LOCAL CONVERGENCE FOR PERMUTATIONS AND LOCAL LIMITS FOR UNIFORM \( \rho \)-AVOIDING PERMUTATIONS WITH \(|\rho| = 3\)

JACOPO BORGA

Abstract. We set up a new notion of local convergence for permutations and we prove a characterization in terms of proportions of consecutive pattern occurrences. We also characterize random limiting objects for this new topology introducing a notion of “shift-invariant” property (corresponding to the notion of unimodularity for random graphs).

We then study two models in the framework of random pattern-avoiding permutations. We compute the local limits of uniform \( \rho \)-avoiding permutations, for \(|\rho| = 3\), when the size of the permutations tends to infinity. The core part of the argument is the description of the asymptotics of the number of consecutive occurrences of any given pattern. For this result we use bijections between \( \rho \)-avoiding permutations and rooted ordered trees, local limit results for Galton–Watson trees, the Second moment method and singularity analysis.

Contents

1. Introduction 2
   1.1. Scaling and local limits for discrete structures 2
   1.2. Overview of our results 4
   1.3. Future projects and open problems 7
   1.4. Outline of the paper 8
2. Local convergence for permutations 8
   2.1. Permutations and patterns 8
   2.2. The set of rooted permutations 9
   2.3. Local convergence for rooted permutations 11
   2.4. Benjamini–Schramm convergence: the deterministic case 14
   2.5. Benjamini–Schramm convergence: the random case 15
   2.6. Characterization of the annealed Benjamini–Schramm limits 24
3. Trees and random trees, a toolbox 26
   3.1. Rooted ordered trees and Galton–Watson trees 27
   3.2. Depth-first traversal 28
   3.3. Binary trees 28
4. The local limit for uniform random 231-avoiding permutations 29
   4.1. Notation and preliminary results 31
   4.2. The annealed version of the Benjamini–Schramm convergence 34
   4.3. The quenched version of the Benjamini–Schramm convergence 42
   4.4. The construction of the limiting object 43
   4.5. Some examples of limiting probabilities 47
5. The local limit for uniform random 321-avoiding permutations 48
   5.1. From uniform 321-avoiding permutations to Galton–Watson trees 49
   5.2. A local limit result for Galton–Watson trees pointed at a uniform vertex 51
   5.3. Contour function for trees 52
   5.4. The proof of the main result 55

2010 Mathematics Subject Classification. 60C05, 05A05.
Key words and phrases. local weak limits, permutation patterns.
1. Introduction

The goal of this article is to introduce a local topology for permutations and to study some initial interesting applications.

1.1. Scaling and local limits for discrete structures. It is well-known that, after a suitable rescaling, a simple random walk on $\mathbb{Z}$ converges to the 1-dimensional Brownian motion: probably this is the most famous scaling limit result in probability theory. In the last thirty years such scaling limits have been intensively studied for different discrete structures. We mention some mathematical areas where this notion has been investigated, without aiming at giving a complete list.

In the framework of random trees and planar maps, the systematic study of scaling limits has been initiated by Aldous with the pioneering series of articles about the well-known Continuum random tree (3, 4, 5). After that, many new results have been proven, in particular about the Brownian map, which is the scaling limit of random planar maps uniformly distributed over the class of all rooted $q$-angulations with $n$ faces, for $q = 3$ and $q \geq 4$ even integer (see for instance 38, 42, 43).

In statistical mechanics, the study of scaling limits of discrete models is also a very active topic. For instance, the scaling limit of the discrete Gaussian free field on lattices, i.e., the continuum GFF, has been proven to be the scaling limit of many others random structures (see for instance 34, 19).

More closely related to this work, in 27 a notion of scaling limits for permutations has been recently introduced, called permutons. They are probability measures on the unit square with uniform marginals, and they represent the scaling limit of the diagram of permutations as the size grows to infinity. This new notion of convergence has been studied (sometimes phrased in other terms) in several works. We mention a few of them.

• The study of the scaling limits of a uniform random permutation avoiding a pattern of length three was initiated by Miner and Pak [44] and Madras and Pehlivan [41]. Next, with a series of two articles, Hoffman, Rizzolo and Slivken (24, 25) strengthened these early results and understood many interesting phenomena that had previously gone unexplained. In particular they explored the connection of these uniform pattern-avoiding permutations to Brownian excursions. All these articles do not use the “permuton language”.

• The first concrete and explicit example of convergence in the “permuton language” has been studied by Kenyon, Kral, Radin and Winkler [35]. They studied scaling limits of random permutations in which a finite number of pattern densities have been fixed.

• Bassino, Bouvel, Féray, Gerin and Pierrot [10] showed that a sequence of uniform random separable permutations of size $n$ converges to the Brownian separable permuton. In this work they introduced this new limiting object that has been later investigated by Maazoun [40].

• In a second work, Bassino, Bouvel, Féray, Gerin, Maazoun and Pierrot [9] showed that the Brownian permuton has a universality property: they consider uniform random permutations in proper substitution-closed classes and study their limiting behavior in the sense
of permutons, showing that the limit is an elementary one-parameter deformation of the Brownian separable permuton.

- With a statistical mechanical approach, Starr [50] investigates the permuton limit of Mallows permutations. This is a non-uniform model where the probability of every permutation is proportional to $q^{\text{number of inversions}}$. Treating the question as a mean-field problem, he is able to calculate the distribution of the permuton limit (after a suitable rescaling of the parameter $q$).

- Rahman, Virag and Vizer [48] generalized the study of scaling limits for permutations to permutation valued processes, i.e., sequences of sequences of permutations.

- Permutons limits have been characterized in terms of convergence of frequencies of pattern occurrences (see [9, Theorem 2.5]). We will come back later to this characterization.

In parallel to scaling limits, a second notion of limits of discrete structures, on which this paper focuses, has been introduced: the local limits. Informally, scaling limits look at the convergence of the objects from a global point of view (after a rescaling of the distances between points of the objects), while local limits look at discrete objects in a neighborhood of a distinguished point (without rescaling distances). As done for scaling limits, we allude to some frameworks where this notion has been studied, again without aiming at giving a complete overview.

Local limit results around the root of random trees were first implicitly proved by Otter [46] and then explicitly by Kesten [36] and Aldous and Pitman [6]. Janson [29] gives a unified treatment of the local limits, as the number of vertices tends to infinity, of simply generated random trees. Recently, Stufler [52] studied also the local limits for large Galton–Watson trees around a uniformly chosen vertex building on previous result of Aldous [2] and Holmgren and Janson [26].

Although implicit in many earlier works, the notion of local convergence around a random vertex (called weak local convergence) for random graphs has been formally introduced by Benjamini and Schramm [11] and Aldous and Steele [7]. One would expect that a limit object for this topology should “look the same” when regarded from any of its vertices (because of the uniform choice of the root). This property is made precise by the notion of unimodular random rooted graph and the weak local limit of any sequence of random graphs is indeed proven to be unimodular (see [11]).

Also in the framework of random planar maps the local convergence has been studied. For example, a well-known result is that the local limit for random infinite triangulations/quadrangulations is the uniform infinite planar triangulation/quadrangulation (UIPT/UIPQ) (see for instance [8, 37, 51]).

In the framework of permutations, to the best of our knowledge, a local limit approach has not been investigated so far. The goal of the current paper is to fill this gap and show several interesting aspect of the local convergence for permutations. More precisely,

- we prove a characterization of the local convergence in terms of proportions of consecutive pattern occurrences (see Theorems 1.3 and 1.5). For the sake of comparison, recall that permutons limits have been characterized in terms of convergence of frequencies of (non-consecutive) pattern occurrences.

- We characterize random limiting objects for this new topology (see Theorem 1.6), introducing a notion of “shift-invariant” property (corresponding to the notion of unimodularity for random graphs).

- We exhibit some concrete applications of the developed theory in the setting of pattern-avoiding permutations (see Theorems 1.7 and 1.8).

Remark 1.1. We wish to mention that Hoffman, Rizzolo and Slivken [23] used local limits result for random trees in the study of some specific properties of permutations, like the number of fixed points (i.e., indexes $i$ such that $\sigma_i = i$) of pattern-avoiding permutations. But they do not study general local limits results for permutations.
Remark 1.2. We also mention that recently Pinsky [47] studied limits of random permutations avoiding patterns of size three considering a different topology. His topology captures the local limit of the permutation diagram around a corner. The two topologies (that from [47] and that from the present article) are not comparable.

The reader not familiar with the terminology of permutation patterns (such as pattern occurrences, pattern avoidance, etc.) can find the necessary background in Section 2.1.

1.2. Overview of our results. Our work can be divided into two main parts.

1.2.1. Local convergence for permutations. In the first part (Section 2) we develop the theory of local convergence for permutations. We will often view a permutation $\sigma$ as a diagram, i.e., the set of points of the Cartesian plane at coordinates $(j,\sigma_j)$ (see Fig. 1 for an example). In the context of local convergence we need to look at permutations with a distinguished entry, called root. Denoting with $S^n$ the set of permutations of size $n$, we say that a pair $(\sigma,i)$ is a finite rooted permutation of size $n$ if $\sigma \in S^n$ and $i \in \{1,\ldots,n\}$.

To a rooted permutation $(\sigma,i)$, we associate a total order $\preceq_{\sigma,i}$ on a finite interval of integers containing 0 denoted by $A_{\sigma,i}$. The total order $(A_{\sigma,i},\preceq_{\sigma,i})$ is simply obtained from the diagram of $\sigma$ (as shown in the left-hand side of Fig. 1). We shift the indices of the $x$-axis in such a way that the column containing the root of the permutation has index zero (the new indices are reported under the columns of the diagram). Then we set $j \preceq_{\sigma,i} k$ if the point in column $j$ is lower than the point in column $k$.

![Figure 1](image-url)

Since this correspondence defines a bijection, we identify every rooted permutation $(\sigma,i)$ with the total order $(A_{\sigma,i},\preceq_{\sigma,i})$. Thanks to this identification, it is natural to call infinite rooted permutation a pair $(A,\preceq)$ where $A$ is an infinite interval of integers containing 0 and $\preceq$ is a total order on $A$.

In order to define a notion of local convergence, we introduce a notion of $h$-restriction around the root. It can be thought as the diagram of the pattern induced by a “vertical strip” of width $2h+1$ around the root of the permutation (see Fig. 1) or equivalently of the restriction of the order $(A,\preceq)$ to $A \cap [-h,h]$. Note that this second equivalent characterization trivially applies also to infinite rooted permutations.

Then we say that a sequence $(A_n,\preceq_n)_{n \in \mathbb{Z}_{>0}}$ of rooted permutations is locally convergent to a rooted permutation $(A,\preceq)$ if, for all $h \in \mathbb{Z}_{>0}$, the $h$-restrictions of the sequence $(A_n,\preceq_n)_{n \in \mathbb{Z}_{>0}}$ converge to the $h$-restriction of $(A,\preceq)$ (see Definition 2.1 for a formal statement).

We extend this notion of local convergence for rooted permutations to (unrooted) permutations, rooting them at a uniformly chosen index of $\sigma$. With this procedure, a fixed permutation $\sigma$ naturally identifies a random variable $(\sigma,i)$ that takes values in the set of finite rooted permutations (here and throughout the paper we denote random quantities using bold characters). In this way, the following
The annealed version of the Benjamini–Schramm convergence is natural: given a sequence of permutations \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\), we say that \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) Benjamini–Schramm converges to a randomly possibly infinite rooted permutation \((A, \xi)\), if the sequence \((\sigma^n, i_n)_{n \in \mathbb{Z}_{>0}}\), where \(i_n\) is a uniform index in \([1, |\sigma^n|]\), converges in distribution to \((A, \xi)\) with respect to the above defined local topology. We point out that our choice for the terminology comes from the analogous notion for graphs (see [11]).

We prove the following characterization in terms of proportions of consecutive pattern occurrences (for a formal and complete statement see Theorem 2.21). For any pattern \(\pi\) of size \(k\), and any permutation \(\sigma\) of size \(n\), we denote by
\[
\tilde{c}\text{-occ}(\pi, \sigma) = \frac{c\text{-occ}(\pi, \sigma)}{n} = \frac{\text{number of consecutive occurrences of } \pi \text{ in } \sigma}{n},
\]
the proportion of consecutive occurrences of \(\pi\) in \(\sigma\). Moreover, we denote with \(\mathcal{S}\) the set of permutations of finite size.

**Theorem 1.3.** For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a permutation of size \(n\). Then the Benjamini–Schramm convergence for the sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) is equivalent to the existence of non-negative real numbers \((\Delta_\pi)_{\pi \in \mathcal{S}}\) such that, for all patterns \(\pi \in \mathcal{S}\),
\[
\tilde{c}\text{-occ}(\pi, \sigma^n) \rightarrow \Delta_\pi.
\]

**Remark 1.4.** We could have introduced local convergence for permutations also from another point of view, closer to the one developed for graphons (see [39]): we could say that a sequence of permutations is locally convergent if for every pattern \(\pi \in \mathcal{S}\) the sequences \(c\text{-occ}(\pi, \sigma^n)\) converge and then characterize the limiting objects. Theorem 1.3 shows that the two approaches are equivalent.

What makes Theorem 1.3 particularly interesting is the intense study of consecutive patterns. This is a very active field both from a combinatorial and a probabilistic point of view, with applications in computer science, biology, and physics. Classical statistics like the number of descents, runs and peaks contained in a permutation can be viewed as particular examples of consecutive patterns in permutations, but more general approaches have been developed, for example concerning the asymptotic behavior of the number of permutations avoiding a consecutive pattern. For a complete overview, we refer to the wonderful survey on the topic by Elizalde [20].

Once the Benjamini–Schramm convergence has been introduced for sequences of deterministic permutations, we then extend this notion to sequences of random permutations \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\). The presence of two sources of randomness, one for the choice of the permutation and one for the (uniform) choice of the root, leads to two non-equivalent possible definitions: the **annealed** and the **quenched** version of the Benjamini–Schramm convergence (see Definitions 2.22 and 2.29). Intuitively, in the second definition, the random permutation is frozen, whereas in the first one, the random permutation and the choice of the root are treated on the same level.

We obtain the following two additional characterizations (for formal and complete statements see Theorems 2.26 and 2.34).

**Theorem 1.5.** For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a random permutation of size \(n\). Then
(a) The annealed version of the Benjamini–Schramm convergence of \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) is equivalent to the existence of non-negative real numbers \((\Delta_\pi)_{\pi \in \mathcal{S}}\) such that
\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, \sigma^n)] \rightarrow \Delta_\pi, \quad \text{for all patterns } \pi \in \mathcal{S}.
\]
(b) The quenched version of the Benjamini–Schramm convergence of \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) is equivalent to the existence of non-negative real random variables \((\Lambda_\pi)_{\pi \in \mathcal{S}}\) such that
\[
(\tilde{c}\text{-occ}(\pi, \sigma^n))_{\pi \in \mathcal{S}} \xrightarrow{(d)} (\Lambda_\pi)_{\pi \in \mathcal{S}},
\]
w.r.t. the product topology (where \(\xrightarrow{(d)}\) indicates the convergence in distribution).

Obviously, the quenched version implies the annealed version.
At this point we want to highlight a remarkable difference with the case of permuton convergence and its characterization mentioned before. Indeed, for the proportion of (non-consecutive) occurrences of patterns, the convergences in (a) and (b) are equivalent as shown in [9, Theorem 2.5]. On the other hand, in Example 2.38 we show that the annealed and quenched version of the Benjamini–Schramm convergence are not equivalent.

Finally, we are also able to characterize random limiting objects for the annealed version of the Benjamini–Schramm convergence introducing a “shift-invariant” property. For an order \((\mathbb{Z}, \preceq)\), its shift \((\mathbb{Z}, \preceq')\) is defined by \(i + 1 \preceq' j + 1\) if and only if \(i \preceq j\). A random infinite rooted permutation, or equivalently a random total order, is said to be shift-invariant if it has the same distribution than its shift (for precise statements see Definition 2.43, Proposition 2.46 and Theorem 2.47).

**Theorem 1.6.** A random infinite rooted permutation is “shift-invariant” if and only if it is the local limit of a sequence of random permutations in the annealed Benjamini–Schramm sense.

In some sense, our “shift-invariant” property corresponds to the notion of unimodularity for random graphs. As said before, it is well-known that the Benjamini–Schramm limit of a sequence of random graphs is unimodular. However, it is an open problem to determine if every unimodular random graph is the local limit of a sequence of finite random graphs rooted uniformly at random.

1.2.2. **Local limits for uniform \(\rho\)-avoiding permutations with \(|\rho| = 3\).** In the second part of the article (Sections 4 and 5), we demonstrate the relevance of this new notion of convergence. As first models for studying local convergence, we consider some classes of pattern avoiding permutations. This choice is motivated by the intense study of permutation classes in the last thirty/forty years. For a great survey we refer to Vatter [53]. In general, they have been deeply studied from a combinatorial point of view, finding the enumeration of specific classes. The most basic result in the field is that the number of \(\rho\)-avoiding permutations of size \(n\), for \(|\rho| = 3\), is given by the \(n\)-th Catalan number.

More recently a new probabilistic approach to the study of permutation classes has been investigated. An example, as explained in the previous section, is the study of scaling limits of uniform random permutations in a fixed pattern-avoiding class. Another example is the general problem of studying the limiting distribution of the number of occurrences (after a suitable rescaling) of a fixed pattern \(\pi\) in a uniform random permutation belonging to a fixed class when the size tends to infinity (see for instance Janson [32, 30, 31] where the author studied this problem in the model of uniform permutations avoiding a fixed family of patterns of size three).

In this work we focus on the classes of \(\rho\)-avoiding permutations, for \(|\rho| = 3\), since several bijections (see for instance [13]) with other combinatorial objects (such as rooted ordered trees and Dyck paths) are known and we expect some connection between the local limit convergence for trees and the local limit convergence for permutations.

Our two main results of this second part are the two following theorems.

**Theorem 1.7.** For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a uniform random 231-avoiding permutation of size \(n\). Then we have the following convergence in probability,

\[
\tilde{c}\text{-occ}(\pi, \sigma^n) \xrightarrow{P} \frac{2^{\text{LRMax}(\pi)} + \text{RLMax}(\pi)}{2^{2|\pi|}}, \quad \text{for all } \pi \in \text{Av}(231),
\]

where LRMax(\(\pi\)) (resp. RLMax(\(\pi\))) denotes the number of left-to-right maxima (resp. right-to-left maxima) in \(\pi\).

**Theorem 1.8.** For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a uniform random 321-avoiding permutation of size \(n\). Then we have the following convergence in probability for all \(\pi \in \text{Av}(321),\)

\[
\tilde{c}\text{-occ}(\pi, \sigma^n) \xrightarrow{P} \begin{cases} 
\frac{|\pi| + 1}{2^{|\pi|}} & \text{if } \pi = 12 \ldots |\pi|, \\
\frac{1}{2^{|\pi|}} & \text{if } \text{c-occ}(21, \pi^{-1}) = 1, \\
0 & \text{otherwise.}
\end{cases}
\]
Observation 1.9. By symmetry, this covers all cases of $\rho$-avoiding permutations with $|\rho| = 3$. Indeed, every permutation of size three is in the orbit of either 231 or 321 by applying reverse (symmetry of the diagram w.r.t the vertical axis) and complementation (symmetry of the diagram w.r.t the horizontal axis). Beware that inverse (symmetry of the diagram w.r.t the principal diagonal) cannot be used since it does not preserve consecutive pattern occurrences.

Since the limits of the random sequences $(\tilde{c}_{\text{occ}}(\pi, \sigma^n))_{n \in \mathbb{Z}_{>0}}$ are deterministic, for all patterns $\pi \in \mathcal{S}$, the convergence in probability in (1) and (2) is equivalent to the convergence in distribution of the vectors $(\tilde{c}_{\text{occ}}(\pi, \sigma^n))_{\pi \in \mathcal{S}}$ for the product topology. Therefore Theorems 1.7 and 1.8 trivially imply the characterization (b) in Theorem 1.5 and so prove that the sequences $(\sigma^n)_{n \in \mathbb{Z}_{>0}}$ converge for the quenched version, i.e., the stronger version, of the Benjamini–Schramm convergence.

Here are some other interesting aspects of Theorems 1.7 and 1.8.

- A first important fact is the concentration phenomenon, namely the fact that the limits of the random sequences $(\tilde{c}_{\text{occ}}(\pi, \sigma^n))_{n \in \mathbb{Z}_{>0}}$ are deterministic, for all pattern $\pi \in \mathcal{S}$. Indeed Janson [32, Remark 1.1] notices that, in some classes, we have concentration for the (non-consecutive) pattern occurrences around their mean, in others not. It would be interesting to understand in a more general setting when this concentration phenomenon does or does not occur (we will come back to this fact in the next section).
- The second important fact is the different behavior of the two models of Theorem 1.7 and Theorem 1.8: the first limiting density has full support on the space of 231-avoiding permutations, whereas the second gives positive measure only to 321-avoiding permutations whose inverse have at most one descent. Indeed, despite having the same enumeration sequence, Av(231) and Av(321) are often considered in the community as behaving really differently. Our results give new evidence of this belief.
- Theorem 1.5 ensures just the existence of the limiting random total orders on $\mathbb{Z}$ for the Benjamini–Schramm convergence, without providing an explicit construction. Nevertheless, in both cases (see Sections 4.4 and 5.5), we are able to provide an explicit construction of them.

Although the two theorems have very similar statements and in both models we use bijections between $\rho$-avoiding permutations and rooted ordered trees, the two proofs involve different techniques.

- For the proof of Theorem 1.7 we use the Second moment method. We study the asymptotic behavior of the first and the second moments of $\tilde{c}_{\text{occ}}(\pi, \sigma^n)$ applying a technique introduced by Janson [30, 28]. Instead of studying uniform trees with $n$ vertices, we focus on specific families of binary Galton–Watson trees (which have some nice independence properties). Then we recover results for the first family of trees using singularity analysis for generating functions.
- For the proof of Theorem 1.8 we use a probabilistic approach for the study of local limits for Galton–Watson trees pointed at a uniform vertex. The bijection between trees and 321-avoiding permutations that we used, strongly depends on the position of the leaves. We therefore study the contour functions of some specific Galton–Watson trees in order to extract information about the positions of the leaves in the neighborhood of a uniform vertex.

1.3. Future projects and open problems. We believe that this new notion of local convergence for permutations is worth being further investigated in several ways. We list our ideas for future projects.

- Motivated by the work of Janson [32] we would like to extend our results to uniform permutations avoiding multiple patterns. We have the impression that, in the case of patterns of size three, this generalization does not contain any difficulty (due to the high level of
independence for the values of a uniform permutation in these classes) and the explicit analysis of the problem will probably be developed in the bachelor thesis of Petrella under the supervision of Barbato at the University of Padua.

- We would like to generalize the proof of Theorem 1.8 for all uniform \(\rho\)-avoiding permutations when \(\rho\) is an increasing or decreasing permutation of size greater than three.
- We studied the local convergence for uniform permutations in some classes. Another interesting direction could be to study non-uniform models. Motivated by the work of Crane, DeSalvo and Elizalde [19] we believe that the local limit of random permutations with Mallows distribution could be investigated. More generally, we would like to investigate different types of biased models.
- In a future work with Bouvel, Féray and Stufler we investigate local limits for uniform permutations in substitution-closed classes. We suspect that also in this setting, the concentration phenomenon for the proportions of consecutive pattern occurrences occurs.
- A challenging problem (motivated by the previous expected results) is to find natural (but non-trivial) models where there is no concentration for the proportion of consecutive pattern occurrences (for a trivial model see Example 2.38).
- In a more theoretical direction we would like to examine in depth the relationship between local convergence and permuton convergence (and other possible intermediate notion of convergence for permutations, as done for example in [22] for graphs).
- As said before we are able to characterize random limiting objects for the annealed version of the Benjamini–Schramm convergence. It would be interesting to find a similar characterization for the stronger quenched version, where the limiting objects are random probability measures.

1.4. Outline of the paper. The paper is organized as follows:

- Section 2 deals with the general theory of local convergence for deterministic and random permutations;
- Section 3 summarizes notation and results on trees and random trees that we need in the two following sections;
- Sections 4 and 5 contain the proof of Theorem 1.7 and Theorem 1.8 respectively and the constructions of the corresponding limiting objects.

2. Local convergence for permutations

2.1. Permutations and patterns. We introduce some notation to be used in the sequel. For any \(n \in \mathbb{Z}_{>0}\), we denote the set of permutations of \([n] = \{1, 2, \ldots, n\}\) by \(\mathcal{S}^n\). We write permutations of \(\mathcal{S}^n\) in one-line notation as \(\sigma = \sigma_1 \sigma_2 \ldots \sigma_n\). For a permutation \(\sigma \in \mathcal{S}^n\) the size \(n\) of \(\sigma\) is denoted by \(|\sigma|\). We let \(\mathcal{S} := \bigcup_{n \in \mathbb{Z}_{>0}} \mathcal{S}^n\) be the set of finite permutations. We write sequences of permutations in \(\mathcal{S}\) as \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\).

If \(x_1 \ldots x_n\) is a sequence of distinct numbers, let \(\text{std}(x_1 \ldots x_n)\) be the unique permutation \(\pi\) in \(\mathcal{S}^n\) that is in the same relative order as \(x_1 \ldots x_n\), i.e., \(\pi_i < \pi_j\) if and only if \(x_i < x_j\). Given a permutation \(\sigma \in \mathcal{S}^n\) and a subset of indices \(I \subset [n]\), let \(\text{pat}_I(\sigma)\) be the permutation induced by \((\sigma_i)_{i \in I}\), namely, \(\text{pat}_I(\sigma) := \text{std}((\sigma_i)_{i \in I})\). For example, if \(\sigma = 87532461\) and \(I = \{2, 4, 7\}\) then \(\text{pat}_{\{2,4,7\}}(87532461) = \text{std}(736) = 312\).

Given two permutations, \(\sigma, \pi \in \mathcal{S}^n\) for some \(n \in \mathbb{Z}_{>0}\) and \(\pi \in \mathcal{S}^k\) for some \(k \leq n\), we say that \(\sigma\) contains \(\pi\) as a pattern if \(\sigma\) has a subsequence of entries order-isomorphic to \(\pi\), that is, if there exists a subset \(I \subset [n]\) such that \(\text{pat}_I(\sigma) = \pi\). In addition, we say that \(\sigma\) contains \(\pi\) as a consecutive pattern if \(\sigma\) has a subsequence of adjacent entries order-isomorphic to \(\pi\), that is, if there exists an interval \(I \subset [n]\) such that \(\text{pat}_I(\sigma) = \pi\). All the intervals considered in the article are to be understood as integer intervals, i.e., intervals contained in \(\mathbb{Z}\), and will be denoted by \([a, b]\), for \(a, b \in \mathbb{Z}\).
Example 2.1. The permutation $\sigma = 1532467$ contains 1423 as a pattern but not as a consecutive pattern and 321 as consecutive pattern. Indeed $\text{pat}_{\{1,2,3,5\}}(\sigma) = 1423$ but no interval of indices of $\sigma$ induces the permutation 1423. Moreover, $\text{pat}_{\{2,4\}}(\sigma) = \text{pat}_{\{2,3,4\}}(\sigma) = 321$.

We say that $\sigma$ avoids $\pi$ if $\sigma$ does not contain $\pi$ as a pattern. We point out that the definition of $\pi$-avoiding permutation refers to patterns and not to consecutive patterns. We denote by $\text{Av}^n(\pi)$ the set of $\pi$-avoiding permutations of size $n$ and by $\text{Av}(\pi) := \bigcup_{n \in \mathbb{Z}_{>0}} \text{Av}^n(\pi)$ the set of $\pi$-avoiding permutations of arbitrary size.

We denote by $c\text{-occ}(\pi, \sigma)$ the number of consecutive occurrences of a pattern $\pi$ in $\sigma$. More precisely

$$c\text{-occ}(\pi, \sigma) := \text{Card}\left\{ I \subseteq [n] \mid \text{Card}(I) = |\pi|, I \text{ is an interval, } \text{pat}_I(\sigma) = \pi \right\},$$

where Card$(\cdot)$ denotes the cardinality of a set. Moreover, we denote by $\tilde{c}\text{-occ}(\pi, \sigma)$ the proportion of consecutive occurrences of a pattern $\pi$ in $\sigma$, that is,

$$\tilde{c}\text{-occ}(\pi, \sigma) := \frac{c\text{-occ}(\pi, \sigma)}{n}.$$ 

Remark 2.2. The natural choice for the denominator of the previous expression should be $n - k + 1$ and not $n$, but we make this choice for later convenience (for example, in order to give a probabilistic interpretation of the quantity $\tilde{c}\text{-occ}(\pi, \sigma)$ as done in Equation $[15]$ p[16]). Moreover, for every fixed $k$, there are no difference in the asymptotics when $n$ tends to infinity.

2.2. The set of rooted permutations. In the first part of this section we define the notion of finite and infinite rooted permutation. Then we introduce a local distance and at the end we show that the space of (possibly infinite) rooted permutations is the natural space to study local limits of permutations. More precisely, we show that this space is complete for the local distance and contains the space of finite rooted permutations as a dense subset.

For the reader convenience, we recall the following fundamental definition.

Definition 2.3. A finite rooted permutation is a pair $(\sigma, i)$, where $\sigma \in \mathcal{S}^n$ and $i \in [n]$ for some $n \in \mathbb{Z}_{>0}$.

We extend the notion of size of a permutation to rooted permutations in the natural way, i.e., $|{(\sigma, i)}| := |\sigma|$. We denote with $\mathcal{S}^n$ the set of rooted permutations of size $n$ and with $\mathcal{S} := \bigcup_{n \in \mathbb{Z}_{>0}} \mathcal{S}^n$ the set of finite rooted permutations. We write sequences of finite rooted permutations in $\mathcal{S}$ as $(\sigma^n, i^n)_{n \in \mathbb{Z}_{>0}}$.

To a rooted permutation $(\sigma, i)$, we associate the pair $(A_{\sigma,i}, \preceq_{\sigma,i})$, where $A_{\sigma,i} := [-i + 1, |\sigma| - i]$ is a finite interval containing 0 and $\preceq_{\sigma,i}$ is a total order on $A_{\sigma,i}$, defined for all $\ell, j \in A_{\sigma,i}$ by

$$\ell \preceq_{\sigma,i} j \iff \text{ and only if } \sigma_{\ell+i} \leq \sigma_{j+i}.$$ 

Clearly this map is a bijection from the space of finite rooted permutations $\mathcal{S}$ to the space of total orders on finite integer intervals containing zero. Consequently and throughout the paper, we identify every rooted permutation $(\sigma, i)$ with the total order $(A_{\sigma,i}, \preceq_{\sigma,i})$.

We want to highlight in the following example that the total order $(A_{\sigma,i}, \preceq_{\sigma,i})$ associated to a rooted permutation indicates how the elements of $\sigma$ at given positions compare to each other.

In most examples, we draw permutations as diagrams. We recall that the diagram of a permutation $\sigma$ of size $n$ is the set of points of the Cartesian plane at coordinates $(j, \sigma_j)$, for $1 \leq j \leq n$. For a rooted permutation $(\sigma, i)$ we draw the root with a bigger red dot at coordinates $(i, \sigma_i)$.

