + \overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}}. \end{cases} \tag{44}$$

Here $\overline{\mathcal{D}}$ is the class of all 2-connected DH-trees, while $\overline{\mathcal{D}}_{\mathcal{K}}$ and $\overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{X}}}$ are the subclasses of $\overline{\mathcal{D}}$, consisting of trees with root of type $\mathcal{K}$ or S_{X}, respectively. The class $\overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}}$ is the class of DH-trees with a root of type S_{C}, such that *no other leaf than the root-leaf* is connected to the center of a star. DH-trees in $\overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}}$ are not 2-connected DH-trees since one of their leaves, namely the root-leaf, is connected to a center of a star. This explains why $\overline{\mathcal{D}}_{\mathcal{S}_{\mathrm{C}}}$ does not appear in the equation defining $\overline{\mathcal{D}}$ above. We nevertheless need to introduce this auxiliary class to write a full system of equations.

6.2 Singularity analysis of the system

As usual, for a class $\overline{\mathcal{D}}_{\alpha}$, we denote by $\overline{D}_{\alpha}$ its exponential generating function. From Eq. (44), we immediately check that, as in the unconstrained case, we have $\overline{D}_{\mathcal{K}} = \overline{D}_{\mathcal{S}_{\mathrm{C}}}$. Also, from the Drmota-Lalley-Woods theorem, all series $\overline{D}$, $\overline{D}_{\mathcal{K}}$ and $\overline{D}_{\mathcal{S}_{\mathrm{X}}}$ have the same radius of convergence ρ_{2c} and square root-singularities.

Again, it is useful to introduce the series

$$F_{2c} = \exp_{\geq 1}(z + \overline{D}_{\mathcal{K}} + \overline{D}_{\mathcal{S}_{\mathrm{X}}}).$$

The system (44) is then rewritten as

$$\begin{cases} \overline{D}_{\mathcal{K}} = & F_{2c} - z - \overline{D}_{\mathcal{K}} - \overline{D}_{\mathcal{S}_{\mathrm{X}}}; \\ \overline{D}_{\mathcal{S}_{\mathrm{X}}} = & 2\,\overline{D}_{\mathcal{K}}\,F_{2c}, \end{cases} \tag{45}$$

which is easily solved as

$$\begin{cases} \overline{D}_{\mathcal{K}} = & \frac{F_{2c}-z}{2+2F_{2c}}; \\ \overline{D}_{\mathcal{S}_{\mathrm{X}}} = & \frac{(F_{2c})^2-zF_{2c}}{1+F_{2c}}. \end{cases} \tag{46}$$

This implies that F_{2c} is solution of an equation of the type $F_{2c} = G(z, F_{2c})$, with

$$G(z,w) = \exp_{\geq 1}\left(z + \frac{w-z}{2+2w} + \frac{w^2 - zw}{1+w}\right). \tag{47}$$

Arguing as in Theorem 3.11, we find, after some elementary computations (the last equality being computed in the companion Maple worksheet), that:

- $\rho_{2c} = 2\,(F_{2c}(\rho_{2c}))^2 + 2\,F_{2c}(\rho_{2c}) - 1$;
- $F_{2c}(\rho_{2c})$ is the unique positive solution of the equation $s = \exp_{\geq 1}(2s - \frac{1}{2})$;
- $F_{2c}(z) = F_{2c}(\rho_{2c}) - \gamma_{F,2c}\sqrt{1 - z/\rho_{2c}} + O\big(1 - z/\rho_{2c}\big)$ with $\gamma_{F,2c} > 0$ and

$$(\gamma_{F,2c})^2 = \frac{\rho_{2c}(1 + F(\rho_{2c}))}{1 + 2F(\rho_{2c})} = \frac{\Big(2\,(F_{2c}(\rho_{2c}))^2 + 2\,F_{2c}(\rho_{2c}) - 1\Big)\Big(1 + F_{2c}(\rho_{2c})\Big)}{1 + 2F_{2c}(\rho_{2c})}.$$

6.3 2-connected DH-trees with a marked leaf

As in the case of general DH-graphs, the next step is to analyze families of 2-connected DH-trees with a marked leaf. For a, b in $\{\mathcal{K}, S_X, S_C\}$, we denote $\overline{\mathcal{D}}_a^b$ the class of DH-tree with a root of type a, a marked leaf of cotype b, and such that no leaf is connected to the center of a star, except possibly the root-leaf or the marked leaf (when a and/or b is equal to S_C). Moreover, we let $\overline{D}_a^b(z, u)$ be the corresponding bivariate (exponential) generating series, where the exponent of the variable z is the size (number of nonmarked nonroot leaves) of the tree and the exponent of u is the number of jumps on the path from the root-leaf to the marked leaf.

These nine series satisfy the following system of equations, whose proof is similar to that of Theorem 4.3:

$$\begin{cases} \overline{D}_{\mathcal{K}}^{\mathcal{K}} = (1 + \overline{D}_{\mathcal{S}_C}^{\mathcal{K}} + \overline{D}_{\mathcal{S}_X}^{\mathcal{K}})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X} + z); \\ \overline{D}_{\mathcal{S}_X}^{\mathcal{K}} = (\overline{D}_{\mathcal{S}_C}^{\mathcal{K}} + \overline{D}_{\mathcal{K}}^{\mathcal{K}})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z) \\ \qquad\qquad + u \cdot (\overline{D}_{\mathcal{K}}^{\mathcal{K}} + \overline{D}_{\mathcal{S}_X}^{\mathcal{K}})\,(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{K}})\ \exp(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z); \\ \overline{D}_{\mathcal{S}_C}^{\mathcal{K}} = (\overline{D}_{\mathcal{K}}^{\mathcal{K}} + \overline{D}_{\mathcal{S}_X}^{\mathcal{K}})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z); \end{cases} \tag{48}$$

$$\begin{cases} \overline{D}_{\mathcal{K}}^{\mathcal{S}_X} = (\overline{D}_{\mathcal{S}_X}^{\mathcal{S}_X} + \overline{D}_{\mathcal{S}_C}^{\mathcal{S}_X})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X} + z); \\ \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_X} = (\overline{D}_{\mathcal{S}_C}^{\mathcal{S}_X} + \overline{D}_{\mathcal{K}}^{\mathcal{S}_X})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z) \\ \qquad\qquad + u \cdot (1 + \overline{D}_{\mathcal{K}}^{\mathcal{S}_X} + \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_X})\,(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{K}})\ \exp(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z); \\ \overline{D}_{\mathcal{S}_C}^{\mathcal{S}_X} = (1 + \overline{D}_{\mathcal{K}}^{\mathcal{S}_X} + \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_X})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z); \end{cases} \tag{49}$$

$$\begin{cases} \overline{D}_{\mathcal{K}}^{\mathcal{S}_C} = (\overline{D}_{\mathcal{S}_C}^{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_C})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X} + z); \\ \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_C} = (1 + \overline{D}_{\mathcal{S}_C}^{\mathcal{S}_C} + \overline{D}_{\mathcal{K}}^{\mathcal{S}_C})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z) \\ \qquad\qquad + u \cdot (\overline{D}_{\mathcal{K}}^{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_C})\,(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{K}})\ \exp(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z); \\ \overline{D}_{\mathcal{S}_C}^{\mathcal{S}_C} = (\overline{D}_{\mathcal{K}}^{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X}^{\mathcal{S}_C})\ \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z). \end{cases} \tag{50}$$

