Voronoi tessellations in the CRT and continuum random maps of finite excess

Louigi Addario-Berry* Omer Angel† Guillaume Chapuy‡ Éric Fusy§
Christina Goldschmidt¶

Abstract

Given a large graph G and k agents on this graph, we consider the Voronoi tessellation induced by the graph distance. Each agent gets control of the portion of the graph that is closer to itself than to any other agent. We study the limit law of the vector $\mathrm{Vor} := (V_1/n, V_2/n, ..., V_k/n)$, whose i'th coordinate records the fraction of vertices of G controlled by the i'th agent, as n tends to infinity. We show that if G is a uniform random tree, and the agents are placed uniformly at random, the limit law of Vor is uniform on the $(k-1)$-dimensional simplex. In particular, when $k = 2$, the two agents each get a uniform random fraction of the territory. In fact, we prove the result directly on the Brownian continuum random tree (CRT), and we also prove the same result for a "higher genus" analogue of the CRT that we call the continuum random unicellular map, indexed by a genus parameter $g \geq 0$. As a key step of independent interest, we study the case when G is a random planar embedded graph with a finite number of faces. The main idea of the proof is to show that Vor has the same distribution as another partition of mass $\mathrm{Int} := (I_1/n, I_2/n, ..., I_k/n)$ where I_j is the contour length separating the i-th agent from the next one in clockwise order around the graph.

1 Introduction

1.1 Motivation and informal presentation of the main result. We are interested in the following problem, which models the partitioning of a territory between $k \geq 2$ agents who are in competition with one another. These agents live on a large graph G, and each agent takes control of the portion of the graph which is closer to itself than to any other agent. We study the k-dimensional vector whose i'th coordinate records the fraction of vertices of G controlled by the i'th agent. More precisely, we will try to understand the law of this vector when the agents are placed uniformly at random on the graph G, and when the number of vertices of the graph becomes very large.

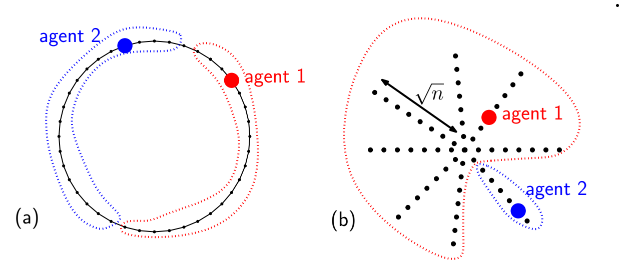

Figure 1: Examples of Voronoi competition between $k = 2$ agents on a deterministic graph. (a) On an n-cycle, each agent gets about half of the graph. (b) On a star made of $\sqrt{n}$ spikes, the winner takes almost all the graph.

Let us start with two instructive examples (Figure 1). If $k = 2$ and if G is an n-cycle, then, regardless of the positions of the agents, each of them controls $\frac{n}{2}$ vertices – up to a constant error term depending on the parity of n and the distance between them. Therefore when n tends to infinity, the "Voronoi partition of mass" converges to the *deterministic* vector $(\frac{1}{2}, \frac{1}{2})$. For the second example, again take $k = 2$ but now G is a star made of $\sqrt{n}$ spikes, each of length $\sqrt{n}$ (we omit integer parts in this informal discussion). Then when n is very large, the agent that is closest to the center of the star gets almost all the graph, namely a proportion $1 - O(n^{-1/2})$ of the vertices. Therefore, the Voronoi partition of mass converges in law to a "winner takes it all" situation, formally to an average of two Dirac laws $\frac{1}{2}\delta_{(1,0)} + \frac{1}{2}\delta_{(0,1)}$.

The two preceeding examples involve rather exceptional graphs. This motivates the idea of studying the problem on random graphs, which is what we do in this

*Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montréal, QC, H3A 0B9, Canada

†Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver, BC, V6T 1Z2, Canada

‡CNRS/IRIF, Université Paris Diderot. 75205 Paris Cedex 13, France.

§CNRS/LIX, École Polytechnique, 91120 Palaiseau, France

¶Department of Statistics and Lady Margaret Hall, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK

paper.

Our first result is an explicit determination of the limiting law in the case of (unrooted) uniform random labeled trees.

THEOREM 1. *Fix $k \geq 2$. Let T_n be a uniform random labeled tree on n vertices and let $v_1, v_2, \dots, v_k$ be k random uniform vertices of T_n. Then the k-dimensional vector whose i'th coordinate is equal to the proportion of vertices of T_n that are closer to v_i than to any other v_j converges to a uniform vector on the $(k-1)$-dimensional simplex, as n goes to infinity.*

This theorem also holds if one takes T_n to be uniform on any of the following classes of (unlabelled) trees: rooted plane trees, rooted unembedded binary trees, rooted unembedded trees, and unrooted unembedded trees; is also holds for "tree-like" families of planar maps such as stacked triangulations and outerplanar maps; and it holds (in the labelled setting) for any subcritical graph family. Indeed, the property in which we are interested is, in fact, a property of the scaling limit of these objects, which is universal and is known[1] to be the Brownian continuum random tree (CRT) for all these models, see [Ald91a, Ald91b, MM11, HM12, Stu14, AM08, Car16, PSW16]. For this reason, after this informal introduction, most of our results will be stated (and our proofs will be carried out) directly in a continuum setting. However, one should keep in mind that our results capture the behaviour of very large, but finite, random graphs.

This work was originally motivated by two rather different settings. The first is the subject of a conjecture of one of the authors, that has its roots at the intersection of mathematical physics and the probability theory of random surfaces. In [Cha16], the third author conjectures that the result of Theorem 1 holds for random plane graphs, or more generally random embedded graphs of any fixed genus g. This conjecture is still wide open, despite a remarkable paper of Guitter who proved it for $k = 2$ and $g = 0$ [Gui17b] (see also [Gui17a]). The conjecture of [Cha16] is especially intruiguing because embedded graphs of different genera g are known to belong to *different* universality classes – they have different scaling limits.

A second, less involved setting which possesses uniform Voronoi partitions is the *stochastic mean-field model of distance*. Take the complete graph K_n on n vertices, and assign independent and identically distributed random lengths to its edges, with Exp(1) distribution. Then we may grow the Voronoi cells of k uniformly chosen agents as follows. Let us work on the event of high probability (for large n) that the k agents are distinct. Cells grow along the edges at speed 1, and their sizes are given by the numbers of vertices they contain. We start with k cells of size 1. Suppose that at some time t, the cells have sizes $s_1, s_2, \dots, s_k$ and a new vertex has just been added to some cell. Then the closest vertex among the remaining $n - s_1 - \cdots - s_k$ vertices to any of the cells is at distance given by the minimum of $(s_1 + \cdots + s_k)(n - s_1 - \cdots - s_k)$ independent Exp(1) random variables (here, we use the memoryless property of the exponential distribution). Moreover, it is closest to cell i with probability $s_i/(s_1 + \cdots + s_k)$. It follows that if we just look at the instants when some cell grows, then the cell sizes perform a k-color Pólya's urn process, run until there are n balls in total in the urn. It is well known that the proportions of the different colors converge, as n tends to infinity, to the uniform distribution on the $(k-1)$-dimensional simplex.

This unexpected coincidence motivates the following question:

Which models of random graphs give rise to uniform random Voronoi partitions of mass, and why?

In this work we show that the CRT has this property. In fact, we show much more. A tree, once embedded in the plane, is nothing but a plane graph with a unique face, which led us to the following generalisation: we study the Voronoi vector in the case when G is a graph embedded in a fixed surface, with a fixed number of faces, and a fixed number of agents in each face – under the uniform measure, when the graph becomes large. We characterize the limit law of this vector in terms of another random vector, related to the interval lengths, which is easier to analyse. In particular, we exhibit an infinite family of models that have uniform Voronoi partitions of mass, namely embedded graphs with one face on any fixed surface. We also obtain the explicit limit law in the case of planar graphs with a bounded number of faces.

Is it worth noticing the striking similarity between our results and the conjecture of [Cha16]. Both papers deal with random embedded graphs, and both involve (proved here and conjecturally there) uniform Voronoi partitions of mass. However, this similarity is very mysterious, since the graphs we study in this paper are *very different* from uniform random embedded graphs.

[1]Strictly speaking, the papers [AM08, Car16, PSW16] only prove that the CRT is the scaling limit of these objects for the Gromov-Hausdorff topology, whereas we need here the Gromov-Hausdorff-Prokhorov topology (see below). However, it is not hard to see that the arguments in these papers also enable one to prove this stronger convergence. We thank a referee for this remark.

Indeed, we consider a regime in which graphs have a finite number of cycles (finite excess), while the random embedded graphs considered in [Cha16] have a linear excess with high probability. Their scaling limits are also very different [LG13, Mie13]. We are not able to explain, not even at a heuristic level, why their Voronoi partitions would behave similarly.

We end this informal introduction here and proceed to a more precise presentation of our results. As explained above, the statements of our results are more transparent at the level of continuum objects, so we will start with a short presentation of the required formalism.