Example 2.4. The diagram of the rooted permutation $(\sigma, i) = (752934861, 4)$ is provided in the left-hand side of Fig. 2. The associated total order is $(A_{\sigma,i}, \preceq_{\sigma,i}) = ([{-3, 5}], \preceq)$, where

$$5 \preceq -1 \preceq 1 \preceq 2 \preceq -2 \preceq 4 \preceq -3 \preceq 3 \preceq 0.$$ 

We now make an observation that will be useful for the proofs of Proposition 2.46 and Theorem 2.47.
Observation 2.5. A total order \((A_{\sigma,i}, \preceq_{\sigma,i})\) (associated to a rooted permutation \((\sigma, i)\)) shifted to the interval \([1, |\sigma|]\) seen as a word coincides with the inverse permutation \(\sigma^{-1}\).

For instance, the total order in Example 2.4 shifted on the interval \([1, 9]\) gives \(9 \preceq 3 \preceq 5 \preceq 6 \preceq 2 \preceq 8 \preceq 1 \preceq 7 \preceq 4\). The word 935628174 corresponding to it coincides with the inverse permutation \(\sigma^{-1}\).

\[
\begin{array}{c c}
(\sigma, i) & \sigma^{-1} \\
\begin{array}{c c c c c c c c}
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\end{array}
& \begin{array}{c c c c c c c c}
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\
\end{array}
\end{array}
\]

Figure 2. Diagram of the rooted permutation \((\sigma, i) = (752934861, 4)\) on the left. The inverse \(\sigma^{-1} = 935628174\) of \(\sigma\) is shown on the right.

Thanks to the identification between rooted permutations and total orders, the following definition of infinite rooted permutation is natural.

Definition 2.6. We call infinite rooted permutation a pair \((A, \preceq)\) where \(A\) is an infinite interval of integers containing 0 and \(\preceq\) is a total order on \(A\). We denote the set of infinite rooted permutations by \(S^\infty\).

We highlight that infinite rooted permutations can be thought of as rooted at 0. We set \(\tilde{S}^\bullet := S^\bullet \cup S^\infty\), namely, the set of finite and infinite rooted permutations and we denote by \(\tilde{S}^\bullet\) as \((A_n, \preceq_n)_{n \in \mathbb{Z}_{>0}}\).

We now introduce the following restriction function around the root defined, for every \(h \in \mathbb{Z}_{>0}\), as follow

\[
(3) \quad r_h : \tilde{S}^\bullet \rightarrow S^\bullet, \quad (A, \preceq) \mapsto (A \cap [-h, h], \preceq).
\]

We can think of restriction functions as a notion of neighborhood around the root. For finite rooted permutations we also have the equivalent description of the restriction functions \(r_h\) in terms of consecutive patterns: if \((\sigma, i) \in S^\bullet\) then \(r_h(\sigma, i) = \text{pat}_{[a,b]}(\sigma)\) where \(a = \max\{1, i - h\}\) and \(b = \min\{|\sigma|, i + h\}\).

In the next example we give a graphical interpretation of the restriction function around the root.

Example 2.7. We continue Example 2.4 with the rooted permutation \((\sigma, i) = (752934861, 4)\). When we consider the restriction \(r_h(\sigma, i)\), we draw in gray a vertical strip “around” the root of width \(2h + 1\) (or less if we are near the boundary of the diagram). An example is provided in Fig. 3 where \(r_2(\sigma, i)\) is computed. In particular \(r_2(\sigma, i) = ([-2, 2], \preceq)\), whith \(-1 \preceq 1 \preceq 2 \preceq -2 \preceq 0\).

We now introduce a last definition.

Definition 2.8. We say that a family \((A_h, \preceq_h)_{h \in \mathbb{Z}_{>0}}\) of elements in \(S^\bullet\) is consistent if

\[r_h(A_{h+1}, \preceq_{h+1}) = (A_h, \preceq_h), \quad \text{for all} \quad h \in \mathbb{Z}_{>0}.\]

We have the following.
Figure 3. Diagram of the permutation $\sigma = 752934861$ rooted at $i = 4$ with the corresponding restriction $r_2(\sigma, i)$.

Observation 2.9. For every infinite rooted permutation $(A, \preceq)$, the corresponding family of restrictions $(r_h(A, \preceq))_{h \in \mathbb{Z} > 0}$ is consistent.

In the next example, we exhibit a concrete infinite rooted permutation, with some of its restrictions and the associated rooted permutations.

Example 2.10. We consider the following total order on $\mathbb{Z}$:

$$-2 \preceq -4 \preceq -6 \preceq \cdots \preceq -1 \preceq -3 \preceq -5 \preceq \cdots \preceq 0 \preceq 1 \preceq 2 \preceq 3 \preceq \ldots,$$

that is, the standard order on the integers except that the order on negative numbers is reversed and negative even numbers are smaller than negative odd numbers. Using the definition given in Equation (3), we have for $h = 4$,

$$r_4(\mathbb{Z}, \preceq) = ([-4, 4], \preceq),$$

with

$$-2 \preceq -4 \preceq -1 \preceq -3 \preceq 0 \preceq 1 \preceq 2 \preceq 3 \preceq 4,$$

which represents the rooted permutation $(\pi^4, 5)$, for $\pi^4 = 241356789$, whose diagram is represented in the left-hand side of Fig. 4. In particular, $r_4(\mathbb{Z}, \preceq) = (A_{\pi^4, 5}, \preceq_{\pi^4, 5})$.

Moreover, for $h = 3$,

$$r_3(\mathbb{Z}, \preceq) = ([-3, 3], \preceq),$$

with

$$-2 \preceq -1 \preceq -3 \preceq 0 \preceq 1 \preceq 2 \preceq 3,$$

which represents the rooted permutation $(\pi^3, 4)$, for $\pi^3 = 3124567$, whose diagram is represented in the right-hand side of Fig. 4. In particular, $r_3(\mathbb{Z}, \preceq) = (A_{\pi^3, 4}, \preceq_{\pi^3, 4})$. We also note that $r_3(\pi^4, 5) = (\pi^3, 4)$ as shown in Fig. 4.

Figure 4. The $r_3$-restriction of the rooted permutation $(\pi^4, 5)$.

2.3. Local convergence for rooted permutations. We are now ready to define a notion of local distance on the set of (possibly infinite) rooted permutations $\tilde{\mathcal{S}}$. Given two rooted permutations $(A_1, \preceq_1), (A_2, \preceq_2) \in \tilde{\mathcal{S}}$, we define

$$d((A_1, \preceq_1), (A_2, \preceq_2)) = 2^{-\sup \{ h \in \mathbb{Z} > 0 : r_h(A_1, \preceq_1) = r_h(A_2, \preceq_2) \}},$$

with the classical conventions that $\sup \emptyset = 0$, $\sup \mathbb{Z} > 0 = +\infty$ and $2^{-\infty} = 0$. It is a basic exercise to check that $d$ is a distance. Actually, $d$ is ultra metric, and so all the open balls of radius $2^{-h}$,
We state the following two easy properties of the distance $d$.

**Definition 2.11.** We say that a sequence $(A_n, \preceq_n)_{n \in \mathbb{Z}_{>0}}$ of rooted permutations in $\mathcal{S}_*$ is **locally convergent** to an element $(A, \preceq) \in \mathcal{S}_*$, if it converges with respect to the local distance $d$. In this case we write $(A_n, \preceq_n) \xrightarrow{\text{loc}} (A, \preceq)$.

We are now going to show that the space $\tilde{\mathcal{S}}_*$ is the right space to consider in order to study the limits of finite rooted permutations with respect to the local distance $d$. We will say that the balls are **clopen**, i.e., closed and open.

Once we have a distance, the following notion of convergence is very natural.

**Observation 2.12.** We state the following two easy properties of the distance $d$ and the function $r_h$:

1. $d((A_1, \preceq_1), (A_2, \preceq_2)) \leq 2^{-h}$ if and only if $r_\ell(A_1, \preceq_1) = r_\ell(A_2, \preceq_2)$ for all $\ell \leq h$.
2. Given a rooted permutation $(\sigma, i)$, it holds that $|r_h(\sigma, i)| = 2h + 1$ if and only if the root of $r_h(\sigma, i)$ is $h + 1$ if and only if $|\sigma| \geq 2h + 1$ and $i \in [h + 1, |\sigma| - h]$.

The next proposition gives a converse to Observation 2.9.

**Proposition 2.13.** Given a consistent family $(A_h, \preceq_h)_{h \in \mathbb{Z}_{>0}}$ of elements in $\mathcal{S}_*$ there exists a unique (possibly infinite) rooted permutation $(A, \preceq) \in \mathcal{S}_*$ such that $r_h(A, \preceq) = (A_h, \preceq_h)$, for all $h \in \mathbb{Z}_{>0}$.

Moreover, $(A_h, \preceq_h) \xrightarrow{\text{loc}} (A, \preceq)$.

**Proof.** We explicitly construct the rooted permutation $(A, \preceq)$. We set $A := \bigcup_{h \in \mathbb{Z}_{>0}} A_h$. From the consistency property of the family $(A_h, \preceq_h)_{h \in \mathbb{Z}_{>0}}$ we immediately deduce that $A_h \subseteq A_{h+1}$, for all $h \in \mathbb{Z}_{>0}$ and so the set $A$ is an interval of $\mathbb{Z}$. Moreover, the set $A$ trivially contains $0$. In order to construct a (possibly infinite) permutation we have to endow the interval $A$ with a total order $\preceq$. For all $i, j \in A$, we set $\tilde{h} = \tilde{h}(i, j) := \min\{h \in \mathbb{Z}_{>0} : i \in A_h, j \in A_h\}$, and we make the following choice:

$$i \preceq j \quad \text{if and only if} \quad i \preceq_{\tilde{h}} j.$$  

The only non-trivial property for this total order is transitivity. Fix $i, j, k \in \mathbb{Z}$ such that $i \preceq j$ and $j \preceq k$. Thanks to the consistency property, the definition in Equation (6) is equivalent to

$$i \preceq j \quad \text{if and only if} \quad i \preceq_h j,$$

and so we have

$$i \preceq_h j \preceq_h k, \quad \text{for all } h \text{ large enough}.$$  

Hence $i \preceq k$ and transitivity is proved.

By construction,

$$r_h(A, \preceq) = (A_h, \preceq_h), \quad \text{for all } h \in \mathbb{Z}_{>0}.$$
We can also conclude that \((A_h, \preceq_h) \xrightarrow{\text{loc}} (A, \preceq)\) since, by the above equation and Equation (5), we have
\[
d((A, \preceq), (A_h, \preceq_h)) \leq 2^{-h}, \quad \text{for all } h \in \mathbb{Z}_{>0}.
\]
Uniqueness follows from the uniqueness of the limit of a sequence in a metric space. \(\square\)

Observation 2.14. We consider the product space \(\prod_{h \in \mathbb{Z}_{>0}} S_h^{\leq 2h+1}\) endowed with the product topology, that is metrizable by the distance
\[
d^\prime((x_i)_{i \in \mathbb{Z}_{>0}}, (y_i)_{i \in \mathbb{Z}_{>0}}) = 2^{-\sup\{H \in \mathbb{Z}_{>0} : r_h(x_i) = r_h(y_i), \forall h \leq H\}}.
\]

We have the following isometric embedding
\[
\mathcal{I}: \tilde{S} \rightarrow \prod_{h \in \mathbb{Z}_{>0}} S_h^{\leq 2h+1}
\]
\[
(A, \preceq) \mapsto (r_h(A, \preceq))_{h \in \mathbb{Z}_{>0}},
\]
with inverse
\[
\mathcal{I}^{-1}: \mathcal{I}(\tilde{S}) \rightarrow \tilde{S}
\]
\[
(A_h, \preceq_h)_{h \in \mathbb{Z}_{>0}} \mapsto \lim_{h \rightarrow \infty} (A_h, \preceq_h).
\]

We note that \(\mathcal{I}^{-1}\) is well-defined and coincides with the inverse of \(\mathcal{I}\) thanks to Proposition 2.13. Moreover, the embedding is isometric since obviously the two distances coincide on \(\tilde{S}\) and \(\mathcal{I}(\tilde{S})\).

Therefore, \(\tilde{S}\) is homeomorphic to \(\mathcal{I}(\tilde{S})\).

Since \(\tilde{S}\) and \(\mathcal{I}(\tilde{S})\) are homeomorphic, then the corresponding notions of convergence are equivalent. An immediate consequence is the following.

Proposition 2.15. Given a sequence \((A_n, \preceq_n)_{n \in \mathbb{Z}_{>0}}\) of rooted permutations in \(\tilde{S}\), the following are equivalent:
(a) There exists \((A, \preceq) \in \tilde{S}\) such that
\[
(A_n, \preceq_n) \xrightarrow{\text{loc}} (A, \preceq);
\]
(b) There exists a family \((B_h, \ll_h)_{h \in \mathbb{Z}_{>0}}\) of finite rooted permutations such that
\[
r_h(A_n, \preceq_n) \xrightarrow{\text{loc}} (B_h, \ll_h), \quad \text{for all } h \in \mathbb{Z}_{>0}.
\]
In particular if one of the two conditions holds (and so both), then \((B_h, \ll_h) = r_h(A, \preceq)\), for all \(h \in \mathbb{Z}_{>0}\).

Remark 2.16. The most interesting way to use the previous proposition is as follow: suppose we know that (b) holds, then there exists a rooted permutation \((A, \preceq) \in \tilde{S}\) such that \((A_n, \preceq_n) \xrightarrow{\text{loc}} (A, \preceq)\). Specifically, we can prove the convergence of the sequence \((A_n, \preceq_n)_{n \in \mathbb{Z}_{>0}}\) without explicitly knowing the limit \((A, \preceq)\) and then reconstruct \((A, \preceq)\) using the relations \((B_h, \ll_h) = r_h(A, \preceq)\) for all \(h \in \mathbb{Z}_{>0}\).

Observation 2.17. Note that for all \(h \in \mathbb{Z}_{>0}\), the restriction functions \(r_h\) are continuous. This is a simple consequence of implication \((a) \Rightarrow (b)\) in the previous proposition.

We end this section with the following.

Theorem 2.18. The metric space \((\tilde{S}, d)\) is a compact Polish space, i.e., compact, separable and complete.
Proof. As said before (see the discussion below Definition 2.11), the countable set of finite rooted permutations is dense in \((\hat{\mathcal{S}}_*, d)\), indeed \(d((A, \xi), r_h(A, \xi)) \leq 2^{-h}\) for all \((A, \xi) \in \hat{\mathcal{S}}_*\), all \(h \in \mathbb{Z}_{>0}\). This proves the separability.

For completeness, consider a Cauchy sequence \((A_n, \xi_n)_{n \in \mathbb{Z}_{>0}}\) in \(\hat{\mathcal{S}}_*\). For all \(h \in \mathbb{Z}_{>0}\), by the Cauchy condition, there exists a positive natural number \(N_h\) such that for all natural number \(n, m > N_h\),

\[
d((A_n, \xi_n), (A_m, \xi_m)) \leq 2^{-h}.
\]

In particular, using Observation 2.12 (1), we have, for all \(n, m > N_h\),

\[
r_h(A_n, \xi_n) = r_h(A_m, \xi_m),
\]

therefore, for all \(h \in \mathbb{Z}_{>0}\), the restrictions \(r_h(A_n, \xi_n) \in \mathcal{S}_*\) are stationary in \(n\) and so convergent. By Proposition 2.15 there exists \((A, \xi) \in \hat{\mathcal{S}}_*\) such that \((A_n, \xi_n) \xrightarrow{\text{loc}} (A, \xi)\). This shows completeness.

Since the embedding \(\mathcal{I}\) in Observation 2.14 is isometric and \(\hat{\mathcal{S}}_*\) is complete then \(\mathcal{I}(\hat{\mathcal{S}}_*)\) is closed in the compact product space \(\prod_{h \in \mathbb{Z}_{>0}} \mathcal{S}_*^{2h+1}\). Therefore \(\mathcal{I}(\hat{\mathcal{S}}_*)\) is compact (since it is a closed subset of a compact set) and so also \(\hat{\mathcal{S}}_*\).

\(\square\)

2.4. Benjamini–Schramm convergence: the deterministic case. In this section we want to define a notion of weak-local convergence for a deterministic sequence of finite permutations \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\). The case of a random sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) will be discussed in Section 2.5.

Some of the results contained in this section are just special cases of the those in Section 2.5.1 about the annealed version of the Benjamini–Schramm convergence. However we choose to present the deterministic case separately with a double purpose: first, in this way, the definitions are clearer, secondly, we believe that the subsequent quenched version of the Benjamini–Schramm convergence (see Section 2.5.2) will be more intuitive.

A possible way to define this weak-local convergence is to use the notion of local distance defined in Equation (4). Therefore, we need to construct a sequence of rooted permutations from the sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\). We can see a fixed permutation \(\sigma\) as an element of \(\mathcal{S}_*\) only after a root \(i\) has been chosen. A natural way to choose a root is to make the choice at random, and uniformly among the indices of \(\sigma\). In this way, a fixed permutation \(\sigma\) naturally identifies a probability measure \(\mu_\sigma\) on the space \(\mathcal{S}_*\), that is

\[
\text{(12)} \quad \text{for all Borel sets } A \subset \mathcal{S}_*, \quad \mu_\sigma(A) = \mathbb{P}((\sigma, i) \in A),
\]

where \(i\) is a uniform index of \(\sigma\). Equivalently, the measure \(\mu_\sigma\) is the law of the random variable \((\sigma, i)\) that takes values in the set \(\mathcal{S}_*\).

Notation 2.19. In order to avoid any confusion, as done in Equation (12), we write random quantities using bold characters to distinguish them from deterministic quantities. Moreover, given a probability measure \(\mu\) we denote with \(\mathbb{E}_\mu\) the expectation with respect to \(\mu\) and given a random variable \(X\), we denote with \(\mathcal{L}(X)\) its law. Given a sequence of random variables \((X_n)_{n \in \mathbb{Z}_{>0}}\) we write \(X_n \xrightarrow{(d)} X\) to denote the convergence in distribution and \(X_n \xrightarrow{p} X\) to denote the convergence in probability. Finally, given an event \(A\) we denote with \(A^c\) the complement event of \(A\).

We are now ready to define our notion of Benjamini–Schramm convergence for sequences of permutations.

Definition 2.20. Given a sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) of elements in \(\mathcal{S}\), we say that \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) Benjamini–Schramm converges to a random (possibly infinite) rooted permutation \((A, \xi)\), if the sequence \((\sigma^n, i_n)_{n \in \mathbb{Z}_{>0}}\), where \(i_n\) is a uniform index of \(|\sigma|\), converges in distribution to \((A, \xi)\) with respect to the local distance \(d\) defined in Equation (4). In this case we simply write \(\sigma^n \xrightarrow{\text{BS}} (A, \xi)\) instead of \((\sigma^n, i_n) \xrightarrow{(d)} (A, \xi)\).
Sometimes, in order to simplify notation, we will denote the Benjamini–Schramm limit of a sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) simply by \(\sigma^\infty\) instead of \((A, \varsigma)\).

Our main theorem in this section deals with the following interesting relation between the Benjamini–Schramm convergence and the convergence of consecutive pattern densities.

**Theorem 2.21.** For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a permutation of size \(n\). Then the following are equivalent:

(a) There exists a random rooted infinite permutation \(\sigma^\infty\) such that
\[
\sigma^n \xrightarrow{\text{BS}} \sigma^\infty.
\]

(b) There exist non-negative real numbers \((\Delta_\pi)_{\pi \in S}\) such that for all patterns \(\pi \in S\),
\[
\bar{c}\text{-occ}(\pi, \sigma^n) \to \Delta_\pi.
\]

Moreover, if one of the two conditions holds (and so both) we have the following relation between the limiting objects,
\[
P(r_h(\sigma^\infty) = (\pi, h + 1)) = \Delta_\pi, \quad \text{for all } h \in \mathbb{Z}_{>0}, \text{ all } \pi \in S^{2h+1}.
\]

We skip for the moment the proof of this theorem since it will be a particular case of the more general Theorem 2.26.

2.5. **Benjamini–Schramm convergence: the random case.** The goal of this section is to characterize the convergence in distribution of a sequence of random permutations \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) with respect to the local distance defined in Equation (4). We have two different natural choices for this definition, one stronger than the other (and not equivalent as shown in Example 2.38). The weaker definition is an analogue of the notion of Benjamini–Schramm convergence for random graphs developed in [11].

Before starting, we give some more explanations about our notation for random quantities. We will use a superscript notation on probability measure \(P\) (and on the corresponding expectation \(E\)) to record the source of randomness. Specifically, given two independent random variables \(X\) and \(Y\) (with values in two spaces \(E\) and \(F\) respectively) and a set \(A \subseteq E \times F\), we write
\[
P^X((X, Y) \in A) := P((X, Y) \in A | X),
\]
and similarly
\[
P^Y((X, Y) \in A) := P((X, Y) \in A | Y).
\]

Moreover, we recall the following standard relation
\[
P((X, Y) \in A) = E[1_{(X,Y)\in A}] = E^X[E^Y[1_{(X,Y)\in A}]] = E^X[P^Y((X, Y) \in A)] = E^Y[P^X((X, Y) \in A)].
\]

(13)

2.5.1. **The annealed version of the Benjamini–Schramm convergence.** Recall that a sequence of deterministic permutations \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) Benjamini–Schramm converges if the corresponding sequence of random variables \((\sigma^n, i_n)_{n \in \mathbb{Z}_{>0}}\) converges in distribution with respect to the local distance \(d\). In analogy with that (and the definition of Benjamini–Schramm convergence for graphs) we can state the following.

**Definition 2.22** (Annealed version of the Benjamini–Schramm convergence). Given a sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) of random permutations in \(S\), let \(i_n\) be a uniform index of \(\sigma^n\) conditionally on \(\sigma^n\). We say that \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) converges in the annealed Benjamini–Schramm sense to a random variable \(\sigma^\infty\) with values in \(\mathcal{S}_\varsigma\) if the sequence of random variables \((\sigma^n, i_n)_{n \in \mathbb{Z}_{>0}}\) converges in distribution to \(\sigma^\infty\) with respect to the local distance. In this case we write \(\sigma^n \xrightarrow{\text{aBS}} \sigma^\infty\) instead of \((\sigma^n, i_n) \xrightarrow{(d)} \sigma^\infty\).

Like in the deterministic case, the limiting object \(\sigma^\infty\) is a random rooted infinite permutation.
Remark 2.23. Even though the above definition is stated for sequences of random permutations of arbitrary size, we will often consider the case when, for all \(n \in \mathbb{Z}_{>0}\), \(|\sigma^n| = n\) almost surely and the random choice of the root \(i_n\) is uniform in \([n]\) and independent of \(\sigma^n\).

Before stating the main theorem of this section we clarify two interesting properties of the statistics \(c\)-occ\((\pi, \sigma)\) introduced in Section 2.1.

Given a permutation \(\pi \in S^k\), we set
\[
\pi^{*m} := \text{std}(\pi_1, \ldots, \pi_k, m - 1/2) \in S^{k+1}, \quad \text{for all } 1 \leq m \leq k + 1.
\]

In the following example we give a graphical interpretation of the operation \(\pi^{*m}\).

Example 2.24. Let \(\pi = 3421 = \begin{array}{cccc}
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\end{array}\) then \(\pi^{*3} = \text{std}(3, 4, 2, 1, 5/2) = \begin{array}{cccc}
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\end{array}\) = 45213,

where we highlight in light blue the entries already presented in the permutation \(\pi = 3421\).

Now, for all \(\pi \in S^k\) and \(\sigma \in S^n\) with \(k < n\), we just observe that
\[
c\text{-occ}(\pi, \sigma) = \sum_{m=1}^{k+1} c\text{-occ}(\pi^{*m}, \sigma) + \mathbb{1}\{\text{pat}_{[i_n-k, i_n]}(\sigma) = \pi\},
\]

and so the convergence of \(\tilde{c}\text{-occ}(\pi, \sigma)\) depends only on the convergence of \(c\text{-occ}(\pi^{*m}, \sigma)\), for all \(1 \leq m \leq k + 1\).

We also give a probabilistic interpretation of \(\tilde{c}\text{-occ}(\pi, \sigma)\) for \(\pi \in S^{2h+1}\) and \(\sigma \in S^n\) with \(n \geq 2h+1\). Namely, let \(i_n\) be uniform in \([n]\), then
\[
\tilde{c}\text{-occ}(\pi, \sigma) = \mathbb{P}(r_h(\sigma, i_n) = (\pi, h + 1)).
\]

We need the following observation.

Observation 2.25. We claim that the set of clopen balls (see the discussion after Equation (4))
\[
\mathcal{A} = \left\{B((A, \infty), 2^{-h}) : h \in \mathbb{Z}_{>0}, (A, \infty) \in \tilde{S}_*\right\}
\]
is a separating class for the space \((\tilde{S}_*, d)\), i.e., if two probability measures agree on \(\mathcal{A}\) then they necessarily agree also on the whole space. This is a trivial consequence of \([14, \text{Theorem 2.3}]\) (with the consecutive discussion) and the fact that the intersection of two balls is either empty or one of them.

We give in the following theorem some characterizations of the annealed version of the Benjamini–Schramm convergence.

Theorem 2.26. For any \(n \in \mathbb{Z}_{>0}\), let \(\sigma^n\) be a random permutation of size \(n\) and \(i_n\) be a uniform random index in \([n]\), independent of \(\sigma^n\). Then the following are equivalent:

(a) There exists a random rooted infinite permutation \(\sigma^\infty\) such that
\[
\sigma^n \xrightarrow{aBS} \sigma^\infty,
\]
i.e., \((\sigma^n, i_n) \xrightarrow{d} \sigma^\infty\) w.r.t. the local distance \(d\) on \(\tilde{S}_*\);

(b) For all \(h \in \mathbb{Z}_{>0}\), there exist non-negative real numbers \((\Gamma^h_\pi)_{\pi \in S^{2h+1}}\) such that
\[
\mathbb{P}(r_h(\sigma^n, i_n) = (\pi, h + 1)) \xrightarrow{n \to \infty} \Gamma^h_\pi, \quad \text{for all } \pi \in S^{2h+1}.
\]

(c) There exist non-negative real numbers \((\Delta_\pi)_{\pi \in S}\) such that for all patterns \(\pi \in S\),
\[
\mathbb{E}[c\text{-occ}(\pi, \sigma^n)] \to \Delta_\pi.
\]
Moreover, if one of the three conditions holds (and so all of them), for every fixed \( h \in \mathbb{Z}_{>0} \), we have the following relations between the limiting objects,
\[
P(r_h(\sigma^\infty) = (\pi, h + 1)) = \Delta_{\pi} = \Gamma_{\pi}^h, \quad \text{for all } \pi \in S^{2h+1}.
\]

Before proving the theorem we point out two important facts.

**Remark 2.27.** Note that the theorem proves the existence of a random rooted infinite permutation \( \sigma^\infty \) but does not furnish any explicit construction of this object as a random total order on \( \mathbb{Z} \).

**Remark 2.28.** For every fixed \( h \in \mathbb{Z}_{>0} \), the condition (b) in the previous theorem considers only the probabilities \( P(r_h(\sigma^n, i_n) = (\pi, j)) \) for \( \pi \in S^{2h+1} \) and \( j = h + 1 \). We remark that all the other cases are trivial. Indeed if \( \pi \in S^{2h+1} \) and \( j \neq h + 1 \), then, using Observation 2.28 (2), for all \( n \in \mathbb{Z}_{>0} \),
\[
P(r_h(\sigma^n, i_n) = (\pi, j)) = 0.
\]

Finally, if \( \pi \in S^{\leq 2h+1} \) with \( |\pi| < 2h + 1 \) then for all \( n \in \mathbb{Z}_{>0} \),
\[
P(r_h(\sigma^n, i_n) = (\pi, j)) \leq \frac{1}{n} \to 0,
\]
where we used that \( r_h(\sigma^n, i_n) \) can be equal to \((\pi, j)\) if and only if either \( j = |\pi| - h \) and \( i_n = j \) or \( j = h + 1 \) and \( i_n = n - |\pi| + h \), and both possibilities cannot be simultaneously satisfied for a fixed \((\pi, j)\) with \(|\pi| < 2h + 1\).

**Proof of Theorem 2.26** (a) \( \Rightarrow \) (b). For all \( h \in \mathbb{Z}_{>0} \), the convergence in distribution of the sequence \( (r_h(\sigma^n, i_n))_{n \in \mathbb{Z}_{>0}} \), follows from the continuity of the functions \( r_h \) (Observation 2.17). We note that the random variable \( r_h(\sigma^n, i_n) \) takes its values in the finite set \( S^{\leq 2h+1} \). Therefore, using condition (b) and Remark 2.28, we can conclude that
\[
(\pi, j) \in S_{\leq 2h+1}.
\]

(b) \( \Rightarrow \) (a). Thanks to Theorem 2.18 \((\mathcal{S}_t, d)\) is a compact Polish space. Therefore, applying Prokhorov’s Theorem, in order to show that \((\sigma^n, i_n) \overset{d}{\to} \sigma^\infty\), for some random infinite permutation \( \sigma^\infty \), it is enough to show that for every pair of convergent subsequences \((\sigma^n_k, i_{n_k})_{k \in \mathbb{Z}_{>0}}\) and \((\sigma^\infty_{l}, i_{n_l})_{l \in \mathbb{Z}_{>0}}\) with limits \( \sigma^\infty_1 \) and \( \sigma^\infty_2 \) respectively, then
\[
(17) \quad \sigma^\infty_1 \overset{d}{=} \sigma^\infty_2.
\]

Thanks to Observation 2.25 it is enough to check that
\[
(18) \quad P(\sigma^\infty_1 \overset{d}{=} B((A, \varepsilon), 2^{-h})) = P(\sigma^\infty_2 \overset{d}{=} B((A, \varepsilon), 2^{-h})) , \quad \text{for all } B((A, \varepsilon), 2^{-h}) \in A.
\]

Since for all \( h \in \mathbb{Z}_{>0} \), all \((A, \varepsilon) \in \mathcal{S}_t\), the ball \( B((A, \varepsilon), 2^{-h}) \) is clopen, then using the assumption that \((\sigma^n_k, i_{n_k})_{k \in \mathbb{Z}_{>0}} \overset{d}{=} \sigma^\infty_1 \) and the Portmanteau Theorem (see [14 Theorem 2.1]), we have
\[
P(\sigma^\infty_1 \overset{d}{=} B((A, \varepsilon), 2^{-h})) = \lim_{n \to \infty} P((\sigma^n_k, i_{n_k}) \overset{d}{=} B((A, \varepsilon), 2^{-h})) = \lim_{n \to \infty} P(r_h(\sigma^n_k, i_{n_k}) = r_h(A, \varepsilon)).
\]

Therefore, using condition (b) and Remark 2.28 we can conclude that
\[
P(\sigma^\infty_1 \overset{d}{=} B((A, \varepsilon), 2^{-h})) = \Gamma_h^A, \quad \text{for } r_h(A, \varepsilon) = (\pi, h + 1) \text{ for some } \pi \in S^{2h+1} \text{ and } \Gamma_h^{r_h(A, \varepsilon)} = 0, \text{ otherwise.}
\]
Similarly we have,
\[
P(\sigma^\infty_2 \overset{d}{=} B((A, \varepsilon), 2^{-h})) = \Gamma_h^A,
\]
Therefore condition \[ (b) \] is satisfied and this concludes the proof.

\[(b) \Rightarrow (c). \] Note that if \( \pi \in S \) then there exist \( h \in \mathbb{Z}_{>0} \) such that either \( \pi \in S^{2h+1} \) or \( \pi \in S^{2h} \). If \( \pi \in S^{2h+1} \), using relations \[ (15) \] and \[ (13) \] with the independence between \( \sigma^n \) and \( i_n \), we have

\[
\mathbb{E} \left[ \frac{\partial \overline{\lim}(\pi, \sigma^n)}{\partial \overline{\lim}(\pi, \sigma^n)} \right]_{\sigma} \mathbb{E}^\sigma = \mathbb{P}^n \left( r_h(\sigma^n, i_n) = (\pi, h+1) \right) \mathbb{P} \left( r_h(\sigma^n, i_n) = (\pi, h+1) \right),
\]

which converges if \( (b) \) holds. Otherwise, if \( \pi \in S^{2h} \) we just use the observation done in Equation \[ (14) \] and so, the convergence of \( \mathbb{E}^\sigma \left[ \frac{\partial \overline{\lim}(\pi, \sigma^n)}{\partial \overline{\lim}(\pi, \sigma^n)} \right] \), for \( \pi \in S^{2h} \), follows.

\[(c) \Rightarrow (b). \] As before, if \( \pi \in S^{2h+1} \), the convergence of \( \mathbb{P} \left( r_h(\sigma^n, i_n) = (\pi, h+1) \right) \) follows from Equation \[ (19) \].

\[ \blacksquare \]

2.5.2. The quenched version of the Benjamini–Schramm convergence. We start by recalling that a permutation \( \sigma \), the associated probability measure \( \mu_\sigma \) is defined by Equation \[ (12) \]. The quenched version of the Benjamini–Schramm convergence is inspired by the following equivalent reformulation of Definition \[ 2.20 \].