Recall that we have $\exp_{\geq 1}(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{S}_X} + z) = \exp_{\geq 1}(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z) = F_{2c}$. Furthermore, using that $\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{K}} = 2\overline{D}_{\mathcal{K}} = \frac{F_{2c}-z}{1+F_{2c}}$, we have

$$(\overline{D}_{\mathcal{S}_C} + \overline{D}_{\mathcal{K}}) \exp(\overline{D}_{\mathcal{S}_X} + \overline{D}_{\mathcal{K}} + z) = F_{2c} - z.$$

After these simplifications, the system is similar to that of Eqs. (20) to (22), except that F is replaced by F_{2c} and uF by $u\,(F_{2c} - z)$. This system is solved as follows (either directly or by substituting F with F_{2c} and uF with $u\,(F_{2c} - z)$ in Eqs. (23) to (28)):

$$\overline{D}_{\mathcal{K}}^{\mathcal{K}} = \frac{F_{2c}}{F_{2c}+1} - \frac{(F_{2c})^2}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}; \tag{51}$$

$$\overline{D}_{\mathcal{S}_X}^{\mathcal{S}_X} = \frac{-1}{F_{2c}+1} + \frac{(1-F_{2c})^2}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}; \tag{52}$$

$$\overline{D}_{\mathcal{S}_C}^{\mathcal{S}_C} = \frac{(F_{2c})^2}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}; \tag{53}$$

$$\overline{D}_{\mathcal{K}}^{\mathcal{S}_X} = \overline{D}_{\mathcal{S}_X}^{\mathcal{K}} = \frac{-F_{2c}}{F_{2c}+1} + \frac{F_{2c}\,(1-F_{2c})}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}; \tag{54}$$

$$\overline{D}_{\mathcal{K}}^{\mathcal{S}_C} = \overline{D}_{\mathcal{S}_C}^{\mathcal{K}} = \frac{(F_{2c})^2}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}; \tag{55}$$

$$\overline{D}_{\mathcal{S}_X}^{\mathcal{S}_C} = \overline{D}_{\mathcal{S}_C}^{\mathcal{S}_X} = \frac{F_{2c}(1-F_{2c})}{(1+F_{2c})(1-2F_{2c}) - u(F_{2c}-z)}. \tag{56}$$

Recalling that F_{2c} depends only on z (not on u), we note that in each case, the series can be written under the form

$$\overline{D}_a^b = \overline{Q}_a^b(z) + \frac{\overline{M}_a^b(z)}{1 - u\,\overline{H}_a^b(z)}, \tag{57}$$

where $\overline{Q}_a^b$, $\overline{M}_a^b$ and $\overline{H}_a^b$ are rational functions in F_{2c} (and z in the case of $\overline{H}_a^b$). For example, looking at Eq. (51), we have

$$\overline{Q}_{\mathcal{K}}^{\mathcal{K}} = \frac{F_{2c}}{F_{2c}+1},\ \overline{M}_{\mathcal{K}}^{\mathcal{K}} = \frac{(F_{2c})^2}{(1+F_{2c})(1-2F_{2c})} \text{ and } \overline{H}_{\mathcal{K}}^{\mathcal{K}} = \frac{F_{2c}-z}{(1+F_{2c})(1-2F_{2c})}.$$

Similar formulas are easily written for other $a, b \in \{\mathcal{K}, \mathcal{S}_C, \mathcal{S}_X\}$, looking at Eqs. (52) to (56). As in the case of unconstrained DH-graphs, the auxiliary series $\overline{H}_a^b = \overline{H}$ do not depend on a and b. At $z = \rho_{2c}$, $\overline{H}$ admits an expansion of the form $\overline{H}(z) = \overline{H}(\rho_{2c}) - \gamma_{H,2c}\sqrt{1 - z/\rho_{2c}} + O\big(1 - z/\rho_{2c}\big)$ with

$$\gamma_{H,2c} = \frac{2\sqrt{(1+2F_{2c}(\rho_{2c}))\cdot(2F_{2c}(\rho_{2c})^3 + 4F_{2c}(\rho_{2c})^2 + F_{2c}(\rho_{2c}) - 1)}}{1 - F_{2c}(\rho_{2c}) - 2F_{2c}(\rho_{2c})^2}, \tag{58}$$

whose numerical estimate is $\gamma_{H,2c} \approx 7.5022$ (see Maple worksheet).

Furthermore we have

$$\overline{H}(\rho_{2c}) = \frac{F_{2c}(\rho_{2c}) - \rho_{2c}}{(1+F_{2c}(\rho_{2c}))(1-2F_{2c}(\rho_{2c}))} = 1,$$

thanks to the relation $\rho_{2c} = 2F_{2c}(\rho_{2c})^2 + 2F_{2c}(\rho_{2c}) - 1$ given at the end of Section 6.2.

6.4 2-connected DH-trees with marked leaves inducing a given subtree

Using the same terminology as in Section 5, we define $\overline{\mathcal{D}}_{t_0}$ to be the class of 2-connected DH-trees with k marked leaves inducing a given k-proper tree t_0. Furthermore, we let $\overline{D}_{t_0}(z, u_0, ..., u_{2k-2})$ be the multivariate (exponential) generating series of $\overline{\mathcal{D}}_{t_0}$, where the exponent of z is the size of the tree and the exponent of u_i the number of jumps in the path corresponding to e_i (in the fixed enumeration $(e_0, e_1, \ldots, e_{2k-2})$ of the edges of t_0).

To write a combinatorial decomposition for $\overline{D}_{t_0}(z, u_0, ..., u_{2k-2})$, we need to introduce a subclass $\overline{\mathcal{J}}_a^{b\,c}$ of $\mathcal{J}_a^{b\,c}$, where the nonmarked (and nonroot) leaves are not allowed to be attached to the center of a star. The generating function of this auxiliary class is given by

$$\begin{aligned}\overline{J}_a^{b\,c}(z) &= (\mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C)\exp(\overline{D}_{\mathcal{S}_\mathrm{X}} + \overline{D}_{\mathcal{K}} + z) + \mathbb{1}_{\mathcal{K}\notin\{a,b,c\}}\exp(\overline{D}_{\mathcal{S}_\mathrm{X}} + \overline{D}_{\mathcal{S}_\mathrm{C}} + z) \\ &\quad + \mathbb{1}_{\mathcal{S}_\mathrm{C}\notin\{a,b,c\}}(\overline{D}_{\mathcal{S}_\mathrm{C}} + \overline{D}_{\mathcal{K}})\exp(\overline{D}_{\mathcal{K}} + \overline{D}_{\mathcal{S}_\mathrm{X}} + z) \\ &= (\mathbb{1}_A + \mathbb{1}_B + \mathbb{1}_C + \mathbb{1}_{\mathcal{K}\notin\{a,b,c\}})(1 + F_{2c}) + \mathbb{1}_{\mathcal{S}_\mathrm{C}\notin\{a,b,c\}}(F_{2c} - z),\end{aligned} \tag{59}$$

where A, B, C are given in Theorem 5.5. Note that the factor $(D_{\mathcal{S}_\mathrm{C}} + D_{\mathcal{K}} + z)$ in the second line of (32) in Lemma 5.5 is here replaced by $(\overline{D}_{\mathcal{S}_\mathrm{C}} + \overline{D}_{\mathcal{K}})$. Indeed, this factor corresponds to a subtree attached to the center of a star, which is not allowed to be a leaf in the 2-connected case.