1.2 On the formalism chosen for graph limits.

In order to deal with Voronoi partitions of mass, it is convenient to think of graphs as *measured metric spaces*. Formally, this means that for us a graph will be a triple $(V(G), d, \mu)$ where $V(G)$ is the set of vertices, d the graph distance on $V(G)$, and μ the uniform measure on $V(G)$. (We will often abuse notation and write (G, d, μ) for the same object.) This point of view enables us to take scaling limits in the sense of the *Gromov-Hausdorff-Prokhorov distance*, see e.g. [ADH13]. For example, if (T_n, d_n, μ_n) is a uniform random tree on n vertices, then we have the convergence in distribution [Ald91a, LG05]:

$$\left(T_n, \frac{1}{\sqrt{n}} d_n, \mu_n\right) \longrightarrow (T, d, \mu),$$

where (T, d, μ) is a *continuum* random metric space called the Brownian continuum random tree (CRT). In all the examples of random graphs considered in this paper, it is straightforward to see that the Voronoi partition of mass of the scaling limit is the limit in distribution of the Voronoi partition of mass of the discrete random graphs. This follows using Proposition 10 of [Mie09], which shows that GHP convergence of a sequence of measured metric spaces implies the convergence (in the appropriate topology) of the same spaces equipped with k independent uniformly-sampled points. For example, Theorem 1 is a direct consequence of Theorem 2 below.

1.3 Main results.

Theorem 2. (Trees) *Let (T, d, μ) be the CRT, with distance d and mass measure μ. Let $X_1, \ldots, X_k \in T$ be k independent samples from μ in the tree T. Partition the tree into k Voronoi cells $C_1, \ldots, C_k$, where C_i is the part of T closer to X_i than to any other X_j. Then the vector $\mathrm{Vor}(T) := (\mu(C_i))_{i \le k}$ is a uniform partition of unity (i.e., uniformly distributed on the $(k-1)$-dimensional simplex).*

We now state the extension to higher genus. The continuum random unicellular map (CRUM) is a generalization of the CRT that is defined for each connected orientable surface. If the surface is the sphere, the CRUM is the CRT. For a general surface S, the CRUM on S is the scaling limit of random embedded graphs on S having a single face, in the same way that the CRT is the scaling limit of random plane trees.

Theorem 3. (Unicellular maps) *Fix an orientable surface S, and let (M, d, μ) be the CRUM on S. Let $X_1, \ldots, X_k \in M$ be k independent samples from μ in M. Partition M into k Voronoi cells $C_1, \ldots, C_k$. Then the vector $\mathrm{Vor}(M) := (\mu(C_i))_{i \le k}$ is a uniform partition of unity.*

In fact, our more general result (Theorem 6 below) enables us to deal with random embedded graphs on any surface *and* with an arbitrary number of faces. It takes an explicit form in genus 0. For $\ell \ge 1$ and $k_1, k_2, \ldots, k_\ell \ge 1$, the *CRM of signature* $(0; k_1, k_2, \ldots, k_\ell)$, denoted $CRM(0; k_1, k_2, \ldots, k_\ell)$ for short, can be thought of as a continuum random plane graph, with ℓ faces numbered from 1 to ℓ, and having k_i marked vertices in the i'th face for each $i \in [1..\ell]$. The $CRM(0; k_1, k_2, \ldots, k_\ell)$ is the scaling limit of random plane graphs with ℓ numbered faces having k_i marked vertices in the i'th face, under the uniform measure when their total number of vertices tends to infinity. We have:

Theorem 4. (Plane graphs) *For $\ell \ge 1$ and for $k_1, \ldots, k_\ell \ge 1$, let (M, d, μ) be the $CRM(0; k_1, k_2, \ldots, k_\ell)$, and let $X_1^{(i)}, \ldots, X_{k_i}^{(i)} \in M$ denote the distinguished points in the i'th face, for $i \in [1..\ell]$. Let*

$$\mathrm{Vor}(M) = (\mu(C_j^{(i)}))_{\substack{1 \le i \le \ell \\ 1 \le j \le k_i}}$$

denote the Voronoi partition of mass induced by the marked points, where $C_j^{(i)}$ is the subset of M formed by points closer to $X_j^{(i)}$ than to any other of the $\sum_{i=1}^{\ell} k_i$ marked points. Then $\mathrm{Vor}(M)$ has the same law as the vector

$$(D_i U_j^{(i)})_{\substack{1 \le i \le \ell \\ 1 \le j \le k_i}}$$

where $(D_1, D_2, \ldots, D_\ell)$ follows a Dirichlet distribution of parameters $\left(k_1 + \frac{1}{2}, \ldots, k_\ell + \frac{1}{2}\right)$ and where, for $i \in [1..\ell]$, the vector $(U_j^{(i)})_{1 \le j \le k_i}$ is a uniform partition of unity, all of these variables being independent.

We emphasize once again that the two previous theorems give the limit law of the Voronoi cells in the corresponding model of *finite* random plane graphs, when their size tends to infinity. We also observe that the case of the CRT is contained in the last theorem, with $\ell = 1$ and $k_1 = k$.

1.4 Plan of the paper. In Section 2 we consider the case of trees. This will also be an introduction to our proof strategy, which consists in showing, by induction on the number k of agents, that the vector $\mathrm{Vor}(T)$ has the same distribution as another vector called the "interval vector". As announced we will work directly on continuum structures, and in fact we will use a reduction to a model of random graphs with exponential edge-lengths. In Section 3 we apply the same program to graphs with more faces embedded on any orientable surface[2]. The strategy is similar, but there are more cases to consider in the induction. Finally in Section 4, we present a bijective proof of the asymptotic equidistribution between the interval and the Voronoi vectors, in the case of trees, that works directly at the discrete level. This bijection more or less follows the arguments given in the proofs on continuum objects, but by constructing it explicitly we hope to emphasize that our results truly apply to finite objects.

1.5 Additional remarks. In a later version of this work, we will present the proofs for non-orientable surfaces, that follow a similar pattern yet require new ideas. A curious consequence of this more general result is that the following classes of uniform random *unembedded* graphs have uniform random Voronoi partitions of mass in the limit: random unicycles, random barbell graphs, and random theta graphs (the case of theta graphs follows already from the present paper, namely from Theorem 3 when S is a torus; the other two follow by considering CRUM's on the projective plane and on the Klein bottle, respectively).

1.6 Acknowledgments. This work was started and continued at the wonderful 11'th and 12'th workshops on Probability and Combinatorics held at the Bellairs Institute of McGill University in Barbados in April 2016 and April 2017. We thank the other participants for interesting discussions, especially Marie Albenque, Igor Kortchemski and Remco van der Hofstad. We also acknowledge the hospitality of the excellent program *combi17: Combinatorics and Interactions* held at IHP in Paris in 2017, where some of the crucial ideas were discovered. LAB was partially supported by NSERC Discovery Grant 341845. OA was supported in part by NSERC. GC's research is supported by the European Research Council, grant ERC-2016-STG 716083 "CombiTop". EF's research is supported by the ANR grant "GATO", ANR-16-CE40-0009-01. CG's research is supported by EPSRC Fellowship EP/N004833/1.

[2]In fact our results hold for arbitrary surfaces, orientable or not, but we prefer to keep the present paper shorter and focus on orientable surfaces only.

2 Voronoi and interval vectors in random trees

Let us start by taking a paragraph to give a formal definition of the CRT. This is standard, but we make rather extensive use of the idea that the CRT may be endowed with a contour order, and we prefer to specify precisely what this means. As alluded to above, the CRT is a random metric space, (T, d, μ). We define it using a standard Brownian excursion. This is a certain random continuous function $\mathbb{e} : [0,1] \to [0,\infty)$ such that $\mathbb{e}(0) = \mathbb{e}(1) = 0$ and $e(t) > 0$ for $t \in (0,1)$; the precise details of its distribution are unimportant for the present discussion, but it may be usefully thought of as the scaling limit of an excursion of simple random walk. First define a pseudo-metric d on $[0,1]$ via

$$d(s,t) = 2\mathbb{e}(s) + 2\mathbb{e}(t) - 4 \min_{s\wedge t \le r \le s\vee t} \mathbb{e}(r).$$

Now define an equivalence relation $\sim$ by declaring $s \sim t$ if $d(s,t) = 0$. Then set T to be $[0,1]/\sim$, endowed with the metric d. We write $p : [0,1] \to T$ for the projection map onto the tree, and let μ be the push-forward of the Lebesgue measure on $[0,1]$ onto T (we refer to this as the uniform measure on T). It is a (non-trivial) fact that points picked from μ are with probability 1 leaves (that is, points whose removal does not disconnect the space). The leaves of the metric space T inherit a natural *clockwise contour order* from this encoding (where we imagine starting at the equivalence class of 0): if $s,t \in [0,1]$ with $s < t$ and $p(s)$ and $p(t)$ are leaves, then we say that $p(s)$ precedes $p(t)$ in the clockwise ordering.