Given a sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) of deterministic elements in \( S \), we can equivalently say that \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) converges to a probability measure \( \mu \in \mathcal{P}(\hat{S}_*) \), if the sequence \((\mu_{\sigma^n})_{n \in \mathbb{Z}_{>0}}\) converges to \( \mu \) with respect to the weak topology induced by the local distance \( d \).

In analogy, we can state the following for the random case.

**Definition 2.29** (Quenched version of the Benjamini–Schramm convergence). Given a sequence \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) of random permutation in \( S \) and a random measure \( \mu^\infty \) on \( \hat{S}_* \), we say that \((\sigma^n)_{n \in \mathbb{Z}_{>0}}\) converges in the quenched Benjamini–Schramm sense to \( \mu^\infty \) if the sequence of random measures \((\mu_{\sigma^n})_{n \in \mathbb{Z}_{>0}}\) converges in distribution to \( \mu^\infty \) with respect to the weak topology induced by the local distance \( d \). In this case we write \( \sigma^n \overset{qBS}{\rightarrow} \mu^\infty \) instead of \( \mu_{\sigma^n} \overset{(d)}{\rightarrow} \mu^\infty \).

Unlike the annealed version of the Benjamini–Schramm converge, the limiting object \( \mu^\infty \) is a random measure on \( \hat{S}_* \).

**Remark 2.30.** Note that, given a random permutation \( \sigma \), the corresponding random measure \( \mu_{\sigma} \) is the conditional law of the random variable \( ( (\sigma, i)|\sigma) \).

We want to remark some important topological facts.

**Remark 2.31.** Given a random permutation \( \sigma \), the associated random uniformly rooted permutation \( (\sigma, i) \) can be viewed as a random variable with values in the Polish compact space \( (\hat{S}_*, d) \). Similarly the associated random probability measure \( \mu_{\sigma} \) given by Equation \[ (12) \] can be viewed as a random variable with values in the set of probability measures \( \mathcal{P}(\hat{S}_*) \). The space \( \mathcal{P}(\hat{S}_*) \) endowed with the weak convergence topology is Polish.

We recall that given a metric space \((E, d)\) endowed with a \( \sigma \)-algebra \( \mathcal{E} \), then \( \mathcal{P}(E) \) is a metric space once equipped with the Prokhorov metric \( \nu \).

Moreover, if \((E, d)\) is a Polish space, the metric \( \nu \) is such that the weak convergence is equivalent to \( \nu \)-convergence and \((\mathcal{P}(E), \nu)\) is also a Polish space (see for instance \[ 14 \], Theorem 6.8, p.73). Finally, if \((E, d)\) is compact, Prokhorov’s theorem ensures that \((\mathcal{P}(E), \nu)\) is also compact (for general results on convergence of measure, we refer to \[ 14 \]).

Thanks to the previous remark we can interpret a random probability measure on \( \hat{S}_* \) as a random variable with values in the compact Polish space \( \mathcal{P}(\hat{S}_*) \).

We now generalize the notion of consecutive pattern density to probability measures on \( \hat{S}_* \).
Definition 2.32. Given a probability measure $\mu$ on $\tilde{S}^*$, we define the consecutive pattern density $\tilde{c}\text{-occ}(\pi, \mu)$ as

$$\tilde{c}\text{-occ}(\pi, \mu) = \mu\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right), \quad \text{for all } \pi \in S^{2h+1}, \text{ all } h \in \mathbb{Z}_{>0},$$

(20)

$$\tilde{c}\text{-occ}(\pi, \mu) = \mu\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right), \quad \text{for all } \pi \in S^{2h}, \text{ all } h \in \mathbb{Z}_{>0},$$

Note that if $\mu \equiv \mu_\sigma$ for some $\sigma \in S$ then

(21) $$\tilde{c}\text{-occ}(\pi, \mu_\sigma) = \tilde{c}\text{-occ}(\pi, \sigma) + O(1/|\sigma|), \quad \text{for all } \pi \in S.$$  

Indeed, if $\pi \in S^{2h+1}$,

$$\tilde{c}\text{-occ}(\pi, \mu_\sigma) = \mu_\sigma\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right) = \mathbb{P}\left((\sigma, i) \in B((\pi, h + 1), 2^{-h})\right) = \mathbb{P}(r_h(\sigma, i) = (\pi, h + 1))$$

and if $\pi \in S^{2h}$, using that the balls $\left\{B((\pi^m, h + 1), 2^{-h})\right\}_{1 \leq m \leq 2h+1}$ are disjoint,

$$\tilde{c}\text{-occ}(\pi, \mu_\sigma) = \mu_\sigma\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right) = \sum_{m=1}^{2h+1} \mathbb{P}\left((\sigma, i) \in B((\pi^m, h + 1), 2^{-h})\right) = \sum_{m=1}^{2h+1} \mathbb{P}(r_h(\sigma, i) \in (\pi^m, h + 1))$$

$$= \sum_{m=1}^{2h+1} \mathbb{P}\left((\sigma, i) \in B((\pi^m, h + 1), 2^{-h})\right) + \sum_{m=1}^{2h+1} \mathbb{P}(r_h(\sigma, i) \in (\pi^m, h + 1)) + O(1/|\sigma|).$$

Before stating our main theorem of this section we state and prove a technical proposition.

Proposition 2.33. For all $\pi \in S$, the function

$$\tilde{c}\text{-occ}(\pi, \cdot): \left(\mathcal{P}(\tilde{S}^*), \nu\right) \rightarrow [0, 1]$$

$$\mu \mapsto \tilde{c}\text{-occ}(\pi, \mu).$$

is continuous.

Proof. By definition, depending on the parity of $|\pi|$, either $\tilde{c}\text{-occ}(\pi, \mu) = \mu\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right)$ or $\tilde{c}\text{-occ}(\pi, \mu) = \mu\left(\bigcup_{m=1}^{2h+1} B((\pi^m, h + 1), 2^{-h})\right)$. Since a finite union of clopen balls is clopen, using the Portmanteau Theorem, if $\mu_n \rightarrow \mu$ in the weak sense then $\tilde{c}\text{-occ}(\pi, \mu_n) \rightarrow \tilde{c}\text{-occ}(\pi, \mu)$. Since weak convergence is equivalent to $\nu$-convergence, then $\tilde{c}\text{-occ}(\pi, \cdot)$ is continuous. \qed

We are now ready to state and prove our main theorem that gives us a characterization of the quenched version of the Benjamini–Schramm convergence.

Theorem 2.34. For any $n \in \mathbb{Z}_{>0}$, let $\sigma^n$ be a random permutation of size $n$ and $i_n$ be a uniform random index in $[n]$, independent of $\sigma^n$. Then the following are equivalent:

(a) There exists a random measure $\mu^\infty$ on $\tilde{S}^*$ such that

$$\sigma^n \overset{qBS}{\rightarrow} \mu^\infty,$$

i.e., $\mu_{\sigma^n} \overset{(d)}{\rightarrow} \mu^\infty$ w.r.t. the weak topology induced by the local distance $d$ on $\tilde{S}^*$;
Proof. The other cases are trivial. For more details see Remark 2.28.

(b) There exists a family of non-negative real random variables \((\Gamma^h_\pi)_{h \in \mathbb{Z}_{>0}, \pi \in \mathcal{S}^{2h+1}}\) such that
\[
\left(\mathbb{P}^n(\tau_h(\sigma^n, j_n) = (\pi, h + 1))\right)_{h \in \mathbb{Z}_{>0}, \pi \in \mathcal{S}^{2h+1}} \xrightarrow{d} \left(\Gamma^h_\pi\right)_{h \in \mathbb{Z}_{>0}, \pi \in \mathcal{S}^{2h+1}},
\]
\(w.r.t.\) the product topology;

(c) There exists an infinite vector of non-negative real random variables \((\Lambda_\pi)_{\pi \in \mathcal{S}}\) such that
\[
\left(c\text{-occ}(\pi, \sigma^n)\right)_{\pi \in \mathcal{S}} \xrightarrow{d} (\Lambda_\pi)_{\pi \in \mathcal{S}}
\]
\(w.r.t.\) the product topology.

In particular, if one of the three conditions holds (and so all of them) then
\[
\left(c\text{-occ}(\pi, \mu^\infty)\right) = \Lambda_\pi, \quad \text{for all } \pi \in \mathcal{S},
\]
and moreover, for every fixed \(h \in \mathbb{Z}_{>0}\),
\[
\left(c\text{-occ}(\pi, \mu^\infty)\right) \equiv \Gamma^h_\pi, \quad \text{for all } \pi \in \mathcal{S}^{2h+1},
\]
\[
\left(c\text{-occ}(\pi, \mu^\infty)\right) \equiv \sum_{m=1}^{2h+1} \Gamma^h_{\pi+m}, \quad \text{for all } \pi \in \mathcal{S}^{2h}.
\]

Before giving the proof, as in Theorem 2.26 we highlight two important facts.

Remark 2.35. Note that the theorem states the existence of a random measure \(\mu^\infty\) but does not furnish any explicit construction of this object.

Remark 2.36. For every fixed \(h \in \mathbb{Z}_{>0}\), the condition (b) in the previous theorem considers only the conditional probabilities \(\mathbb{P}^n(\tau_h(\sigma^n, j_n) = (\pi, j))\) for \(\pi \in \mathcal{S}^{2h+1}\) and \(j = h + 1\). We remark that all the other cases are trivial. For more details see Remark 2.28.

Proof. (a) \(\Rightarrow\) (c) Let \(\pi_1, \ldots, \pi_r\) be a finite sequence of patterns. By Proposition 2.33 the map \(\mu \mapsto (c\text{-occ}(\pi_i, \mu))_{1 \leq i \leq r}\) is continuous. Therefore, the convergence of \(\sigma^n \xrightarrow{q_{BS}} \mu^\infty\) implies the convergence of \((c\text{-occ}(\pi_i, \mu^\infty))_{1 \leq i \leq r} \xrightarrow{d} (c\text{-occ}(\pi_i, \mu^\infty))_{1 \leq i \leq r}\). Since \(c\text{-occ}(\pi_i, \mu^\infty) = c\text{-occ}(\pi_i, \sigma^n) + O(1/n)\) we have the convergence in distribution of all the finite-dimensional marginals of the infinite vector \((c\text{-occ}(\pi, \sigma^n))_{\pi \in \mathcal{S}}\), and this proves (c) (see for instance [14, ex. 2.4, p.19]).

(c) \(\Rightarrow\) (b). Note that for all \(\pi \in \mathcal{S}^{2h+1}, h \in \mathbb{Z}_{>0}\) we have by Equation (15),
\[
\mathbb{P}^n(\tau_h(\sigma^n, i_n) = (\pi, h + 1)) = c\text{-occ}(\pi, \sigma^n).
\]

We can conclude that if (c) holds then the vector \(\left(\mathbb{P}^n(\tau_h(\sigma^n, i_n) = (\pi, h + 1))\right)_{h \in \mathbb{Z}_{>0}, \pi \in \mathcal{S}^{2h+1}}\) converges in distribution.

(b) \(\Rightarrow\) (a). Thanks to Remark 2.31 \(\mu^\sigma_n\) is a random variable with values in the compact Polish space \((\mathcal{P}(\hat{\mathcal{S}}), \nu)\). Therefore, applying Prokhorov’s theorem to \((\mathcal{P}(\hat{\mathcal{S}}), \nu)\), in order to show that \(\mu^\sigma_n \xrightarrow{d} \mu^\infty\) for some random measure \(\mu^\infty\), it is enough to show that for every pair of convergent subsequences \((\mu^\sigma_{n_k})_{k \in \mathbb{Z}_{>0}}\) and \((\mu^\sigma_{n_l})_{l \in \mathbb{Z}_{>0}}\) with limits \(\mu^\infty_1\) and \(\mu^\infty_2\) respectively, it holds that
\[
\mu^\infty_1 \equiv \mu^\infty_2.
\]

Thanks to [33, Theorem 2.2], in order to prove the above equality in distribution it is enough to show that
\[
(\mu^\infty_1(B_1), \ldots, \mu^\infty_1(B_n)) \equiv (\mu^\infty_2(B_1), \ldots, \mu^\infty_2(B_n)), \quad \text{for all } n \in \mathbb{Z}_{>0}, B_1, \ldots, B_n \in \mathcal{A},
\]
where \(\mathcal{A}\) is the separating class defined in Observation 2.25.
Fix \( n \in \mathbb{Z}_{>0} \) and for all \( 1 \leq i \leq n \), let \( B_i = B((A_i, \xi_i), 2^{-h_i}) \) for some \( (A_i, \xi_i) \in \mathcal{S}_\bullet \), \( h_i \in \mathbb{Z}_{>0} \).

Since by assumption \( \mu_{\sigma^n} \xrightarrow{(d)} \mu_1^\infty \) then applying [33] Theorem 4.11 (which is a generalization of the Portmanteau theorem for random measure) we have

\[
(\mu_1^\infty(B_i))_{1 \leq i \leq n} \xrightarrow{(d)} \lim_{k \to \infty} (\mu_{\sigma^n_k}(B_i))_{1 \leq i \leq n}
\]

\[
\xrightarrow{(d)} \lim_{k \to \infty} \left( \mathbb{P}^{i_{nk}}(r_{h_i}(\sigma^{nk}, i_{nk}) = r_{h_i}(A_i, \xi_i)) \right)_{1 \leq i \leq n}.
\]

Therefore, using condition (c) and Remark 2.36, we can conclude that

\[
(\mu_1^\infty(B_i))_{1 \leq i \leq n} \xrightarrow{(d)} (\Gamma_{\sigma^{hi}}(A_i, \xi_i))_{1 \leq i \leq n},
\]

where \( \Gamma_{r_{hi}}(A_i, \xi_i) = \Gamma_{\pi}, \) if \( r_{hi}(A_i, \xi_i) = (\pi, h_i + 1) \) for some \( \pi \in S^{2h_i+1} \) and \( \Gamma_{\sigma^{hi}}(A_i, \xi_i) = 0, \) otherwise. Similarly we have the following equality in distribution,

\[
(\mu_2^\infty(B_i))_{1 \leq i \leq n} \xrightarrow{(d)} (\Gamma_{\sigma^{hi}}(A_i, \xi_i))_{1 \leq i \leq n}.
\]

Therefore \( \mu_1^\infty \xrightarrow{(d)} \mu_2^\infty. \]

2.5.3. Relation between the annealed and the quenched versions of the Benjamini–Schramm convergence. We recall that the intensity (measure) \( \mathbb{E}[\mu] \) of a random measure \( \mu \) is defined as the expectation \( \mathbb{E}[\mu](A) := \mathbb{E}[\mu(A)] \) for all measurable sets \( A \). We have the following expected implication.

**Proposition 2.37.** For any \( n \in \mathbb{Z}_{>0} \), let \( \sigma^n \) be a random permutation of size \( n \) and \( i_n \) be a uniform random index in \([n]\), independent of \( \sigma^n \). If \( \sigma^n \xrightarrow{qBS} \mu^\infty \), for some random measure \( \mu^\infty \) on \( \mathcal{S}_\bullet \), then

\[
\sigma^n \xrightarrow{qBS} \sigma^\infty,
\]

where \( \sigma^\infty \) is the random rooted infinite permutation with law \( \mathcal{L}(\sigma^\infty) = \mathbb{E}[\mu^\infty]. \)

**Proof.** The first part of the statement, in particular the existence of the random rooted infinite permutation \( \sigma^\infty \), is a trivial consequence of Theorem 2.26 and Theorem 2.34. For example, condition (c) in the second theorem trivially implies the condition (c) in the first theorem.

Therefore we just have to show that \( \mathcal{L}(\sigma^\infty) = \mathbb{E}[\mu^\infty] \). For every clopen ball \( B \) contained in the separating class \( A \) (see Observation 2.25) we have

\[
\mathbb{E}[\mu^\infty](B) = \mathbb{E}[\mu^\infty(B)] = \lim_{n \to \infty} \mathbb{E}[\mu_{\sigma^n}(B)] = \lim_{n \to \infty} \mathbb{E}[\mathbb{P}^{i_n}(\{\sigma^n, i_n\} \in B)]
\]

\[
\xrightarrow{[13]} \lim_{n \to \infty} \mathbb{P}(\{\sigma^n, i_n\} \in B) = \mathbb{P}(\sigma^\infty \in B) = \mathcal{L}(\sigma^\infty)(B),
\]

where in the second equality we used that \( \sigma^n \xrightarrow{qBS} \mu^\infty \) and in the fifth equality that \( \sigma^n \xrightarrow{qBS} \sigma^\infty \).

Since the two measures agree in a separating class we can conclude that they are equal. \( \square \)

We now show in the next example that in general the two versions of Benjamini–Schramm converge are not equivalent.

**Example 2.38.** For all \( n \in \mathbb{Z}_{>0} \), let \( \sigma^n \) be the random permutation defined by

\[
\mathbb{P}(\sigma^n = 12\ldots n) = \frac{1}{2} = \mathbb{P}(\sigma^n = n - 1\ldots 1),
\]

and \( \tau_n \) be the deterministic permutation (see also Fig. 5)

\[
\tau_n = 135\ldots n-3\ldots 2, \quad \text{if } n \text{ is odd},
\]

\[
\tau_n = 135\ldots n-1\ldots n-2\ldots 2, \quad \text{if } n \text{ is even}.
\]
We are going to show that the two sequences \((\sigma^n)_{n \in \mathbb{Z}^+}\) and \((\tau^n)_{n \in \mathbb{Z}^+}\) have two different quenched Benjamini–Schramm limits but the same annealed Benjamini–Schramm limit. Therefore, the alternating sequence between \(\sigma^n\) and \(\tau^n\) converges in the annealed Benjamini–Schramm sense but does not converge in the quenched Benjamini–Schramm sense.

![Figure 5. Some elements of the sequence \((\tau_n)_{n \in \mathbb{Z}^+}\).](image)

We note that, for all \(k \in \mathbb{Z}^+\):

1. If \(\pi = 12 \ldots k\) then for \(n > k\),
   \[
   \mathbb{P}(\widetilde{c-occ}(\pi, \sigma^n) = \frac{n - k + 1}{n}) = \mathbb{P}(\sigma^n = 12 \ldots n) = \frac{1}{2};
   \]
   \[
   \mathbb{P}(\widetilde{c-occ}(\pi, \sigma^n) = 0) = \mathbb{P}(\sigma^n = n n-1 \ldots 1) = \frac{1}{2};
   \]

2. If \(\pi = k k-1 \ldots 1\) then for \(n > k\),
   \[
   \mathbb{P}(\widetilde{c-occ}(\pi, \sigma^n) = \frac{n - k + 1}{n}) = \mathbb{P}(\sigma^n = n n-1 \ldots 1) = \frac{1}{2};
   \]
   \[
   \mathbb{P}(\widetilde{c-occ}(\pi, \sigma^n) = 0) = \mathbb{P}(\sigma^n = 12 \ldots n) = \frac{1}{2};
   \]

3. If \(\pi \in S^k, \pi \neq 12 \ldots k\) and \(\pi \neq k k-1 \ldots 1\) then for \(n > k\),
   \[
   \mathbb{P}(\widetilde{c-occ}(\pi, \sigma^n) = 0) = 1;
   \]

Therefore,

\[
(\widetilde{c-occ}(\pi, \sigma^n))_{\pi \in S} \overset{(d)}{\rightarrow} (X_\pi)_{\pi \in S},
\]

where, taking a Bernoulli random variable \(Y\) with parameter \(1/2\), \(X_\pi = Y\) if \(\pi = 12 \ldots k\) for some \(k \in \mathbb{Z}^+\), \(X_\pi = 1 - Y\) if \(\pi = k k-1 \ldots 1\) for some \(k \in \mathbb{Z}^+\) and \(X_\pi = 0\) otherwise.

On the contrary,

\[
(\widetilde{c-occ}(\pi, \tau^n))_{\pi \in S} \overset{(d)}{\rightarrow} (Z_\pi)_{\pi \in S},
\]

where \(Z_\pi = \frac{1}{2}\) if \(\pi = 12 \ldots k\), or \(\pi = k k-1 \ldots 1\) and \(Z_\pi = 0\) otherwise.

We can conclude, using condition (c) in Theorem 2.34, that the two sequences \((\sigma^n)_{n \in \mathbb{Z}^+}\) and \((\tau^n)_{n \in \mathbb{Z}^+}\) have two different quenched Benjamini–Schramm limits (they are different by Equation (22)). On the contrary, noting that \(\lim_{n \to \infty} \mathbb{E}[\widetilde{c-occ}(\pi, \sigma^n)] = \lim_{n \to \infty} \mathbb{E}[\widetilde{c-occ}(\pi, \tau^n)]\) for all \(\pi \in S\), then, using condition (c) in Theorem 2.26, we conclude that the two sequences \((\sigma^n)_{n \in \mathbb{Z}^+}\) and \((\tau^n)_{n \in \mathbb{Z}^+}\) have the same annealed Benjamini–Schramm limit.

We now analyze the particular case when the limiting objects \((\Lambda_\pi)_{\pi \in S}\) (or \((\Gamma_{h,\pi})_{h \in \mathbb{Z}^+, \pi \in S^{2h+1}}\)) in Theorem 2.34 are deterministic, i.e., when there is a concentration phenomenon (this will be the case of Sections 4 and 5). Before stating our result we need the following.
Remark 2.39. When a random measure $\mu$ on $\tilde{S}_*$ is almost surely equal to a deterministic measure $\mu$ on $\tilde{S}_*$, we will simply denote it with $\mu$. In particular, if a sequence of random permutations $(\sigma^n)_{n \in \mathbb{Z}_+}$ converges in the quenched Benjamini–Schramm sense to a deterministic measure $\mu$ (instead of random measure) on $\tilde{S}_*$, we will simply write $\sigma^n \xrightarrow{qBS} \mu$.

Corollary 2.40. For any $n \in \mathbb{Z}_+$, let $\sigma^n$ be a random permutation of size $n$ and $i_n$ be a uniform random index in $[n]$, independent of $\sigma^n$. Then the following are equivalent:

(a) There exists a (deterministic) measure $\mu$ on $\tilde{S}_*$ such that

$$\sigma^n \xrightarrow{qBS} \mu;$$

(b) There exists an infinite vector of non-negative real numbers $(\Lambda_\pi)_{\pi \in S}$ such that

$$\left(\text{c-occ}(\pi, \sigma^n)\right)_{\pi \in S} \xrightarrow{P} (\Lambda_\pi)_{\pi \in S}$$

w.r.t. the product topology;

(c) There exists an infinite vector of non-negative real numbers $(\Lambda'_\pi)_{\pi \in S}$ such that for all $\pi \in S$,

$$\text{c-occ}(\pi, \sigma^n) \xrightarrow{P} \Lambda'_{\pi};$$

(d) For all $h \in \mathbb{Z}_+$, there exists a family of non-negative real numbers $(\Gamma^h_{\pi})_{\pi \in S^{2h+1}}$ such that

$$\mathbb{P}(r_h(\sigma^n, i_n) = (\pi, h + 1)) \xrightarrow{P} \Gamma^h_{\pi}, \quad \text{for all } \pi \in S^{2h+1}.$$

In particular, if one of the four conditions holds (and so all of them) we have the same relations as in Theorem 2.34 with the additional relation $\Lambda_\pi = \Lambda'_\pi$ for all $\pi \in S$.

Remark 2.41. Thanks to Proposition 2.37 note that if (a) holds then $\sigma^n \xrightarrow{aBS} \sigma^\infty$, where $\sigma^\infty$ is the random rooted infinite permutation with $\mathcal{L}(\sigma^\infty) = \mu$.

Proof. (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Leftrightarrow$ (d). The first and the last relations are simple consequences of Theorem 2.34, where the determinism is given by Equations (22) and (23). The second relation is trivial.

(c) $\Rightarrow$ (a). Since when the limits are deterministic, the pointwise convergence in distribution is equivalent to the converge in distribution for the product topology, assumption (c) implies condition (b). Therefore, using Theorem 2.34 and Proposition 2.37, we obtain that there exist a random measure $\mu^\infty$ on $\tilde{S}_*$ and a random rooted permutation $\sigma^\infty$ (with $\mathcal{L}(\sigma^\infty) = \mathbb{E}[\mu^\infty]$) such that

$$\sigma^n \xrightarrow{qBS} \mu^\infty \quad \text{and} \quad \sigma^n \xrightarrow{aBS} \sigma^\infty.$$

In order to conclude, we need only to show that, under our assumption, we have $\mu^\infty = \mathcal{L}(\sigma^\infty)$ a.s. (this obviously implies that $\mu^\infty$ is deterministic). Using again [33, Theorem 2.2] and Observation 2.25 in order to prove the above equality it is enough to show that

$$(\mu^\infty(B_1), \ldots, \mu^\infty(B_k)) = (\mathcal{L}(\sigma^\infty)(B_1), \ldots, \mathcal{L}(\sigma^\infty)(B_k)), \quad \text{for all } k \in \mathbb{Z}_+, B_1, \ldots, B_k \in \mathcal{A}.$$

Fix $k \in \mathbb{Z}_+$ and for all $1 \leq i \leq k$ let $B_i = B((A_i, \xi_i), 2^{-h_i})$ for some $(A_i, \xi_i) \in \tilde{S}_*$, $h_i \in \mathbb{Z}_+$. Since $\mu_{\sigma^n} \xrightarrow{(d)} \mu^\infty$ then applying [33, Theorem 4.11] we have

$$\left(\mu^\infty(B_i)\right)_{1 \leq i \leq k} \xrightarrow{(d)} \lim_{n \to \infty} \left(\mu_{\sigma^n}(B_i)\right)_{1 \leq i \leq k} = \lim_{n \to \infty} \left(\mathbb{P}^{i_n}(r_{h_i}(\sigma^n, i_n) = r_{h_i}(A_i, \xi_i))\right)_{1 \leq i \leq k} \xrightarrow{(d)} \left(\mathbb{P}^{i_n}(r_{h_i}(A_i, \xi_i))\right)_{1 \leq i \leq k},$$

where in the last equality we use assumption (d) that is equivalent to (c). We also use again the fact that when the limits are deterministic, the pointwise convergence in distribution is equivalent.
Moreover we set \( O := \text{occc}(\pi, \sigma^n) \). We have seen that in order to prove the convergence in the quenched Benjamini–Schramm sense when the limiting object is deterministic, it is enough to show the pointwise convergence for the vector \( (\text{occc}(\pi, \sigma^n))_{\pi \in S} \). This is not true when the limiting object is random. Indeed it is quite easy to explicitly construct a counter example.

For every \( n \in \mathbb{Z}_{\geq 0} \), we can consider the three permutations

\[
\sigma^n = 12\ldots n, \quad \tau^n = n\ldots 21, \quad \rho^n = \begin{array}{c}
\vdots \\
\cdot \\
\cdot \\
\cdot \\
\cdot \\
\end{array},
\]

and we can define the two following random permutations

\[
\pi^n := \begin{cases}
\sigma^{2n} \oplus \tau^n, & \text{with probability } 1/3, \\
\sigma^n \oplus \rho^{2n}, & \text{with probability } 1/3, \\
\tau^{2n} \oplus \rho^n, & \text{with probability } 1/3,
\end{cases} \quad \bar{\pi}^n := \begin{cases}
\sigma^n \oplus \tau^{2n}, & \text{with probability } 1/3, \\
\sigma^{2n} \oplus \rho^n, & \text{with probability } 1/3, \\
\tau^n \oplus \rho^{2n}, & \text{with probability } 1/3,
\end{cases}
\]

where \( \oplus \) denotes the direct sum of two permutations, i.e., for \( \pi \in S^k \) and \( \sigma \in S^n \),

\[
\pi \oplus \sigma = \pi_1 \ldots \pi_k (\sigma_1 + k) \ldots (\sigma_n + k).
\]

For each pattern \( \pi \in S \), both \( \text{occc}(\pi, \pi^n) \) and \( \text{occc}(\pi, \bar{\pi}^n) \) converge to the same limit. However the joint vectors \( (\text{occc}(\pi, \pi^n))_{\pi \in S} \) and \( (\text{occc}(\pi, \bar{\pi}^n))_{\pi \in S} \) have different limits in distribution for the product topology.

### 2.6. Characterization of the annealed Benjamini–Schramm limits

We now characterize the annealed Benjamini–Schramm limiting objects. More precisely we show that such a limiting object satisfies a “shift-invariant” property (see Definition 2.43 below). Conversely, we also prove in Theorem 2.47 that every random “shift-invariant” infinite rooted permutation is the annealed Benjamini–Schramm limit of a sequence of random finite rooted permutations.

We start by considering for all patterns \( \pi \) of size \( k \) and every shift \( s \in \mathbb{Z} \), the following sets

\[
O^s(\pi) = \{(A, \xi) \in \tilde{S}_s : [1 + s, k + s] \subseteq A \text{ and } \pi_1 + s \leq \pi_2 + s \leq \cdots \leq \pi_k + s\}.
\]

Moreover we set \( O(\pi) := O^0(\pi) \).

**Definition 2.43.** We say that a random rooted permutation \( (A, \xi) \in \tilde{S}_s \) is **shift-invariant** if for all patterns \( \pi \in S \),

\[
P( (A, \xi) \in O^s(\pi) ) = P( (A, \xi) \in O(\pi) ), \quad \text{for all } s \in \mathbb{Z}.
\]

**Observation 2.44.** Note that for a random rooted shift-invariant permutation \( (A, \xi) \) we have \( A = \mathbb{Z} \) a.s. Indeed,

\[
P( A = \mathbb{Z} ) = P\left( \bigcap_{s \in \mathbb{Z}_{\geq 0}} \{s, -s \subseteq A\} \right) = \lim_{s \to \infty} P(\{s, -s \subseteq A\}),
\]
where we used that $A$ is an interval and the continuity from above for probability measures. Rewriting the last term as
\[
P\{\{s, -s\} \subset A\} = P((A, \varsigma) \in O^{(s-1)}(1), (A, \xi) \in O^{-{(s-1)}}(1))
\]
and noting that trivially $P((A, \varsigma) \in O^{(-1)}(1)) = 1$, then using the shift-invariant property we conclude that $P(A = Z) = 1$.