At this stage, there is a small difference with the case of unconstrained DH-trees. In a 2-connected DH-tree, the root cannot have type S_C and no leaves (in particular the marked ones) can have cotype S_C. Therefore in the combinatorial decomposition of Fig. 12, the piece corresponding to e_0 has a root type different from S_C, and pieces corresponding to leaf-edges of t_0 have a marked leaf with a cotype different from S_C as well. This is easily captured in equations by defining

$$\overline{\mathcal{D}}_a^{\bullet} = \overline{\mathcal{D}}_a^{\mathcal{K}} + \overline{\mathcal{D}}_a^{\mathcal{S}_\mathrm{X}};$$
$$\overline{\mathcal{D}}_{\bullet}^{b} = \overline{\mathcal{D}}_{\mathcal{K}}^{b} + \overline{\mathcal{D}}_{\mathcal{S}_\mathrm{X}}^{b}.$$

In the unconstrained case, each of these equations had an extra term corresponding to the type (or cotype) S_C. With these definitions, Theorem 5.6 is still valid when replacing each series by its 2-connected counterpart. The asymptotic analysis in the 2-connected case is then identical to that of the unconstrained case, up to the verification of the identities

$$\gamma_{H,2c}^2 = 2\rho_{2c}\,\nu_{2c} \text{ and } \gamma_{H,2c}\mu_{2c} = \gamma_{D,2c},$$

where ν_{2c} and μ_{2c} are defined via the obvious analogs of Eqs. (41) and (42). Verifying these identities is done in the companion Maple worksheet. We therefore have the following analog of Theorem 5.7.

Proposition 6.2 *Let $(\boldsymbol{T}_n, \boldsymbol{\ell})$ be a uniform random 2-connected DH-tree of size n with k marked leaves (not counted in the size). Fix a k-proper tree t_0 and real numbers*

$x_0, \ldots, x_{2k-2} > 0$. *We set* $a_i = \lfloor x_i \sqrt{n} \rfloor$. *Then*

$$\mathbb{P}\big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \ldots, a_{2k-2})\big] \sim \frac{\gamma_{H,2c}^{2k}}{2^k \sqrt{n}^{2k-1}} s \exp\left(\frac{-\gamma_{H,2c}^2 s^2}{4}\right), \tag{60}$$

where $s = \sum_i x_i$. *Moreover, this estimate is uniform for* $x_0, \ldots, x_{2k-2}$ *in any compact subset of* $(0, +\infty)^{2k-1}$.

From here, the convergence to the Brownian CRT in Gromov–Prohorov topology, *i.e.* the second case of Theorem 1.1, follows using the same arguments as in the case of unconstrained DH-graphs.

7 The case of 3-leaf power graphs

The goal of this section is to prove the convergence of a uniform random 3-leaf power graph to the Brownian CRT, *i.e.* the case $f = 3\ell$ in Theorem 1.1. We start by recalling a characterization of 3-leaf power graphs through their associated (reduced) clique-star tree, given in [23]. The proof of convergence then follows essentially the same steps as in the two other cases. There is however one notable difference. As we shall see, in this model, first common ancestors of marked leaves are of type and cotype $\mathcal{S}_X$ with probability tending to 1; therefore, we only need to consider two types of trees with one marked leaf, simplifying significantly the analysis.

7.1 Definition and combinatorial analysis of 3-leaf power graphs

This section follows closely [14, Section 2].

Definition 7.1 *Let T be a tree and L its set of leaves. The k-leaf power graph G of T has by definition vertex set L, and ℓ and ℓ' are connected in G if they are at distance at most k in T. And a graph is a k-leaf power graph if it is the k-leaf power graph of some tree.*

We are interested in the case $k = 3$. An example of a tree, together with its 3-leaf power graph, is given on Fig. 13. It is known, see *e.g.*, [14, Section 2] that 3-leaf power graphs form a subclass of distance-hereditary graphs, and that they can be characterized on the clique-star trees as follows (see [23, Section 3.3]).

Proposition 7.2 *A distance hereditary graph G is a 3-leaf power graph if and only if its reduced clique star-tree τ satisfies the following properties:*

- *the set of star nodes forms a connected subtree of τ;*
- *no edge connects two centers of star nodes.*

In the sequel, we call 3-*leaf power trees* the DH-trees corresponding to (rooted) 3-*leaf power graphs*. Let $\mathcal{E}$ be the combinatorial class of 3-leaf power trees. To get a combinatorial decomposition of this class, it is convenient to introduce the following subclasses:

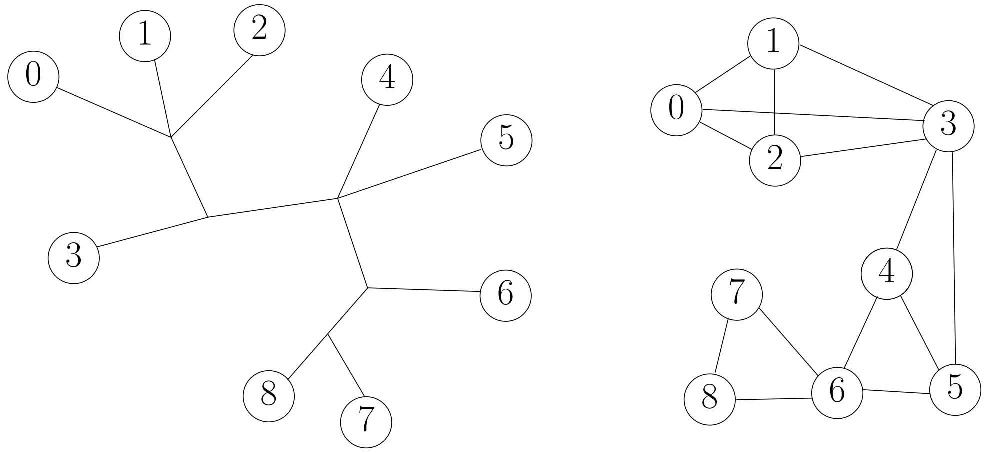

Figure 13: A tree (on the left) and the associated 3-leaf power graph (on the right). For instance, the leaves 0 and 3 are at distance 3 from each other in the tree, and hence connected by an edge in the graph. On the other hand, the leaves 5 and 8 are at distance 4 in the tree, and not connected by an edge in the graph.