Let $A = \{X_1,\ldots,X_k\}$, where $X_1,\ldots,X_k$ are k independent samples from μ in T, listed in contour order. Concretely, let $U_1,\ldots,U_k$ be the order statistics of k independent and identically distributed uniform random variables on $[0,1]$, and set $X_i = p(U_i)$ for $1 \le i \le k$. For $1 \le i \le k$, let C_i be the Voronoi cell of X_i:

$$C_i = \{y \in T : \ d(y,A) = d(y,X_i)\}.$$

We let $\mathrm{Vor}(T)$ be the vector $(\mu(C_1),\ldots,\mu(C_k)\}$, called the *Voronoi vector* of T.

On the other hand the clockwise contour of T is split into k "intervals" $I_1,\ldots,I_k$, where I_i is the part of the contour between X_i and X_{i+1}, namely $I_i = p([U_i,U_{i+1}])$ for $1 \le i \le k-1$ and $I_k = p([U_k,1] \cup [0,U_1])$. We let $\mathrm{Int}(T)$ be the vector $(\mu(I_1),\ldots,\mu(I_k))$, called the *interval vector* of T, and note that $\mu(I_i) = U_{i+1} - U_i$ for $1 \le i \le k-1$ and $\mu(I_k) = 1 - U_k + U_1$. The main result we obtain in this section is:

THEOREM 5. *Let (T,μ,d) be the Brownian CRT. Let $X_1,\ldots,X_k$ be k independent samples from μ in T, listed* *in contour order. Then* $\mathrm{Vor}(T)$ *and* $\mathrm{Int}(T)$ *have the same distribution.*

Theorem 5 easily implies Theorem 2 (hence also Theorem 1), since it is clear that the interval vector is a uniform partition of unity.

REMARK 1. *At the level of discrete structures, let* $\mathcal{T}_n^k$ *be the set of plane trees* T *with* n *edges and* k *marked corners of respective labels* $1, \ldots, k$*, the label-ordering being consistent with the ordering of a clockwise walk around the tree (starting from the corner of label 1). For* $1 \leq i \leq k$ *the corner of label* i *is denoted* c_i *and its incident vertex is denoted* a_i*. The vertices* $a_1, \ldots, a_k$ *are called the* marked vertices *of* T*, and the set of these vertices is denoted by* A*.*

The contour of T *is split into* k *intervals* $I_1, \ldots, I_k$ *(with* I_i *the part of the contour between* c_i *and* c_{i+1}*) and we let* $u_i := \mathrm{length}(I_i)/(2n)$*, and set* $\mathrm{Int}(T) := (u_1, \ldots, u_k)$*. On the other hand, we define* C_i *to be the Voronoi cell of the marked vertex* a_i*, i.e., the set of vertices* p *such that* $d(p, a_i) = d(p, A)$*; and we let* $v_i := |C_i|/n$*, and set* $\mathrm{Vor}(T) := (v_1, \ldots, v_k)$*. Then Theorem 5 entails that, for* T *a uniformly random tree in* $\mathcal{T}_n^k$*, as* $n \to \infty$*,* $\mathrm{Vor}(T)$ *and* $\mathrm{Int}(T)$ *are asymptotically equidistributed. (Note that here we are picking uniform corners rather than uniform vertices; this is equivalent to picking points uniformly in the domain of the discrete contour process. Since the discrete contour process, suitably rescaled, converges in distribution to* $2\mathbb{e}$*, it is easy to see that this also yields uniform vertices in the limit.) We will present in Section 4 a bijection on* $\mathcal{T}_n^k$ *that explains this result.*

Our first proof of Theorem 5, presented in this section, proceeds on the k-leaf skeletons of the CRT rather than on the CRT itself. As we briefly recall here, these can be seen as the scaling limits of the kernels of random trees in $\mathcal{T}_n^k$ as $n \to \infty$. For $T \in \mathcal{T}_n^k$, the *core* of T is the tree $\xi(T)$ obtained from T by greedily deleting every non-marked leaf (together with its incident edge) until only marked leaves remain. The *kernel* of T is then the tree $\kappa(T)$ obtained from $\xi(T)$ by erasing every non-marked vertex of degree 2, see Figure 2 for an example. Note that to every edge $e \in \kappa(T)$ corresponds a path P in $\xi(T)$ such that all the non-extremal vertices of P are non-marked of degree 2, and each of the two extremities of P is either marked or of degree greater than 2.

We now define a *tree-skeleton* to be a plane tree S with all vertices of degree 1 or 3, with the k leaves carrying distinct labels in $\{1, \ldots, k\}$ and such that the labels $1, \ldots, k$ occur in clockwise order around S. The leaf of label i is denoted a_i. We let $\mathcal{S}^k$ be the set of tree-skeletons with k leaves (note that, for $k \geq 3$, a tree in $\mathcal{S}^k$ can be turned bijectively into a rooted binary tree with $k-2$ nodes, upon considering the node connected to the leaf of label 1 as the root-node, hence $|\mathcal{S}^k|$ is the $(k-2)$'th Catalan number). A *tree-skeleton with edge-lengths* is defined to be a tree-skeleton S where every edge e is assigned a positive real value called its *length*, and denoted $\lambda(e)$. We denote by $\overline{\mathcal{S}}^k$ the set of tree-skeletons with edge-lengths and k leaves, $a_1, \ldots, a_k$. Any $S \in \overline{\mathcal{S}}^k$ is a 1-dimensional metric space, each edge $e \in S$ being considered as a line-segment of length $\lambda(e)$. The total length of S is denoted $L(S)$. Let $A = \{a_1, \ldots, a_k\}$ be the set of leaves. For $1 \leq i \leq k$ the *Voronoi cell* of a_i is the (connected) set $C_i := \{p \in S : \ d(p, a_i) = d(p, A)\}$. We let $w_i := \lambda(C_i)$ be the total length of C_i; the vector $(w_1, \ldots, w_k)$ is called the *Voronoi vector* of S and is denoted by $\mathrm{Vor}(S)$ (note that if there is no 1-dimensional intersection of Voronoi cells, then the components of $\mathrm{Vor}(S)$ add up to $L(S)$). On the other hand, the contour of S is split into k intervals $I_1, \ldots, I_k$, with I_i the part of the contour between a_i and a_{i+1}. We let $z_i := \lambda(I_i)$ be the total length of I_i; the vector $(z_1, \ldots, z_k)$ is called the *Interval vector* of S and is denoted $\mathrm{Int}(S)$ (note that the components of $\mathrm{Int}(S)$ add up to $2L(S)$).

We will consider random tree-skeletons under some specific distributions. The *Exp-uniform* distribution on $\overline{\mathcal{S}}^k$ consists in picking a tree-skeleton $S \in \mathcal{S}^k$ uniformly at random, and then assigning to each edge $e \in S$ an independent length $\lambda(e)$ following an Exp(1) distribu-

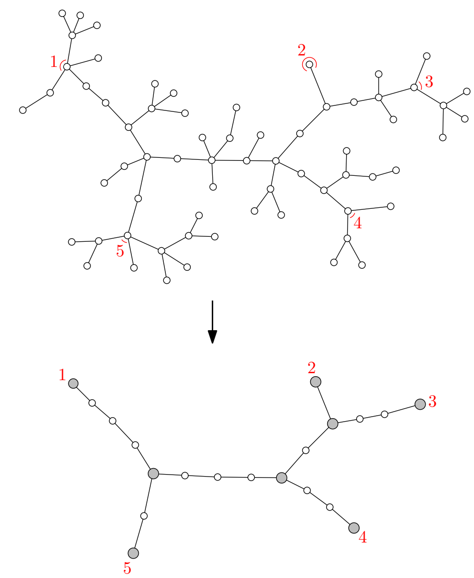

Figure 2: Top: a plane tree with 5 marked corners. Bottom: the associated core (the vertices that also belong to the kernel are colored gray).

tion. Equivalently, if the edges of S are (arbitrarily) ordered as $e_1, \dots, e_{2k-3}$, then we may sample the total length L following a Gamma law of parameter $2k-3$, and then sample a uniformly random split of L, as $L = \lambda_1 + \dots + \lambda_{2k-3}$, and take λ_i to be the length of e_i. The *Tree-uniform distribution* on $\overline{\mathcal{S}}^k$ is defined similarly, except that L is now distributed according to the law of density

$$(2.1) \qquad f_m(x) = \frac{2}{\Gamma(\frac{m+1}{2})} x^m e^{-x^2}, \quad x \geq 0,$$

with $m = 2k-3$ the number of edges in the skeleton (this law also corresponds to the square-root of a Gamma law of parameter $k-1$).

Let T be a tree chosen uniformly at random from $\mathcal{T}_n^k$, and let $\kappa(T)$ be its kernel (where each edge of the kernel carries a length-parameter to record the length of the corresponding path in the core). Then it is well known (see equation (49) of [Ald93]) that, as $n \to \infty$ and the edge-lengths in $\kappa(T)$ are divided by $\sqrt{n}$, $\kappa(T)$ converges to a random tree-skeleton in $\overline{\mathcal{S}}^k$ with the Tree-uniform distribution. At the continuous level, one can also directly extract the skeleton $S \in \overline{\mathcal{S}}^k$ out of the CRT T, and alternatively one can also directly generate S (under the Tree-uniform distribution) by the so-called *line-breaking* construction [Ald91a] (which will not be used here).