**Observation 2.45.** We recall that in the introduction (see the discussion before Theorem 1.6) we said that a random infinite rooted permutation, or equivalently a random total order, is shift-invariant if it has the same distribution than its shift. Note that the probabilities in Equation (25), when $s = 1$, are the finite-dimensional distributions of the random infinite rooted permutation $(A, \varsigma)$ and its shift. Therefore, the two definitions coincide since a random variable on a product space is completely determined by its finite-dimensional distributions (see for instance [14] ex. 1.2, pp. 9-11).

**Proposition 2.46.** Let $\sigma^\infty$ be the annealed Benjamini–Schramm limit of a sequence $(\sigma^n)_{n \in \mathbb{Z}_+}$ of random elements in $S$ with $|\sigma^n| = n$ a.s. (and $i_n$ a uniform random index in $[n]$, independent of $\sigma^n$) then $\sigma^\infty$ is shift-invariant.

**Proof.** Fix a pattern $\tau \in S^k$, a shift $s \in \mathbb{Z}$ and suppose that $\sigma^n \overset{aBS}{\rightarrow} \sigma^\infty$. Since $O^s(\pi)$ is clopen, by the Portmanteau theorem, we have,
\[
P(\sigma^\infty \in O^s(\pi)) = \lim_{n \to \infty} P((\sigma^n, i_n) \in O^s(\pi)) = \lim_{n \to \infty} E_{\sigma^n} P^{i_n}((\sigma^n, i_n) \in O^s(\pi)).
\]
We note that $(\sigma^n, i_n) \in O^s(\pi)$ if and only if
\[
[1 + s, k + s] \subseteq A_{\sigma^n, i_n} \quad \text{and} \quad \sigma^n_{\rho_1 + s + i_n} \leq \sigma^n_{\rho_2 + s + i_n} \leq \cdots \leq \sigma^n_{\rho_h + s + i_n}.
\]
The first event has probability that tends to 1 since $|\sigma^n| \to \infty$. The second event (when well-defined, i.e., conditionally on the first) is equivalent to $\text{pat}_{[i_n + s + 1, i_n + s + k]}(\sigma^n) = \pi^{-1} + [1 + s, k + s] \subseteq A_{\sigma^n, i_n}$.

We obtain that
\[
P(\sigma^\infty \in O^s(\pi)) = \lim_{n \to \infty} E_{\sigma^n} \left[P^{i_n}(\sigma^n, i_n) \in O^s(\pi)\right]
= \lim_{n \to \infty} E_{\sigma^n} \left[P^{i_n}(\text{pat}_{[i_n + s + 1, i_n + s + k]}(\sigma^n) = \pi^{-1}, [1 + s, k + s] \subseteq A_{\sigma^n, i_n})\right]
= \lim_{n \to \infty} E_{\sigma^n} \left[\text{oc}(\pi^{-1}, \sigma^n)\right]
\]
and the last term is independent of $s$. This concludes the proof. \qed

We now prove the inverse statement of the previous proposition.

**Theorem 2.47.** Let $(\mathbb{Z}, \varsigma)$ be a random shift-invariant rooted permutation. Then the sequence of random permutations $(\sigma^n)_{n \in \mathbb{Z}_+}$ defined, for all $n \in \mathbb{Z}_+$, by
\[
(27) \quad P(\sigma^n = \rho) = P((\mathbb{Z}, \varsigma) \in O(\rho^{-1})), \quad \text{for all } \rho \in S^n,
\]
(rooted at a uniform random index $i_n$ in $[n]$, independent of $\sigma^n$) converges in the annealed Benjamini–Schramm sense to $(\mathbb{Z}, \varsigma)$.

**Proof.** Thanks to Theorem 2.26 it is enough to prove that for all $h \in \mathbb{Z}_+$,
\[
(28) \quad P(r_h(\sigma^n, i_n) = (\pi, h + 1)) \xrightarrow{n \to \infty} P(r_h(\mathbb{Z}, \varsigma) = (\pi, h + 1)), \quad \text{for all } \pi \in S^{2h + 1},
\]
where $i_n$ is uniform in $[n]$.

Fix $h \in \mathbb{Z}_+$. Note that if $\pi \in S^{2h + 1}$ then for $n \geq |\pi|$,
$$P(r_h(\sigma^n, i_n) = (\pi, h + 1)) = \mathbb{E}^{i_n} \left[ P(\sigma^n (r_h(\sigma^n, i_n)) = (\pi, h + 1)) \right]$$
$$= \mathbb{E}^{i_n} \left[ \sum_{\rho \in S^n} P(\sigma^n = \rho) \mathbb{1}_{(r_h(\rho, i_n) = (\pi, h + 1))} \right]$$
$$= \sum_{\rho \in S^n} P(\sigma^n = \rho) P(i_n = (\pi, h + 1))$$
$$= \sum_{\rho \in S^n} P(Z, \emptyset) \in O(\rho^{-1}) c-\text{occ}(\pi, \rho),$$

where in the last equality we used Equation (27). In particular, if the limits exist,

$$\lim_{n \to \infty} P(r_h(\sigma^n, i_n) = (\pi, h + 1)) = \lim_{n \to \infty} \sum_{\rho \in S^n} P((Z, \emptyset)) \in O(\rho^{-1}) c-\text{occ}(\pi, \rho).$$

In order to prove the existence of the previous limit we rewrite \( P(r_h(Z, \emptyset) = (\pi, h + 1)) \) in a more convenient way. Before proceeding with this, we note that for \( n \) large enough, we have

$$P((Z, \emptyset)) \in O(\rho^{-1}) = \sum_{\rho \in S^n} P((Z, \emptyset)) \in O(\rho) \mathbb{1}_{\text{pat}_{\{1,2h+1\}(\rho^{-1}) = \pi}}$$

and more generally, for all \( s \in \mathbb{Z}_{>0}, \)

$$P((Z, \emptyset)) \in O^s(\rho^{-1}) = \sum_{\rho \in S^n} P((Z, \emptyset)) \in O(\rho) \mathbb{1}_{\text{pat}_{\{s+1, s+2h+1\}(\rho^{-1}) = \pi}}.$$

We start by noting that,

$$P(r_h(Z, \emptyset) = (\pi, h + 1)) = P(([-h, h], \emptyset) = (\pi, h + 1)) = P((Z, \emptyset)) \in O^-(h + 1)(\pi^{-1}),$$

where we used in the last equality that \([[-h, h], \emptyset) = (\pi, h + 1)\) if and only if

$$\pi_1 - (h + 1) \preceq \pi_2 - (h + 1) \preceq \ldots \preceq \pi_{2h+1} - (h + 1).$$

Using the shift-invariant property we can rewrite the last term of Equation (31) as follow,

$$P((Z, \emptyset)) \in O^-(h + 1)(\pi^{-1}) = \sum_{s=0}^{n-2h-1} \frac{P((Z, \emptyset)) \in O^s(\pi^{-1})}{n-2h}$$

$$= \sum_{\rho \in S^n} P((Z, \emptyset)) \in O(\rho) \sum_{s=0}^{n-2h-1} \frac{\mathbb{1}_{\text{pat}_{\{s+1, s+2h+1\}(\rho^{-1}) = \pi}}}{n-2h}.$$

where in the last equality we applied Equation (30).

Moreover, noting that \( \sum_{s=0}^{n-2h-1} \mathbb{1}_{\text{pat}_{\{s+1, s+2h+1\}(\rho^{-1}) = \pi}} = c-\text{occ}(\pi, \rho^{-1}) \) and combining Equations (31) and (32), we obtain, for all \( n \) large enough,

$$P(r_h(Z, \emptyset) = (\pi, h + 1)) = \sum_{\rho \in S^n} P((Z, \emptyset)) \in O(\rho) \frac{c-\text{occ}(\pi, \rho^{-1})}{n-2h}.$$

The right-hand side coincides in the limit for \( n \) that goes to infinity with the limit in Equation (29) and so, since the left-hand side is independent of \( n \), we proved the existence of the limit. In particular, this proves Equation (28) and concludes the proof. \( \square \)

3. Trees and random trees, a toolbox

In this section we summarize notation and results on trees and random trees that we need in the two following sections. The reader comfortable with this topic may skip this section.
3.1. Rooted ordered trees and Galton–Watson trees. We briefly recall in this section the definition of Galton–Watson trees (for a complete introduction, see e.g. [41]). We recall Neveu’s formalism (see [45]) for rooted ordered trees. We set
\[ U := \bigcup_{n \geq 0} (\mathbb{Z}_{>0})^n, \]
the set of finite sequences of positive integers with the convention \((\mathbb{Z}_{>0})^0 = \{\emptyset\}\). For \(n \geq 1\) and \(u = u_1 \ldots u_n \in U\), we set \(|u| = n\) the length of \(u\) with the convention \(|\emptyset| = 0\). If \(u\) and \(v\) are two sequences of \(U\), we denote by \(uv\) the concatenation of the two sequences. Let \(u, v \in U\). We say that \(v\) is an ancestor of \(u\) and write \(v \prec u\), if there exists \(w \in U\), \(w \neq \emptyset\) such that \(u = vw\). The set of ancestors of \(u\) is denoted by \(A_u\).

**Definition 3.1.** A rooted ordered tree \(T\) is a subset of \(U\) that satisfies:
- \(\emptyset \in T\) (this vertex is called root);
- If \(u \in T\), then \(A_u \subseteq T\);
- For every \(u \in T\), there exists \(d^+_u(T) \in \mathbb{Z}_{>0} \cup \{+\infty\}\) such that, for every \(i \in \mathbb{Z}_{>0}\), \(u_i \in T\) if and only if \(1 \leq i \leq d^+_u(T)\). The number \(d^+_u(T)\) represents the number of children of the vertex \(u \in T\).

We always think of (and draw) a rooted ordered tree \(T\) as a graph (embedded in the plane) where the vertices are the elements of \(T\) and the edges connect vertices \(u\) and \(u_i\), for all \(i \in \mathbb{Z}_{>0}\), \(u, u_i \in T\).

We embed rooted ordered trees in the plane with the root at the top of the trees in such a way that the trees grow downward and the labels of the \(n\)-th generation (i.e., of all the vertices with length \(n\)) increase from left to right. An example is given in Fig. 6.

**Figure 6.** A rooted ordered tree embedded in the plane.

We denote with \(\mathbb{T}\) the set of rooted ordered trees with no vertex having infinitely many children,
\[ \mathbb{T} := \{T : d^+_u(T) < \infty, \forall u \in T\}, \]
and with \(\mathbb{T}^n \subset \mathbb{T}\) the set of rooted ordered trees with \(n\) vertices. We often denote the root of the tree also by \(r\) instead of \(\emptyset\). We recall that given \(T \in \mathbb{T}\), for each vertex \(v \in T\), the height of \(v\) is defined as \(h(v) := |v|\). We denote with \(d\) the classical distance on trees, i.e., for every \(u, v \in T\), \(d(u, v)\) is equal to the number of edges in the unique path between \(u\) and \(v\). Note that \(h(v) = d(r, v)\). We also denote by \(|T|\) the number of vertices of \(T\). Moreover we set \(\mathcal{L}(T) := \{v \in T : d^+_u(T) = 0\}\), namely \(\mathcal{L}(T)\) is the set of leaves in \(T\). We denote by \(a(v)\) the parent of \(v\) in \(T\), i.e., the largest ancestor of \(v\) (for the order \(\prec\)) and we also write \(a^i(v), i \geq 0\), for the \(i\)-th ancestor of \(v\) in \(T\), i.e., \(a^i(v) = a \circ a \circ \cdots \circ a(v)\) (with the convention that if \(r = a^k(v)\) then \(a^i(v) = r\) for all \(i \geq k\)).

We denote with \(c(u, v)\) the youngest common ancestor of \(u\) and \(v\), namely \(c(u, v)\) is the largest vertex \(w\) (for the order \(\prec\)) such that \(u = wu'\) and \(v = wv'\) for some \(u', v' \in U\). We also denote by...
\[ f(T,v) \] the fringe subtree of \( T \) rooted at \( v \), i.e., the subtree of \( T \) containing all the vertices \( u \in \mathcal{U} \) such that \( vu \in T \).

Finally, we recall the following definition of a standard model for random trees.

**Definition 3.2.** A \( \mathbb{T} \)-valued random variable \( T \) is said to have the branching property if for \( n \in \mathbb{Z}_{>0} \), conditionally on \( \{d^+_\emptyset(T) = n\} \), the \( n \) fringe subtrees, rooted at the \( n \) children of the root, are independent and distributed as the original tree \( T \). A \( \mathbb{T} \)-valued random variable \( T^\eta \) is a Galton–Watson tree with offspring distribution \( \eta \) if it has the branching property and the distribution of \( d^+_\emptyset(T) \) is \( \eta \).

### 3.2. Depth-first traversal

We now recall two classical ways of visiting the vertices of a rooted ordered tree that go under the name of depth-first traversal. These traversals originally developed for algorithms searching vertices in a tree, work as follows. Let \( T \) be a rooted ordered tree with root \( r \).

**Definition 3.3 (Pre-order).** If \( T \) consists only of \( r \), then \( r \) is the pre-order traversal of \( T \). Otherwise, suppose that \( T_1, T_2, \ldots, T_d \) are the fringe subtrees in \( T \) respectively rooted at the \( m \)-th child of \( r \) from left to right. The pre-order traversal begins by visiting \( r \). It continues by traversing \( T_1 \) in pre-order, then \( T_2 \) in pre-order, and so on, until \( T_d \) is traversed in pre-order.

**Definition 3.4 (Post-order).** If \( T \) consists only of \( r \), then \( r \) is the post-order traversal of \( T \). Otherwise, suppose that \( T_1, T_2, \ldots, T_d \) are the fringe subtrees in \( T \) respectively rooted at the \( m \)-th child of \( r \) from left to right. The post-order traversal begins by traversing \( T_1 \) in post-order, then \( T_2 \) in post-order, and so on, until \( T_d \) is traversed in post-order. It ends by visiting \( r \).

We will also say that \( T \) is labeled with the pre-order (resp. post-order) labeling starting from \( k \), if we associate the tag \( i + k - 1 \) to the \( i \)-th visited vertex by the pre-order (resp. post-order) traversal. Moreover, if we simply say that \( T \) is labeled with the pre-order (resp. post-order), we will assume that we are starting from \( k = 1 \).

We will show an example in Fig. 8 in the case of binary trees.

### 3.3. Binary trees

We now introduce binary trees. We set
\[
\mathcal{U}_{1,2} := \bigcup_{n \geq 0} \{1, 2\}^n
\]
the set of finite binary words with the convention \( \{1, 2\}^0 = \{\emptyset\} \). We use the same notation as in the case of ordered trees.

**Definition 3.5.** A binary tree \( T \) is a subset of \( \mathcal{U}_{1,2} \) that satisfies:

- \( \emptyset \in T \);
- If \( u \in T \), then \( A_u \subseteq T \).

**Remark 3.6.** Note that a binary tree is not a particular rooted ordered tree. Indeed the third condition in Definition 3.1 is omitted in the case of binary trees. This implies that, if a vertex of a binary tree has only one child then it is either a left vertex (if it is marked with 1) or a right vertex (if it is marked with 2). See Fig. 7 for an example.

Given a vertex \( v \) in the tree, we denote by \( T^v_L \) and \( T^v_R \) the left and the right fringe subtrees of \( T \) hanging below the vertex \( v \). We simply write \( T^v_L \) and \( T^v_R \) if \( v = \emptyset \) namely, for the left and the right fringe subtrees below the root. Finally we let \( \mathbb{T}_b \) be the set of all binary trees.

The notions of pre-order and post-order traversal extend immediately to binary trees. Moreover, in the case of binary trees, the in-order traversal is also available. Let \( T \) be a binary tree with root \( r \).
Definition 3.7 (In-order). If \( T \) consists only of \( r \), then \( r \) is the in-order traversal of \( T \). Otherwise the in-order traversal begins by traversing \( T_L \) in in-order, then visits \( r \), and concludes by traversing \( T_R \) in in-order.

The notion of in-order labeling is defined similarly to the pre-order labeling or post-order labeling. We exhibit an example of the three different labelings in Fig. 8.

We finally extend the notion of Galton–Watson trees to binary Galton–Watson trees.

Definition 3.8. A binary Galton–Watson tree \( T^{\eta} \) with offspring distribution \( \eta = (\eta_0, \eta_L, \eta_R, \eta_2) \), is a \( \mathbb{T}_1 \)-valued random variable with the branching property and such that

\[
\begin{align*}
\mathbb{P}(\text{The root has 0 child}) &= \eta_0, \\
\mathbb{P}(\text{The root has only a left child}) &= \eta_L, \\
\mathbb{P}(\text{The root has only a right child}) &= \eta_R, \\
\mathbb{P}(\text{The root has 2 children}) &= \eta_2.
\end{align*}
\]

4. The local limit for uniform random 231-avoiding permutations

The goal of this section is to prove that uniform random 231-avoiding permutations converge in the quenched Benjamini–Schramm sense. Before stating our precise result we fix some notation. Given a permutation \( \pi \in \mathcal{S} \) we denote with \( \text{LRMax}(\pi) \) the set of indices of the left-to-right maxima of \( \pi \), i.e., the values \( i \in |\pi| \) such that \( \pi_j < \pi_i \) for every \( j < i \). Similarly we denote with \( \text{RLMax}(\pi) \) the set of indices of the right-to-left maxima of \( \pi \), i.e., values \( i \in |\pi| \) such that \( \pi_i > \pi_j \) for every
Moreover, we set $|\text{LRMax}(\pi)| := \text{Card}(\text{LRMax}(\pi))$ and $|\text{RLMax}(\pi)| := \text{Card}(\text{RLMax}(\pi))$. We also fix $\text{Max}(\pi) := \text{LRMax}(\pi) \cup \text{RLMax}(\pi)$, and similarly we set $|\text{Max}(\pi)| := \text{Card}(\text{Max}(\pi)))$. Note that $|\text{Max}(\pi)| = |\text{LRMax}(\pi)| + |\text{RLMax}(\pi)| - 1$ since $\text{LRMax}(\pi) \cap \text{RLMax}(\pi)$ contains only the index of the maximum of $\pi$.

For all $k > 0$, we define the following probability distribution on $\text{Av}^k(231)$,

$$P_{231}(\pi) := \frac{2^{\text{LRMax}(\pi)} + |\text{RLMax}(\pi)|}{2^{|\pi|}}, \quad \text{for all } \pi \in \text{Av}^k(231).$$

Remark 4.1. The fact that the above equation defines a probability distribution on $\text{Av}^k(231)$ is a consequence of Theorem 4.4 below, noting that in Equation (33), summing over all $\pi \in \text{Av}^k(231)$ the left-hand side, we trivially obtain $\frac{n - k + 1}{n}$ that tends to 1. We also present in Appendix A a generating function proof of this fact.

Example 4.2. We consider the permutation

$$\pi = 132985476,$$

where we draw in green the left-to-right maxima, in blue the right-to-left maxima, and in orange the maximum of the permutation, which is both a left-to-right maximum and a right-to-left maximum. We have $\text{LRMax}(\pi) = \{1, 2, 4\}$ and $\text{RLMax}(\pi) = \{9, 8, 5, 4\}$. Thus $P_{231}(\pi) = \frac{3 + 4}{2^{18}} = \left(\frac{1}{2}\right)^{11}$.

Before stating our main result, we recall the following trivial fact.

Observation 4.3. If $\pi \notin \text{Av}(231)$ and $\sigma \in \text{Av}(231)$ then obviously $c\text{-occ}(\pi, \sigma) = 0$. Therefore, in what follow, we simply analyze the consecutive occurrences densities $\tilde{c}\text{-occ}(\pi, \sigma)$ for $\pi \in \text{Av}(231)$.

Theorem 4.4. Let $\sigma^n$ be a uniform random permutation in $\text{Av}^n(231)$ for all $n \in \mathbb{Z}_{>0}$. For all $\pi \in \text{Av}(231)$, we have the following convergence in probability,

$$(33) \quad \tilde{c}\text{-occ}(\pi, \sigma^n) \overset{P}{\rightarrow} P_{231}(\pi).$$

Since the limiting frequencies $(P_{231}(\pi))_{\pi \in \text{Av}(231)}$ are deterministic, using Corollary 2.40 we have the following.

Corollary 4.5. Let $\sigma^n$ be a uniform random permutation in $\text{Av}^n(231)$ for all $n \in \mathbb{Z}_{>0}$. There exists a random infinite rooted permutation $\sigma_{231}^\infty$ such that

$$\sigma^n \overset{qBS}{\rightarrow} \mathcal{L}(\sigma_{231}^\infty).$$

Remark 4.6. We highlight that Corollary 4.5 proves the existence of a random infinite rooted permutation $\sigma_{231}^\infty$ without explicitly constructing this limiting object. Nevertheless we are able to provide a construction of $\sigma_{231}^\infty$ in Section 4.4.

Section 4 is structured as follow.

- In Section 4.1 we introduce a well-known bijection between binary trees and 231-avoiding permutations. Moreover we present a technique due to Janson [30]. With this tool we show that in order to study $\mathbb{E}[\tilde{c}\text{-occ}(\pi, \sigma^n)]$, it is enough to study a similar expectation in terms of a specific binary Galton–Watson tree.
- In Section 4.2 we study the behavior of $\mathbb{E}[\tilde{c}\text{-occ}(\pi, \sigma^n)]$ (see in particular Proposition 4.14) using the results from the previous section.
In Section 4.3 we prove Theorem 4.4 with similar techniques we study the second moment $E[\tilde{c}\text{-occ}(\pi, \sigma^m)^2]$ and then we conclude the proof applying the Second moment method.

In Section 4.4 we give an explicit construction of the limiting object $\sigma_{231}^\infty$.

In Section 4.5 we give some explicit examples of limiting probabilities.

4.1. Notation and preliminary results.

4.1.1. Notation. We introduce some more notation. When convenient, we extend in the trivial way the various notion introduced for permutations to arbitrary sequences of distinct numbers. For example, we extend the notion of $c\text{-occ}(\pi, \sigma)$ for two arbitrary sequences of distinct numbers $x_1 \ldots x_n$ and $y_1 \ldots y_k$ as

$$c\text{-occ}(y_1 \ldots y_k, x_1 \ldots x_n) := c\text{-occ}(\std(y_1 \ldots y_k), \std(x_1 \ldots x_n)).$$

Given $\sigma \in S^n$, let $\indmax(\sigma)$ be the index of the maximal value $n$, i.e., $\sigma_{\indmax(\sigma)} = n$. If $\ell = \indmax(\sigma)$, we set $\sigma_L := \sigma_1 \ldots \sigma_{\ell - 1}$ and $\sigma_R := \sigma_{\ell + 1} \ldots \sigma_n$, respectively the (possibly empty) left and right subsequences of $\sigma$, before and after the maximal value. In particular we have $\sigma = \sigma_L \sigma_{\ell} \sigma_R$, where we point out that $\sigma_L \sigma_R$ is not the composition of permutations but just the concatenation of $\sigma_L, \sigma_R$ seen as words.

We will write $\pat_{b(k)}(\sigma)$ for $\pat_{[1, k]}(\sigma)$ namely, for the pattern occurring in the first $k$ positions of $\sigma$. Similarly we will write $\pat_{e(k)}(\sigma)$ for $\pat_{[|\sigma| - k + 1, |\sigma|]}(\sigma)$ namely, for the pattern occurring in the last $k$ positions of $\sigma$. Note that $b(k)$ and $e(k)$ stands for beginning and end. Moreover, if either $k = 0$ or $k > |\sigma|$ we set

$$\pat_{b(k)}(\sigma) := \emptyset \text{ and } \pat_{e(k)}(\sigma) := \emptyset.$$

To simplify notation, sometimes, if it is clear from the context, we will simply write $\pat_b(\sigma) = \pi$ or $\pat_e(\sigma) = \pi$, instead of $\pat_{b(\pi)}(\sigma) = \pi$ or $\pat_{e(\pi)}(\sigma) = \pi$.

We finally recall that a Laurent polynomial over $\mathbb{R}$ is a linear combination of positive and negative powers of the variable with coefficients in $\mathbb{R}$, i.e., of the form

$$P(x) = a_{-m}x^{-m} + \cdots + a_{-1}x^{-1} + a_0 + a_1x + \cdots + a_nx^n,$$

where $m, n \in \mathbb{Z}_{\geq 0}$ and $a_i \in \mathbb{R}$ for all $-m \leq i \leq n$. We denote by $O(x^{\alpha})$, $\alpha \in \mathbb{Z}_{>0}$, an arbitrary Laurent polynomial in $x$ of valuation at least $-\alpha$, i.e., of the form $a_{-\alpha}x^{-\alpha} + \cdots + a_nx^n$. Moreover we denote by $O(x^\alpha)$, $\alpha \in \mathbb{Z}_{\geq 0}$, a classical polynomial in $x$ of valuation at least $\alpha$, i.e., of the form $a_0x^\alpha + \cdots + a_nx^n$. We remark that this is not the classical definition of $O(\cdot)$ since we are adding the additional hypothesis that the elements of $O(\cdot)$ are Laurent polynomials and not general functions.

4.1.2. A bijection between binary trees and 231-avoiding permutations. We present in this section a well-known bijection between 231-avoiding permutations and the set of binary trees (see e.g., [16]) which sends the size of the permutation to the number of vertices of the tree.

This bijection is based on the following result (see e.g., [15]).

Observation 4.7. Let $\sigma$ be a permutation and $\ell = \indmax(\sigma)$. The permutation $\sigma$ avoids 231 if and only if $\sigma_L$ and $\sigma_R$ both avoid 231 and furthermore $\sigma_i < \sigma_j$ whenever $i < \ell$ and $j > \ell$.

Given a permutation $\sigma \in \Av(231)$ we build the following binary tree $T = \varphi(\sigma)$: if $\sigma$ is empty then $T$ is the empty tree. Otherwise we add the root in $T$, which corresponds to the maximal element of $\sigma$ and we split $\sigma$ in $\sigma_L$ and $\sigma_R$; then the left subtree of $T$ will be the tree induced by $\sigma_L$, i.e., $T_L = \varphi(\sigma_L)$ and similarly, the right subtree of $T$ will be the tree induced by $\std(\sigma_R)$, i.e., $T_R = \varphi(\std(\sigma_R))$.

Conversely, given a binary tree $T$ we construct the corresponding permutation $\sigma = \psi(T)$ in $\Av(231)$ as follow (cf. Fig. 9 below): if $T$ is empty then $\sigma$ is the empty permutation. Otherwise we split $T$ in $T_L$ and $T_R$ and we set $n := |T|$, the number of vertices of $T$, $S_L = \psi(T_L)$ and $S_R = \psi(T_R)$. Then we define $\sigma$ as $\sigma = \sigma_1 \ldots \sigma_n$ where $\sigma_L$ is simply $S_L$ and $\sigma_R$ is obtained by shifting all entries of $S_R$ by $|T_L|$, i.e., $(\sigma_R)_j = (S_R)_j + |T_L|$, for all $j \leq |T_R|$. 
It is clear that \( \varphi \) and \( \psi \) are inverse of each other, hence providing the desired bijection.

\[
\begin{array}{c}
T_L \\
\downarrow \\
\psi \\
\uparrow \\
T_R
\end{array}
\]

**Figure 9.** A binary tree and the corresponding 231-avoiding permutation represented as a diagram.

From now until the end, we denote by \( T_\sigma \) the binary tree associated to \( \sigma \) and, analogously, by \( \sigma_T \) the permutation associated to the binary tree \( T \).

Moreover, we define
\[
c\text{-occ}(\pi, T) := c\text{-occ}(\pi, \sigma_T), \quad \text{for all } \pi \in \mathcal{S},
\]
and similarly
\[
\text{pat}_I(T) := \text{pat}_I(\sigma_T), \quad \text{for all finite interval } I \subset \mathbb{Z}_{>0}.
\]

**Observation 4.8.** By construction, the left-to-right maxima in a permutation \( \sigma \) correspond to the vertices in the left branch of \( T_\sigma \), i.e., to the vertices of the form \( v = 1 \ldots 1 \) in the Neveu’s formalism. Similarly, the right-to-left maxima correspond to the vertices in the right branch of \( T_\sigma \), i.e., to the vertices of the form \( v = 2 \ldots 2 \). Obviously the maximum corresponds to the root of the tree. An example is given in Fig. 10.

\[
\begin{array}{c}
12122212 \\
\downarrow
\end{array}
\]

**Figure 10.** Correspondence between maxima of a permutation and vertices on the left and right branches in the associated tree. We highlight in green the left-to-right maxima, in blue the right-to-left maxima and in orange the maximum.

We make the following final remark that is not strictly necessary in this section but useful for comparison with the next one.

**Remark 4.9.** It is not difficult to prove the following equivalent characterization of the bijection between trees and permutations based on the use of the tree traversals. Given a binary tree \( T \) with \( n \) vertices we reconstruct the corresponding permutation in \( \text{Av}(231) \) setting \( \sigma_i \) to be equal to the label given by the post-order to the \( i \)-th vertex visited by the in-order. Namely, we are using the in-order labeling for the indices of the permutation and the post-order labeling for the values. An example is provided in Fig. 11.
Figure 11. A binary tree and the corresponding 231-avoiding permutation. We are using the in-order labeling (in green) for the indices of the permutation and the post-order labeling (in blue) for the values.

4.1.3. From uniform 231-avoiding permutations to binary Galton–Watson tree. Thanks to the bijection between binary trees and 231-avoiding permutations, instead of studying \( E[\tilde{c}\text{-occ}(\pi, \sigma^n)] \) we can equivalently study the expectation \( E[\tilde{c}\text{-occ}(\pi, T_n)] \) where \( T_n \) denotes a uniform random binary tree with \( n \) vertices. A method for studying study the behavior of the latter expectation is to analyze the expectation \( E[\tilde{c}\text{-occ}(\pi, T_\delta)] \) for a binary Galton–Watson tree \( T_\delta \) defined as follow.

For \( \delta \in (0,1) \), we set \( p = \frac{1-\delta}{2} \), and we consider a binary Galton–Watson tree \( T_\delta \) with offspring distribution \( \eta \) defined by:

\[
\begin{align*}
\eta_0 &= \mathbb{P}(\text{The root has 0 child}) = (1-p)^2, \\
\eta_L &= \mathbb{P}(\text{The root has only one left child}) = p(1-p), \\
\eta_R &= \mathbb{P}(\text{The root has only one right child}) = p(1-p), \\
\eta_2 &= \mathbb{P}(\text{The root has 2 children}) = p^2. 
\end{align*}
\]

(34)

Remark 4.10. Note that with this particular offspring distribution, having a left child is independent from having a right child.

We now state two lemmas due to Svante Janson (see [30 Lemmas 5.2-5.4]). The first one is a result that rewrites the expectation of a function of \( T_\delta \) in terms of uniform random binary trees with \( n \) vertices \( T_n \), using the classical fact that \( \mathcal{L}(T_n) = \mathcal{L}(T_\delta | |T_\delta| = n) \). More precisely,

Lemma 4.11. Let \( F \) be a functional from the set of binary trees \( \mathcal{T}_b \) to \( \mathbb{R} \) such that \( |F(T)| \leq C \cdot |T|^m \) for some constants \( C \) and \( m \) (this guarantees that all expectations and sums below converge). Then

\[
\mathbb{E}[F(T_\delta)] = \frac{1 + \delta}{1 - \delta} \sum_{n=1}^{+\infty} \mathbb{E}[F(T_n)] \cdot C_n \cdot \left(\frac{1 - \delta^2}{4}\right)^n,
\]

where \( C_n \) is the \( n \)-th Catalan number.