- $\mathcal{E}_{\mathcal{S}_X}$, $\mathcal{E}_{\mathcal{S}_C}$ and $\mathcal{E}_{\mathcal{K}}$ are the subclasses of $\mathcal{E}$, where the root-node is required to have type $\mathcal{S}_X$, $\mathcal{S}_C$ and $\mathcal{K}$ respectively;
- $\mathcal{L}$ is the class containing the tree restricted to a single leaf, and trees consisting of a single internal node, of type $\mathcal{K}$, with at least two pending leaves.

Recall also that $\mathcal{Z}$ is a class with only one element, which is of size 1, representing a leaf. The following set of equations, characterizing all these classes, is obtained easily:

$$\begin{cases} \mathcal{L} = \mathcal{Z} + \mathrm{Set}_{\geq 2}(\mathcal{Z}); \\ \mathcal{E}_{\mathcal{S}_X} = \mathcal{L} \times \mathrm{Set}_{\geq 1}(\mathcal{L} + \mathcal{E}_{\mathcal{S}_X}); \\ \mathcal{E}_{\mathcal{S}_C} = \mathrm{Set}_{\geq 2}(\mathcal{L} + \mathcal{E}_{\mathcal{S}_X}); \\ \mathcal{E}_{\mathcal{K}} = \mathcal{L} + (\mathcal{E}_{\mathcal{S}_X} + \mathcal{E}_{\mathcal{S}_C})\,\mathrm{Set}_{\geq 1}(\mathcal{Z}); \\ \mathcal{E} = \mathcal{E}_{\mathcal{S}_X} + \mathcal{E}_{\mathcal{S}_C} + \mathcal{E}_{\mathcal{K}}. \end{cases} \tag{61}$$

In terms of generating series (with the usual convention that $Y(z)$ is the exponential generating function of a class $\mathcal{Y}$), the first equation implies $L = e^z - 1$. The second equation yields:

$$E_{\mathcal{S}_X} = L \cdot \big(\exp(L + E_{\mathcal{S}_X}) - 1\big) = (e^z - 1) \cdot \big(\exp(E_{\mathcal{S}_X} + e^z - 1) - 1\big). \tag{62}$$

We note that other equations of the system Eq. (61) are nonrecursive and simply express $E_{\mathcal{S}_C}$, $E_{\mathcal{K}}$ and E in terms of $E_{\mathcal{S}_X}$. It is thus not surprising that most of the asymptotic analysis reduces to that of $E_{\mathcal{S}_X}$. We first prove the following result.

Proposition 7.3 *The series $E_{\mathcal{S}_X}$ is Δ-analytic at $\rho_{3\ell} = \log(1 + e^{-1})$ and admits the following singular expansion around $\rho_{3\ell}$:*

$$E_{\mathcal{S}_X} = (1 - e^{-1}) - \gamma_{E,3\ell}\sqrt{1 - z/\rho_{3\ell}} + \mathcal{O}(1 - z/\rho_{3\ell}), \tag{63}$$

where

$$\gamma_{E,3\ell} = \sqrt{2(1+e)\log(1+e^{-1})}, \tag{64}$$

whose numerical estimate is $\gamma_{E,3\ell} \approx 1.5263$ *(see Maple worksheet).*

Proof. As in the proof of Theorem 3.11 we use the smooth implicit-function schema. We write $E_{\mathcal{S}_X} = G(z, E_{\mathcal{S}_X})$ where

$$G(z,w) = (e^z - 1)\,(\exp(w + e^z - 1) - 1)$$

which is analytic in z, w on the whole complex plane and which has nonnegative coefficients. The characteristic system $\{G(r,s) = s; G_w(r,s) = 1\}$ is easily solved and the unique solution is

$$r = \log(1+e^{-1}), \qquad s = 1 - e^{-1}.$$

We apply [22, Theorem VII.3] and obtain Eq. (63). □

With system (61) and the singular expansion of $E_{\mathcal{S}_X}$ given above, we find that of the other series. In particular,

$$E = (e - e^{-1} - e^{-2}) - (e+1)\gamma_{E,3\ell}\sqrt{1 - z/\rho_{3\ell}} + \mathcal{O}(1 - z/\rho_{3\ell}), \tag{65}$$

which will be useful later.

7.2 3-leaf power trees with a marked leaf

We now consider families of 3-leaf power trees with a marked leaf. It turns out that the only classes relevant for the asymptotic analysis are the classes $\mathcal{E}_{\mathcal{S}_X}^{\mathcal{S}_X}$ and $\mathcal{E}_{\mathcal{S}_X}^{\bullet}$ defined as follows: we let $\mathcal{E}_{\mathcal{S}_X}^{\mathcal{S}_X}$ (respectively $\mathcal{E}_{\bullet}^{\mathcal{S}_X}$ or $\mathcal{E}_{\mathcal{S}_X}^{\bullet}$) be the subclass of 3-leaf power trees with a marked leaf such that the root has type S_X and the marked leaf has cotype S_X (respectively with no constraints on the type of the root or on the cotype of the marked leaf). As above, we consider the associated exponential bivariate generating series $E_{\mathcal{S}_X}^{\mathcal{S}_X}$, $E_{\bullet}^{\mathcal{S}_X}$ and $E_{\mathcal{S}_X}^{\bullet}$, where the exponent of z is the size of the tree (number of nonmarked nonroot leaves) and that of u is the number of jumps on the path from the root-leaf to the marked leaf. By symmetry we have $E_{\bullet}^{\mathcal{S}_X} = E_{\mathcal{S}_X}^{\bullet}$.

We also let $\mathcal{L}^\bullet$ be the class of objects in $\mathcal{L}$ with a marked leaf. Its generating series is $L^\bullet = e^z$ (there is no jumps in such objects). An easy combinatorial analysis yields the following equations:

$$\begin{aligned} E_{\mathcal{S}_X}^{\mathcal{S}_X} &= u \cdot (1 + E_{\mathcal{S}_X}^{\mathcal{S}_X}) \cdot L \cdot \mathrm{Set}(L + E_{\mathcal{S}_X}); \\ E_{\mathcal{S}_X}^{\bullet} &= L^\bullet \cdot \mathrm{Set}_{\geq 1}(L + E_{\mathcal{S}_X}) + u \cdot (L^\bullet + E_{\mathcal{S}_X}^{\bullet}) \cdot L \cdot \mathrm{Set}(L + E_{\mathcal{S}_X}). \end{aligned}$$

This is a 2 × 2 linear system of equations in the unknown series $E_{\mathcal{S}_X}^{\mathcal{S}_X}$ and $E_{\mathcal{S}_X}^{\bullet}$. The solutions can be put under a form similar to Eq. (29):

$$\begin{aligned} E_{\mathcal{S}_X}^{\mathcal{S}_X}(u) &= -1 + \frac{1}{1 - u\,P} \\ E_{\mathcal{S}_X}^{\bullet}(u) &= -e^z + \frac{e^z\,\exp(e^z - 1 + E_{\mathcal{S}_X})}{1 - u\,P}, \end{aligned}$$

where

$$P = P(z) = (e^z - 1)\exp(e^z - 1 + E_{\mathcal{S}_\mathrm{X}}).$$

Using Eq. (63), we immediately see that P has radius of convergence $\rho_{3\ell}$, is Δ-analytic and admits the following singular expansion for z near $\rho_{3\ell}$:

$$P = 1 - \gamma_{E,3\ell}\sqrt{1 - z/\rho_{3\ell}} + \mathcal{O}(1 - z/\rho_{3\ell}). \tag{66}$$

7.3 3-leaf power trees with k marked leaves

Let us fix a k-proper tree t_0. We consider the following class of marked 3-leaf power trees.