With this useful formalism introduced, we can now proceed with the proof of Theorem 5, which we split into 3 lemmas. The first lemma says that it is enough to prove equidistribution at the level of the skeletons, while the second one enables us to work under the Exp-uniform distribution. The core of our argument is contained in the proof of the third lemma.

Lemma 1. *Showing Theorem 5 reduces to showing the following statement:*

($\star$) *For $S \in \overline{\mathcal{S}}^k$ under the Tree-uniform distribution, the vectors* $2\mathrm{Vor}(S)$ *and* $\mathrm{Int}(S)$ *are equidistributed.*

Proof. Let T be the CRT with its k ordered uniform random leaves $X_1, \dots, X_k$. Let S be the skeleton extracted from T, which is in $\overline{\mathcal{S}}^k$ under the Tree-uniform distribution. $T \setminus S$ consists of a (countably infinite) collection of subtrees. Let $I_1, \dots, I_k$ be the interval vector of S and, for $1 \leq i \leq k$, let C_i be the Voronoi cell of X_i in S. Then it is easy to see that the i'th interval of T may be recovered by just adding back in the subtrees of T attached to S at some point of I_i, and the i'th Voronoi cell may be recovered by adding the subtrees attached to S at some point of C_i. Now, the subtrees of T attached to S have exchangeable masses and are attached independently to uniformly random points of S (this is perhaps most easily seen using Aldous' line-breaking construction [Ald91a]) and so if the vectors $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ have the same distribution, so do the vectors they induce in T.

The next lemma follows from the fact that multiplying the distances in a tree by an independent random constant has the same effect on both the Voronoi and interval vectors.

Lemma 2. *The statement ($\star$) in Lemma 1 is equivalent to the following statement:*

($\star\star$) *For $S \in \overline{\mathcal{S}}^k$ under the Exp-uniform distribution, the vectors* $2\mathrm{Vor}(S)$ *and* $\mathrm{Int}(S)$ *are equidistributed.*

It now remains to show the following (which is actually the core of the proof):

Lemma 3. *The statement ($\star\star$) in Lemma 2 holds true.*

Proof. The proof is by induction on $k \geq 2$. For $k = 2$ the result is trivial. Indeed in that case, S consists of a single edge of length L (following an $\mathrm{Exp}(1)$ distribution) connecting the two leaves, and we have $\mathrm{Vor}(S) = (L/2, L/2)$ and $\mathrm{Int}(S) = (L, L)$.

Now assume $k \geq 3$. We refer to Figure 3 for an illustration of our discussion. Let $S \in \overline{\mathcal{S}}^k$ be sampled from the Exp-uniform distribution. Let e be the shortest edge incident to a leaf. The edge e is called the *fuse-edge*, and its extremity of degree 1 (resp. 3) is called the *fuse-leaf* (resp. *fuse-node*); we denote by ν the fuse-node. The two edges after e in counterclockwise order around ν are denoted e', e''; and the corner between e' and e'' is called the *fuse-end corner*.

Consider the operation of simultaneously "burning" (starting from the leaf) the extremity of length x of each of the k edges connected to a leaf, i.e., the length of every such edge is decreased by x as in Figure 3. The fuse-edge is completely burned and the fuse-node ν has its degree decreased to 2. We now cut at ν (i.e., we split it into two leaves) so that we get two components U, V that are tree-skeletons with edge-lengths, where U (resp. V) is the component containing e' (resp. e''). The *root-leaf* of U (resp. V) is the one resulting from splitting ν. Let p be the number of leaves of U and q the number of leaves of V (note that $p, q \geq 2$ and $p + q = k + 1$, so that p and q are smaller than k).

Consider the root-leaves of U, V to have label 1. Since the fuse-leaf is a uniformly random leaf of S, every tree-skeleton in $\mathcal{S}^p$ (resp. $\mathcal{S}^q$) is equally likely for U (resp. for V). Moreover by the memoryless property of the exponential law, the edge-lengths in U and in V are independent and follow an $\mathrm{Exp}(1)$ distribution.

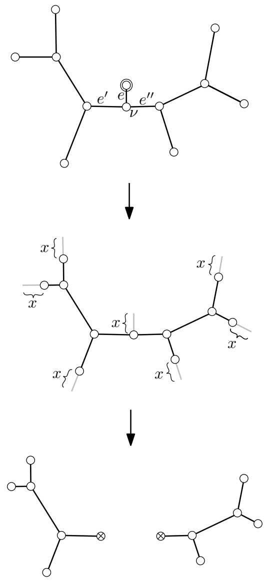

Figure 3: Top: a tree-skeleton in $\overline{\mathcal{S}}^7$, where the fuse-leaf is surrounded (leaf labels are not indicated). Middle: burning a part x of every edge incident to a leaf, where x is the length of the fuse-edge. Bottom: the two components resulting from cutting at the fuse-node (the root-leaf in each component is crossed).

Hence U (resp. V) follows the Exp-uniform distribution on $\overline{\mathcal{S}}^p$ (resp. $\overline{\mathcal{S}}^q$). Let $\mathrm{Vor}(U) = (u_1, \ldots, u_p)$ and $\mathrm{Vor}(V) = (v_1, \ldots, v_q)$, and let ℓ be the label (in S) of the fuse-leaf. Note that $\mathrm{Vor}(U)$ and $\mathrm{Vor}(V)$ are independent when conditioning only on p and ℓ. Let us introduce the following notation: for $\vec{z} = (z_1, \ldots, z_r)$ a vector and $i \in [1..r]$, the i-shift of $\vec{z}$ is the cyclic shift of $\vec{z}$ that puts z_1 in the ith position; for $x \in \mathbb{R}$, $x + \vec{z}$ denotes the vector $(z_1 + x, \ldots, z_r + x)$. Clearly, if the fuse-leaf has label ℓ, then $\mathrm{Vor}(S)$ is the ℓ-shift of the vector $x + (u_1 + v_1, v_2, \ldots, v_q, u_2, \ldots, u_p)$.

Regarding interval vectors, we change a little our conventions to better establish the parallel with Voronoi vectors. We now call V the component containing e', and let q be its number of leaves, and call U the component containing e'', and let p be its number of leaves. We still consider the root-leaf of V to have label 1 within V, but the root-leaf of U is now considered to have label 2 (so that it is at the end of the 1st interval of U). This is just a single shift of the labels in U and does not change the fact that U follows the Exp-uniform distribution on $\overline{\mathcal{S}}^p$ (and V also follows the Exp-uniform distribution on $\overline{\mathcal{S}}^q$). Let $\ell \in [1..k]$ be the label of the interval (in S) that contains the fuse-end corner (i.e., this interval starts at the leaf of label ℓ). Let $\mathrm{Int}(U) = (\tilde{u}_1, \ldots, \tilde{u}_p)$ and $\mathrm{Int}(V) = (\tilde{v}_1, \ldots, \tilde{v}_q)$ (note that $\mathrm{Int}(U)$ and $\mathrm{Int}(V)$ are independent when conditioning only on p and ℓ). Then it is easy to see that $\mathrm{Int}(S)$ is the ℓ-shift of the vector $2x + (\tilde{u}_1 + \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_q, \tilde{u}_2, \ldots, \tilde{u}_p)$.

Hence, if for $\ell \in [1..k]$ and $p \in [2..k-1]$ we let $\mathcal{E}_{p,\ell}$ be the event that the component containing e' has p leaves and the fuse-leaf has label ℓ, and let $\tilde{\mathcal{E}}_{p,\ell}$ be the event that the component containing e'' has p leaves and the interval containing the fuse-end corner is the ℓ-th interval, then we find that $\mathrm{Vor}(S)$ conditioned on $\mathcal{E}_{p,\ell}$ has the same distribution as $\mathrm{Int}(S)$ conditioned on $\tilde{\mathcal{E}}_{p,\ell}$ (we use induction on the number of leaves, which guarantees that $2\,\mathrm{Vor}(U)$ is distributed as $\mathrm{Int}(U)$ and $2\,\mathrm{Vor}(V)$ is distributed as $\mathrm{Int}(V)$). Since the events $\mathcal{E}_{p,\ell}$ and $\tilde{\mathcal{E}}_{p,\ell}$ clearly have the same probability (which is equal to $\frac{|\mathcal{S}^p||\mathcal{S}^q|}{k|\mathcal{S}^k|}$), we conclude that $2\,\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed on $\overline{\mathcal{S}}^k$ under the Exp-uniform distribution.

