THE NUMBER OF OCCURRENCES OF PATTERNS IN A RANDOM TREE OR FOREST PERMUTATION

SVANTE JANSON

ABSTRACT. The classes of tree permutations and forest permutations were defined by Acan and Hitczenko (2016). We study random permutations of a given length from these classes, and in particular the number of occurrences of a fixed pattern in one of these random permutations. The main results show that the distributions of these numbers are asymptotically normal.

The proof uses representations of random tree and forest permutations that enable us to express the number of occurrences of a pattern by a type of U-statistics; we then use general limit theorems for the latter.

1. Introduction

A number of authors have studied properties of random permutations drawn uniformly from all permutations of a given (large) length in some given class of permutations. The chosen class of permutations is often a *pattern class*, i.e., is the class of all permutations avoiding a certain set of one or several given patterns; equivalently, the class is closed under taking patterns (subpermutations). (See Section 2 for definitions of various terms used here and below.) Several different properties have been studied; in the present paper we consider the asymptotic distribution of the number of occurences of some fixed pattern. For this problem (and many others), it seems impossible to give general results valid for all such permutation classes. (See e.g. Garrabrant and Pak [4] for some related impossibility results supporting this.) Therefore, typically these classes are studied one by one, with methods depending on the knowledge of some structure theorem for permutations in that particular class. See e.g. [2] and [10] for some results of this type.

The present paper continues this line of research by studying the number of occurences of a given pattern in a random *tree permutation* or *forest permutation*. These classes of permutations were defined by Acan and Hitczenko [1] as follows.

Definition 1.1. For a permutation π of $[n]$, its permutation graph G_π is the (labelled, undirected) graph with vertex set $[n]$, and an edge ij for every inversion (i, j) in π, i.e., for every pair (i, j) such that $i < j$ and $\pi(i) > \pi(j)$.

A permutation π is a *tree permutation* if G_π is a tree, and a *forest permutation* if the graph G_π is a forest (i.e., acyclic).

Thus, every tree permutation is a forest permutation.

Acan and Hitczenko [1] noted also the following characterization, showing that the forest permutations form a pattern class.

Proposition 1.2 ([1]). *The forest permutations are precisely the permutations avoiding the patterns* 321 *and* 3412*.*

Date: 8 March, 2022; typo corrected 9 March, 2022.

Supported by the Knut and Alice Wallenberg Foundation.

However, the class of tree permutations is *not* a pattern class, since a subpermutation of a tree permutation may be a forest permutation with a disconnected permutation graph. (For example, 312 is a tree permutation, but its subpermutation 12 is not.)

The structures of tree permutations and forest permutations were studied in [1]; see Section 4. Using this, and results on (conditioned) U-statistics, we will show that the number of occurences of a fixed pattern in a random tree or forest permutation is asymptotically normal, as the length tends to ∞; precise results are stated in Section 3, and proved in the remainder of the paper. Section 5 defines the versions of U-statistics that are used in the paper, and cites some results for them from [9] and [11]. Tree and forest permutations are studied in Sections 6–12, leading to a representation of random forest permutations in Section 7 and a, quite different, representation of random tree permutations in Section 10; these representations both enable us to count patterns by U-statistics, which eventually yields proofs of the theorems.

Remark 1.3. Although we use similar methods for patterns in random tree permutations and in random forest permutations, the details are quite different, and we see no direct relation between the results for the two cases. Note that a random forest permutation is a (random) sum of tree permutations, but most of these are very small (see (7.19) and (7.5)); hence there is no reason to expect a relation between asymptotics for large forest permutations and large tree permutations. □

2. Definitions and notation

2.1. **Permutations.** Let $\mathfrak{S}_n$ be the set of permutations of $[n] := \{1, \dots, n\}$, and $\mathfrak{S}_* := \bigcup_n \mathfrak{S}_n$. Similarly, let $\mathfrak{F}_n$ be the set of all forest permutations of length n and $\mathfrak{T}_n$ the subset of tree permutations, and let $\mathfrak{F}_* := \bigcup_n \mathfrak{F}_n$ and $\mathfrak{T}_* := \bigcup_n \mathfrak{T}_n$. Thus $\mathfrak{T}_n \subseteq \mathfrak{F}_n \subseteq \mathfrak{S}_n$.

We denote the length of a permutation π by $|\pi|$.