Definition 7.4 *We let $\mathcal{E}_{t_0}$ be the labeled combinatorial class of 3-leaf power trees T with k marked leaves $\boldsymbol{\ell}$ such that:*

i) the subtree of T induced by $\boldsymbol{\ell}$ is t_0;

ii) no two essential vertices of T are neighbors of each other;

iii) every internal essential vertex of T has type and cotype S_X and its children which are roots of subtrees containing marked leaves are also of type S_X.

As above, we fix an enumeration $(e_0, \ldots, e_{2k-2})$ of the edges of t_0 such that the edge adjacent to the root-leaf is labeled with e_0; here, we additionally require that the edges incidents to leaves of t_0 get labels $e_1, \ldots, e_k$. Recall that we defined $\mathrm{jp}_e(T;\boldsymbol{\ell})$ as the number of jumps on the path corresponding to e, with the convention that essential vertices are not counted as jumps (but the root-node of T can be a jump). We consider the following multivariate generating series for $\mathcal{E}_{t_0}$:

$$E_{t_0}(u_0, \ldots, u_{2k-2}) = \sum_{(T,\boldsymbol{\ell})\in\mathcal{E}_{t_0}} \frac{z^{|(T,\boldsymbol{\ell})|}}{|(T;\boldsymbol{\ell})|!} u_0^{\mathrm{jp}_{e_0}(T,\boldsymbol{\ell})} \cdots u_{2k-2}^{\mathrm{jp}_{e_{2k-2}}(T,\boldsymbol{\ell})},$$

where we recall that $|(T,\boldsymbol{\ell})|$ is the number of non-root unmarked leaves of $(T,\boldsymbol{\ell})$.

Moreover, let $\mathcal{I}_a^{b\,c}$ be the set of 3-leaf power trees T with two marked leaves such that

- the two marked leaves are children of the root-node;
- if T_1 is a 3-leaf power of type b, one can glue T_1 on the first marked leaf of T (merging the marked leaf and the root-node of T_1) such that the tree obtained is a 3-leaf power tree;
- the same condition holds with gluing a 3-leaf power tree of type c on the second marked leaf;
- additionally, if T_0 is a 3-leaf power tree with a marked leaf of cotype a, one can glue T on the marked leaf of T_0 obtaining a 3-leaf power tree.

Lemma 7.5 *The generating function of* $\mathcal{I}_{S_X}^{S_X\, S_X}$ *is*

$$I_{S_X}^{S_X\, S_X}(z) = L \exp(E_{\mathcal{S}_X} + L). \tag{67}$$

Proof. Because of the allowed adjacencies between nodes of various types in 3-leaf power trees, a 3-leaf power tree in $\mathcal{I}_{S_X}^{S_X\, S_X}$ necessarily has a root-node r of type S_X. The factor L then accounts for the tree pending under the center of the star labeling r, and $\exp(E_{\mathcal{S}_X} + L)$ accounts for the trees pending under its extremities which do not correspond to the marked leaves. □

Proposition 7.6 *We have*

$$E_{t_0}(u_0, \ldots, u_{2k-2}) = \left(\prod_{i=0}^{k} E^{\bullet}_{\mathcal{S}_X}(u_i)\right) \cdot \left(\prod_{i=k+1}^{2k-2} E^{\mathcal{S}_X}_{\mathcal{S}_X}(u_i)\right) \cdot \left(\mathcal{I}_{S_X}^{S_X\, S_X}\right)^{k-1}. \tag{68}$$

Proof.[Sketch of proof] We use the same decomposition as in the proof of Theorem 5.6. The main difference is that, because of item iii) in Theorem 7.4 above, all types tp_i and cotypes ct_i which do not correspond to the marked leaves or the root-leaf must be equal to S_X. Consequently the tuple $(tp_i, ct_i)_{0 \le i \le 2k-2}$ can only take one possible value and we get

$$E_{t_0}(u_0, \ldots, u_{2k-2}) = E^{\mathcal{S}_X}_{\bullet}(u_0) \left(\prod_{i=1}^{k} E^{\bullet}_{\mathcal{S}_X}(u_i)\right) \cdot \left(\prod_{i=k+1}^{2k-2} E^{\mathcal{S}_X}_{\mathcal{S}_X}(u_i)\right) \cdot \left(\mathcal{I}_{S_X}^{S_X\, S_X}\right)^{k-1}.$$

We conclude using the symmetry $E^{\bullet}_{\mathcal{S}_X} = E^{\mathcal{S}_X}_{\bullet}$. □

Proposition 7.7 *Let* $(\boldsymbol{T}_n, \boldsymbol{\ell})$ *be a uniform random 3-leaf power tree of size* n *with* k *marked leaves (not counted in the size). Fix a* k*-proper tree* t_0 *and real numbers* $x_0, \ldots, x_{2k-2} > 0$. *We set* $a_i = \lfloor x_i \sqrt{n} \rfloor$. *Then*

$$\mathbb{P}\Big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \ldots, a_{2k-2}) \wedge (\boldsymbol{T}_n, \boldsymbol{\ell}) \in \mathcal{E}_{t_0}\Big] \sim \frac{\gamma_{E,3\ell}^{2k}}{2^k \sqrt{n}^{2k-1}} s \exp\left(\frac{-\gamma_{E,3\ell}^2 s^2}{4}\right), \tag{69}$$

where $s = \sum_i x_i$. *Moreover, this estimate is uniform for* $x_0, \ldots, x_{2k-2}$ *in any compact subset of* $(0, +\infty)^{2k-1}$.

Proof. Writing $p(n) := \mathbb{P}\big[r(\boldsymbol{T}_n, \boldsymbol{\ell}) = (t_0, a_0, \ldots, a_{2k-2}) \wedge (\boldsymbol{T}_n, \boldsymbol{\ell}) \in \mathcal{E}_{t_0}\big]$, we have

$$p(n) = \frac{[z^n u_0^{a_0} \ldots u_{2k-2}^{a_{2k-2}}]\ E_{t_0}(z, u_0, \ldots, u_{2k-2})}{[z^n] E^{(k)}(z)}. \tag{70}$$

We first analyze the denominator. From Eq. (65), routine computations yield:

$$[z^n] E^{(k)}(z) = \frac{(1+e)\, \gamma_{E,3\ell}}{2\sqrt{\pi}} \rho_{3\ell}^{-k-n} n^{k-3/2}. \tag{71}$$