Applying singularity analysis (see [21 Theorem VI.3]) one obtains:

Lemma 4.12. If \( \mathbb{E}[F(T_\delta)] = a \cdot \delta^{-m} + O(\delta^{-(m-1)}) \), for \( \delta \to 0 \), where \( m \geq 1 \) and \( a \neq 0 \), then

\[
\mathbb{E}[F(T_n)] \sim \frac{\Gamma(1/2)}{\Gamma(m/2)} \cdot a \cdot n^{m+1/2}, \quad \text{as} \quad n \to \infty.
\]
Remark 4.13. We recall that since by definition the elements of $O(\delta^{-(m+1)})$ are Laurent polynomials (and not general functions) then they are $\Delta$–analytic in $z = \frac{1-\delta^2}{4}$ (for a precise definition see Definition VI.1.) and so the standard hypothesis to apply singularity analysis is fulfilled.

Lemma 4.12 shows us why it is enough to study the behavior of $\mathbb{E}[\hat{\text{c-occ}}(\pi, T_n)]$ in order to derive information about the behavior of $\mathbb{E}[\hat{\text{c-occ}}(\pi, T_n)]$. This will be the goal of the next section.

4.2. The annealed version of the Benjamini–Schramm convergence. In this section we are going to prove the following weaker version of Theorem 4.4.

**Proposition 4.14.** Let $\sigma^n$ be a uniform random permutation in $\text{Av}^n(231)$ for all $n \in \mathbb{Z}_{>0}$. It holds that

$$\mathbb{E}[\hat{\text{c-occ}}(\pi, \sigma^n)] \rightarrow P_{231}(\pi), \text{ for all } \pi \in \text{Av}(231).$$

4.2.1. A basic recursion. Observation 4.7 gives us a basic recursion for $\text{c-occ}(\pi, \sigma)$ as we will show in the next lemmas.

**Lemma 4.15.** Let $\pi \in \text{Av}^k(231)$ with $k \geq 1$. Set $m = \text{indmax}(\pi)$. Then, for any permutation $\sigma \in \text{Av}^n(231)$ with $n \geq 1$,

$$\text{c-occ}(\pi, \sigma) = \text{c-occ}(\pi, \sigma_L) + \text{c-occ}(\pi, \sigma_R) + \mathbb{1}_{\{\text{pat}_{[\ell-m+1, \ell+k-m]}(\sigma) = \pi\}},$$

where as always $\ell = \text{indmax}(\sigma)$.

**Proof.** Recall that consecutive occurrences of $\pi$ in $\sigma$ correspond to intervals $I \subseteq [n]$ such that $\text{pat}_I(\sigma) = \pi$. We have three different possibilities for $I$.

1. If $I \subseteq [1, \ell - 1]$ then $\pi$ is a consecutive occurrence in $\sigma_L$.
2. If $I \subseteq [\ell + 1, n]$ then $\pi$ is a consecutive occurrence in $\sigma_R$.
3. If $I \subseteq [\ell]$ then $I$ has to be the interval $[\ell - m + 1, \ell + k - m]$. Indeed $\text{Card}(I) = k$ and the value induced in $\text{pat}_I(\pi)$ by $\sigma_\ell$ has to be $\pi_m$.

This is enough to prove Equation (35). \qed

We can translate this result in terms of trees.

**Lemma 4.16.** Let $\pi \in \text{Av}^k(231)$ with $k \geq 1$ and define $m$ as in the previous lemma. Then, for every binary tree $T$, denoting $\ell = \text{indmax}(\sigma_T)$,

$$\text{c-occ}(\pi, T) = \text{c-occ}(\pi, T_L) + \text{c-occ}(\pi, T_R) + \mathbb{1}_{\{\text{pat}_{[\ell-m+1, \ell+k-m]}(T) = \pi\}},$$

4.2.2. The behavior of $\mathbb{E}[\hat{\text{c-occ}}(\pi, \sigma^n)]$. We now focus on the behavior of $\mathbb{E}[\hat{\text{c-occ}}(\pi, T_n)]$ for $\pi \in \text{Av}(231)$, and then we will recover results for $\mathbb{E}[\hat{\text{c-occ}}(\pi, \sigma^n)]$ using Lemma 4.12 and the bijection between binary trees and 231-avoiding permutations as explained in Section 4.1.3. In order to simplify notation we set $T := T_3$. Thanks to Lemma 4.16 we know that, for all $\pi \in \text{Av}(231)$,

$$\text{c-occ}(\pi, T) = \text{c-occ}(\pi, T_L) + \text{c-occ}(\pi, T_R) + \mathbb{1}_{\{\text{pat}_J(T) = \pi\}},$$

where $J = [\ell - m + 1, \ell + k - m]$, $\ell = \text{indmax}(\sigma_T)$ and $m = \text{indmax}(\pi)$. Taking the expectation in Equation (37) we obtain,

$$\mathbb{E}[\hat{\text{c-occ}}(\pi, T)] = \mathbb{E}[\hat{\text{c-occ}}(\pi, T_L)] + \mathbb{E}[\hat{\text{c-occ}}(\pi, T_R)] + \mathbb{P}(\text{pat}_J(T) = \pi).$$

Since $T_L$ is an independent copy of $T$ with probability $p$ and empty with probability $1 - p$, and the same holds for $T_R$, we have,

$$\mathbb{E}[\hat{\text{c-occ}}(\pi, T)] = \frac{\mathbb{P}(\text{pat}_J(T) = \pi)}{1 - 2p} = \delta^{-1} \cdot \mathbb{P}(\text{pat}_J(T) = \pi),$$

where in the last equality we used that $p = \frac{1-\delta}{2}$. We now focus on the term $\mathbb{P}(\text{pat}_J(T) = \pi)$.
Lemma 4.17. Let $\pi \in \text{Av}(231)$. Using notation as above and decomposing as usual $\pi$ in
\[ \pi = \pi_L \pi_m \pi_R, \]
we have
\[
\mathbb{P}(\text{pat}_J(T) = \pi) = \begin{cases} 
\rho^2 \cdot \mathbb{P}(\text{pat}_e(T) = \pi_L) \cdot \mathbb{P}(\text{pat}_b(T) = \pi_R), & \text{if } \pi_L \neq \emptyset, \pi_R \neq \emptyset, \\
\rho \cdot \mathbb{P}(\text{pat}_e(T) = \pi_L), & \text{if } \pi_L \neq \emptyset, \pi_R = \emptyset, \\
\rho \cdot \mathbb{P}(\text{pat}_b(T) = \pi_R), & \text{if } \pi_L = \emptyset, \pi_R \neq \emptyset, \\
1, & \text{if } \pi = 1.
\end{cases}
\]

Proof. First of all note that
\[
\mathbb{P}(\text{pat}_J(T) = \pi) = \mathbb{P}(\text{pat}_e(T_L) = \pi_L, \text{pat}_b(T_R) = \pi_R).
\]
We recall that if $\pi_L$ (resp. $\pi_R$) is empty then $\text{pat}_e(T_L)$ (resp. $\text{pat}_b(T_R)$) is empty by definition.
We consider the case when $\pi_L \neq \emptyset, \pi_R \neq \emptyset$. We have,
\[
\mathbb{P}(\text{pat}_J(T) = \pi) = \rho^2 \cdot \mathbb{P}(\text{pat}_e(T) = \pi_L) \cdot \mathbb{P}(\text{pat}_b(T) = \pi_R),
\]
since in Equation (40) the random tree $T_L$ is an independent copy of $T$ with probability $\rho$ and empty with probability $1 - \rho$ and obviously the same hold also for $T_R$. The other three cases are similar. \qed

In view of Lemma 4.17, we now focus on $\mathbb{P}(\text{pat}_e(T) = \pi)$ (the analysis for $\mathbb{P}(\text{pat}_b(T) = \pi)$ following by symmetry). We want to rewrite the event $\{\text{pat}_e(T) = \pi\}$ conditioning on the position of the maximum among the last $|\pi|$ values of $\sigma_T$. Using Observation 4.8 we know that this maximum is always reached at an index of $\sigma_T$ corresponding to a vertex of $T$ of the form
\[
v = \frac{2}{n}, \qquad \text{for some } n \in \mathbb{Z}_{\geq 0},
\]
with the convention $v_{20} := \emptyset$. For an example, see Fig. 12.

Therefore, defining the following events, for all $n \in \mathbb{Z}_{\geq 0}$, all $\pi \in \text{Av}(231)$,
\[
M^n_\pi := \{v_{2n} \text{ is the vertex in } T \text{ corresponding to the maximum among the last } |\pi| \text{ values of } \sigma_T\}
\]
and using the formula of total probability, we have
\[
\mathbb{P}(\text{pat}_e(T) = \pi) = \sum_{n=0}^{\infty} \mathbb{P}(\text{pat}_e(T) = \pi, M^n_\pi).
\]
We also introduce the events, for all $n \in \mathbb{Z}_{\geq 0}$,
\[
R^n := \{\text{The vertex } v_{2n} \text{ is in } T\}.
\]
Note that
\begin{equation}
\sum_{n=0}^{\infty} P(R^n) = \sum_{n=0}^{\infty} p^n = \frac{1}{1-p}.
\end{equation}

The following lemma computes recursively \( P(\text{pat}_e(T) = \pi) \) and \( P(\text{pat}_b(T) = \pi) \).

**Lemma 4.18.** Let \( \pi \in \text{Av}(231) \). Using notation as above,
\begin{equation}
P(\text{pat}_e(T) = \pi) = \begin{cases}
p^2 \cdot P(\text{pat}_e(T) = \pi_L, T = T_{\pi_R}), & \text{if } \pi_L \neq \emptyset, \pi_R \neq \emptyset, \\
p \cdot P(\text{pat}_e(T) = \pi_L, T = T_{\pi_R}), & \text{if } \pi_L \neq \emptyset, \pi_R = \emptyset, \\
\frac{p}{1-p} \cdot P(T = T_{\pi_R}), & \text{if } \pi_L = \emptyset, \pi_R \neq \emptyset, \\
1, & \text{if } \pi = 1,
\end{cases}
\end{equation}

and
\begin{equation}
P(\text{pat}_b(T) = \pi) = \begin{cases}
p^2 \cdot P(T = T_{\pi_L}, \text{pat}_b(T) = \pi_R), & \text{if } \pi_L \neq \emptyset, \pi_R \neq \emptyset, \\
p \cdot P(T = T_{\pi_L}, \text{pat}_b(T) = \pi_R), & \text{if } \pi_L \neq \emptyset, \pi_R = \emptyset, \\
\frac{p}{1-p} \cdot P(\text{pat}_b(T) = \pi_R), & \text{if } \pi_L = \emptyset, \pi_R \neq \emptyset, \\
1, & \text{if } \pi = 1.
\end{cases}
\end{equation}

**Proof.** We start with the study of \( P(\text{pat}_e(T) = \pi) \). We distinguish again four different cases according to the structure of \( \pi_L \) and \( \pi_R \).

- \( \pi_L \neq \emptyset, \pi_R \neq \emptyset \): By Equation \([41]\), we know that
  \[ P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_0} P(\text{pat}_e(T) = \pi, M^n_\pi) \]
  \[ = \sum_{n \in \mathbb{Z}_0} P(\text{pat}_e(T^{v_{2n}}) = \pi_L, T^{v_{2n}}_{\pi_R} = T_{\pi_R}, M^n_\pi), \]

  where in the last equality we used that conditioning on \( v_{2n} \) being the vertex in \( T \) corresponding to the maximum among the last \(|\pi|\) values of \( \sigma_T \), then \( \text{pat}_e(\pi)(T) = \text{pat}_e(\pi_{\{n\}})(T^{v_{2n}}) \).

Since the event \( \{ \text{pat}_e(T^{v_{2n}}) = \pi_L \} \cap \{ T^{v_{2n}}_{\pi_R} = T_{\pi_R} \} \) is contained both in \( M^n_\pi \) and in \( R^n \), then
\[ P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_0} P(\text{pat}_e(T^{v_{2n}}) = \pi_L, T^{v_{2n}}_{\pi_R} = T_{\pi_R}, R^n). \]

Using the independence between \( T^{v_{2n}}_L \) and \( T^{v_{2n}}_R \) conditionally on \( R^n \) and continuing the sequence of equalities,
\[ P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_0} P(\text{pat}_e(T^{v_{2n}})_L = \pi_L|R^n) \cdot P(\text{pat}_e(T^{v_{2n}})_R = \pi_{\pi_{\pi}}|R^n) \cdot P(R^n). \]

Since, conditionally on \( R^n \), \( T^{v_{2n}}_L \) is an independent copy of \( T \) with probability \( p \) and empty with probability \( 1-p \) and the same obviously holds for \( T^{v_{2n}}_R \), we can rewrite the last term as
\[ P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_0} p^2 \cdot P(\text{pat}_e(T) = \pi_L) \cdot P(T = T_{\pi_R}) \cdot P(R^n) \]
\[ = p^2 \cdot P(\text{pat}_e(T) = \pi_L) \cdot P(T = T_{\pi_R}) \cdot \sum_{n \in \mathbb{Z}_0} P(R^n) \]
\[ = \frac{p^2}{1-p} \cdot P(\text{pat}_e(T) = \pi_L) \cdot P(T = T_{\pi_R}), \]

where in the last equality we used Equation \([42]\).
• \(\pi_L \neq \emptyset\), \(\pi_R = \emptyset\): Similarly as before

\[
P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_{\geq 0}} P(\text{pat}_e(T_L^{v_{2n}}) = \pi_L, T_R^{v_{2n}} = \emptyset | R^n) \cdot P(R^n).
\]

Noting that with similar arguments as before we have

\[
P(\text{pat}_e(T_L^{v_{2n}}) = \pi_L, T_R^{v_{2n}} = \emptyset | R^n) = p \cdot P(\text{pat}_e(T) = \pi_L) \cdot (1 - p).
\]

and using again Equation (42), we can conclude that

\[
P(\text{pat}_e(T) = \pi) = p \cdot P(\text{pat}_e(T) = \pi_L).
\]

• \(\pi_L = \emptyset\), \(\pi_R \neq \emptyset\): Again

\[
P(\text{pat}_e(T) = \pi) = \sum_{n \in \mathbb{Z}_{\geq 0}} P(\text{pat}_e(T_L^{v_{2n}}) = \emptyset, T_R^{v_{2n}} = T_{\pi_R} | R^n) \cdot P(R^n).
\]

Noting that \(\text{pat}_e(0)(T_L^{v_{2n}})\) is empty by definition and using similar arguments as before, we have

\[
P(\text{pat}_e(T_L^{v_{2n}}) = \emptyset, T_R^{v_{2n}} = T_{\pi_R} | R^n) = p \cdot P(T = T_{\pi_R}).
\]

Therefore, using again Equation (42), we can write

\[
P(\text{pat}_e(T) = \pi) = \frac{p}{1 - p} \cdot P(T = T_{\pi_R}).
\]

• \(\pi_L = \emptyset\), \(\pi_R = \emptyset\): Since \(\pi = 1\) then \(P(\text{pat}_e(T) = \pi) = 1\).

Since the result for \(P(\text{pat}_b(T) = \pi)\) follows by symmetry, we can conclude the proof. \(\square\)

We now continue the analysis of \(P(\text{pat}_f(T) = \pi)\). In order to do that, we need a formula for \(P(T = T)\), for a given tree \(T\), because such probabilities appear in Equations (43) and (44).

**Observation 4.19.** In a binary tree with \(n\) vertices, every vertex has two potential children. Out of these \(2n\) potential children \(n - 1\) exist and \(n + 1\) do not exist. Hence

\[
P(T = T) = p^{|T| - 1} \cdot (1 - p)^{|T| + 1}.
\]

Using together Lemmas 4.17 and 4.18 and the above observation we have an explicit recursion to compute the probability \(P(\text{pat}_f(T) = \pi)\). We show an example of the recursion obtained for an explicit pattern \(\pi\).

**Example 4.20.** Let \(\pi\) be the following 231-avoiding permutation,

\[
\pi = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \ 11 \ 12 \ 16 \ 13 \ 15 \ 14 = \begin{array}{c}
\end{array}
\]

where, as before, we draw in green the left-to-right maxima, in blue the right-to-left maxima, and in orange the maximum.

We now make the explicit computation for \(P(\text{pat}_f(T) = \pi)\) using Lemmas 4.17 and 4.18. First of all, we recursively split our permutation as shown in Fig. 13.
\[ \pi = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \ 11 \ 12 \ 16 \ 13 \ 15 \ 14 \]

\[ \pi_L = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \ 11 \ \underline{12} \]
\[ \pi_R = 1 \ 3 \ 2 \ 16 \ 13 \ 15 \ 14 \]

\[ \pi_{LL} = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \ 11 \]
\[ \pi_{RR} = 1 \ 3 \ 2 \]
\[ \pi_{RL} = 1 \]
\[ \pi_{RR} = 1 \]
\[ \pi_{LL} = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \]
\[ \pi_{LLLL} = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \]
\[ \pi_{LLLL} = 4 \ 1 \ 3 \ 2 \]
\[ \pi_{LLLL} = 4 \ 1 \ 3 \ 2 \]
\[ \pi_{LLLL} = 1 \ 3 \ 2 \]
\[ p^2 \]
\[ (1 - p)^2 \]
\[ p^2 \]

**Figure 13.** We draw the recursive decompositions in the left and right part (w.r.t. the position of the maximum underlined in orange) for the permutation \( \pi = 4 \ 1 \ 3 \ 2 \ 6 \ 5 \ 7 \ 10 \ 8 \ 9 \ 11 \ 12 \ 16 \ 13 \ 15 \ 14 \). Moreover, we write in green the factors \( p^\alpha \cdot (1 - p)^\beta \) that we are adding at each step (coming from Lemmas 4.17 and 4.18).

Then, using Equation (39) with the decomposition of \( \pi \) in \( \pi_L \pi_m \pi_R \) (shown at the root of the tree in Fig. 13), we have

\[ \mathbb{P}(\text{pat}_J(T) = \pi) = p^2 \cdot \mathbb{P}(\text{pat}_e(T) = \pi_L) \cdot \mathbb{P}(\text{pat}_b(T) = \pi_R). \]

We continue decomposing \( \pi_L \) and \( \pi_R \) around their maxima (which correspond to the left and right children of the root in the tree in Fig. 13). Using Equation (43) for the left part, we obtain

\[ \mathbb{P}(\text{pat}_e(T) = \pi_L) = p \cdot \mathbb{P}(\text{pat}_e(T) = \pi_{LL}), \]

and using Equation (44) for the right part, we obtain

\[ \mathbb{P}(\text{pat}_b(T) = \pi_R) = \frac{p^2}{1 - p} \cdot \mathbb{P}(T = T_{\pi_{RL}}) \cdot \mathbb{P}(\text{pat}_b(T) = 1). \]

where the first probability in the right-hand side is computed using Equation (45). Therefore, summing up the last three equations and then proceeding similarly through the left subtree in Fig. 13, we deduce that

\[ \mathbb{P}(\text{pat}_J(T) = \pi) = p^2 \cdot p \cdot \mathbb{P}(\text{pat}_e(T) = \pi_{LL}) \cdot \frac{p^2}{1 - p} \cdot (1 - p)^2 = \ldots \]

\[ = p^2 \cdot p \cdot \frac{p^2}{1 - p} \cdot p \cdot \frac{p^2}{1 - p} \cdot \frac{p^2}{1 - p} \cdot p(1 - p)^4 \cdot (1 - p)^2 \cdot p(1 - p)^3 \cdot \frac{p^2}{1 - p} \cdot (1 - p)^2 \]

\[ = p^{15}(1 - p)^{11} = \left( \frac{1 - \delta}{2} \right)^{15} \cdot \left( \frac{1 + \delta}{2} \right)^7 = \left( \frac{1}{2} \right)^{22} + O(\delta), \]

and this concludes Example 4.20.
We now proceed with the analysis of the general case. Using the recursion obtained by combining Lemmas 4.17 and 4.18 and Observation 4.19, we immediately realize that

\[ P(\text{pat}_J(T) = \pi) = p^\alpha \cdot (1 - p)^\beta, \quad \text{for some } \alpha, \beta \in \mathbb{Z}_{\geq 0}. \]

(46)

Note that \( \alpha \geq 0 \) since in Lemmas 4.17 and 4.18 and Observation 4.19 the \( p^* \) factors always appear with positive exponent. Moreover, \( \beta \geq 0 \) since in Equations (43) and (44), each time a \( (1 - p)^{-1} \) factor appears, there is also a \( p^* \) factor that contains a \( (1 - p)^\gamma \) factor with \( \gamma \geq 2 \).

Since \( p = \frac{1 - \delta}{2} \), it follows that

\[ P(\text{pat}_J(T) = \pi) = \left( \frac{1 - \delta}{2} \right)^\alpha \cdot \left( \frac{1 + \delta}{2} \right)^\beta = \left( \frac{1}{2} \right)^{\alpha + \beta} + O(\delta), \quad \text{for some } \alpha, \beta \in \mathbb{Z}_{\geq 0}. \]

We now compute the value \( \alpha + \beta \). In order to do that, we have to compute how many factors \( p \) and \( (1 - p) \) appear in \( P(\text{pat}_J(T) = \pi) \).

Before stating our result we need to introduce a notion of distance between maxima. Given an index \( j \in \text{LRMax}(\pi) \) which is not the maximum of \( \pi \), we define its distance from the following left-to-right maximum as

\[ \partial_\pi(j) := \min \{ a \geq 0 : j + 1 + a \in \text{LRMax}(\pi) \}. \]

Analogously, given an index \( j \in \text{RLMax}(\pi) \) which is not the maximum of \( \pi \), we define its distance from the following right-to-left maximum as

\[ \partial_\pi(j) := \min \{ a \geq 0 : j - 1 - a \in \text{RLMax}(\pi) \}. \]

By convention, if \( j \) is the index of the maximum of \( \pi \), then \( \partial_\pi(j) := 0 \).

**Example 4.21.** We continue Example 4.2 where we considered the permutation

\[ \pi = 132985476 = \]

\[ \begin{array}{cccc}
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\bullet & \bullet & \bullet & \bullet \\
\end{array} \]

For \( \text{LRMax}(\pi) = \{1, 2, 4\} \) we have \( \partial_\pi(1) = 0, \partial_\pi(2) = 1, \partial_\pi(4) = 0 \) and for \( \text{RLMax}(\pi) = \{9, 8, 5, 4\} \) we have \( \partial_\pi(9) = 0, \partial_\pi(8) = 2, \partial_\pi(5) = 0 \) and \( \partial_\pi(4) = 0 \).

We are now ready to state the key result of this section.

**Proposition 4.22.** Let \( \pi \in \text{Av}(231) \) and \( T = T_\delta \) be a Galton–Watson tree defined as above, then

\[ P(\text{pat}_J(T) = \pi) = \frac{2^{\lfloor \text{LRMax}(\pi) \rfloor + \lfloor \text{RLMax}(\pi) \rfloor}}{2^{2|\pi|}} + O(\delta) = P_{231}(\pi) + O(\delta). \]

(47)

Note that Proposition 4.22 claims that for the specific pattern \( \pi \) given in Example 4.20 we have,

\[ P(\text{pat}_J(T) = \pi) = \frac{2^{7+3}}{2^{32}} + O(\delta) = \left( \frac{1}{2} \right)^{22} + O(\delta), \]

(48)

as previously computed. With this example in our hands, we can now use the same ideas in order to prove the proposition.

**Proof of Proposition 4.22.** Suppose that we are able to prove that

\[ P(\text{pat}_J(T) = \pi) = \left( \frac{1}{2} \right)^{|\text{Max}(|\pi|) - 1 + 2 \sum_{j \in \text{Max}(\pi)} \partial_\pi(j)} + O(\delta). \]

(49)

Then noting that \( |\text{Max}(\pi)| + \sum_{j \in \text{Max}(\pi)} \partial_\pi(j) = |\pi| \) and that

\[ |\text{Max}(\pi)| + 1 = |\text{LRMax}(\pi)| + |\text{RLMax}(\pi)|, \]

we complete the proof.
we immediately derives the desired expression
\[
\mathbb{P}(\text{pat}_\mathcal{J}(T) = \pi) = \frac{2^{\lfloor \text{LRMax}(\pi) \rfloor + \lfloor \text{RLMax}(\pi) \rfloor}}{2^{\|\pi\|}} + O(\delta).
\]

We now prove Equation (49). As explained in the discussion below Equation (46), we just need to count how many factors \( p \) and \( (1 - p) \) appear in the formula for \( \mathbb{P}(\text{pat}_\mathcal{J}(T) = \pi) \). We can suppose that \( |\pi| \geq 2 \) (if \( |\pi| = 1 \) then the statement is trivial). Then by Equation (39),

\[
\mathbb{P}(\text{pat}_\mathcal{J}(T) = \pi) = \begin{cases} 
p^2 \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_L) \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_R), & \text{if } \pi_L \neq \emptyset, \pi_R \neq \emptyset, 
p \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_L), & \text{if } \pi_L \neq \emptyset, \pi_R = \emptyset, 
p \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_R), & \text{if } \pi_L = \emptyset, \pi_R \neq \emptyset, 
1, & \text{if } \pi_L = 1.
\end{cases}
\]

(50) and so we obtain either a factor \( p \sim \frac{1}{2} \) (if the maximum of \( \pi \) is reached at the first or at the last index of \( \pi \)) or a factor \( p^2 \sim \left( \frac{1}{2} \right)^2 \) (otherwise). We recall that the asymptotics are for \( \delta \to 0 \).

We now look at the factors of the form \( \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_L) \). By Equation (43),

\[
\mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_L) = \begin{cases} 
\frac{p^2}{1-p} \cdot \mathbb{P}(T = T_{LR}) \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_{LL}), & \text{if } \pi_{LL} \neq \emptyset, \pi_{LR} \neq \emptyset, 
p \cdot \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_{LL}), & \text{if } \pi_{LL} \neq \emptyset, \pi_{LR} = \emptyset, 
p \cdot \mathbb{P}(T = T_{LR}), & \text{if } \pi_{LL} = \emptyset, \pi_{LR} \neq \emptyset, 
1, & \text{if } \pi_L = 1.
\end{cases}
\]

(51) We are going to focus on the factors that are not of the form \( \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_{LL}) \). We call them prefactors. For example, when \( \pi_{LL} \neq \emptyset \) and \( \pi_{LR} \neq \emptyset \) the prefactor is \( \frac{p^2}{1-p} \cdot \mathbb{P}(T = T_{LR}) \).

Therefore, denoting by \( j := \text{indmax}(\pi_L) \), we obtain:

(1) For \( |\pi_L| \geq 2 \):
- If \( j \neq 1 \) and \( j \neq |\pi_L| \), then we have the following prefactor,
  \[
p^2 \cdot p^{\partial_\pi(j)-1} \cdot (1 - p)^{\partial_\pi(j)+1} \cdot \left( \frac{1}{2} \right)^{1+2\partial_\pi(j)}.
\]
- If \( j = |\pi_L| \) (in this case \( \partial_\pi(j) = 0 \)), then we have the following prefactor,
  \[
p \sim \frac{1}{2} = \left( \frac{1}{2} \right)^{1+2\partial_\pi(j)};
\]
- If \( j = 1 \) (in this case \( \partial_\pi(j) = |\pi_{LR}| \)), then we have the following prefactor,
  \[
p \cdot p^{\partial_\pi(j)-1} \cdot (1 - p)^{\partial_\pi(j)+1} \cdot \left( \frac{1}{2} \right)^{2\partial_\pi(j)};
\]

(2) For \( |\pi_L| = 1 \), i.e., \( \pi_L = 1 \):
- In this case the index of the trivial maximum is \( j = 1 \) and obviously \( \partial_\pi(j) = 0 \), hence we have the following prefactor,
  \[
1 = \left( \frac{1}{2} \right)^{2\partial_\pi(j)}.
\]

Summing up, and noting that the analysis of the term \( \mathbb{P}(\text{pat}_\mathcal{P}(T) = \pi_R) \) is symmetric, we conclude that we have the following correspondence between maxima of \( \pi \) and \( \left( \frac{1}{2} \right)^* \) factor in the asymptotic expansion of \( \mathbb{P}(\text{pat}_\mathcal{J}(T) = \pi) \) :
(a) If $j$ is the index of the maximum of $\pi$, by the comment below Equation (50), its contribution is
\[
\left(\frac{1}{2}\right)^2, \quad \text{if } j \neq 1 \text{ and } j \neq |\pi|,
\]
\[
\frac{1}{2}, \quad \text{otherwise};
\]
(b) If $j \in \text{Max}(\pi)$, $j \neq 1$ and $j \neq |\pi|$, but $j$ is not the index of the maximum, by the previous analysis, its contribution is
\[
\left(\frac{1}{2}\right)^{1+2\delta_e(j)};
\]
(c) If $j \in \text{Max}(\pi)$, $j = 1$ or $j = |\pi|$, but $j$ is not the index of the maximum, again by the previous analysis, its contribution is
\[
\left(\frac{1}{2}\right)^{2\delta_e(j)}.
\]
Finally, we note that,

- If $\text{indmax}(\pi) = 1$ or $\text{indmax}(\pi) = |\pi|$ then
  \[
  \mathbb{P}(\text{pat}_J(T) = \pi) = \left(\frac{1}{2}\right)^{1+|\text{Max}(\pi)|-2+2\sum_{j \in \text{Max}(\pi)} \delta_e(j)} + O(\delta),
  \]
  where the term 1 in the exponent comes from $(a)$ and the term $|\text{Max}(\pi)| - 2$ comes from $(b)$;
- If $\text{indmax}(\pi) \neq 1$ and $\text{indmax}(\pi) \neq |\pi|$ then
  \[
  \mathbb{P}(\text{pat}_J(T) = \pi) = \left(\frac{1}{2}\right)^{2+|\text{Max}(\pi)|-3+2\sum_{j \in \text{Max}(\pi)} \delta_e(j)} + O(\delta),
  \]
  where the term 2 in the exponent comes from $(a)$ and the term $|\text{Max}(\pi)| - 3$ comes from $(b)$.

This proves Equation (49) and so it concludes the proof.

\[\square\]

**Observation 4.23.** For later goals (see the proof of Proposition 4.29), we note that from the analysis below Equation (51), it also follows easily that
\[
\mathbb{P}(\text{pat}_b(T) = \pi) = \frac{2^{|\text{RLMax}(\pi)|+1}}{2^{2|\pi|}}, \quad \text{for all } \pi \in \text{Av}(231),
\]
\[
\mathbb{P}(\text{pat}_e(T) = \pi) = \frac{2^{|\text{LRMax}(\pi)|+1}}{2^{2|\pi|}}, \quad \text{for all } \pi \in \text{Av}(231).
\]

We can finally prove Proposition 4.14.