Remark 2. *This concludes the proof of Theorem 2. The proof of Lemma 3 can be formulated as a correspondence on $\overline{\mathcal{S}}^k$ (preserving the Exp-uniform distribution) that transforms* $2\mathrm{Vor}(S)$ *into* $\mathrm{Int}(S)$*, while involving $k-1$ cutting/merging operations. As we will explain in Section 4 the analogous mapping on discrete trees can be formulated as an effective bijection that maps (asymptotically) the Voronoi vector to the interval vector.*

3 Voronoi and interval vectors in random maps of fixed signature

We now generalize the results of the previous section to the context of maps with fixed genus and a fixed number of faces. After introducing the framework of maps of fixed signature and describing their scaling limits, we will establish the equidistribution of the Voronoi and interval vectors in this setting.

3.1 Maps of fixed signature and their scaling limits. A *map* M is a connected graph G (possibly with loops and multiple edges) embedded on a compact orientable surface Σ, such that all components of $\Sigma \backslash G$ are homeomophic to a topological disk; these components are called the *faces* of M. The *genus* of M is the genus of the surface Σ on which it embeds. In the following, we will rather use the *ribbon-graph* representation of a map, i.e., a map is just a connected graph where a cyclic order is specified for the incident half-edges around every vertex. This information is enough to retrieve the contours of the faces, and also the genus of the map (using the Euler relation).

Let $\sigma = (g; k_1, \ldots, k_r)$ be a sequence of integers, where $g \geq 0$, and the k_i's are all positive. A map

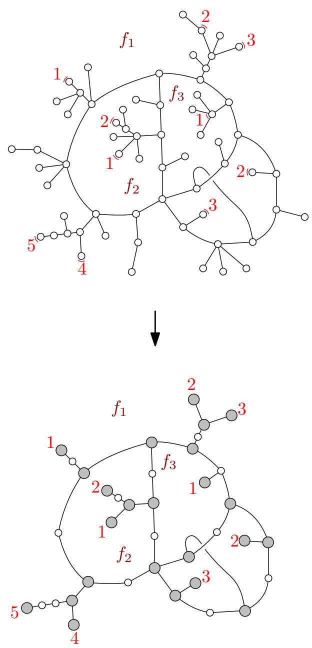

Figure 4: Top: a map of signature $\sigma = (1; 5, 2, 3)$ (in the ribbon graph representation). Bottom: the associated core (the vertices that also belong to the kernel are drawn gray).

M is said to be of *signature* σ if it has genus g and r faces labelled as $f_1, \ldots, f_r$, such that there are k_i marked corners $c_{i,1}, \ldots, c_{i,k_i}$ in f_i (among the $\deg(f_i)$ corners), labelled from 1 to k_i so that these corners occur in ccw order around f_i, see Figure 4 (left part) for an example. The *marked vertices* of such a map are the vertices incident to at least one marked corner. Let $\mathcal{M}_n^\sigma$ be the set of maps of signature σ with n edges.

As in the previous section, for $M \in \mathcal{M}_n^\sigma$, the *core* $\xi(M)$ is the map obtained from M by greedily deleting every non-marked leaf (and the incident edge) until all leaves are marked, and the *kernel* $\kappa(M)$ is obtained from $\xi(M)$ by erasing all non-marked vertices of degree 2 (see Figure 4). Every edge of $\kappa(M)$ is naturally equiped with a *length-parameter* to record the length of the corresponding path in $\xi(M)$. We now define a *map-skeleton* of signature σ as a map of signature σ where all vertices have degree in $\{1, 3\}$, and the marked vertices are all the leaves. The set of map-skeletons of signature σ is denoted by $\mathcal{S}^\sigma$. By the Euler relation the number $m \equiv m(\sigma)$ of edges of every $S \in \mathcal{S}^\sigma$ is given by

$$m(\sigma) = 6g - 6 + 3r + 2(k_1 + \cdots + k_r).$$

A *map-skeleton with edge-lengths* is a map-skeleton S where every edge is assigned a positive value called its length. The total length over all edges is denoted $L(S)$, and the set of map-skeletons of signature σ with edge-lengths is denoted $\overline{\mathcal{S}}^\sigma$. A distribution on $\overline{\mathcal{S}}^\sigma$ is called *length-uniform* if it is obtained by:

- picking $S \in \mathcal{S}^\sigma$ uniformly at random, and ordering its m edges arbitrarily as $e_1, \ldots, e_m$;
- drawing (independently of S) the total length L of S under a given distribution with support on $\{x > 0\}$;
- drawing a random split of L into m parts, as $L = \lambda_1 + \cdots + \lambda_m$, and assigning length λ_i to e_i for $1 \le i \le m$.

The *Exp-uniform* distribution on $\overline{\mathcal{S}}^\sigma$ is the one where L is drawn from a Gamma law of parameter m; this is equivalent to drawing independently the length of every edge according to an Exp(1) law. On the other hand, the *Map-uniform distribution* on $\overline{\mathcal{S}}^\sigma$ is the one where L is drawn under the probability density $f_m(x)$ given by (2.1). When $n \to \infty$ and M is drawn uniformly at random in $\mathcal{M}_n^\sigma$, the kernel $\kappa(M)$, with its edge-lengths divided by $\sqrt{n}$, converges to a random map-skeleton $S \in \overline{\mathcal{S}}^\sigma$ under the Map-uniform distribution.

It is also possible to describe directly the scaling limit of M. (We do not know of a reference to precisely define this scaling limit theorem in the literature, but it may be deduced from results of [ABBG10].) The so-called *continuum random map* of signature σ (abbreviated to CRM$^\sigma$) is described as follows:

- draw a skeleton $S \in \mathcal{S}^\sigma$ uniformly at random, and order (arbitrarily) its edges as $e_1, \ldots, e_m$;
- independently sample a vector $(X_1, \ldots, X_m)$ with Dirichlet distribution of parameters $(1/2, 1/2, \ldots, 1/2)$;
- independently sample m independent CRTs $T_1, \ldots, T_m$, each with two uniformly chosen points;
- rescale the tree T_i by X_i i.e., multiply the mass measure by X_i and multiply distances by $X_i^{1/2}$;
- for $1 \le i \le m$, replace the edge e_i of S by the rescaled version of T_i, so that the two uniform points match with the extremities of e_i.

3.2 Equidistribution of the Voronoi and Interval vectors.

Let (M, d, μ) be the CRM$^\sigma$. For $1 \le i \le r$ and $1 \le j \le k_i$ let $a_{i,j}$ be the leaf of label j in the face f_i. Let $I_{i,j}$ be the part of M between $a_{i,j}$ and $a_{i,j+1}$ in the ccw contour of f_i. We let $\mathrm{Int}(M, i) := (\mu(I_{i,1}), \ldots, \mu(I_{i,k_i}))$, called the ith interval vector of M. The concatenation $(\mathrm{Int}(M, 1), \ldots, \mathrm{Int}(M, r))$ of these vectors is called the *interval-vector* of M and is denoted by $\mathrm{Int}(M)$. Moreover, we let $A := \{a_{i,j},\ 1 \le i \le r,\ 1 \le j \le k_i\}$ be the set of all labelled leaves and, for $1 \le i \le r$ and $1 \le j \le k_i$ we denote by

$$C_{i,j} := \{p \in M,\ d(p, a_{i,j}) = d(p, A)\}$$

the Voronoi cell of the leaf $a_{i,j}$. We let $\mathrm{Vor}(M, i) := (\mu(C_{i,1}), \ldots, \mu(C_{i,k_i}))$, called the ith Voronoi vector of M. The concatenation $(\mathrm{Vor}(M, 1), \ldots, \mathrm{Vor}(M, r))$ of these vectors is called the *Voronoi-vector* of M and is denoted by $\mathrm{Vor}(M)$.

Our main result is the following generalization of Theorem 5:

THEOREM 6. *Let $\sigma = (g; k_1, \ldots, k_r)$ where $g \ge 0$, $k_1, \ldots, k_r$ are positive, and $g + k_1 + \cdots + k_r \ge 2$. Let (M, d, μ) be the CRM^σ. Then $\mathrm{Vor}(M)$ and $\mathrm{Int}(M)$ have the same distribution.*

Note that the law of $\mathrm{Int}(M)$ (and thus also of $\mathrm{Vor}(M)$) can be described as follows. If we let $g^\sigma(t_1, \ldots, t_r)$ be the joint probability density function (with support on $\{t_1 + \cdots + t_r = 1\}$) of the contour-lengths of $f_1, \ldots, f_r$ in the CRM^σ, then the law of $\mathrm{Int}(M)$ is as follows:

- $L_1, \ldots, L_r$ is drawn according to the density $g^\sigma(t_1, \ldots, t_r)$,
- for $1 \le i \le r$, $\mathrm{Int}(M, i)$ is drawn as a uniform random split of L_i into k_i parts.

In principle, it should be possible to compute algorithmically the density $g^\sigma(t_1, \ldots, t_r)$ for any fixed σ. When $g = 0$, this density is easily obtained from exact enumeration results (for example from Tutte's slicings formula [Tut62]) and this gives Theorem 4.

When $r = 1$, the CRM^σ is a continuum random map with only one face, that we call the *continuum random unicellular map of genus g*, or CRUM. Of course the density g^σ is trivial when $r = 1$ (it is a Dirac distribution at 1). In that case, for $\sigma = (g; k)$, M corresponds to a CRUM of genus g with k random leaves, and $\mathrm{Vor}(S)$ is thus a uniform split of 1 into k parts, which yields Theorem 3.