2.2. **Occurrence of patterns.** If $\sigma = \sigma_1 \cdots \sigma_m \in \mathfrak{S}_m$ and $\pi = \pi_1 \cdots \pi_n \in \mathfrak{S}_n$, then an *occurrence* of σ in π is a subsequence $\pi_{i_1} \cdots \pi_{i_m}$, with $1 \leqslant i_1 < \cdots < i_m \leqslant n$, that has the same order as σ, i.e., $\pi_{i_j} < \pi_{i_k} \iff \sigma_j < \sigma_k$ for all $j, k \in [m]$. In this context, σ is often called a *pattern*; we may also say that σ is a *subpermutation* of π. We let $\operatorname{occ}_\sigma(\pi)$ be the number of occurrences of σ in π, and note that

$$\sum_{\sigma \in \mathfrak{S}_m} \operatorname{occ}_\sigma(\pi) = \binom{n}{m}, \tag{2.1}$$

for every $\pi \in \mathfrak{S}_n$ and every m. For example, an inversion is an occurrence of 21, and thus $\operatorname{occ}_{21}(\pi)$ is the number of inversions in π.

We say that a permutation π *avoids* another permutation τ if $\operatorname{occ}_\tau(\pi) = 0$; otherwise, π *contains* τ.

2.3. **Sums and decompositions of permutations.** If $\sigma \in \mathfrak{S}_m$ and $\tau \in \mathfrak{S}_n$, their (direct) *sum* $\sigma \oplus \tau \in \mathfrak{S}_{m+n}$ is defined by letting τ act on $[m+1, m+n]$ in the natural way; more formally, $\sigma \oplus \tau = \pi \in \mathfrak{S}_{m+n}$ where $\pi_i = \sigma_i$ for $1 \leqslant i \leqslant m$, and $\pi_{j+m} = \tau_j + m$ for $1 \leqslant j \leqslant n$. It is easily seen that $\oplus$ is an associative operation. We say that a permutation $\pi \in \mathfrak{S}_*$ is *decomposable* if $\pi = \sigma \oplus \tau$ for some $\sigma, \tau \in \mathfrak{S}_*$, and *indecomposable* otherwise; we also call an indecomposable permutation a *block*. See further e.g. [3, Exercise VI.14].

It is easy to see that any permutation $\pi \in \mathfrak{S}_*$ has a unique decomposition $\pi = \pi_1 \oplus \cdots \oplus \pi_\ell$ into indecomposable permutations (blocks) $\pi_1, \dots, \pi_\ell$ (for some, unique, $\ell \geqslant 1$); we may call these the *blocks of* π

If $i < j < k$ and ik is an edge in the permutation graph G_π (i.e., an inversion), then at least one of ij and jk is also an edge. It follows that the components of the graph G_π are intervals in $[n]$, and then it is easy to see that they correspond to the blocks of π; in particular, G_π is connected if and only if π is indecomposable.

2.4. **Random permutations.** $\boldsymbol{\tau}_n$ will always denote a uniformly random tree permutation of length n; similarly, $\boldsymbol{\pi}_n$ is a uniformly random forest permutation of length n. In other words, these are uniformly random elements of $\mathfrak{T}_n$ and $\mathfrak{F}_n$, respectively.

$\widetilde{\boldsymbol{\tau}}$ denotes a certain random tree permutation of random length defined in Section 7, see (7.5); $\widetilde{\boldsymbol{\tau}}_1, \widetilde{\boldsymbol{\tau}}_2, \dots$ will denote independent copies of $\widetilde{\boldsymbol{\tau}}$. Similarly, $\boldsymbol{\tau}^*_m$ is another random tree permutation of random length, defined in Section 10.

2.5. **Some further notation.** Convergence in distribution is denoted by $\overset{\mathrm{d}}{\longrightarrow}$, and convergence in probability by $\overset{\mathrm{p}}{\longrightarrow}$. We let $\overset{\mathrm{d}}{=}$ denote equality in distribution.

Given sequences of random variables X_n and constants $a_n > 0$, and a fixed exponent $q > 0$, we let $X_n = O_{L^q}(a_n)$ mean $\mathbb{E}\,|X_n/a_n|^q = O(1)$. Moreover, we write $X_n = O_{L^*}(a_n)$ if $X_n = O_{L^q}(a_n)$ for every $q < \infty$.

By "convergence of all moments" we mean both ordinary and absolute moments, including centered versions.

We find it convenient to express some explicit constants using

$$\phi := \frac{1+\sqrt{5}}{2}, \tag{2.2}$$

the golden ratio. Recall that $\phi^2 = \phi + 1$. We will also let $p := \phi^{-2}$, see (7.1)–(7.3).

Unspecified limits are as $n \to \infty$.