**Proof of Proposition 4.14** Summing up all the results and recalling that $T = T_\delta$, we obtain,
\[
\mathbb{E}[c-\text{occ}(\pi, T_\delta)] \overset{(52)}{=} \delta^{-1} \mathbb{P}(\text{pat}_J(T_\delta) = \pi) \overset{(52)}{=} \delta^{-1} \left(P_{231}(\pi) + O(\delta)\right) = \delta^{-1} \cdot P_{231}(\pi) + O(1).
\]

Applying Lemma 4.12 and using the bijection between 231-avoiding permutations and binary trees, we conclude that for $n \to \infty$,
\[
\mathbb{E}[c-\text{occ}(\pi, \sigma^n)] \sim P_{231}(\pi) \cdot n, \quad \text{for all } \pi \in \text{Av}(231).
\]

Dividing by $n$ yields Proposition 4.14. \[\square\]
4.3. The quenched version of the Benjamini–Schramm convergence. In this section we finally prove Theorem 4.14. We are going to use a standard technique in probability theory called the Second moment method. We have to study the second moment \( \mathbb{E}[\tilde{c}\text{-occ}(\pi, \sigma^n)^2] \). As before we start with a result regarding trees and then we will extend it to permutations.

**Proposition 4.24.** Using notation as before and setting \( T = T_\delta \), we have,

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T)^2] = \frac{P_{231}(\pi)^2}{2} \cdot \delta^{-2} + O(\delta^{-2}), \quad \text{for all } \pi \in \text{Av}(231).
\]

**Proof.** We use again the decomposition given by Lemma 4.16,

\[
\tilde{c}\text{-occ}(\pi, T) = \tilde{c}\text{-occ}(\pi, T_L) + \tilde{c}\text{-occ}(\pi, T_R) + 1\{\text{pat}_j(T) = \pi\}.
\]

Taking the expectation of the square of the previous equation we obtain,

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T)^2] = \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)^2] + \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_R)^2] + 2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)\tilde{c}\text{-occ}(\pi, T_R)] + 2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_R)1\{\text{pat}_j(T) = \pi\}] + \mathbb{P}(\text{pat}_j(T) = \pi).
\]

For the first two terms of the right-hand side of the above equation we use that \( T_L \) is an independent copy of \( T \) with probability \( p \) and empty with probability \( 1 - p \) (and similarly for \( T_R \)). Therefore we can rewrite the above expression as

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T)^2] = \delta^{-2} \left(2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)\tilde{c}\text{-occ}(\pi, T_R)] + 2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)1\{\text{pat}_j(T) = \pi\}] + 2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_R)1\{\text{pat}_j(T) = \pi\}] + \mathbb{P}(\text{pat}_j(T) = \pi)\right).
\]

(53)

where we have also used that \( 1 - 2p = \delta \). We now analyze the four terms that appear on the right-hand side of the above equation.

First of all, using that \( T_L \) is independent of \( T_R \), we obtain

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)\tilde{c}\text{-occ}(\pi, T_R)] = p^2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T)^2].
\]

Applying Equation (52) and recalling that \( p = \frac{1-\delta}{2} \), we conclude that

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)\tilde{c}\text{-occ}(\pi, T_R)] = \frac{P_{231}(\pi)^2}{4} \cdot \delta^{-2} + O(\delta^{-1}).
\]

(54)

Obviously, using Proposition 4.22 we have

\[
\mathbb{P}(\text{pat}_j(T_\delta) = \pi) = P_{231}(\pi) + O(\delta).
\]

(55)

Therefore we have only to study the remaining two similar terms in Equation (53). Note that, if \( \pi_L \neq \emptyset \) and \( \pi_R \neq \emptyset \), using the same arguments that we already used in a lot of different steps, we have

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)1\{\text{pat}_j(T) = \pi\}] = \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)1\{\text{pat}_s(T_L) = \pi_L\}1\{\text{pat}_b(T_R) = \pi_R\}]
\]

\[
= \mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)1\{\text{pat}_s(T_L) = \pi_L\}] \mathbb{E}[1\{\text{pat}_b(T_R) = \pi_R\}]
\]

\[
= p^2 \cdot \mathbb{E}[\tilde{c}\text{-occ}(\pi, T)1\{\text{pat}_s(T) = \pi_L\}] \cdot \mathbb{P}(\text{pat}_b(T) = \pi_R).
\]

(56)

By Equation (52), \( \mathbb{E}[\tilde{c}\text{-occ}(\pi, T)] = O(\delta^{-1}) \). Therefore we can conclude that

\[
\mathbb{E}[\tilde{c}\text{-occ}(\pi, T_L)1\{\text{pat}_j(T) = \pi\}] = O(\delta^{-1}).
\]
Similarly,
\begin{equation}
\mathbb{E}[c\text{-occ}(\pi, T_R) \mathbb{I}_{\{\text{pat}, j(\pi) = \pi\}}] = O(\delta^{-1}).
\end{equation}
Substituting Equations (54), (55), (56) and (57) in Equation (53), we finally have
\begin{equation}
\mathbb{E}[c\text{-occ}(\pi, T)^2] = \frac{P_{231}(\pi)^2}{2} \cdot \delta^{-3} + O(\delta^{-2}).
\end{equation}

We can finally prove our main theorem.

**Proof of Theorem 4.4.** Applying again Lemma 4.12 and using our bijection between 231-avoiding permutations and binary trees, we conclude that for \( n \to \infty \),
\begin{equation}
\mathbb{E}[c\text{-occ}(\pi, \sigma^n)^2] \sim P_{231}(\pi)^2 \cdot n^2, \quad \text{for all } \pi \in \text{Av}(231).
\end{equation}
This, with Proposition 4.14 implies that
\begin{equation}
\text{Var}(c\text{-occ}(\pi, \sigma^n)) \to 0, \quad \text{for all } \pi \in \text{Av}(231).
\end{equation}
We can apply the Second moment method and deduce that
\begin{equation}
c\text{-occ}(\pi, \sigma^n) \overset{P}{\to} P_{231}(\pi), \quad \text{for all } \pi \in \text{Av}(231).
\end{equation}
Indeed by Chebyschev’s inequality, one has, for any fixed \( \varepsilon > 0 \),
\begin{equation}
\mathbb{P}\left(|c\text{-occ}(\pi, \sigma^n) - \mathbb{E}[c\text{-occ}(\pi, \sigma^n)]| \geq \varepsilon \right) \leq \frac{1}{\varepsilon^2} \cdot \text{Var}(c\text{-occ}(\pi, \sigma^n)),
\end{equation}
and the right-hand side tends to zero. \( \square \)

4.4. **The construction of the limiting object.** We now exhibit an explicit construction of the limiting object \( \sigma_{231}^\infty \) as a random order \( \mathfrak{S}_{231} \) on \( \mathbb{Z} \).

**Remark 4.25.** The intuition behind the construction that we are going to present comes from the local limit for uniform binary trees (for more details, see [52]). When the size of a uniform binary tree tends to infinity, looking around a uniform distinguished vertex, we see an infinite upward spine. Each vertex in this spine is the left or the right child of the previous one with probability \( 1/2 \). Moreover, attached to this infinite spine there are some independent copies of binary Galton–Watson trees. Using this idea and the bijection between binary trees and 231-avoiding permutations, we are going to construct the limiting random total order.

This intuition is not formally needed in the following since we present the construction of the limiting object using the permutations point of view.

We have to introduce some notation. We define two operations from \( \mathfrak{S}^k_k \times \mathfrak{S}^\ell^\ell \) to \( \mathfrak{S}^{k+\ell+1}_{k+\ell+1} \), for \( k > 0 \) and \( \ell \geq 0 \) (\( \mathfrak{S}^0 \) is the set containing the empty permutation): let \( (\sigma, i) \in \mathfrak{S}^k_k \) be a rooted permutation and \( \pi \in \mathfrak{S}^\ell^\ell \) be another (unrooted) permutation, we set
\begin{align*}
(\sigma, i) \ast^R \pi &:= (\sigma_1 \ldots \sigma_k(\ell + k + 1)(\pi_1 + k)\ldots(\pi_\ell + k), i), \\
(\sigma, i) \ast^L \pi &:= (\pi_1 \ldots \pi_\ell(\ell + k + 1)(\sigma_1 + \ell)\ldots(\sigma_k + \ell), \ell + i + 1).
\end{align*}

In words, from a graphical point of view, the diagram of \( (\sigma, i) \ast^R \pi \) (resp. \( (\sigma, i) \ast^L \pi \)) is obtained starting from the diagram of the rooted permutation \( (\sigma, i) \), adding on the top-right (resp. bottom-left) the diagram of \( \pi \) and adding a new maximal element between the two diagrams. We give an example below.

**Example 4.26.** Let \( (\sigma, i) = (132, 3) \) and \( \pi = 21 \) then
\begin{align*}
(\sigma, i) \ast^R \pi &= \begin{array}{c}
\bullet \\
\bullet \\
\bullet \\
\bullet
\end{array} = (132621, 3) \quad \text{and} \quad (\sigma, i) \ast^L \pi &= \begin{array}{c}
\bullet \\
\bullet
\end{array} = (216354, 6).
\end{align*}
Observation 4.27. Note that, given a rooted permutation \((\sigma, i)\), for all permutations \(\pi\) (possibly empty), the rooted permutation \((\sigma, i) \ast_R \pi\) (resp. \((\sigma, i) \ast_L \pi\)) contains at least one more element on the right (resp. left) of the root than \((\sigma, i)\).

We now consider two families of random variables: let \(\{\mathbf{S}^i\}_{i \geq 0}\) be i.i.d. random variables with values in \(\Av(231) \cup \{\emptyset\}\) and Boltzman distribution equal to

\[
\mathbb{P}(\mathbf{S}^i = \pi) = \frac{1}{2} \left(\frac{1}{4}\right)^{|\pi|}, \quad \text{for all } \pi \in \Av(231) \quad \text{and} \quad \mathbb{P}(\mathbf{S}^i = \emptyset) = \frac{1}{2}.
\]

This is a probability distribution since the generating function for the class \(\Av(231)\) is given by

\[
A(z) = \frac{1 - \sqrt{1 - 4z}}{2z} \quad \text{see e.g. Appendix A}.
\]

Moreover, let \(\{X_i\}_{i \geq 1}\) be i.i.d. random variables (also independent of \(\{\mathbf{S}^i\}_{i \geq 0}\)) with values in the set \(\{R, L\}\) such that

\[
\mathbb{P}(X_i = R) = \frac{1}{2} = \mathbb{P}(X_i = L).
\]

Observation 4.28. Recalling that \(T_0\) is the binary Galton–Watson tree with offspring distribution \(\eta(\delta)\) defined by Equation (34), we note that, for \(\delta = 0\), \(\sigma_{T_0}^{(d)} = (S^i | S^i \neq \emptyset)\).

We are now ready to construct our random order \(\mathfrak{s}_{231}\) on \(\mathbb{Z}\). Set \(\tilde{S}^0 := (\tilde{S}^0, m)\), where \(\tilde{S}^0\) is the random variable \(S^0\) conditioned to be non empty, and \(m\) is the (random) index of the maximum of \(S^0\). Then we set,

\[
\sigma^\infty_\bullet := ((\ldots ((\tilde{S}^0 \ast X_1 \mathbf{S}^1) \ast X_2 \mathbf{S}^2) \ldots \mathbf{S}^{n-1}) \ast X_n \mathbf{S}^n), \quad \text{for all } n \in \mathbb{Z}_{>0}.
\]

Note that, for every fixed \(h \in \mathbb{Z}_{>0}\), the sequence \(\{r_h(\sigma^d_n)\}_{n \in \mathbb{Z}_{>0}}\) is a.s. stationary (hence a.s. convergent) by Observation 4.27 since asymptotically almost surely at least \(h\) random variables \(X_i\) are equal to \(R\) and at least \(h\) are equal to \(L\). Let \(\tau_h\) be the limit of this sequence. The family \(\{\tau_h\}_{h \in \mathbb{Z}_{>0}}\) is a.s. consistent. Therefore, applying Proposition 2.13, the family \(\{\tau_h\}_{h \in \mathbb{Z}_{>0}}\) determines a.s. a unique random total order \((\mathbb{Z}, \mathfrak{s}_{231})\) such that, for all \(h \in \mathbb{Z}_{>0}\),

\[
r_h(\mathbb{Z}, \mathfrak{s}_{231}) = \tau_h = \lim_{n \to \infty} r_h(\sigma^d_n), \quad \text{a.s.} .
\]

Proposition 4.29. Let \((\mathbb{Z}, \mathfrak{s}_{231})\) be the random total order defined above and \(\sigma^\infty_{231}\) be the limiting object defined in Corollary 4.5. Then

\[
(\mathbb{Z}, \mathfrak{s}_{231}) \overset{(d)}{=} \sigma^\infty_{231}.
\]

Before proving the proposition we need to introduce some more notation in order to split a permutation in several different parts. In what follows, we suggest to compare the various definitions with the example in Fig. 14.

Given a pattern \(\pi \in \Av_{h+1}^{2h+1}(231)\) such that \(\text{indmax}(\pi) > h + 1\), and

\[
\text{LRMax}(\pi) \cap [h + 2, 2h + 1] = \{M_1, \ldots, M_k\},
\]

with \(M_1 < \cdots < M_k\), then we can decompose it as

\[
(\pi, h + 1) = \left(\ldots \left((\rho^0(\pi) \ast_R \rho_R^1(\pi)) \ast_R \ldots \right) \ast_R \rho_R^{k-1}(\pi)\right) \ast R \rho_R^h(\pi), \quad \text{if } h + 1 \in \text{LRMax}(\pi)
\]

(60)

\[
(\pi, h + 1) = \left(\ldots \left((\rho^0(\pi) \ast_L \rho_L^0(\pi)) \ast_R \rho_R^1(\pi) \ast R \ldots \right) \ast_R \rho_R^{k-1}(\pi)\right) \ast_R \rho_R^h(\pi), \quad \text{if } h + 1 \notin \text{LRMax}(\pi)
\]

(61)

where the different terms in the decomposition are precisely described in what follows.
We set,
\[
\rho^0_0(\pi) := \begin{cases} 
  \text{pat}_{1,RM-1}(\pi) & \text{if } h+1 \in \text{LRMax}(\pi), \\
  \text{pat}_{LM+1,1,RM-1}(\pi) & \text{if } h+1 \notin \text{LRMax}(\pi),
\end{cases}
\]
where \( LM = \max\{0 < j < h+1 : j \in \text{LRMax}(\pi)\} \) and \( RM = \min\{h+1 < j < 2h+2 : j \in \text{LRMax}(\pi)\} \). We also set
\[
\rho^0_0(\pi) := \begin{cases} 
  (\rho^0_0(\pi), h+1) & \text{if } h+1 \in \text{LRMax}(\pi), \\
  (\rho^0_0(\pi), h+1 - LM) & \text{if } h+1 \notin \text{LRMax}(\pi),
\end{cases}
\]
i.e., in both cases, we are rooting \( \rho^0_0(\pi) \) at the index corresponding to the root of \((\pi, h+1)\).

Remark 4.30. Note that, in general, if \( h+1 \not\in \text{LRMax}(\pi) \) then \( \text{indmax}(\rho^0_0(\pi)) \neq h+1 - LM \), i.e., the index of the maximum of \( \rho^0_0(\pi) \) does not correspond to the root of \( \rho^0_0(\pi) \).

Then, analyzing the left side of the pattern \( \pi \), if \( h+1 \notin \text{LRMax}(\pi) \), we set
\[
\rho^e_L(\pi) := \text{pat}_{1,LM-1}(\pi).
\]
Finally, analyzing the right side of the pattern \( \pi \) (regardless of \( h+1 \) being or not in \( \text{LRMax}(\pi) \)), if
\[
\text{LRMax}(\pi) \cap [h+2, 2h+1] = \{M_1, \ldots, M_k\}, \quad \text{with } M_1 < \cdots < M_k,
\]
we set
\[
\rho^e_R(\pi) := \text{pat}_{[M_{i+1}, M_i+1-1]}(\pi), \quad \text{for all } 1 \leq i \leq k-1,
\]
(with the convention that, if \( M_{i+1} - M_i = 1 \) then \( \rho^e_R(\pi) = \emptyset \)) and
\[
\rho^b_R(\pi) := \text{pat}_{[M_k+1, 2h+1]}(\pi),
\]
(with the convention that, if \( M_k = 2h+1 \) then \( \rho^b_R(\pi) = \emptyset \)).

**Figure 14.** Two diagram representations of the splitting of a pattern \( \pi \in \text{Av}^{2h+1}(231) \). On the left-hand side the case when \( h+1 \in \text{LRMax}(\pi) \), on the right-hand side the case when \( h+1 \notin \text{LRMax}(\pi) \). We highlight in red the value corresponding to the index \( h+1 \). Moreover, we highlight in green the left-to-right maxima of the permutations \( \pi \).
Proof of Proposition 4.29. Since by Observation 2.25 the set of clopen balls
\[ A = \left\{ B((A, \ell), 2^{-h}) : h \in \mathbb{Z}_{>0}, (A, \ell) \in \hat{S}_* \right\} \]
is a separating class for the space \((\hat{S}_*, d)\), it is enough to prove that for all \( h \in \mathbb{Z}_{>0} \), for all \( \pi \in \text{Av}^{2h+1}(231) \),
\[
\mathbb{P}(r_h(Z, \xi_{231}) = (\pi, h + 1)) = P_{231}(\pi).
\] (62)
We fix a pattern \( \pi \in \text{Av}^{2h+1}(231) \). We first suppose that:
- \( \text{indmax}(\pi) > h + 1 \);
- \( h + 1 \notin \text{LRMax}(\pi) \) and \( \text{LRMax}(\pi) \cap [h + 2, 2h + 1] = \{M_1, \ldots, M_k\} \), with \( M_1 < \cdots < M_k \);
- \( \rho^0_\pi(\pi) \) is rooted at its maximum.

Using the notation explained before, \( \pi \) rewrites as
\[
(\pi, h + 1) = \left( \left( \ldots \left( (\rho^0_\pi(\pi) * L \rho^1_\pi(\pi)) * R \rho^1_\pi(\pi) \right) * R \ldots \right) * R \rho^{k-1}_\pi(\pi) \right) * R \rho^b_\pi(\pi).
\]
Recalling that
\[
\sigma^n_\pi := \left( (\ldots ((\hat{S}_0 * X_1 S_1) * X_2 S_2) \ldots S_{n-1}) * X_n S_n \right), \quad \text{for all } n \in \mathbb{Z}_{>0},
\]
and since by assumption \( \rho^0_\pi(\pi) \) is rooted at its maximum, by some easy combinatorial arguments, the event \( \{r_h(Z, \xi_{231}) = (\pi, h + 1)\} \) is equal to
\[
\{ \hat{S}_0 = \rho^0_\pi(\pi), X_1 = L, \text{pat}_e(S^1) = \rho^1_\pi(\pi), S^j_\ell = \rho^j_\pi(\pi), 1 \leq j \leq k - 1, \text{pat}_b(S^j) = \rho^b_\pi(\pi) \},
\]
where \( \ell_j \) denotes the (random) index of the \( j \)-th \( X_i \) equal to \( R \) (note that \( \ell_j \) is almost surely well-defined). Using the independence of \( \{S^i\}_{i \geq 0} \) and \( \{X_i\}_{i \geq 1} \) we have,
\[
\mathbb{P}(r_h(Z, \xi_{231}) = (\pi, h + 1))
= \mathbb{P}(\hat{S}_0 = \rho^0_\pi(\pi)) \cdot \mathbb{P}(X_1 = L) \cdot \mathbb{P}(\text{pat}_e(S^1) = \rho^1_\pi(\pi))
\cdot \prod_{1 \leq j \leq k - 1} \mathbb{P}(S^j_\ell = \rho^j_\pi(\pi)) \cdot \mathbb{P}(\text{pat}_b(S^j) = \rho^b_\pi(\pi))
= \left( \frac{1}{4} \right)^{|ho^0_\pi(\pi)|} \cdot \frac{1}{2} \cdot \mathbb{P}(\text{pat}_e(S^1) = \rho^1_\pi(\pi)) \cdot \prod_{1 \leq j \leq k - 1} \left( \frac{1}{4} \right)^{|ho^j_\pi(\pi)|} \cdot \mathbb{P}(\text{pat}_b(S^j) = \rho^b_\pi(\pi)).
\]
Using Observation 4.23 and the fact that, for \( \delta = 0 \), \( \sigma^0_T \stackrel{(d)}{=} (S^i|S^i \neq \emptyset) \), we have that,
\[
\mathbb{P}(\text{pat}_b(S^i) = \rho) = \frac{2^{(|\text{LRMax}(\rho)|)}}{2^{|ho|}}, \quad \text{for all } \rho \in \text{Av}(231),
\]
\[
\mathbb{P}(\text{pat}_e(S^i) = \rho) = \frac{2^{(|\text{LRMax}(\rho)|)}}{2^{|ho|}}, \quad \text{for all } \rho \in \text{Av}(231).
\]
Therefore, we can conclude that
\[
\mathbb{P}(r_h(Z, \xi_{231}) = (\pi, h + 1)) = \left( \frac{1}{4} \right)^{|ho^0_\pi(\pi)|} \cdot \frac{1}{2} \cdot \frac{2^{(|\text{LRMax}(\rho^e_\pi(\pi))|)}}{2^{|ho^e_\pi(\pi)|}} \cdot \prod_{1 \leq j \leq k - 1} \left( \frac{1}{4} \right)^{|ho^j_\pi(\pi)|} \cdot \frac{2^{(|\text{LRMax}(\rho^b_\pi(\pi))|)}}{2^{|ho^b_\pi(\pi)|}}
= \frac{2^{(|\text{LRMax}(\rho^e_\pi(\pi))|)+(|\text{LRMax}(\rho^b_\pi(\pi))|)}}{2^{2(|\pi|-(k+1))2^k}} = \frac{2^{(|\text{LRMax}(\pi)|)+(|\text{LRMax}(\pi)|)}}{2^{2(|\pi|}}} = P_{231}(\pi).
\]
This concludes the proof of Equation (62) when \( \text{indmax}(\pi) > h + 1 \), \( h + 1 \notin \text{LRMax}(\pi) \) and \( \rho^0_\pi(\pi) \) is rooted at its maximum. For the other cases we can proceed as follows.
• If \( \text{indmax}(\pi) > h + 1 \), \( h + 1 \notin \text{LRMax}(\pi) \) but \( \rho^0(\pi) \) is not rooted at its maximum, we have to repeat the decomposition explained in Equations (60) and (61) for the permutation \( \rho^0(\pi) \). We iterate this procedure until \( \rho^0(\rho^0(\ldots(\rho^0(\pi))\ldots)) \) is rooted at its maximum (this procedure finishes since at each step \( |\rho^0(\rho^0(\pi))| < |\rho^0(\pi)| \)). In this way we obtain a new decomposition of \((\pi, h + 1)\). Finally, with similar computations as before we can verify Equation (62).

• If \( \text{indmax}(\pi) > h + 1 \), \( h + 1 \in \text{LRMax}(\pi) \) then the proof of Equation (62) is similar to the one in the first case substituting the event \( \{\tilde{S}^0 = \rho^0(\pi), X_1 = L, \text{pat}_e(\tilde{S}^1) = \rho^L(\pi)\} \) with the event \( \{\text{pat}_e(\tilde{S}^0) = \rho^0(\pi)\} \) and then arguing exactly in the same way as before.

• If \( \text{indmax}(\pi) < h + 1 \), then Equation (62) follows from the previous cases by a symmetry argument.

• Finally, if \( \text{indmax}(\pi) = h + 1 \),

\[
P(\rho^h(Z, \leq_{231}) = (\pi, h + 1)) = P(\rho^h(\tilde{S}^0) = (\pi, h + 1)) = P_{231}(\pi),
\]

where in the last equality we used Proposition 4.22 and the fact that, for \( \delta = 0 \), \( \sigma_{T^0}^{(d)} = \tilde{S}^0 \).

The analysis of all these cases concludes the proof. \( \square \)

4.5. **Some examples of limiting probabilities.** We explicitly compute in Fig. 15 the probabilities \( P_{231}(\pi) \) for 231-avoiding permutations \( \pi \) of size 3, 4 and 5, organizing the various permutations in different classes according to their number of maxima.

| \(|\pi|\) | \(\pi\) | \(|\text{LRMax}(\pi)|\) | \(|\text{RLMax}(\pi)|\) | \(P_{231}(\pi)\) |
|---|---|---|---|---|
| 3 | 132, 123, 312 | 3 | 4 | 1/2^2 |
| 312, 213 | 2 | 3 | 1/3^2 |
| 4 | 1234, 1243, 1432, 4321 | 4 | 5 | 1/2^3 |
| 1324, 1423, 2134, 2143, 4132, 4312 | 6 | 4 | 1/2^4 |
| 3124, 3214, 4123, 4213 | 4 | 3 | 1/2^5 |
| 5 | 12345, 12354, 12543, 15432, 54321 | 5 | 6 | 1/2^4 |
| 12435, 12534, 13245, 13254, 15243, 15423, 21345, 21354, 21543, 51432, 54132, 54312 | 12 | 5 | 1/2^5 |
| 14235, 14325, 15234, 15324, 21435, 21534, 31245, 31254, 32145, 32154, 51243, 52143, 53214, 54123, 54213, 54321, | 15 | 4 | 1/2^6 |
| 41235, 41325, 42135, 43125, 43215, 51234, 51423, 51324, 52134, 53124 | 10 | 3 | 1/2^7 |

**Figure 15.** Probabilities \( P_{231}(\pi) \) for 231-avoiding permutations of size 3, 4 and 5. In the dotted column we report how many permutations of fixed size have the same limiting probability, or equivalently the same total numbers of left-to-right and right-to-left maxima.
5. The local limit for uniform random 321-avoiding permutations

The goal of this section is to prove that uniform random 321-avoiding permutations converge in the quenched Benjamini–Schramm sense. We start by introducing, for all \( k > 0 \), the following probability distribution on \( \text{Av}^k(321) \),

\[
P_{321}(\pi) := \begin{cases} 
\frac{|\pi|+1}{2^{|\pi|}} & \text{if } \pi = 12 \ldots |\pi|, \\
\frac{1}{2^{|\pi|}} & \text{if } c\text{-occ}(21, \pi^{-1}) = 1, \\
0 & \text{otherwise}.
\end{cases}
\]  

Remark 5.1. We note that the second condition \( c\text{-occ}(21, \pi^{-1}) = 1 \) (i.e., \( \pi^{-1} \) has exactly one descent, or equivalently, \( \pi \) has exactly one inverse descent) is equivalent to saying that we can split the permutation \( \pi \) in exactly two increasing subsequences such that all the values of one subsequence are smaller than the values of the other one. Formally, if \( \pi \in \text{Av}^k(321) \) and \( c\text{-occ}(21, \pi^{-1}) = 1 \) there exists a unique bipartition of \([m]\) as \([m] = L(\pi) \sqcup U(\pi)\) such that for all \((i,j) \in L(\pi) \times U(\pi)\) then \( \pi_i < \pi_j \) and \( \text{pat}_{L(\pi)}(\pi) \) and \( \text{pat}_{U(\pi)}(\pi) \) are increasing.

Given the diagram of a permutation in \( \text{Av}^k(321) \) with exactly one inverse descent, we call separating line the horizontal line in the diagram between these two increasing subsequences. An example is given in Fig. 16 in the case of the permutation \( \pi = 14526738 \), where we highlight in orange the separating line between the two increasing subsequences.

![Figure 16. Diagram of the permutation 14526738 with the separating line highlighted in orange.](image)

Remark 5.2. That Equation (63) defines a probability distribution on \( \text{Av}^k(321) \) is a consequence of Theorem 5.4 below: in Equation (64), summing over all \( \pi \in \text{Av}^k(321) \) the left-hand side, we obtain \( \frac{n-k+1}{n} \) which trivially tends to 1. It is however easy to give a direct explanation. Indeed the number of permutations in \( \text{Av}^k(321) \) with exactly one inverse descent and separating line at height \( \ell \in \{1, \ldots, m-1\} \) is \( \binom{k}{\ell} - 1 \) (where the binomial coefficient \( \binom{k}{\ell} \) counts the possible choices of points that are below the separating line and the factor \(-1\) stands for the identity permutation which has no inverse descent). Therefore the number of permutations in \( \text{Av}^k(321) \) with exactly one inverse descent is

\[
\sum_{\ell=1}^{k-1} \left( \binom{k}{\ell} - 1 \right) = 2^k - (k + 1),
\]

which entails \( \sum_{\pi \in \text{Av}^k(321)} P_{321}(\pi) = 1 \).

We now recall, similarly to the previous section, the following trivial fact.

Observation 5.3. If \( \pi \notin \text{Av}(321) \) and \( \sigma \in \text{Av}(321) \) then obviously \( c\text{-occ}(\pi, \sigma) = 0 \). Therefore, in what follow, we simply analyze the consecutive occurrences densities \( c\text{-occ}(\pi, \sigma) \) for \( \pi \in \text{Av}(321) \).

Theorem 5.4. Let \( \sigma^n \) be a uniform random permutation in \( \text{Av}^n(321) \) for all \( n \in \mathbb{Z}_{>0} \). We have the following convergence in probability,

\[
\widetilde{c\text{-occ}}(\pi, \sigma^n) \xrightarrow{P} P_{321}(\pi), \quad \text{for all } \pi \in \text{Av}(321).
\]
Since the limiting frequencies \( (P_{321}(\pi))_{\pi \in \text{Av}(321)} \) are deterministic, using Corollary 2.40 we have the following.

**Corollary 5.5.** Let \( \sigma^n \) be a uniform random permutation in \( \text{Av}^n(321) \) for all \( n \in \mathbb{Z}_{\geq 0} \). There exists a random infinite rooted permutation \( \sigma_{321}^\infty \) such that

\[
\sigma^n \xrightarrow{q_{BS}} \mathcal{L}(\sigma_{321}^\infty).
\]

**Remark 5.6.** We note that, as in the case of Corollary 4.5, Corollary 5.5 proves the existence of a random infinite rooted permutation \( \sigma_{321}^\infty \) without constructing explicitly this limiting object. Nevertheless, also in this case, we are able to provide an explicit construction of \( \sigma_{321}^\infty \) (see Section 5.5).

The proof of Theorem 5.4 is structured as follows.

- In Section 5.1 we introduce a well-known bijection between rooted ordered trees and 321-avoiding permutations. Moreover, we explain why it is enough to study a specific family of Galton–Watson trees.
- In Section 5.2 we present two results, due to Stufler [52] and Janson [29], that characterize the local limit for Galton–Watson trees pointed at a uniform vertex.
- In Section 5.3 we introduce the notion of contour of a tree and we state some technical results about the characterization of the shape of some specific contours.
- In Section 5.4 we state two propositions (whose the proofs are postponed to Sections 5.4.1 and 5.4.2) that are the key results for the proof of Theorem 5.4. The proofs of the two propositions are based on the results presented in the previous Sections 5.2 and 5.3.
- Finally, in Section 5.5 we give an explicit construction of the limiting object \( \sigma_{321}^\infty \).
- In Section 5.6 we give some explicit examples of limiting probabilities.

### 5.1. From uniform 321-avoiding permutations to Galton–Watson trees.

It is well known that 321-avoiding permutations can be broken into two increasing subsequences, one weakly above the diagonal and one strictly below the diagonal (see e.g. [17, Section 4.2]). Therefore, defining for every permutation \( \sigma \in \text{Av}^n(321) \)

\[
E^+ = E^+(\sigma) := \{ i \in [n] | \sigma_i \geq i \},
\]

\[
E^- = E^-(\sigma) := \{ i \in [n] | \sigma_i < i \},
\]

we have that \( \text{pat}_{E^+}(\sigma) \) and \( \text{pat}_{E^-}(\sigma) \) are the increasing patterns. Note that by convention the indices of fixed points are elements of \( E^+ \).