With the same reasoning as in Theorem 5.7, we have

$$[z^n u_0^{a_0} \dots u_{2k-2}^{a^{2k-2}}]\ E_{t_0}(z, u_0, \dots, u_{2k-2}) = [z^n]\, N(z)P(z)^{a_0+\dots+a_{2k-2}}, \tag{72}$$

where

$$N(z) = \Big(e^z \exp\left(e^z - 1 + E_{\mathcal{S}_X}\right)\Big)^{k+1} \cdot \Big((e^z - 1)\cdot \exp\left(e^z - 1 + E_{\mathcal{S}_X}\right)\Big)^{k-1}. \tag{73}$$

Applying the Semi-large powers Theorem (Theorem A.1) and using Eq. (66), we have

$$[z^n]\, N(z)P(z)^{a_0+\dots+a_{2k-2}} \sim \frac{s\gamma_{E,3\ell}}{2} \exp\left(\frac{-s^2\,\gamma_{E,3\ell}^2}{4}\right) \frac{1}{n\rho_{3\ell}^n\sqrt{\pi}} N(\rho_{3\ell}).$$

And since $N(\rho_{3\ell}) = (e+1)^{k+1}$, it follows that

$$p(n) \sim \frac{\frac{s\gamma_{E,3\ell}}{2}\exp\left(\frac{-s^2\,\gamma_{E,3\ell}^2}{4}\right)\frac{1}{n\rho_{3\ell}^n\sqrt{\pi}}\,(e+1)^{k+1}}{\frac{(1+e)\gamma_{E,3\ell}}{2\sqrt{\pi}}\rho_{3\ell}^{-k-n} n^{k-3/2}} = (e+1)^k \rho_{3\ell}^k\, n^{-k+1/2}\, s \exp\left(\frac{-s^2\,\gamma_{E,3\ell}^2}{4}\right).$$

Using the explicit expressions of $\gamma_{E,3\ell}$ and $\rho_{3\ell}$ given in Theorem 7.3, we observe that $\gamma_{E,3\ell}^2 = 2(1+e)\rho_{3\ell}$. Thus the above formula coincides with (69). ☐

From here, the convergence to the Brownian CRT in Gromov–Prohorov topology, *i.e.* the third case of Theorem 1.1, follows using the same arguments as in the case of unconstrained DH-graphs.

Appendix A The Semi-large powers Theorem

In order to prove Theorems 5.7, 6.2 and 7.7, we need to estimate quantities of the form $[z^n]M(z)H(z)^p$ where p is of order $\sqrt{n}$. The following statement is essentially the Semi-large powers Theorem ([22, Theorem IX.16] for $\lambda = 1/2$; see also [5] for the original reference) which deals with the case $M(z) = 1$. As we will see, there are no particular difficulties in generalizing the proof.

Theorem A.1 *Let $\rho > 0$ and let M, H be Δ-analytic functions at ρ. Assume that*

i) H has a square-root singularity at ρ: for some $\sigma, h > 0$

$$H(z) = \sigma - h\sqrt{1 - z/\rho} + \mathcal{O}(1 - z/\rho).$$

ii) M converges at ρ.

Let us fix a compact subset K of $(0, +\infty)$ and a constant $C > 0$. Take (a_n) a sequence of real numbers such that

$$|a_n - x\sqrt{n}| \le C, \tag{74}$$

for some x in K. Then we have

$$[z^n]M(z)H(z)^{a_n} \overset{n\to+\infty}{\sim} \mathsf{Ray}\left(\frac{xh}{\sigma}\right)\frac{1}{n\rho^n\sqrt{\pi}}M(\rho)\sigma^{a_n},$$

where $\mathsf{Ray}(x) = \frac{x}{2}e^{-x^2/4}$ *is the Rayleigh density.*

The error term in the above convergence is uniform for all $x \in K$, and all sequences (a_n) satisfying (74), but depends on M, H, K and C.

Proof of Theorem A.1

We mimic the proof of [22, Theorem IX.16]. By assumption there exists $R_1 > \rho$ and $\theta > 0$ such that M, H are analytic on

$$\{z : |z| < R_1 \text{ et } |\operatorname{Arg}(z-\rho)| > \theta\}.$$

Fix R in (ρ, R_1) and write

$$[z^n]M(z)H(z)^{a_n} = \frac{1}{2i\pi}\int_\gamma M(z)H(z)^{a_n}\frac{dz}{z^{n+1}},$$

where $\gamma = \gamma_0 \cup \overline{\gamma_0} \cup \gamma_1 \cup \gamma_2$ is a closed counter-clockwise contour surrounding 0 consisting of the following pieces (see Fig. 14):

- γ_0 is a line segment starting at $\rho + i/n$, with a slope θ and stopping when it reaches the circle $\{z : |z| = R\}$;
- $\overline{\gamma_0}$ is its complex conjugate (in reverse direction);
- γ_1 is a semi-circle centered at ρ of radius $1/n$ from $\rho - i\rho/n$ to $\rho + i\rho/n$;
- γ_2 is an arc of circle of radius R closing γ.

The modulus of the integral on γ_2 is easily bounded by $\mathcal{O}(B^{\sqrt{n}}R^{-n})$ for some constant B. On the remainder of the contour, we set $z = \rho(1-t/n)$, *i.e.* $t = n(1-z/\rho)$ and get

$$[z^n]M(z)H(z)^{a_n} = \frac{1}{2i\pi}\int_g \frac{-\rho dt}{n\rho^{n+1}(1-t/n)^{n+1}}M(\rho(1-\tfrac{t}{n}))H(\rho(1-\tfrac{t}{n}))^{a_n},$$

where g is the image of $\gamma_0 \cup \overline{\gamma_0} \cup \gamma_1$ by the change of variable.

We write $g = g_s \cup g_b$ (for small and big), where $g_s = \{t \in g;\ \mathrm{Re}(t) < -n^{3/5}\}$ (this set is not connected) and $g_b = \{t \in g;\ \mathrm{Re}(t) \geq -n^{3/5}\}$; see again Fig. 14.

On g_s, we have

$$|1 - \tfrac{t}{n}|^{n+1} \geq (1 + n^{-2/5})^{n+1} = \Theta(e^{n^{3/5}}).$$

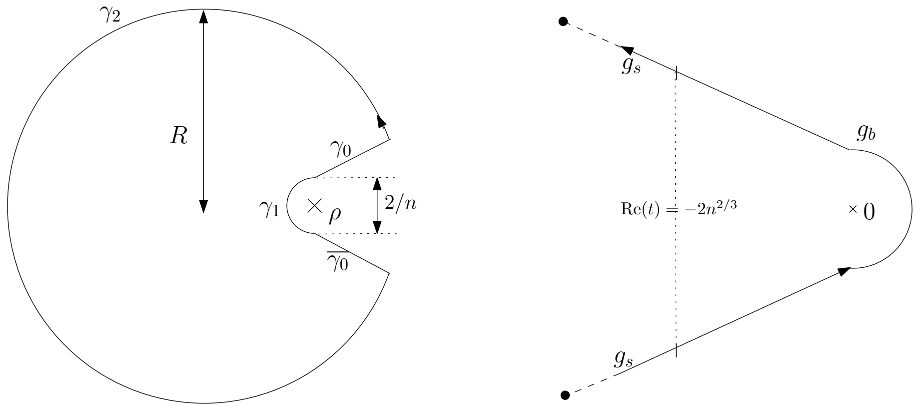

Figure 14: Left: The contour γ. Right: The contour g, which is the image of $\gamma_0 \cup \overline{\gamma_0} \cup \gamma_1$ in the variable t, where $t = n(1 - z/\rho)$.