Proof of Theorem 6. Similarly to the tree case, and by the same arguments as in Lemma 1, we can first reduce proving Theorem 6 to proving the following statement:

($\star$) for $S \in \overline{\mathcal{S}}^\sigma$ under the Map-uniform distribution, the random vectors $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed.

As in Lemma 2, the latter is equivalent to:

($\star\star$) for $S \in \overline{\mathcal{S}}^\sigma$ under the Exp-uniform distribution, the random vectors $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed.

As for trees (but with more cases) we prove ($\star\star$) by induction on the number $m = m(\sigma)$ of edges, the *size* of the signature. In the case $m = 1$, the signature has to be $(0; 2)$ (tree with two leaves) and the property has already been proved in Lemma 3. We now assume that $m \ge 2$ (thus all leaves are adjacent to a vertex of degree 3), and that ($\star\star$) holds for all signatures of smaller size. For smaller size it will be convenient to allow the set F of face-labels to be any given set of integers (not necessarily of the form $\{1, \ldots, h\}$), which does not affect the equidistribution property.

Pick up $S \in \overline{\mathcal{S}}^\sigma$ under the Exp-uniform distribution. The strategy of the proof is again to look for the shortest edge e among those incident to leaves. We keep the same terminology and notation as for tree-skeletons, i.e., the length of e is denoted by x, the extremity of e of degree 1 (resp. 3) is called the *fuse-leaf* (resp. the *fuse-node* ν); the two edges after e in ccw order around ν are denoted e', e''; and the corner between e' and e'' is called the *fuse-end corner*. The face containing the fuse-leaf is called the *fuse-face*. We then consider the *burning operation* that consists in shortening by x the length of every edge incident to a leaf; this way the fuse-leaf is merged with ν, which gets of degree 2, and we can then cut at ν to turn it into two leaves (one at e', the other at e''). We let $\hat{S}$ be the resulting 'map' (it might not be connected) with edge-lengths. We have now 3 possible situations, as shown in Figure 5 (for trees, only the first case was possible):

Cut: $\hat{S}$ has two connected components (in that case ν is necessarily incident to a single face in S).

Split: the fuse-node ν is incident to a single face f in S, but $\hat{S}$ is still connected. In that case, cutting at ν has the effect of splitting f into two faces, while decreasing the genus by 1.

Merge: the fuse-node ν is incident to two different faces f, f' (where by convention f is the one incident to the fuse-leaf). In that case, $\hat{S}$ is still connected, the genus remains unchanged, and cutting at ν has the effect of merging f' into f.

We are going to show that conditioned on each of the 3 cases, the vectors $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed.

Proof in the cut case. This case is similar to trees, with some more notation. We let i_0 be the label of the fuse-face f of S, and let ℓ be the label of the fuse-leaf within f. Let U, V be the two connected components of $\hat{S}$,

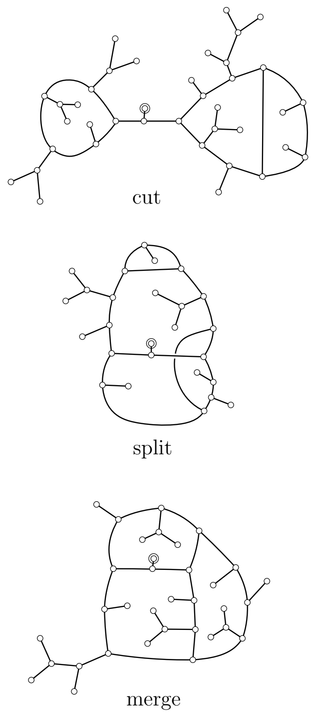

Figure 5: The 3 possible cases when cutting at the fuse-node ν (in each case the fuse-leaf is surrounded, labels are not indicated).

with U (resp. V) the one containing e' (resp. e''). In U and V the *root-face* is the face resulting from splitting f, and the *root-leaf* is the one resulting from splitting ν. In U and V the face-labels are retained to be the same as in S, but the leaves in the root-face are relabelled starting from 1 and increasing by 1 in ccw order around the face, with label 1 at the root-leaf.

Let σ_1 be the signature of U and σ_2 the signature of V. Since the fuse-leaf acts as a random leaf of S, and by the memoryless property of the exponential law, U (resp. V) is in $\overline{\mathcal{S}}^{\sigma_1}$ (resp. in $\overline{\mathcal{S}}^{\sigma_2}$) under the Exp-uniform distribution. Let $B \subset [1..r]\backslash\{i_0\}$ be the set of labels of the non-root faces that appear in U. Then clearly, for $i \neq i_0$, we have (with the notation $x + \vec{v}$ introduced in the proof of Lemma 3):

$$\mathrm{Vor}(S, i) = x + \mathrm{Vor}(U, i) \quad \text{if } i \in B,$$

$$\mathrm{Vor}(S, i) = x + \mathrm{Vor}(V, i) \quad \text{if } i \notin B.$$

Let p (resp. q) be the number of leaves of U (resp. V) within its root-face. Let $(u_1, \ldots, u_p) := \mathrm{Vor}(U, i_0)$ and $(v_1, \ldots, v_q) := \mathrm{Vor}(V, i_0)$. Then, as for trees, it is easy to see that (with the definition of ℓ-shift given in the proof of Lemma 3)

$$\mathrm{Vor}(S, i_0) = \ell\text{−shift of } x + (u_1 + v_1, v_2, \ldots, v_q, u_2, \ldots, u_p).$$

Regarding interval vectors, as for trees, we proceed similarly except that ℓ is now the label of the interval containing the fuse-end corner, and U is the component containing e'', and V the one containing e'. Another difference is that the root-leaf is given label 2 in U (so that it is at the end of the first interval in the root-face of U). Again if σ_1 (resp. σ_2) denotes the signature of U (resp. V) then U is in $\overline{\mathcal{S}}^{\sigma_1}$ (resp. in $\overline{\mathcal{S}}^{\sigma_2}$) under the Exp-uniform distribution. Letting $B \subset [1..r]\backslash\{i_0\}$ be the set of labels of the non-root faces that appear in U, we get for $i \neq i_0$

$$\mathrm{Int}(S, i) = 2x + \mathrm{Int}(U, i) \quad \text{if } i \in B,$$

$$\mathrm{Int}(S, i) = 2x + \mathrm{Int}(V, i) \quad \text{if } i \notin B.$$

Moreover, with p (resp. q) the number of leaves of U (resp. V) in the root-face, and with $(\tilde{u}_1, \ldots, \tilde{u}_p) := \mathrm{Int}(U, i_0)$ and $(\tilde{v}_1, \ldots, \tilde{v}_q) := \mathrm{Int}(V, i_0)$, we have

$$\mathrm{Int}(S, i_0) = \ell\text{−shift of } 2x + (\tilde{u}_1 + \tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_q, \tilde{u}_2, \ldots, \tilde{u}_p).$$

Hence, for Vor (resp. for Int), if we condition on the label of the fuse-face and on the signature of the component U containing e' (resp. e''), on the set of face-labels in U and on the label ℓ of the fuse-leaf (resp. of the interval containing the fuse-end corner), we find (using the fact that the equidistribution property holds for smaller sizes) that the conditioned vector $2\mathrm{Vor}(S)$ is distributed as the conditioned vector $\mathrm{Int}(S)$. Summing over all possible conditioning events we conclude that $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed in the cut-case.

Proof in the split case. The proof is similar to the cut case (in both cases a face is split into two faces, the only difference is that in the cut case we get two connected components). Regarding the Voronoi vectors, we let i_0 be the label (in S) of the fuse-face f, and ℓ the label of the fuse-leaf. The face f is split into two faces f', f'', with f the one containing e' and f'' the one containing e''. We keep label i_0 for f' and give label 0 to f''. The number of leaves in f' (resp. f'') is denoted p (resp. q) and the leaf of label 1 in f' (resp. f'') is the one incident to e' (resp. e''). Let $\hat{\sigma} = (g-1, q, k_1, \ldots, p, \ldots, k_r)$ (with p replacing k_{i_0}) be the induced signature of $\hat{S}$. Then, as before, $\hat{S}$ is in $\overline{\mathcal{S}}^{\hat{\sigma}}$ under the Exp-uniform distribution. We have

$$\mathrm{Vor}(S, i) = x + \mathrm{Vor}(\hat{S}, i) \quad \text{for } i \in \{1, \ldots, r\}\backslash\{i_0\}.$$

We let $(u_1, \ldots, u_p) := \mathrm{Vor}(\hat{S}, i_0)$ and $(v_1, \ldots, v_q) := \mathrm{Vor}(\hat{S}, 0)$. Then we have

$$\mathrm{Vor}(S, i_0) = \ell\text{−shift of } x + (u_1 + v_1, v_2, \ldots, v_q, u_2, \ldots, u_p).$$

For the interval vectors, we now let ℓ be the label of the interval containing the fuse-end corner, give label i_0 to

f'' and label 0 to f', and let p be the number of leaves in f'' and q the number of leaves in f'. In f' we still give label 1 to the leaf at e', but in f'' we now give label 2 to the leaf at e''. As before $\hat{S}$ follows the Exp-uniform distribution for the induced signature $\hat{\sigma}$. We have

$$\mathrm{Int}(S,i) = 2x + \mathrm{Int}(\hat{S},i) \quad \text{for } i \in \{1,\ldots,r\}\backslash\{i_0\}.$$

Let $(\tilde{u}_1,\ldots,\tilde{u}_p) := \mathrm{Int}(\hat{S},i_0)$ and $(\tilde{v}_1,\ldots,\tilde{v}_q) := \mathrm{Int}(\hat{S},0)$. Then we have

$$\mathrm{Int}(S,i_0) = \ell\text{−shift of } 2x + (\tilde{u}_1+\tilde{v}_1,\tilde{v}_2,\ldots,\tilde{v}_q,\tilde{u}_2,\ldots,\tilde{u}_p).$$

Similarly to the cut case, by summing over all possible conditioning events and using induction, we conclude that $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed in the split case.