We now describe a slightly different, but equivalent (see Remark 4.8), version of a well-known bijection between 321-avoiding permutations (of size \( n \)) and rooted ordered trees (with \( n+1 \) vertices) – see e.g. [23].

We recall that \( \mathbb{T}^n \) denotes the set of rooted ordered trees with \( n \) vertices (see Section 3 for notation about trees and random trees). Suppose \( T \in \mathbb{T}^{n+1} \) has \( k \) leaves and let \( \ell(T) = (\ell_1, \ldots, \ell_k) \) be the list of leaves of \( T \), listed in order of appearance in the pre-order traversal of \( T \), i.e., from left to right. Moreover let \( s_i \) and \( q_i \) be the labels given to the leaf \( \ell_i \) respectively by the pre-order labeling of \( T \) starting from zero and the post-order labeling of \( T \) starting from 1. We set \( S(T) := \{ s_i \}_{1 \leq i \leq k} \) and \( Q(T) := \{ q_i \}_{1 \leq i \leq k} \). An example is given in the left-hand side of Fig. 17.

We start by constructing the permutation \( \sigma_T \) associated to the tree \( T \). We define \( \sigma_T \) pointwise on \( Q(T) = \{ q_i \}_{1 \leq i \leq k} \) as

\[
\sigma_T(q_i) = s_i, \quad \text{for all} \quad 1 \leq i \leq k,
\]

and then we extend \( \sigma_T \) on \( [n] \setminus Q(T) \) by assigning values of \( [n] \setminus S(T) \) in increasing order. Clearly \( \sigma_T \) is a 321-avoiding permutation. An example is given in Fig. 17.

We highlight (without proof, for more details see [23]) that this construction implies \( E^+(\sigma_T) = Q(T) \) and \( \sigma(E^+(\sigma_T)) = S(T) \). It can also be proved that this construction is invertible: for a
permutation $\sigma \in \text{Av}^n(321)$ there is a unique rooted ordered tree $T_\sigma$ with $n + 1$ vertices such that $Q(T_\sigma) = E^+(\sigma)$ and $S(T_\sigma) = \sigma(E^+(\sigma))$.

![Diagram of a tree](image)

**Figure 17.** On the left of the picture we see a tree $T$ with 11 vertices and 7 leaves: we record in red on the left of each vertex the pre-order label and in blue on the right of each vertex the post-order label. On the right of the picture we see the associated 321-avoiding permutation of size 10: in the diagram, for every leaf in the tree, we draw a gray dot using the post-order label for the $x$-coordinate and the pre-order label for the $y$-coordinate. Then we complete the diagram with the unique increasing subsequence fitting in the empty rows and columns (shown with black dots on the picture).

**Observation 5.7.** Let $\eta \sim \text{Geom}(1/2)$ i.e., $\eta_k := \mathbb{P}(\eta = k) = 2^{-k-1}$, for all $k \geq 0$. We recall that a uniform random rooted ordered tree with $n$ vertices $T_n$ has the same distribution as a Galton–Watson tree $T_n^\eta$ with offspring distribution $\eta$ conditioned on having $n$ vertices (see [29, Example 9.1]), namely

$$T_n \overset{(d)}{=} \left( T_n^\eta \mid |T_n^\eta| = n \right).$$

Therefore, thanks to the bijection previously introduced, in order to prove Theorem 5.4 we can study a conditioned Galton–Watson tree $T_n^\eta := \left( T_n^\eta \mid |T_n^\eta| = n \right)$ instead of a uniform 321-avoiding permutation $\sigma^n$. We also point out that $E^+(\sigma^n) = Q(T_{n+1})$ a.s.

**Remark 5.8.** At first sight, it seems that the construction presented here is different from that of [23], we argue here that they are in fact the same. In [23] the authors consider the same numbers $\{s_i\}_{1 \leq i \leq k}$ but different numbers $\{q'_i\}_{1 \leq i \leq k}$ (denoted by $\{q_i\}_{1 \leq i \leq k}$ in [23]) which are defined as follow

$$q'_i = s_i - h(\ell_i) + 1, \quad \text{for all } 1 \leq i \leq k.$$

Note that $q'_1 = q_1 = 1$. Analyzing the difference between two consecutive values in $\{q'_i\}_{1 \leq i \leq k}$, we have

$$q'_{i+1} - q'_i = (s_{i+1} - h(\ell_{i+1}) + 1) - (s_i - h(\ell_i) + 1)$$

$$= (s_{i+1} - s_i) - (h(\ell_{i+1}) - h(\ell_i))$$

$$= (d(\ell_{i+1}, c(\ell_{i+1}, \ell_i))) - (d(\ell_{i+1}, c(\ell_{i+1}, \ell_i)) - d(\ell_i, c(\ell_{i+1}, \ell_i)))$$

$$= d(\ell_i, c(\ell_{i+1}, \ell_i)) = q_{i+1} - q_i.$$

where we recall that $c(a, b)$ is the first common ancestor of $a$ and $b$. We conclude that $q_i = q'_i$, for all $1 \leq i \leq k$, so our notation is consistent with that of [23].
5.2. A local limit result for Galton–Watson trees pointed at a uniform vertex. In this section we present a combination of two results due to Stufler [52] and Janson [29] that characterize the local limit for Galton–Watson trees $T^n_\eta$ pointed at a uniform vertex.

We begin by recalling the recent construction, due to Stufler, of the local limit tree $T^*$. We provide the construction in the special case where $\eta \sim \text{Geom}(1/2)$ (cf. Fig. 18 below).

For each $i \geq 1$, we let a vertex $u_i$ receive offspring according to an independent copy of $\hat{\eta}$, that is a random variable with size-biased distribution $\hat{\eta}_k = k \cdot \eta_k$. The vertex $u_{i-1}$ gets identified with a child of $u_i$ chosen uniformly at random. All other children of $u_i$ become the root of independent copies of the Galton–Watson tree $T^n_\eta$. Finally $u_0$ becomes the root of another independent copy of the Galton–Watson tree $T^n_\eta$. We always think of $T^*$ as a tree with distinguished vertex $u_0$.

![Figure 18. A scheme of the “local limit tree” $T^*$.

We point out that $T^*$ is not a tree in the classical Neveu’s formalism. We have to imagine that “the root” of the tree is at the end of the infinite backward spine composed by the vertices $u_i, i \geq 0$. For a formal construction of the set of these trees with infinite spine see [52, Section 4.1, p.6]. In what follow we will refer to this type of trees as trees with an infinite spine.

We now state the following easy reformulation of a result due to Janson [29]. First we recall that a rooted ordered tree $T$ together with a distinguished vertex $v_0$ is called a pointed rooted ordered tree and it is denoted by $T = (T, v_0)$. Given $T = (T, v_0)$ a pointed rooted ordered tree, we denote with $f_h^\bullet(T) = f_h^\bullet(T, v_0)$ the pointed fringe subtree rooted at the $\ell$-th ancestor of the distinguished vertex $v_0 \in T$ (if this is not well-defined because $v_0$ has height smaller than $\ell$, we set $f_h^\bullet(T, v_0) = \diamond$, for some fixed placeholder value $\diamond$). Similarly we denote with $f_h^\bullet(T^*)$ the fringe subtree of $T^*$ rooted at $u_\ell$ with distinguished vertex $u_0$. Despite $T^*$ is not a tree in Neveu’s formalism, $f_h^\bullet(T^*)$ is almost surely a pointed rooted ordered tree.

Finally, we denote with $T^*_h$ the set of pointed rooted ordered trees whose distinguished vertex has height $h \geq 0$.

**Proposition 5.9.** Let $T^n_\eta$ be as before. Let $T^* = (T, v_0) \in T^*_h$. Set also

$$N_{T^*}(T^n_\eta) := \left| \{ v \in T^n_\eta : f_h^\bullet(T^n_\eta, v) = T^* \} \right|.$$ 

Then it holds that, for all $h \geq 0$, all $T^* \in T^*_h$,

$$\frac{N_{T^*}(T^n_\eta)}{n} \xrightarrow{P} \mathbb{P}(f_h^\bullet(T^*) = T^*).$$

**Proof.** We note that

$$N_{T^*}(T^n_\eta) := \left| \{ v \in T^n_\eta : f_h^\bullet(T^n_\eta, v) = T^* \} \right| = \left| \{ u \in T^n_\eta : f(T^n_\eta, u) = T \} \right|.$$
Indeed if \( f_h^*(T_n^0, v) = T^* \) then \( f(T_n^0, a_h(v)) = T \). Conversely, given \( u \) such that \( f(T_n^0, u) = T \), there exists a unique choice for \( v \) such that \( f_h^*(T_n^0, v) = T^* \) : using the Neveu’s formalism, \( v \) must be \( v = uw \), where \( w \) is the pointed vertex of \( T^* \).

Therefore by [29] Theorem 7.12 (ii), we have,

\[
\frac{N_{T^*}(T_n^0)}{n} \xrightarrow{P} \mathbb{P}(T^0 = T).
\]

So, proving that \( \mathbb{P}(T^0 = T) = \mathbb{P}(f_h^*(T^*) = T^*) \) will be enough to conclude the proof. This follows easily, since

\[
\mathbb{P}(T^0 = T) = \prod_{u \in T} \eta_{d_u^+(T)},
\]

and denoting with \( P \) the path from the root to the parent of the distinguished vertex in \( T^* \),

\[
\mathbb{P}(f_h^*(T^*) = T^*) = \prod_{u \in P} \left( \frac{\eta_{d_u^+(T^*)}}{\eta_{d_u^-(T^*)}} \right) \prod_{u \in T^*, u \notin P} \eta_{d_u^+(T^*)} = \prod_{u \in T^*} \eta_{d_u^+(T^*)}.
\]

\[\square\]

**Corollary 5.10.** Let \( T_n^0 \) be as before and \( v_0 \) be a uniform vertex of \( T_n^0 \) chosen independently of \( T_n^0 \). Let also \( h \geq 0 \) and \( A \subseteq T_n^* \). Then

\[
\mathbb{P}^{v_0}(f_h^*(T_n^0, v_0) \in A) \xrightarrow{P} \mathbb{P}(f_h^*(T^*) \in A).
\]

**Proof.** Note that \( N_{T^*}(T_n^0) \) can be rewritten as the following conditional probability,

\[
\frac{N_{T^*}(T_n^0)}{n} \xrightarrow{(d)} \mathbb{P}(f_h^*(T_n^0, v_0) = T^* \mid T_n^0) = \mathbb{P}^{v_0}(f_h^*(T_n^0, v_0) = T^*).
\]

The statement follows from Proposition 5.9 and the fact that \( A \subseteq T_n^* \) is countable.

\[\square\]

**Remark 5.11.** Corollary 5.10 implies the local convergence of the rooted tree \((T_n^0, v_0)\) in the quenched sense, i.e., the convergence of the conditional random variable \((T_n^0, v_0) \mid T_n^0\).

### 5.3. Contour function for trees.

In this section we recall the notion of contour function for trees and we generalize this definition for the local limit tree \( T^* \). We start with the following.

**Definition 5.12.** A **Dyck path** of length \( 2n \) is a lattice path from \((0,0)\) to \((2n,0)\), consisting of \( n \) up-steps (along the vector \((1,1)\)) and \( n \) down-steps (along the vector \((1,-1)\)), such that the path never goes below the \( x \)-axis. We encode a Dyck path by a word \( w_1 \cdots w_{2n} \) containing \( n \) letters \( U \) and \( n \) letters \( D \). The condition “the path never goes below the \( x \)-axis” is equivalent to “every prefix \( w_1 \cdots w_k \) contains at least as many \( U \)’s as \( D \)’s”.

**Definition 5.13.** Given a rooted ordered tree \( T \) with \( n+1 \) vertices, we define the **contour** \( C(T) \) of \( T \) as the Dyck path of length \( 2n \) obtained by turning around \( T \) from left to right (as for the depth-first traversal), starting from the root, and recording, for each traversed edge, either an up-step if we are visiting the edge from the left to top or a down-step if we are visiting the edge on the right from bottom to top.

An example of contour of a tree is given in Fig. 19.

**Remark 5.14.** From now until the end of Section 5, when we refer to labels, we always assume that we are speaking about labels given by the post-order labeling starting from 1. Moreover, given a rooted ordered tree and a label \( j \), we use the abuse of notation \((T,j)\) for denoting the pointed tree with distinguished vertex labeled by \( j \). Similarly, we will simply say “the vertex \( j \)” instead of “the vertex labeled by \( j \)”.

We now make a connection between contours of trees and the labels of their leaves. We first introduce some notation for words over the alphabet \( \{U,D\} \). We denote by \( U^+ \) (resp. \( D^+ \)) the set of finite words of length at least one composed only by letters \( U \) (resp. \( D \)). Given two words \( w,u, \)
we write $wu$ for the concatenation of the two words. More generally, given a family $\{w_i\}_{1 \leq i \leq n}$ of words, we write $\prod_{i=1}^{n} w_i$ for their consecutive concatenation.

**Lemma 5.15.** Let $A = \{x_1, x_2, \ldots, x_m\} \subset \mathbb{Z}_{>0}$ (with $x_1 = 1$) be a set of labels. Then

$$\{ T \in \mathbb{T} : Q(T) = A \} = \left\{ T \in \mathbb{T} : C(T) \in \left( \prod_{j=1}^{m-1} U^+ D^{x_{j+1}-x_j} \right) U^+ D^+ \right\}.$$ 

Moreover, the shape of the contour of the tree between two consecutive leaves $\ell_j, \ell_{j+1}$ respectively labeled by $x_j, x_{j+1}$ is of the form $U^+ D^{x_{j+1}-x_j}$.

**Proof.** First note that for a tree $T \in \mathbb{T}$ (cf. Fig. 19):

- the $k$-th down-step in $C(T)$ corresponds to the right side of the edge between the vertex $v$ of label $k$ and its parent;
- this down-step is preceded by an up-step if and only $v$ is a leaf.

Therefore, given a tree $T$ such that $Q(T) = A$, the only down-steps in its contour preceded by an up-step are the first, $x_2$-th, $\ldots$ and $x_m$-th down-steps. Hence, the contour starts with an arbitrary long sequence of up-steps, then followed by a down-step (corresponding to the first element $x_1 = 1$ of $A$) and then by an additional sequence of $x_2 - x_1 - 1$ down-steps (otherwise there would be an element $x$ in $A$ such that $x_1 < x < x_2$). Considering the next element $x_2$ of $A$, we proceed in the same way as before, and find, in the contour, a new sequence of up-steps before the down-step corresponding to $x_2$. We iterate this procedure for all the elements in $A$, finding factors of the form $U^+ D^{x_{j+1}-x_j}$ for each $x_j \in A, j \neq m$. The final sequence of down-steps corresponding to $x_m$ has length determined by the size of the tree. Therefore $C(T) \in \left( \prod_{j=1}^{m-1} U^+ D^{x_{j+1}-x_j} \right) U^+ D^+$.

Actually, the same reasoning also proves that given a tree $T$ such that

$$C(T) \in \left( \prod_{j=1}^{m-1} U^+ D^{x_{j+1}-x_j} \right) U^+ D^+$$

then $Q(T) = A$, concluding the proof. \hfill \Box

We now extend the notion of contour to the local limit tree $\mathbf{T}^*$. We introduce some more notation for words over the alphabet $\{U, D\}$. We denote by $\{U, D\}^\infty_R$ the set of right-infinite words over the alphabet $\{U, D\}$, i.e., of infinite words of the form $w_1 w_2 \cdots w_n \cdots$, and similarly we denote by $\{U, D\}^\infty_L$ the set of left-infinite words over the alphabet $\{U, D\}$, i.e., of infinite words of the form $\cdots w_{-n} \cdots w_{-2} w_{-1}$.
Definition 5.16. A double-infinite pointed path is an infinite word in \( \{U, D\}^\infty \hat{w} \{U, D\}^\infty \), where \( \hat{w} \in \{\hat{U}, \hat{D}\} \).

Definition 5.17. The contour of a tree \( T \) with an infinite spine is the double-infinite pointed path \( C(T) = \cdots w_{-n} \cdots w_{-1} \hat{w}_0 w_1 \cdots w_n \cdots \), obtained turning around the tree \( T \) (cf. Fig. 20) in the following way.

- We start the visit on the right of the vertex \( u_0 \) in the direction of the vertex \( u_1 \) (in the picture we start at the red square in the direction of the blue dashed line) and we set \( \hat{w}_0 = \hat{D} \) (i.e., \( \hat{w}_0 \) is the down-step corresponding to the right-side of the edge between \( u_0 \) and \( u_1 \)).
- We proceed turning around \( T \), in the same direction (following the blue dashed line) setting, for all \( i \geq 1 \), \( w_i = U \) if the \( i \)-th visited edge is visited from top to bottom, otherwise setting \( w_i = D \).
- We conclude coming back to the starting point (the red square) and turning around \( T \) in the other direction (following the green dotted line), setting, for all \( i \geq 1 \), \( w_{-i} = U \) if the \( i \)-th visited edge is visited from bottom to top, otherwise setting \( w_{-i} = D \).

In particular, the contour of the local limit tree \( T^* \) is the random double-infinite pointed path \( C(T^*) = \cdots w_{-n} \cdots w_{-1} \hat{w}_0 w_1 \cdots w_n \cdots \), obtained turning around the tree \( T^* \) as explained before.

Observation 5.18. Note that the contour of the local limit tree \( T^* \) is consistent with the definition of contour on finite trees: the contour of a fringe subtree \( f^*_h(T^*) \) (forgetting the distinguished vertex) is a factor \( w_{-k} \cdots w_{-1} w_0 w_1 \cdots w_{k'} \) of \( C(T^*) \) (where \( w_0 \) has been replaced by \( w_0 = D \)). Note also that \( w_{-k} \cdots w_{-1} \) contains at least \( h \) up-steps and \( w_0 w_1 \cdots w_{k'} \) contains at least \( h \) down-steps (corresponding to the edges \( (u_i, u_{i+1}) \), for \( 0 \leq i \leq h - 1 \)).

Figure 20. Scheme for the construction of the contour of a tree with an infinite spine.

Lemma 5.19. Let \( C(T^*) = \cdots w_{-n} \cdots w_{-1} \hat{w}_0 w_1 \cdots w_n \cdots \) be the contour of \( T^* \). Then \( \{w_i\}_{i \in \mathbb{Z} \setminus \{0\}} \) is a sequence of i.i.d Bernoulli random variables on the set \( \{U, D\} \).

Proof. We fix \( i > 0 \). Note that if \( w_i = D \) and \( w_i \) is the down-step corresponding to the right side of an edge in \( T^* \) between the vertices \( v \) and \( u = a(v) \) then:
Proposition 5.21. For all $n \in \mathbb{Z}_{>0}$, let $\sigma^n$ be a uniform random permutation in $Av^n(321)$. Fix $k \in \mathbb{Z}_{>0}$ and let $i_n$ be uniform in $[n]$ and independent of $\sigma^n$. Then, for every subset $A \subseteq [-k, k]$, as $n \to \infty$,

$$\mathbb{P}^{i_n}(E_{i_n,k}^+(\sigma^n) = A) \to \left(\frac{1}{2}\right)^{2k+1}.$$

Informally, conditionally on $\sigma^n$, every index of $r_k(\sigma^n, i_n)$ is asymptotically equiprobably in $E^+$ or $E^-$. Moreover, these events are asymptotically independent.

Before stating our second proposition, we introduce some more notation (cf. Fig. 21 below). Given a permutation $\sigma \in Av(321)$ we define, for $i \in [\|\sigma\|]$, $k \in \mathbb{Z}_{>0}$,

$$m_{i,k}^+(\sigma) = \min \{E^+(\sigma) \cap [i-k, i+k] \},$$

$$M_{i,k}^-(-\sigma) = \max \{E^-(\sigma) \cap [i-k, i+k] \},$$

$$w_{i+1} = D,$$ if $v$ is the rightmost child of $u$;

$$w_{i+1} = U,$$ if $v$ is not the rightmost child of $u$.

We claim that, conditioning on $w_i = D$, the event “$v$ is the rightmost child of $u$” is independent of the shape of the contour of $T^*$ before the vertex $v$, namely is independent of $\{w_j\}_{j<i}$. Indeed, the number of children of each vertex of the tree $T^*$ is geometrically distributed (with parameter 1/2) or size-biased geometrically distributed (with parameter 1/2), and both distribution have the well-known memoryless property. This proves the claim. Moreover, thanks to this observation,

$$\mathbb{P}(v \text{ is the rightmost child of } u | w_i = D \text{ and knowing } \{w_j\}_{j<i}) = \frac{1}{2},$$

$$\mathbb{P}(v \text{ is not the rightmost child of } u | w_i = D \text{ and knowing } \{w_j\}_{j<i}) = \frac{1}{2}.$$
with the conventions that $\min \emptyset = +\infty$ and $\max \emptyset = -\infty$.

We also define, for a random 321-avoiding permutation $\sigma$ and a uniform index $i \in [|\sigma|]$, the following events, for all $k \in \mathbb{Z} > 0$,

$$S_k(\sigma, i) = \{ \sigma(m_{i,k}^-(\sigma)) > \sigma(M_{i,k}^-(\sigma)), m_{i,k}^+(\sigma) \neq +\infty, M_{i,k}^+(\sigma) \neq -\infty \}$$

$$\cup \{ m_{i,k}^+(\sigma) = +\infty \} \cup \{ M_{i,k}^-(\sigma) = -\infty \}.$$  

This is the event that the random rooted permutation $r_k(\sigma, i)$ splits into two (possibly empty) increasing subsequences with separated values, namely, the minimum of the upper subsequence is greater than the maximum of the lower subsequence. We will say, if this events holds, that “a separating line exists” (in Fig. 21 this separating line is dashed in orange).

**Proposition 5.22.** For all $n \in \mathbb{Z}_{>0}$, let $\sigma^n$ be a uniform random permutation in $\text{Av}^n(321)$. Fix $k \in \mathbb{Z}_{>0}$ and let $i_n$ be uniform in $[n]$ and independent of $\sigma^n$. Then, as $n \to \infty$,

$$\mathbb{P}^{i_n}(S_k(\sigma^n, i_n)) \xrightarrow{P} 1.$$  

Informally, conditionally on $\sigma^n$, asymptotically almost surely there exists a separating line in $r_k(\sigma^n, i_n)$.

With these two propositions in our hands, we can prove Theorem 5.4.

**Proof of Theorem 5.4.** In order to prove Equation (64) on page 48, we can check the equivalent condition (d) of Corollary 2.40, i.e., we have to prove that, for all $h \in \mathbb{Z}_{>0}$, for all $\pi \in \text{Av}^{2h+1}(321)$,

$$\mathbb{P}^{i_n}(r_h(\sigma^n, i_n) = (\pi, h+1)) \xrightarrow{P} P_{321}(\pi).$$

We fix $h \in \mathbb{Z}_{>0}$ and we distinguish three different cases for $\pi \in \text{Av}^{2h+1}(321)$:

\text{-occ}(21, \pi^{-1}) = 1 : We recall (see Remark 5.1) that, in this case, $\pi$ splits in a unique way into two increasing subsequences whose indices are denoted by $L(\pi)$ and $U(\pi)$. Note that, in general, these two sets are different from $E^-(\pi)$ and $E^+(\pi)$ (specifically when the permutation has some fixed points at the beginning, for example as in the case of Fig. 16 on page 48). We set $U^*(\pi) := \{ j \in [-h, h] : j + h \in U(\pi) \}$, i.e. $U^*(\pi)$ denotes the shifting of the set $U(\pi)$ inside the interval $[-h, h]$. Conditioning on $S_h(\sigma^n, i_n)$, namely assuming there is a separating line, $r_h(\sigma^n, i_n)$ is the union of
two increasing subsequences with $U^*(r_h(\sigma^n, i_n)) = E_{i_{n,h}}^+(\sigma^n)$. In particular, using Proposition 5.22 asymptotically we have that,
\begin{equation}
\mathbb{P}^i_n(r_h(\sigma^n, i_n) = (\pi, h+1)) = \mathbb{P}^i_n(S_h(\sigma^n, i_n), E_{i_{n,h}}^+(\sigma^n) = U^*(\pi)) + o(1).
\end{equation}
By Proposition 5.21 and 5.22 we have that
\begin{equation}
\mathbb{P}^i_n(S_h(\sigma^n, i_n), E_{i_{n,h}}^+(\sigma^n) = U^*(\pi)) \xrightarrow{P} \frac{1}{2^{2h+1}} = P_{321}(\pi),
\end{equation}
and so Equation (66) holds.

$\pi = 12 \ldots |\pi|$: In this case $\pi$ does not split in a unique way into two increasing subsequences. We have exactly $2h+2$ different choices for this splitting, namely setting either $U(\pi) = \emptyset$ or $U(\pi) = [k, h]$, for some $k \in [-h, h]$. Therefore, using Proposition 5.22 asymptotically we obtain that
\begin{equation}
\mathbb{P}^i_n(r_h(\sigma^n, i_n) = (\pi, h+1)) = \mathbb{P}^i_n(S_h(\sigma^n, i_n), E_{i_{n,h}}^+(\sigma^n) \in \{\emptyset\} \cup \bigcup_{k \in [-h, h]} \{[k, h]\}) + o(1).
\end{equation}
Again by Proposition 5.21 and 5.22 we have that
\begin{equation}
\mathbb{P}^i_n(S_h(\sigma^n, i_n), E_{i_{n,k}}^+(\sigma^n) \in \{\emptyset\} \cup \bigcup_{k \in [-h, h]} \{[k, h]\}) \xrightarrow{P} \frac{2h + 2}{2^{2h+1}} = P_{321}(\pi),
\end{equation}
and so Equation (66) holds.

c-occ(21, π⁻¹) > 1: In this case $\pi$ does not present a separating line. Since by Proposition 5.22 we know that asymptotically almost surely there exists a separating line in $r_h(\sigma^n, i_n)$ we can immediately conclude that,
\begin{equation}
\mathbb{P}^i_n(r_h(\sigma^n, i_n) = (\pi, h+1)) \leq \mathbb{P}^i_n(S_h(\sigma^n, i_n))^c \xrightarrow{P} 0 = P_{321}(\pi).
\end{equation}

5.4.1. The proof of Proposition 5.21. We recall that $T^\eta_n$ denotes a Galton–Watson tree with offspring distribution $\eta \sim \text{Geom}(1/2)$ conditioned on having $n$ vertices and $i_n$ is a uniform index in $[n]$ independent of the other random objects. We label the tree $T^\eta_n$ with the post-order labeling starting from 1 and, as said in Remark 5.14, we denote with $(T^\eta_n, i_n)$ the random pointed rooted ordered tree obtained from $T^\eta_n$ by distinguishing the vertex $i_n$. We also recall that $A \subseteq [-k, k]$.

Observation 5.23. Thanks to Observation 5.7 and Definition 5.20
\begin{equation}
\mathbb{P}^i_n(E_{i_{n,k}}^+(\sigma^n) = A) = \mathbb{P}^i_n(Q_{i_{n,k}}(T^\eta_{n+1}) = A) = \mathbb{P}^i_{n+1}(Q_{i_{n+1,k}}(T^\eta_{n+1}) = A) + O(1/n),
\end{equation}
where in the last equality we shifted the index $i_n$ to $i_{n+1}$ for later convenience. Therefore instead of studying the event $\{E_{i_{n,k}}^+(\sigma^n) = A\}$ that appears in the statement of Proposition 5.21 we can study the event $\{Q_{i_{n,k}}(T^\eta_n) = A\}$.

We need to introduce some technical results in order to study the event $\{Q_{i_{n,k}}(T^\eta_n) = A\}$ using Corollary 5.10. Informally, this event is not “local” in full generality, but it is “local” asymptotically almost surely.

Definition 5.24. Given a tree $T$ we say that a vertex $v$ is before (resp. after) a vertex $u$ if the label of $v$ is smaller (resp. greater) than the label of $u$.

For all $H > 0$, we consider the following events,
\begin{equation}
E_H(T^\eta_n, i_n) = \{f_H^i(T^\eta_n, i_n) \text{ has at least } k \text{ vertices before and } k \text{ vertices after the vertex } i_n\}
\end{equation}
and
\begin{equation}
E_H(T^\eta) = \{f_H^i(T^\eta) \text{ has at least } k \text{ vertices before and } k \text{ vertices after the vertex } u_0\}.
\end{equation}
**Lemma 5.25.** We have the following:

(a) For all $H > 0$, $\mathbb{P}^{i_n} \left( E_H(T^n_{i_n}) \right) \xrightarrow{P} \mathbb{P}(E_H(T^*))$, as $n \to \infty$;

(b) $\mathbb{P}(E_H(T^*)) \xrightarrow{H \to \infty} 1$.

In what follow, given a word $w = w_1 \cdots w_n$ we set, for $1 < k < n$, $w_{[1,k]} := w_1 \cdots w_k$ and $w_{[k,n]} := w_k \cdots w_n$.

**Proof.** (a) This is a trivial application of Corollary 5.10.

(b) Let $\ell$ be the index of the down-step in $C(f^*_H(T^*))$ corresponding to the right-side of the edge between the vertices $u_0$ and $u_1$ of $T^*$. The event $E_H(T^*)$ contains the intersection of the following two events:

\[
L_H(T^*) := \{ C(f^*_H(T^*))_{[1,\ell-1]} \text{ contains at least } k \text{ down-steps} \},
\]

\[
R_H(T^*) := \{ C(f^*_H(T^*))_{[\ell,|C(f^*_H(T^*)|]} \text{ contains at least } k \text{ down-steps} \}.
\]

Therefore

\[
\mathbb{P}(E_H(T^*)) \geq \mathbb{P}(L_H(T^*) \cap R_H(T^*)).
\]

Note that, using Observation 5.18 for $C(f^*_H(T^*))_{[\ell,|C(f^*_H(T^*)|]}$, we obtain that $\mathbb{P}(R_H(T^*)) = 1$, for all $H \geq k$. Moreover, recalling that $C(T^*) = \cdots \cdot w_{-n} \cdots \cdot w_{-1} \cdot w_0 \cdot w_1 \cdots$, using again Observation 5.18 and denoting with $m$ the smallest index such that $w_{-m} \cdots w_{-1}$ contains $H + 1$ up-steps, we have

\[
\mathbb{P}(L_H(T^*)) \geq \mathbb{P}(w_{-m} \cdots w_{-1} \text{ contains } k \text{ down-steps}) \xrightarrow{H \to \infty} 1,
\]

where the limit is justified by Lemma 5.19 and the fact that an infinite sequence of i.i.d Bernoulli random variables contains a.s. at least $k$ down-steps.