Since H and M are bounded on the integration path, for a_n as in (74) , we have

$$\int_{g_s} M(\rho(1-\tfrac{t}{n}))H(\rho(1-\tfrac{t}{n}))^{a_n} \frac{-\rho dt}{n\rho^{n+1}(1-t/n)^{n+1}} = \mathcal{O}\big(\rho^{-n}\ e^{B'\sqrt{n}-n^{3/5}}\big),$$

for some constant B'. For t in g_b, a simple computation gives

$$M(\rho(1-\tfrac{t}{n}))H(\rho(1-\tfrac{t}{n}))^{a_n} \sim M(\rho)\sigma^{a_n} \exp\left(-\tfrac{xh\sqrt{t}}{\sigma}\right).$$

Besides $(1-\frac{t}{n})^{-n-1} = \exp(t + \mathcal{O}(\frac{t^2}{n})) = \exp(t + o(t))$. We obtain

$$\frac{1}{2i\pi}\int_{g_b} M(\rho(1-\tfrac{t}{n}))H(\rho(1-\tfrac{t}{n}))^{a_n} \frac{-\rho dt}{n\rho^{n+1}(1-t/n)^{n+1}} = -M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{2i\pi n}\int_{g_b} e^{A(n,t)}dt,$$

where

$$e^{A(n,t)} = e^t e^{\frac{-hx\sqrt{t}}{\sigma}} e^{\mathcal{O}(\frac{t^2}{n})}.$$

The big-$\mathcal{O}$ term above is uniform for x in any compact subinterval of $(0, +\infty)$. Expanding $e^{-hx\sqrt{t}/\sigma}$, putting $u = -t$, and using Hankel's formula for the Gamma function [22, Eq.(13) p. 745] yields (writing MH^a for $M(\rho(1-\frac{t}{n}))H(\rho(1-\frac{t}{n}))^{a_n}$)

$$\begin{aligned}\frac{1}{2i\pi}\int_{g_b} MH^a \frac{dz}{z^{n+1}} &\sim -M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{2i\pi n}\sum_{\ell\geq 0}\frac{(-xh/\sigma)^\ell}{\ell!}\int_{-g_b} e^{-u}(-u)^{\ell/2}(-du)\\ &\sim -M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{n}\sum_{\ell\geq 0}\frac{(-xh/\sigma)^\ell}{\ell!}\times\left(\frac{1}{\Gamma(-\ell/2)}\right).\end{aligned}$$

Now we use the complement formula $\Gamma(s)\Gamma(-s) = -\frac{\pi}{s\sin(\pi s)}$ to rewrite the RHS:

$$\frac{1}{\Gamma(-\ell/2)} = \frac{-\sin(\pi\ell/2)\Gamma(\ell/2+1)}{\pi}.$$

The sine factor vanishes for even ℓ so we are left with

$$\begin{aligned}\frac{1}{2i\pi}\int_{g_b} MH^a \frac{dz}{z^{n+1}} &\sim -M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{n}\sum_{m\geq 0}\frac{(-xh/\sigma)^{2m+1}}{\pi(2m+1)!}(-1)^m\Gamma(m+3/2)\\ &\sim M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{n}\sum_{m\geq 0}\frac{(xh/\sigma)^{2m+1}}{\sqrt{\pi}2^{2m+1}m!}(-1)^m\\ &\sim M(\rho)\sigma^{a_n}\frac{\rho^{-n}}{n}\frac{1}{\sqrt{\pi}}\mathsf{Ray}\Big(\frac{xh}{\sigma}\Big).\end{aligned}$$

We observe that the integral on g_s does not contribute to the asymptotics and we get the announced result. □

Appendix B DH-graphs form a subcritical block-stable class

We start by recalling the definition of blocks and block-stable graph classes; we refer to [32] for details. Recall first that the notion of cut-vertices and of 2-connected graphs have been defined at the beginning of Section 6. A *block* in a graph G is a maximal induced subgraph B of G without cut-vertices (of itself). A class $\mathcal{C}$ of graphs is called *block-stable* if the following holds: a graph G is in $\mathcal{C}$ if and only if all its blocks are in $\mathcal{C}$.

We now argue that DH-graphs form a block-stable class of graphs. It is known, see *e.g.* [12, Theorem 10.1], that DH-graphs are the graphs avoiding as induced subgraphs the following graphs: the house, the holes, the gem and the domino. Since all these graphs are 2-connected, an induced copy of any of them in a graph G is necessarily included in a single block of G. Therefore the avoidance of these induced subgraphs can be checked for each block separately, and the class of DH-graphs is indeed block-stable.

Since the class is block-stable, the generating series $D(z)$ of rooted DH-graphs or, equivalently of DH trees, satisfies the equation

$$D(z) = z\exp(B'(D(z))), \tag{75}$$

where B is the generating series of (unrooted) blocks (this is eq.(14) in [32]). It remains to check that the class of DH-graphs is subcritical, *i.e.* that $D(\rho)$ is smaller than the radius of convergence ρ_B of B. (Recall that D and ρ were defined in Section 3.4.)

For this, we recall that in general, blocks are either 2-connected graphs, or restricted to a single vertex or to two vertices with a single edge. Hence the series B' of rooted blocks of DH-graphs coincide, up to the coefficients of 1 and z, with the generating series $\overline{D}$ of rooted 2-connected DH-graphs. In particular, $\rho_B = \rho_{2c}$ and our analysis in Section 6.2 shows that $B''(\rho_B) = +\infty$. Consequently, there exists $\tau < \rho_B$ such that $\tau B''(\tau) = 1$. This implies that $D(z)$ belongs to the smooth inverse-function schema in the sense of [22, Definition VII.3, p.453]. From [22, Theorem VII.2], we have that $D(\rho) = \tau$. Therefore $D(\rho) < \rho_B$ as wanted, and the class of DH-graphs is indeed subcritical.

Appendix C Gromov–Hausdorff–Prohorov convergence in [32]

The main result of [32] is the convergence for the Gromov–Hausdorff topology of a uniform random graph in a subcritical block-stable class to the Brownian CRT. We argue here that without further effort, the authors could have proven convergence for the stronger Gromov–Hausdorff–Prohorov topology. We use here notation from [32]. The proof compares a uniform random graph $C_n^\bullet$ in the class and its block decomposition tree T_n. It uses the fact that the identity map from T_n to $C_n^\bullet$ does not modify much distances. Obviously this identity map brings the uniform distribution on vertices of T_n to that on vertices of $C_n^\bullet$. Therefore, using [31, Prop.6, p.763], we see that $C_n^\bullet$ and T_n are close for the GHP topology. Besides, since T_n has the distribution of a conditioned Galton–Watson tree, it is known that T_n converges to the Brownian CRT for the GHP topology (for the GH topology, a classical reference is [28]; for the GHP topology, a much stronger result is given in [26]). We conclude that $C_n^\bullet$ also converges to the Brownian CRT for the GHP topology, as claimed.