Proof in the merge case. In this case, the fuse-face f is merged with the face f' on the other side of ν (the face containing the fuse-end corner). Regarding the Voronoi vectors, we let (in S) i_0 be the label of f, i_1 the label of f', ℓ_0 the label of the fuse-leaf (within f), ℓ_1 the label of the interval containing the fuse-end corner (within f'), p the number of leaves in f and q the number of leaves in f'. In $\hat{S}$ we give label i_0 to the merged face, which has $k = p+q+1$ leaves; the leaf of label 1 within the merged face is chosen to be the one at the edge e''. Again $\hat{S}$ follows the Exp-uniform distribution for the induced signature $\hat{\sigma}$. We have

$$\mathrm{Vor}(S,i) = x + \mathrm{Vor}(\hat{S},i) \quad \text{for } i \in \{1,\ldots,r\}\backslash\{i_0,i_1\}.$$

If we let $(u_1,\ldots,u_k) := \mathrm{Vor}(\hat{S},i_0)$ then we have

$$\mathrm{Vor}(S,i_0) = \ell_0\text{−shift of } x + (u_1+u_{p+1},u_2,\ldots,u_p),$$
$$\mathrm{Vor}(S,i_1) = \ell_1\text{−shift of } x + (u_k,u_{p+2}\ldots,u_{k-1}).$$

Regarding the interval vectors, we now let (in S) i_0 be the label of f' and ℓ_0 the label of the interval containing the fuse-end corner, let i_1 be the label of f and ℓ_1 the label of the interval (in f) along the edge e', and let p be the number of leaves of f' and q the number of leaves of f. In $\hat{S}$ we give label i_0 to the merged face, that has $k = p+q+1$ leaves; the leaf of label 1 within the merged face is chosen to be the one at the edge e'. Again $\hat{S}$ follows the Exp-uniform distribution for the induced signature $\hat{\sigma}$. We have

$$\mathrm{Int}(S,i) = 2x + \mathrm{Int}(\hat{S},i) \quad \text{for } i \in \{1,\ldots,r\}\backslash\{i_0,i_1\}.$$

If we let $(\tilde{u}_1,\ldots,\tilde{u}_k) := \mathrm{Int}(\hat{S},i_0)$ then we have

$$\mathrm{Int}(S,i_0) = \ell_0\text{−shift of } 2x + (\tilde{u}_1+\tilde{u}_{p+1},\tilde{u}_2,\ldots,\tilde{u}_p),$$
$$\mathrm{Int}(S,i_1) = \ell_1\text{−shift of } 2x + (\tilde{u}_k,\tilde{u}_{p+2}\ldots,\tilde{u}_{k-1}).$$

As in the other cases, we conclude (summing over all conditioning events) that $2\mathrm{Vor}(S)$ and $\mathrm{Int}(S)$ are equidistributed in the merge-case.

This concludes the proof of Theorem 6. ☐

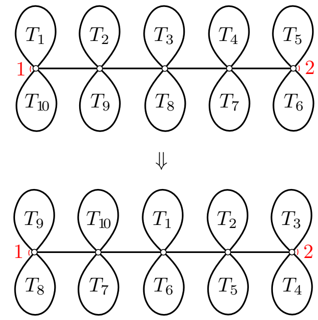

Figure 6: Example of the bijection for $k = 2$ ($s = 5$ here).

4 A bijective proof of Theorem 5

We present here a proof of Theorem 5 as an explicit bijection on (a subfamily of) $\mathcal{T}_n^k$ that asymptotically maps (with the definitions given after the statement of Theorem 5) the Voronoi vector to the interval vector. Let us mention that a bijective proof of Theorem 6 can also be obtained, along very similar lines.

Before describing the construction, we introduce a bit of terminology. For $T \in \mathcal{T}_n^k$ and $A = \{a_1,\ldots,a_k\}$ the set of marked vertices, the *strict Voronoi cell* of a_i is the set of vertices $v \in T$ that are not in the core of T, and which are such that $d(v,a_i) < d(v,a_j)$ for $j \neq i$. The *strict i'th interval* of T is the set of vertices in the i'th interval that are not in the core of T.

We prove the following:

Theorem 7. *For every fixed $k \geq 2$ there is a subfamily $\mathcal{A}_n^k \subset \mathcal{T}_n^k$, satisfying $|\mathcal{A}_n^k|/|\mathcal{T}_n^k| = 1 - O(n^{-1/2})$, and a bijection Φ on $\mathcal{A}_n^k$ that preserves the size (number of vertices) of the core, and such that for $T \in \mathcal{A}_n^k$ and $R = \Phi(T)$, every vertex in the strict Voronoi cell of a_i in T corresponds to a vertex in the strict i'th interval of R, for each $1 \leq i \leq k$.*

The bijection (which essentially emulates the construction on tree-skeletons given in the proof of Lemma 3) works by induction on $k \geq 2$; the subfamily $\mathcal{A}_n^k$ is also defined by induction (simultaneously with the bijection) and satisfies $\mathcal{A}_n^2 = \mathcal{T}_n^2$. We present it first for $k = 2$ (where the construction is easier) and then for $k \geq 3$.

4.1 The bijection for $k = 2$.

Let $T \in \mathcal{T}_n^2$. Let P be the path connecting the two marked vertices a_1, a_2, let $s := \mathrm{length}(P) + 1$. Starting at the marked corner c_1 and turning clockwise around P we see a sequence of $2s$ attached subtrees $T_1,\ldots,T_s,T_{s+1},\ldots,T_{2s}$. Then $R = \Phi(T)$ is obtained by keeping P untouched and "sliding" the sequence of attached subtrees along, to the effect that T_i replaces $T_{(i+\lfloor s/2\rfloor) \bmod 2s}$, see Figure 6. Note that the vertices of the strict Voronoi cell (in T) of a_1 are those in the subtrees $T_{2s-\lfloor s/2\rfloor+1}, \ldots, T_{2s}, T_1, \ldots, T_{\lfloor s/2\rfloor}$. The sliding operation lets these trees replace the trees $T_1, \ldots, T_{2\lfloor s/2\rfloor}$, which are in the first interval of T. Similarly the vertices of the strict Voronoi cell of a_2 are sent to vertices in the strict second interval of R. This construction thus proves Theorem 7 for $k = 2$.

4.2 The bijection for $k \geq 3$. The construction for $k \geq 3$, illustrated in Figure 7, works by induction on k, so we assume that Theorem 7 already holds for all values in $\{2, \ldots, k-1\}$. Let $T \in \mathcal{T}_n^k$, and let $\xi(T)$ be the core of T, and $\kappa(T)$ the kernel of T, where every edge e is equipped with a length $L(e)$ to record the length of the associated path in $\xi(T)$. If $\kappa(T)$ has vertices of degree not in $\{1, 3\}$ or if there is a marked vertex that is not a leaf of $\kappa(T)$, then we report a failing situation and abort the contruction (the subfamily $\mathcal{A}_n^k$ will be specified as the subfamily of $\mathcal{T}_n^k$ where no failing situation has occured). Otherwise, we let e be the shortest edge, called the *fuse-edge*, among the k edges of $\kappa(T)$ that are connected to a leaf (if there is a tie, we again report a failing situation and abort). The extremity of e of degree 1 (resp. 3) is called the *fuse-leaf* (resp. *fuse-node* ν). We let ℓ be the label of the fuse-leaf. The two edges (in $\xi(T)$) after e in ccw order around ν are denoted e', e''; and the corner of $\xi(T)$ between e' and e'' is called the *fuse-end corner*. The subtree attached at that corner is denoted Z.