We note that, for all $H, n > 0$, conditioning on the event $E_H(T^n_{i_n})$, we can rewrite the event $\left\{ Q_{i_n,k}(T^n_{i_n}) = A \right\}$ as

\[
\left\{ f^*_H(T^n_{i_n})_{i_n} \in T_k(A) \right\},
\]

where $T_k(A)$ is the set of all pointed rooted ordered trees $T^* = (T,j)$ with at least $k$ vertices before and $k$ vertices after the vertex $j$ such that $Q_{j,k}(T) = A$. Therefore, using Corollary 5.10 for all $H > 0$,

\[
\mathbb{P}^{i_n} \left( \left\{ Q_{i_n,k}(T^n_{i_n}) = A \right\} \cap E_H(T^n_{i_n}) \right) \xrightarrow{P} \mathbb{P}(\left\{ f^*_H(T^*) \in T_k(A) \right\} \cap E_H(T^*)).
\]

**Lemma 5.26.** With the same notation as above,

\[
\mathbb{P}(\left\{ f^*_H(T^*) \in T_k(A) \right\} \cap E_H(T^*)) \xrightarrow{H \to \infty} \left( 1 - \frac{1}{2} \right)^{2k+1}.
\]

**Proof.** Suppose that $A = \{ x_m, \ldots, x_1, y_1, \ldots, y_\ell \}$, where $m, \ell \leq k$, $x_{i+1} < x_i < 0$ for all $1 \leq i \leq m-1$ and $0 \leq y_i < y_{i+1}$ for all $1 \leq i \leq \ell - 1$. For all $H > 0$, using the same ideas as in Lemma 5.15 on the right and left part of the contour of $T^*$, the event $\{ f^*_H(T^*) \in T_k(A) \}$ rewrites as

\[
\left\{ C(T^*) \in \{ U, D \}_{L}^{\infty} D_{[1-x_m]|+1} \prod_{p=m-1}^{1} w_p U^+ D_{|x_1|-1} \hat{D} D_{y_i} U^+ \prod_{q=1}^{\ell-1} w_q D_{y_q-1} \{ U, D \}_{R}^{\infty} \right\},
\]

where $w_p = U^+ D_{[p+1-x_p]}$, for all $1 \leq p \leq m - 1$ and $w_q = D_{y_q+1} U^+$, for all $1 \leq q \leq \ell - 1$ (see Fig. 22 for a representation of this event).
Therefore, almost surely,

\[
\begin{align*}
\text{Proof of Proposition 5.21.} & \quad \text{By Equation (67), for all } H > 0, \text{ we have that} \\
\Pr(C(T^*) \in \{U, D\}_L \sum_{k=x_{m+1}}^{|x_1-1|} w_{k+y^1} - D^{|x_1-1|} U + \sum_{q=1}^{k-y^1} w_q D^{k-y^1+1} \{U, D\}_R) \\
& = \left(\frac{1}{2}\right)^{k+x_{m+1}} \sum_{p=m-1}^{k+x_{m+1}} \left(\frac{1}{2}\right)^{x_{p-x_{p+1}}} \left(\frac{1}{2}\right)^{-x_{1-1}+y_1} \sum_{q=1}^{k-y^1} \left(\frac{1}{2}\right)^{y_{q+1}-y_q} \left(\frac{1}{2}\right)^{k-y^1+1} = \left(\frac{1}{2}\right)^{2k+1}.
\end{align*}
\]

Since by Lemma 5.25 (b), \(\Pr(E_H(T^*)) \to 1\), the statement is proved.

With these results in our hands we can finish the proof of Proposition 5.21.

**Proof of Proposition 5.21.** We start studying the conditional probability \(\Pr(Q_{i_n,k}(T_n^0) = A)\), which, we recall, is a random variable. For all \(H > 0\), almost surely,

\[
\Pr(Q_{i_n,k}(T_n^0) = A) \geq \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)) + \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n) + \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)^c.
\]

Therefore, almost surely,

\[
\begin{align*}
\Pr(Q_{i_n,k}(T_n^0) = A) & \geq \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)), \\
\Pr(Q_{i_n,k}(T_n^0) = A) & \leq \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n) + \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)^c.
\end{align*}
\]

By Equation (67), for all \(H > 0\), the following limit for the lower bound in Equation (68) holds

\[
\Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)) \to \Pr(\{f_H(T^* \in T_k(A) \cap E_H(T^*)).
\]

Combining Equation (67) and Lemma 5.25 (a), for all \(H > 0\), the following limit for the upper bound in Equation (68) holds

\[
\begin{align*}
\Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n) + \Pr(Q_{i_n,k}(T_n^0) = A) \cap E_H(T_n^0, i_n)^c \\
\to \Pr(\{f_H(T^* \in T_k(A) \cap E_H(T^*) + \Pr(E_H(T^*)^c).
\end{align*}
\]

Since the limiting objects are deterministic, the limits in Equations (69) and (70) hold jointly for all \(H > 0\). Moreover, applying Skorokhod’s representation theorem in the product space over all \(H > 0\), we can also assume that there exists a coupling such that these limits hold also almost surely jointly for all \(H > 0\). Therefore, almost surely, for all \(H > 0\),

\[
\begin{align*}
\liminf_{n \to \infty} \Pr(Q_{i_n,k}(T_n^0) = A) & \geq \Pr(\{f_H(T^* \in T_k(A) \cap E_H(T^*)
\limsup_{n \to \infty} \Pr(Q_{i_n,k}(T_n^0) = A) & \leq \Pr(\{f_H(T^* \in T_k(A) \cap E_H(T^*) + \Pr(E_H(T^*)^c).
\end{align*}
\]

**Figure 22.** A scheme of the contour of \(C(T^*)\). We used dashed lines for sequences of down-steps with a fixed size (explicitly reported) and dotted lines for sequences of up-steps with arbitrary size at least 1.
Finally, since by Lemma 5.26
\[ P\left( \big\{ f^*_H(T^*) \in \mathcal{T}_k(A) \big\} \cap E_H(T^*) \right) \xrightarrow{H \to \infty} \left( \frac{1}{2} \right)^{2k+1} \]
and by Lemma 5.25 (a),
\[ P(E_H(T^*)) \xrightarrow{H \to \infty} 0, \]
we can conclude, from Equation (71), that the limit in probability of \( \mathbb{P}^n(Q_{i_n,k}(T^n_j) = A) \) exists and is equal to \( \left( \frac{1}{2} \right)^{2k+1} \). Finally, thanks to Observation 5.23,
\[ \lim_{n \to \infty} \mathbb{P}^n(E_{i_n,k}^+(\sigma^n) = A) = \lim_{n \to \infty} \mathbb{P}^n(Q_{i_n,k}(T^n_j) = A) = \left( \frac{1}{2} \right)^{2k+1}. \]

5.4.2. The proof of Proposition 5.22. It remains to prove the following:
\[ \mathbb{P}^n(S_k(\sigma^n, i_n)) \xrightarrow{P} 1. \]

Proof of Proposition 5.22. In order to simplify notation we fix \( m^+ = m_{i_n,k}^+(\sigma^n) \) and \( M^- = M_{i_n,k}^-(\sigma^n) \).

We split the probability \( \mathbb{P}(S_k(\sigma^n, i_n)) \) into two different terms as follow,
\[ \mathbb{P}(S_k(\sigma^n, i_n)) = \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
\[ + \mathbb{P}\left( \{ m^+ \neq +\infty, M^- \neq -\infty \}^c \right). \]

We analyze the first term,
\[ \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
\[ = \mathbb{P}(\sigma^n(m^+) - m^+ - (\sigma^n(M^-) - M^-) > M^- - m^+, m^+ \neq +\infty, M^- \neq -\infty). \]

Setting for all \( i \in [n], \Delta_i(\sigma^n) = \sigma^n(i) - i, \) we have
\[ \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
\[ = \mathbb{P}(\Delta_{m^+}(\sigma^n) - \Delta_{M^-}(\sigma^n) > M^- - m^+, m^+ \neq +\infty, M^- \neq -\infty). \]

We now analyze the behavior of \( \Delta_{m^+}(\sigma^n) \) and \( \Delta_{M^-}(\sigma^n) \). By an easy application (stated in [31, p.10]) of a result due to Hoffman, Rizzolo and Slivken (see [24]) we can assume that there exist a coupling for \( \{ (\sigma^n)_{n \in \mathbb{Z}_{>0}}, e \} \), where \( e \) is a Brownian excursion, such that, almost surely,
\[ \Delta_j(\sigma^n) = \sqrt{2n} \cdot e(j/n) + \varepsilon^n_j(j, e, \sigma^n), \quad \text{for all } j \in E^+(\sigma^n), \]
\[ \Delta_j(\sigma^n) = -\sqrt{2n} \cdot e(j/n) + \varepsilon^n_j,j, e, \sigma^n), \quad \text{for all } j \in E^-(\sigma^n), \]

where \( \varepsilon^n_j(j, e, \sigma^n) \) are error terms in probability of order \( o(n^{1/2}) \) uniformly in \( j \), i.e., \( \sup_j \{ \varepsilon^n_j(j, e, \sigma^n) \} \xrightarrow{n \to \infty} 0 \) (w.r.t. the probability given by the coupling of \( \{ (\sigma^n)_{n \in \mathbb{Z}_{>0}}, e \} \)).

Therefore, setting \( j = m^+ \) or \( j = M^- \), we obtain, almost surely,
\[ \Delta_{m^+}(\sigma^n) = \sqrt{2n} \cdot e(m^n/n) + \varepsilon^n_{m^+}(m^+, e, \sigma^n), \]
\[ \Delta_{M^-}(\sigma^n) = -\sqrt{2n} \cdot e(M^-/n) + \varepsilon^n_{M^-}(M^-, e, \sigma^n), \]

where \( \varepsilon^n_{m^+}(m^+, e, \sigma^n) \xrightarrow{n \to \infty} 0 \) (w.r.t. the probability given by the coupling of \( \{ (\sigma^n)_{n \in \mathbb{Z}_{>0}}, e \} \)) and almost surely, \( \varepsilon^n_{m^+}(m^+, e, \sigma^n) \leq \sup_j \{ \varepsilon^n(j, e, \sigma^n) \}. \) The same holds also for \( \varepsilon^n_{M^-}(M^-, e, \sigma^n). \)

In what follows all the probabilities and the error terms have to be interpreted in the same way, i.e. w.r.t. the probability given by the coupling of \( \{ (\sigma^n)_{n \in \mathbb{Z}_{>0}}, e, \{ i_n \}_{n \in \mathbb{Z}_{>0}} \} \). Moreover, we will simply denote the errors terms of order less than \( n^a \) with \( o_P(n^a) \).
We can rewrite the last term in Equation (72) as
\[ \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
(73)
\[ = \mathbb{P}\left(e(m^+/n) + e(M^-/n) > \frac{M^- - m^+}{\sqrt{2n}} + o_p(1), m^+ \neq +\infty, M^- \neq -\infty\right). \]
Noting that, if \( m^+ \neq +\infty \) and \( M^- \neq -\infty \), then \( -k \leq m^+ - i_n, M^- - i_n \leq k \), a.s., we have that
\[ \frac{1}{n}|m^+ - i_n| = o_p(1) \quad \text{and} \quad \frac{1}{n}|M^- - i_n| = o_p(1). \]

Since \( e \) is a continuous process, applying Lemma B.1 in the appendix, we can rewrite Equation (73) as
\[ \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
(74)
\[ = \mathbb{P}\left(e(i_n/n) > \frac{M^- - m^+}{2\sqrt{2n}} + o_p(1), m^+ \neq +\infty, M^- \neq -\infty\right). \]
Noting that, if \( m^+ \neq +\infty, M^- \neq -\infty \), than \( M^- - m^+ \leq 2k \), a.s., we can obtain the following bound,
\[ \mathbb{P}(\sigma^n(m^+) > \sigma^n(M^-), m^+ \neq +\infty, M^- \neq -\infty) \]
(75)
\[ \geq \mathbb{P}\left(e(i_n/n) + o_p(1) > \frac{k}{\sqrt{2n}}, m^+ \neq +\infty, M^- \neq -\infty\right). \]

From now until the end of the proof we denote with \( \{\varepsilon_n\}_{n \geq 1} \) the family of random error terms that appears in the previous equation. Since \( i_n/n \xrightarrow{(d)} U \), where \( U \) is a uniform random variable in \((0, 1)\), by Skorokhod’s representation theorem, we can assume that the convergence holds almost surely. Since \( e \) is a.s. continuous, we can deduce that
\[ e(i_n/n) + \varepsilon_n \xrightarrow{(d)} e(U). \]
In particular, since \( e \) and \( U \) are independent (\( i_n \) is independent of \( \sigma^n \), hence of \( e \)), \( e(U) \) admits a density (called Rayleigh distribution, see [12] p.5) and so the distribution function \( F_{e(U)}(t) \) is continuous. From Equation (76), we can conclude that the convergence
\[ F_{e(i_n/n) + \varepsilon_n}(t) \xrightarrow{} F_{e(U)}(t), \]
is uniform in \( \mathbb{R} \) as stated in [13] exercise 14.8, p.198. Therefore,
\[ \mathbb{P}\left(e(i_n/n) + \varepsilon_n > \frac{k}{\sqrt{2n}}\right) = 1 - F_{e(i_n/n) + \varepsilon_n}\left(\frac{k}{\sqrt{2n}}\right) \rightarrow 1 - F_{e(U)}(0) = 1. \]
Finally, summing up, we obtain that
\[ \lim_{n \to \infty} \mathbb{P}(S_k(\sigma^n, i_n)) \geq \lim_{n \to \infty} \mathbb{P}(m^+ \neq +\infty, M^- \neq -\infty) + \mathbb{P}\left(\{m^+ \neq +\infty, M^- \neq -\infty\}^c\right) = 1. \]
Noting that \( \mathbb{P}(S_k(\sigma^n, i_n)) = \mathbb{E}\sigma^n\left[p_{i_n}(S_k(\sigma^n, i_n))\right] \), and using Markov’s inequality for the probability \( 1 - \mathbb{P}_{i_n}(S_k(\sigma^n, i_n)) \), we can conclude that,
\[ \mathbb{P}_{i_n}(S_k(\sigma^n, i_n)) \xrightarrow{p} 1. \]
\[ \blacksquare \]

5.5. The construction of the limiting object. We now exhibit an explicit construction of the limiting object \( \sigma_{321}^\infty \) as a random order \( \sigma_{321}^\infty \) on \( \mathbb{Z} \).

We consider the set of integer numbers \( \mathbb{Z} \), and we label, uniformly and independently, each integer with a “+” or a “-”: namely, for all \( x \in \mathbb{Z} \),
\[ \mathbb{P}(x \text{ has label “+”}) = \frac{1}{2} = \mathbb{P}(x \text{ has label “-”}). \]
We set \( A^+ := \{ x \in \mathbb{Z} : x \text{ has label } + \} \) and \( A^- := \{ x \in \mathbb{Z} : x \text{ has label } - \} \). Then we define a random total order \( \prec_{321} \) on \( \mathbb{Z} \) saying that, for all \( x, y \in \mathbb{Z} \), \( x \prec_{321} y \) if either \( x < y \) and \( x, y \in A^- \), or \( x < y \) and \( x, y \in A^+ \), or \( x \in A^- \) and \( y \in A^+ \). An example is given in Fig. 23.

![Figure 23. A sample of the random total order \((\mathbb{Z}, \prec_{321})\). On the top, the standard total order on \( \mathbb{Z} \) with the integers labeled by random “-” and “+” signs. We paint in orange the integers labeled by “-” and in blue the integers labeled by “+.” Then, in the bottom part of the picture, we move the orange numbers at the beginning of the new random total order and the blue numbers at the end. Reading the bottom line from left to right gives the random total order \( \prec_{321} \) on \( \mathbb{Z} \).](image)

**Proposition 5.27.** Let \((\mathbb{Z}, \prec_{321})\) be the random total order defined above and \( \sigma_{321}^\infty \) be the limiting object defined in Corollary 5.5. Then

\[
(\mathbb{Z}, \prec_{321}) \overset{(d)}{=} \sigma_{321}^\infty.
\]

**Proof.** Since by Observation 2.25 the set of clopen balls

\[
A = \left\{ B((A, \preceq), 2^{-h}) : h \in \mathbb{Z}_{>0}, (A, \preceq) \in \mathcal{S}_\ast \right\}
\]

is a separating class for the space \((\mathcal{S}_\ast, d)\), it is enough to prove that for all \( h \in \mathbb{Z}_{>0} \), for all \( \pi \in \text{Av}^{2h+1}_{\infty}(321) \),

\[
\mathbb{P}(r_h(\mathbb{Z}, \prec_{321}) = (\pi, h + 1)) = P_{321}(\pi).
\]

We fix \( h \in \mathbb{Z}_{>0} \) and, as in the proof of Theorem 5.4, we distinguish three different cases for \( \pi \in \text{Av}^{2h+1}_{\infty}(321) \):

1. **c-occ\((21, \pi^{-1}) = 1\):** We recall (see Remark 5.1) that, in this case, \( \pi \) splits in a unique way into two increasing subsequences whose indices are denoted by \( L(\pi) \) and \( U(\pi) \). Recalling that \( U^*(\pi) := \{ j \in [-h, h] : j + h \in U(\pi) \} \), we have

\[
\mathbb{P}(r_h(\mathbb{Z}, \prec_{321}) = (\pi, h + 1)) = \mathbb{P}(A^+ \cap [-h, h] = U^*(\pi)) = \frac{1}{2^{2h+1}} = P_{321}(\pi).
\]

2. **\( \pi = 12 \ldots |\pi|\):** In this case \( \pi \) does not split in a unique way into two increasing subsequences. We have exactly \( 2h + 2 \) different choices for this splitting, namely setting either \( U^*(\pi) = \emptyset \) or \( U^*(\pi) = [k, h] \), for some \( k \in [-h, h] \). Therefore, we obtain that

\[
\mathbb{P}(r_h(\mathbb{Z}, \prec_{321}) = (\pi, h + 1)) = \mathbb{P}(A^+ \cap [-h, h] \in \{ \emptyset \} \cup \bigcup_{k \in [-h, h]} \{ [k, h] \}) = \frac{2h + 2}{2^{2h+1}} = P_{321}(\pi).
\]

3. **c-occ\((21, \pi^{-1}) > 1\):** Since, by construction, \( r_h(\mathbb{Z}, \prec_{321}) \) contains a.s. at most one inverse descent, then

\[
\mathbb{P}(r_h(\mathbb{Z}, \prec_{321}) = (\pi, h + 1)) = 0 = P_{321}(\pi).
\]

The analysis of this three cases concludes the proof. \( \square \)

**5.6. Some examples of limiting probabilities.** We explicitly compute in the table in Fig. 24 the probabilities \( P_{321}(\pi) \) for 321-avoiding permutations \( \pi \) of size 3,4 and 5, distinguishing the various permutations in different classes according to their number of inverse descents. We suggest to compare this result with the one in Section 4.5.
|
|\[|\pi|\] | \[\pi\] | \[c - \text{occ}(21, \pi^{-1})\] | \[P_{321}(\pi)\] |
|---|---|---|---|
|3 | 123 | 1 | 0 | 1/2 |
| | 132, 213, 231, 312 | 4 | 1 | 1/2 |
| | \emptyset | 0 | \geq 2 | 0 |
|4 | 1234 | 1 | 0 | 5/2 |
| | 1243, 1324, 1342, 1423, 2134, 2314, 2341, 3124, 3142, 3412, 4123 | 11 | 1 | 1/2 |
| | 2143, 2413 | 2 | \geq 2 | 0 |
|5 | 12345 | 1 | 0 | 3/2 |
| | 12354, 12435, 12453, 12534, 13245, 13425, 13452, 14235, 14253, 14523, 15234, 21345, 23145, 23415, 23451, 31452, 34125, 34512, 41235, 41523, 45123, 51234 | 26 | 1 | 1/2 |
| | 13254, 13524, 21354, 21435, 21453, 24135, 24153, 24513, 31254, 32154, 31524, 35124 | 15 | \geq 2 | 0 |

**Figure 24.** Probabilities \(P_{321}(\pi)\) for 321-avoiding permutations \(\pi\) of size 3, 4 and 5. In the dotted column we report how many permutations of fixed size have the same density.

**Appendix A. A generating function proof that \(P_{231}(\pi)\) is a probability distribution**

We give here a generating function proof of the fact that

\[
P_{231}(\pi) := \frac{2^{|LRMax(\pi)| + |RLMax(\pi)|}}{2^{|\pi|}}, \quad \text{for all} \quad \pi \in \text{Av}(231),
\]

defines a probability distribution on \(\text{Av}^k(231)\), for all \(k \geq 1\).

Using our bijection between 231-avoiding permutations and binary trees, we can consider the generating function

\[
C(z, x, y) = \sum_{T \in \mathcal{T}_b} z^{|T|} x^{|LRMax(\sigma_T)|} y^{|RLMax(\sigma_T)|},
\]

where \(\mathcal{T}_b := \mathcal{T} \cup \emptyset\) is the set of (possibly empty) binary trees. Using the classical decomposition of the binary trees in root, left subtree and right subtree, together with Observation 4.8, we obtain the following equation for our generating function,

\[
C(z, x, y) = 1 + zxy \cdot C(z, x, 1) \cdot C(z, 1, y),
\]

where the term 1 corresponds to the empty tree.

Now setting \(x = y = 1\) in Equation (78) and \(A(z) := C(z, 1, 1)\), we obtain the classical relation for the generating function of Catalan numbers:

\[
A(z) = 1 + z \cdot A(z)^2,
\]

that can be algebraically solved to yield

\[
A(z) = \frac{1 - \sqrt{1 - 4z}}{2z}.
\]

Substituting \(y = 1\) in Equation (78) and setting \(B(z, x) := C(z, x, 1)\), we have

\[
B(z, x) = 1 + zx \cdot B(z, x) \cdot A(z)
\]
and so
\[ B(z, x) = \frac{1}{1 - zx \cdot A(z)} = \frac{2}{x\sqrt{1 - 4z} - x + 2}. \]

Finally, noting the symmetry \( C(z, x, 1) = C(z, 1, x) \), we obtain
\[
C(z, x, y) = 1 + zxy \left( \frac{2}{x\sqrt{1 - 4z} - x + 2} \right) \left( \frac{2}{y\sqrt{1 - 4z} - y + 2} \right)
\]
\[ = 1 + \frac{4zxy}{(2 - x + x\sqrt{1 - 4z})(2 - y + y\sqrt{1 - 4z})}. \]

Once we have the explicit expression for \( C(z, x, y) \), in order to prove that Equation (77) defines a probability distribution on \( \text{Av}^k(231) \), it suffices to notice that
\[ C(z, 2, 2) = \frac{1}{1 - 4z}, \]
and so \( [z^n]C(z, 2, 2) = 2^n \).

### Appendix B. A Small Lemma

**Lemma B.1.** Let \((X_n)_{n \in \mathbb{Z}_0}, (\tilde{X}_n)_{n \in \mathbb{Z}_0}\) be two sequences of real random variables with values in \([0, 1]\) such that, for all \(n \in \mathbb{Z}_0\),
\[ X_n - \tilde{X}_n = o_P(1), \]
i.e., \( X_n - \tilde{X}_n \overset{P}{\rightarrow} 0 \). Then, for every real continuous random process \( Y = (Y_t)_{t \in [0, 1]} \) (i.e., a random variable with values in \( C([0, 1], \mathbb{R}) \)) we have
\[ Y_{X_n} - Y_{\tilde{X}_n} = o_P(1). \]

**Proof.** Since \( Y = (Y_t)_{t \in [0, 1]} \) is a.s. continuous on a compact set, it is also a.s. uniformly continuous. Therefore the continuity modulus \( \omega(Y, \delta) := \sup_{s, t \in [0, 1], |s - t| \leq \delta} |Y_s - Y_t| \) tends a.s. to zero as \( \delta \) tends to zero. By Skorokhod’s representation theorem we can assume that there exist a coupling such that \( X_n - \tilde{X}_n \overset{a.s.}{\rightarrow} 0 \), therefore
\[ |Y_{X_n} - Y_{\tilde{X}_n}| \leq \omega(Y, X_n - \tilde{X}_n) \overset{a.s.}{\rightarrow} 0. \]
Hence we can conclude that \( Y_{X_n} - Y_{\tilde{X}_n} = o_P(1) \) in the original probability space. \( \square \)

### Acknowledgements

The author is very grateful to Mathilde Bouvel and Valentin Féray for introducing him to the fantastic world of random permutations. He also thanks them for the constant and stimulating discussions and suggestions.

The author also warmly thanks Benedikt Stufler for precious suggestions about the terminology for Benjamini–Schramm convergence and for some explanations about local results for random trees.

Finally, he thanks Tommaso Padovan for precious help in the realization of various simulations.

This work was completed with the support of the SNF grant number 200021_172536, “Several aspects of the study of non-uniform random permutations”.

### References

[1] R. Abraham and J.-F. Delmas. An introduction to Galton-Watson trees and their local limits. arXiv:1506.05571, 2015.

[2] D. Aldous. Asymptotic fringe distributions for general families of random trees. The Annals of Applied Probability, pages 228–266, 1991.

[3] D. Aldous. The continuum random tree I. The Annals of Probability, pages 1–28, 1991.

[4] D. Aldous. The continuum random tree II: an overview. Stochastic analysis, 167:23–70, 1991.

[5] D. Aldous. The continuum random tree III. The Annals of Probability, pages 248–289, 1993.
[6] D. Aldous and J. Pitman. Tree-valued Markov chains derived from Galton-Watson processes. In *Annales de l'Institut Henri Poincaré (B) Probability and Statistics*, volume 34, pages 637–686. Elsevier, 1998.

[7] D. Aldous and J. M. Steele. The objective method: probabilistic combinatorial optimization and local weak convergence. In *Probability on discrete structures*, pages 1–72. Springer, 2004.

[8] O. Angel and O. Schramm. Uniform infinite planar triangulations. *Communications in Mathematical Physics*, 241(2-3):191–213, 2003.

[9] F. Bassino, M. Bouvel, V. Féray, L. Gerin, M. Maazoun, and A. Pierrot. Universal limits of substitution-closed permutation classes. *arXiv:1706.08333*, 2017.

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

[11] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electronic Journal of Probability*, 6, 2001.

[12] J. Bertoin and J. Pitman. Path transformations connecting Brownian bridge, excursion and meander. *Bulletin des sciences mathématiques*, 118(2):147–166, 1994.

[13] P. Billingsley. *Probability and measure*. John Wiley & Sons, 2008.

[14] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.

[15] M. Bóna. The absence of a pattern and the occurrences of another. *Discrete Mathematics and Theoretical Computer Science*, 12(2):89–102, 2010.

[16] M. Bóna. Surprising symmetries in objects counted by Catalan numbers. *the Electronic Journal of Combinatorics*, 19(1):Paper 62, 11, 2012.

[17] M. Bóna. *Combinatorics of permutations*. Chapman and Hall/CRC, 2016.

[18] A. Claesson and S. Kitaev. Classification of bijections between 321-and 132-avoiding permutations. *Séminaire Lotharingien de Combinatoire*, 60:B60d, 2008.

[19] H. Crane, S. DeSalvo, and S. Elizalde. The probability of avoiding consecutive patterns in the Mallows distribution. *Random Structures & Algorithms*, 2018.

[20] S. Elizalde. A survey of consecutive patterns in permutations. In *Recent trends in combinatorics*, pages 601–618. Springer, 2016.

[21] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge University press, 2009.

[22] P. Frenkel. Convergence of graphs with intermediate density. *Transactions of the American Mathematical Society*, 370(5):3363–3404, 2018.

[23] C. Hoffman, D. Rizzolo, and E. Slivken. Fixed points of 321-avoiding permutations. *arXiv:1607.08742*, 2016.

[24] C. Hoffman, D. Rizzolo, and E. Slivken. Pattern-avoiding permutations and Brownian excursion part i: shapes and fluctuations. *Random Structures & Algorithms*, 50(3):394–419, 2017.

[25] C. Hoffman, D. Rizzolo, and E. Slivken. Pattern-avoiding permutations and Brownian excursion, part ii: fixed points. *Probability Theory and Related Fields*, 169(1-2):377–424, 2017.

[26] C. Holmgren and S. Janson. Fringe trees, Crump–Mode–Jagers branching processes and m-ary search trees. *Probability Surveys*, 14:53–154, 2017.

[27] C. Hoppen, Y. Kohayakawa, C. G. Moreira, B. Ráth, and R. M. Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, 2013.

[28] S. Janson. The Wiener index of simply generated random trees. *Random Structures & Algorithms*, 22(4):337–358, 2003.

[29] S. Janson. Simply generated trees, conditioned Galton–Watson trees, random allocations and condensation. *Probability Surveys*, 9:103–252, 2012.

[30] S. Janson. Patterns in random permutations avoiding the pattern 132. *Combinatorics, Probability and Computing*, 26(1):24–51, 2017.

[31] S. Janson. Patterns in random permutations avoiding the pattern 321. *arXiv:1709.08427*, 2017.

[32] S. Janson. Patterns in random permutations avoiding some sets of multiple patterns. *arXiv:1804.06071*, 2018.

[33] O. Kallenberg. *Random measures, theory and applications*. Springer, 2017.

[34] R. Kenyon. Dominos and the Gaussian free field. *Annals of probability*, pages 1128–1137, 2001.

[35] R. Kenyon, D. Kral, C. Radin, and P. Winkler. Permutations with fixed pattern densities. *arXiv:1506.02340*, 2015.

[36] H. Kesten. Subdiffusive behavior of random walk on a random cluster. *Ann. Inst. H. Poincaré Probab. Statist*, 22(4):425–487, 1986.

[37] M. Krikun. Local structure of random quadrangulations. *arXiv preprint math/0512304*, 2005.

[38] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *The Annals of Probability*, 41(4):2880–2960, 2013.

[39] L. Lovász. *Large networks and graph limits*, volume 60. American Mathematical Soc., 2012.

[40] M. Maazoun. On the Brownian separable permuto. *arXiv:1711.08986*, 2017.
[41] N. Madras and L. Pehlivan. Structure of random 312-avoiding permutations. *Random Structures & Algorithms*, 49(3):599–631, 2016.
[42] J.-F. Marckert and A. Mokkadem. Limit of normalized quadrangulations: the Brownian map. *The Annals of Probability*, 34(6):2144–2202, 2006.
[43] G. Miermont. The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta mathematica*, 210(2):319–401, 2013.
[44] S. Miner and I. Pak. The shape of random pattern-avoiding permutations. *Advances in Applied Mathematics*, 55:86–130, 2014.
[45] J. Neveu. Arbres et processus de Galton-Watson. *Annales de l’Institut Henri Poincaré. Probabilités et Statistique*, 22(2):199–207, 1986.
[46] R. Otter. The multiplicative process. *The Annals of Mathematical Statistics*, pages 206–224, 1949.
[47] R. Pinsky. The infinite limit of random permutations avoiding patterns of length three. *arXiv:1806.07669*, 2018.
[48] M. Rahman, B. Virag, and M. Vizer. Geometry of permutation limits. *arXiv:1609.03891*, 2016.
[49] S. Sheffield. Gaussian free fields for mathematicians. *Probability theory and related fields*, 139(3-4):521–541, 2007.
[50] S. Starr. Thermodynamic limit for the Mallows model on $S_n$. *Journal of mathematical physics*, 50(9):095208, 2009.
[51] R. Stephenson. Local convergence of large critical multi-type Galton–Watson trees and applications to random maps. *Journal of Theoretical Probability*, 31(1):159–205, 2018.
[52] B. Stufler. Local limits of large Galton-Watson trees rerooted at a random vertex. *arXiv:1611.01048*, 2016.
[53] V. Vatter. Permutation classes. In *Handbook of enumerative combinatorics*, Discrete Math. Appl. (Boca Raton), pages 753–833. CRC Press, Boca Raton, FL, 2015.

*Institut für Mathematik, Universität Zürich, Winterthurerstrasse 190, CH-8057, Zürich*

*E-mail address*: jacopo.borga@math.uzh.ch