Acknowledgements

MB has been partially supported by the Swiss National Science Foundation, under grant number 200021-172536.

The authors are grateful to Anita Winter for explanations and bibliographic pointers on the relation between convergence of distance matrices, convergence for the Gromov–Prohorov distance and convergence for the Gromov–Hausdorff–Prohorov distance. We are also grateful to Éric Fusy for suggesting the argument given in Section B to justify that the class of DH-graphs is subcritical.

The comments of the referees have been particularly constructive and helpful to improve the presentation of the paper, and the authors are really grateful for their work.

Last but not least, this project started as a collaboration with Mickaël Maazoun and the authors are indebted to him for his input at the beginning.

References

[1] M. Albenque and J.-F. Marckert, Some families of increasing planar maps, *Electron. J. Probab.* 13 (2008) art.#56, 1624–1671.

[2] D. Aldous, The Continuum Random Tree III, *Ann. Probab.* 21(1) (1993), 248–289.

[3] S. Athreya, W. Löhr and A. Winter, The gap between Gromov-vague and Gromov–Hausdorff vague topology, *Stochastic Process. Appl.* 126(9) (2016), 2527–2553.

[4] M. Bahrani and J. Lumbroso, Enumerations, forbidden subgraph characterizations, and the split-decomposition, *Electron. J. Combin.* 25(4) (2018), #P4.47.

[5] C. Banderier, Ph. Flajolet, G. Schaeffer, M. Soria, Random maps, coalescing saddles, singularity analysis, and Airy phenomena, *Random Struct. Algorithms*, 19(3-4) (2001), p.194–246.

[6] F. Bassino, M. Bouvel, V. Féray, L Gerin and A. Pierrot, The Brownian limit of separable permutations, *Ann. Probab.* 46(4) (2018), 2134–2189.

[7] F. Bassino, M. Bouvel, V. Féray, L Gerin, M. Maazoun and A. Pierrot, Universal limits of substitution-closed permutation classes, *J. Eur. Math. Soc.* 22(11) (2020), 3565–3639.

[8] F. Bassino, M. Bouvel, V. Féray, L Gerin, M. Maazoun and A. Pierrot, Scaling limits of permutation classes with a finite specification: a dichotomy, *Adv. Math.* 405 (2022), art. 108513.

[9] F. Bassino, M. Bouvel, V. Féra, L. Gerin, M. Maazoun and A. Pierrot, Random cographs: Brownian graphon limit and asymptotic degree distribution, *Random Struct. Algorithms* 60(2) (2022), 166–200.

[10] J. Bettinelli, Scaling limit of random planar quadrangulations with a boundary, *Ann. Inst. Henri Poincaré (B) Probab. Stat.* 51(2) (2015), 432–477.

[11] J. Borga, M. Bouvel, V. Féray and B. Stufler, A decorated tree approach to random permutations in substitution-closed classes, *Electron. J. Probab.* 25 (2020) art.#67, 1–52.

[12] A. Brandstäd, V. B. Le and J. Spinrad, *Graph Classes: A Survey*, SIAM Monographs on Discrete Mathematics and Applications (1999).

[13] A. Caraceni, The scaling limit of random outerplanar maps, *Ann. Inst. Henri Poincaré (B) Probab. Stat.* 52(4) (2016), 1667–1686.

[14] C. Chauve, É. Fusy and J. Lumbroso, An exact enumeration of distance-hereditary graphs, in: *2017 Proc. Fourteenth Workshop on Analytic Algorithmics and Combinatorics* (ANALCO'17), 31–45.

[15] O. Cogis and E. Thierry, Computing maximum stable sets for distance-hereditary graphs, *Discrete Optim.* 2(2) (2005), 185–188.

[16] N. Curien, B. Haas and I. Kortchemski, The CRT is the scaling limit of random dissections, *Random Struct. Algorithms* 47(2) (2015), 304–327.

[17] W. Cunningham, Decomposition of directed graphs, *SIAM J. Algebraic Discrete Methods* 3 (1982), 214–228.

[18] M. Drmota, É. Fusy, M. Kang, V. Kraus and J. Rué, Asymptotic study of subcritical graph classes, *SIAM J. Discrete Math.* 25(4) (2011), 1615–165.

[19] M. Drmota and M. Noy, Extremal parameters in sub-critical graph classes, In *Proc. 2013 Meeting on Analytic Algorithmics and Combinatorics (ANALCO 13)*, 1–7.

[20] M. Drmota, "Random trees: an interplay between combinatorics and probability", Springer, 2009.

[21] Ph. Duchon, Ph. Flajolet, G. Louchard and G. Schaeffer, Boltzmann Samplers for the Random Generation of Combinatorial Structures, *Comb. Probab. Comput.* 13(4-5) (2004), 577–625.

[22] Ph. Flajolet and R. Sedgewick, "Analytic combinatorics", Cambridge University Press, 2009.

[23] E. Gioan and C. Paul, Split decomposition and graph-labelled trees: characterizations and fully dynamic algorithms for totally decomposable graphs, *Discrete Appl. Math.* 160(6) (2012), 708–733.

[24] X. Gourdon, Largest component in random combinatorial structures, *Discrete Math.* 180(1-3) (1998), 185–209.

[25] A. Greven, P. Pfaffelhuber and A. Winter, Convergence in distribution of random metric measure spaces (Λ-coalescent measure trees), *Probab. Theory Relat. Fields* 145 (1–2) (2009), 285–322.

[26] H. He and M. Winkel, Gromov–Hausdorff–Prokhorov convergence of vertex cut-trees of n-leaf Galton–Watson trees, *Bernoulli* 25(3) (2019), 2301–2329.

[27] S. Janson and S. O. Stefánsson, Scaling limits of random planar maps with a unique large face, *Ann. Probab.* 43(3) (2015), 1045–1081.

[28] J.-F. Le Gall, Random trees and applications, *Probab. Surv.* 2 (2005), 245–311.

[29] W. Löhr, Equivalence of Gromov–Prohorov and Gromov's $\underline{\square}_\lambda$-metric on th space of metric measure spaces, *Electron. Commun. Probab.* 18(17) (2013), 1–10.

[30] G. Miermont, Invariance principles for spatial multitype Galton-Watson trees, *Ann. Inst. Henri Poincaré (B) Probab. Stat.* 44(6) (2008), 1128–1161.

[31] G. Miermont, Tessellations of random maps of arbitrary genus, *Ann. Sci. Éc. Norm. Sup.* Sér.4, 42(5) (2009), 725–781.

[32] K. Panagiotou, B. Stufler and K. Weller, Scaling limits of random graphs from subcritical classes, *Ann. Probab.* 44(5) (2016), 3291–3334.

[33] B. Stufler, Graphon convergence of random cographs, *Random Struct. Algorithms*, 59(3) (2021), 464–491.

[34] Z. Varin, The golf model on $\mathbb{Z}/n\mathbb{Z}$ and on $\mathbb{Z}$, *Electron. J. Probab.* 30 (2025), 1–58.

(Received 17 May 2023; revised 17 Jan 2025, 6 June 2025)