Let b be the length of e; and for $i \in \{1, \ldots, k\}$ let P_i be the path of length b in $\xi(T)$ that ends at the leaf a_i; we let G_i (resp. H_i) be the set of subtrees that are attached at the b positions that precede (resp. follow) c_i in a cw walk along $\xi(T)$ (the first tree in G_ℓ is denoted Y and will be used later). We then delete P_i and the attached groups G_i, H_i, and update the marked corner c_i accordingly. This way ν gets degree 2 in the core, and it has a marked corner (inherited from c_ℓ). We then cut at ν, doing the incision at the marked corner inherited from c_ℓ, and so that Z sits in the right-hand part, as shown in Figure 7 (2nd picture). After cutting, we get two trees U, V, where U is the one containing e' (resp. e''). In each of the two trees there is a marked corner resulting from cutting at ν, which we consider as marked and call the root-corner. We move the retained subtree Y to the position just after the root-corner of U; and we then relabel the marked corners in U (resp. V) in the unique way such that the root-corner gets label 1.

Let p (resp. q) be the number of marked corners in U (resp. V); note that $p, q \geq 2$ and $p+q = k-1$, so p and q are smaller than k and we can use induction. If $U \notin \mathcal{A}^p$ or $V \notin \mathcal{A}^q$ (with the general notation $\mathcal{A}^j = \cup_n \mathcal{A}_n^j$) then

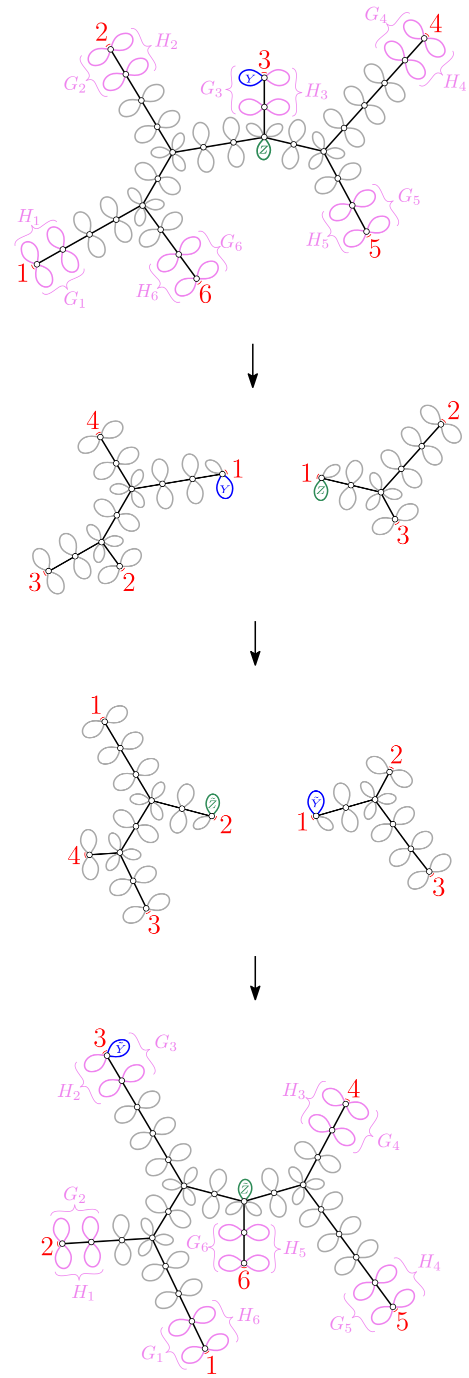

Figure 7: Example of the bijection for $k \geq 3$ (here $k = 6$ and $b = 2$, attached subtrees on the core are represented as blobs).

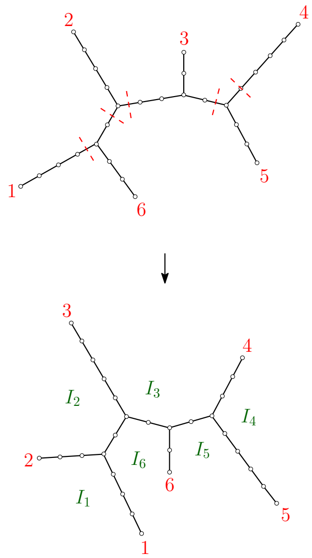

Figure 8: Property of the bijection: if T is a tree and R its image by the bijection, then the 1-dimensional Voronoi vector of the core of T (here equal to $(3.5, 5, 6, 4, 4.5, 5)$, the dashed lines indicating the frontiers between cells), when multiplied by two, matches with the interval vector of the core of R (here equal to $(7, 10, 12, 8, 9, 10)$).

we report a failing situation and abort. Otherwise we consider the images $\Phi(U)$ and $\Phi(V)$ of U and V by the bijection. We let $\tilde{Z}$ be the attached subtree just before corner 2 in $\Phi(U)$, and $\tilde{Y}$ be the attached subtree just after corner 1 in $\Phi(V)$, which we detach from $\Phi(V)$. As shown in the right-hand part of Figure 7, we can then merge corner 2 of $\Phi(U)$ with corner 1 of $\Phi(V)$ (so that their first intervals get merged), and let a path of length b grow out of every marked corner and out of the newly formed vertex (this path is grown in the opposite direction from the one it grew in T), and then we label the marked corners so that the merged interval gets label ℓ (hence the label of this interval, which is 3 in Figure 7, matches with the label of the fuse-leaf of T). Finally we reinsert onto the grown paths the attached trees that were taken from T, in such a way that the groups G_i and H_i (those in the ith Voronoi cell of T) sit in the ith interval of the obtained tree (for the special case of the first tree in G_ℓ, we put the detached tree $\tilde{Y}$ there). We declare the obtained tree $R \in \mathcal{T}_n^k$ to be the image of T by Φ; and we let $\mathcal{A}_n^k$ be the set of trees in $\mathcal{T}_n^k$ where the construction succeeds (no failing situation occurs).

It is easy to derive the inverse mapping and check (by induction on k) that Φ is a bijection that maps $\mathcal{A}_n^k$ to itself. By induction on k one can also check that attached subtrees (on the core of T) contributing to the strict i'th Voronoi cell of T are sent to attached subtrees (on the core of R) contributing to the strict i'th interval of R. To that effect a closely related property (also readily checkable by induction on k) is that if we consider the core of T to be a 1-dimensional metric space (as we did with skeletons in Section 2), then the vector giving the total lengths of the respective cells matches (upon multiplication by 2) the vector giving the lengths of the intervals of the core of R; see Figure 8.

References

[ABBG10] L. Addario-Berry, N. Broutin, and C. Goldschmidt. Critical random graphs: limiting constructions and distributional properties. *Electron. J. Probab.*, 15:no. 25, 741–775, 2010.

[ADH13] R. Abraham, J.-F. Delmas, and P. Hoscheit. A note on the Gromov-Hausdorff-Prokhorov distance between (locally) compact metric measure spaces. *Electron. J. Probab.*, 18:no. 14, 21pp, 2013.

[AM08] M. Albenque and J.-F. Marckert. Some families of increasing planar maps *Electron. J. Probab.*, 13(56):1624–1671, 2008.

[Ald91a] D. Aldous. The continuum random tree. I. *Ann. Probab.*, 19(1):1–28, 1991.

[Ald91b] D. Aldous. The continuum random tree. II. An overview. In *Stochastic analysis (Durham, 1990)*, volume 167 of *London Math. Soc. Lecture Note Ser.*, pages 23–70. Cambridge University Press, Cambridge, 1991.

[Ald93] D. Aldous. The continuum random tree. III. *Ann. Probab.*, 21(1):248–289, 1993.

[Car16] A. Caraceni. The scaling limit of random outerplanar maps. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1667–1686, 2016.

[Cha16] G. Chapuy. On tessellations of random maps and the t_g-recurrence. arXiv:1603.07714 [math.PR], 2016.

[Gui17a] E. Guitter. A universal law for Voronoi cell volumes in infinitely large maps. arXiv:1706.08809 [math.CO], 2017.

[Gui17b] E. Guitter. On a conjecture by Chapuy about Voronoi cells in large maps. *J. Stat. Mech. Theory Exp.*, 2017(10):103401, 2017.

[HM12] B. Haas and G. Miermont, Scaling limits of Markov branching trees with applications to Galton–Watson and random unordered trees. *Ann. Probab.*, 40(6):2589–2666, 2012.

[LG05] J.-F. Le Gall. Random trees and applications. *Probab. Surv.*, 2:245–311 (electronic), 2005.

[LG13] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.*, 41(4):2880–2960, 2013.

[MM11] J.-F. Marckert and G. Miermont. The CRT is the scaling limit of unordered binary trees. *Random Structures Algorithms*, 38:467–501, 2011.

[Mie09] G. Miermont. Tessellations of random maps of arbitrary genus. *Ann. Sci. Éc. Norm. Supér. (4)*, 42(5):725–781, 2009.

[Mie13] G. Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.*, 210(2):319–401, 2013.

[PSW16] K. Panagiotou, B. Stufler, and K. Weller. Scaling limits of random graphs from subcritical classes. *Ann. Probab.*, 44(5):3291–3334, 2016.

[Stu14] B. Stufler. The continuum random tree is the scaling limit of unlabelled unrooted trees. arXiv:1412.6333 [math.PR], 2014.

[Tut62] W. T. Tutte. A census of slicings. *Canad. J. Math.*, 14:708–722, 1962.