3. Main results

Our main results are the following; the proofs are given later. In both cases, note that if σ is not a forest permutation, then $\mathrm{occ}_\sigma(\boldsymbol{\pi}_n) = 0$. Note also that we may assume $|\sigma| \geqslant 2$, since the case $\sigma = 1$ is utterly trivial with $\mathrm{occ}_1(\pi) = n$ for every $\pi \in \Sigma_n$. Moreover, if $\tau \in \mathfrak{T}_n$ is a tree permutation, then $\mathrm{occ}_{21}(\tau) = n - 1$, since the number of inversions equals the number of edges in the tree G_τ.

Theorem 3.1. *Let $\boldsymbol{\tau}_n$ be a uniformly random tree permutation of length n, and let σ be a fixed forest permutation with block decomposition $\sigma = \sigma_1 \oplus \cdots \oplus \sigma_d$. Then, as $n \to \infty$, for some $\gamma^2 = \gamma^2_\sigma \geqslant 0$,*

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\tau}_n) - n^d/d!}{n^{d-1/2}} \overset{\mathrm{d}}{\longrightarrow} \mathsf{N}\big(0, \gamma^2\big), \tag{3.1}$$

with convergence of all moments. Moreover, $\gamma^2 > 0$ unless $|\sigma_i| \leqslant 2$ for every i, i.e., unless each block σ_i is either 1 *or* 21.

We state the special case $d = 1$ separately.

Corollary 3.2. *Let $\boldsymbol{\tau}_n$ be a uniformly random tree permutation of length n, and let σ be a fixed tree permutation. Then, as $n \to \infty$, for some $\gamma^2 = \gamma_\sigma^2 \geqslant 0$,*

$$\frac{\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) - n}{\sqrt{n}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0, \gamma^2\bigr), \tag{3.2}$$

with convergence of all moments. Moreover, $\gamma^2 > 0$ except in the trivial cases $|\sigma| \leqslant 2$, when $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n)$ is deterministic (n or $n-1$).

Furthermore, when σ is a tree permutation, we give an exact formula for $\mathbb{E}\operatorname{occ}_\sigma(\boldsymbol{\tau}_n)$ in Theorem 6.5; this expectation depends on n and $|\sigma|$ only.

The asymptotic variance γ_σ^2 in Theorem 3.1 and Corollary 3.2 can be found from our proof, but we do not know any explicit formula; we evaluate it for some simple cases in Example 12.3. Note that Example 12.3 shows that γ_σ^2 in Corollary 3.2 really depends on σ, and, moreover, that it is not simply a function of $|\sigma|$.

Remark 3.3. If σ is a foresst permutation with $d \geqslant 2$ blocks σ_i, all of lengths $|\sigma_i| \leqslant 2$, then $\gamma^2 = 0$ in (3.1), but $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n)$ is, in general, not deterministic. We conjecture that $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n)$ is asymptotically normal in this case too, with a variance of smaller order than in Theorem 3.1, but we have not pursued this and leave it as an open problem. (Cf. Theorem 3.6 below for random forest permutations π_n.) $\square$

Problem 3.4. Find a combinatorial explanation for the surprising fact that the asymptotic expectation n in (3.2) is the same for all tree permutations σ. (We will see in the proof that this is equivalent to the fact that the expectation in (11.23) is the same for all tree permutations σ.)

More generally, find a combinatorial explanation for the fact that the asymptotic expectation $n^d/d!$ (or, equivalently, $\binom{n}{d}$) in (3.1) depends only on the the number of blocks d in σ.

Moreover, as just mentioned, Theorem 6.5 shows that for two tree permutations σ_1 and σ_2 of the same length, the expectations $\mathbb{E}\operatorname{occ}_{\sigma_1}(\boldsymbol{\tau}_n)$ and $\mathbb{E}\operatorname{occ}_{\sigma_2}(\boldsymbol{\tau}_n)$ are equal for every n. (This obviously requires $|\sigma_1| = |\sigma_2|$, since $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) = 0$ if $n < |\sigma|$.) Again, we do not know a simple proof of this fact, although the proof of Theorem 6.5 gives a kind of combinatorial reason. Also, we do not know whether the equality extends to two forest permutations with the same length and the same number of blocks.

We turn to patterns in a random forest permutation.

Theorem 3.5. *Let $\boldsymbol{\pi}_n$ be a uniformly random forest permutation of length n, and let σ be a fixed forest permutation with block decomposition $\sigma = \sigma_1 \oplus \cdots \oplus \sigma_d$. Let λ be the number of blocks σ_i of length $|\sigma_i| = 1$, and let*

$$\widetilde{\mu}_\sigma := \frac{1}{d!}(\phi+2)^{\lambda-d}\phi^{4d-3\lambda-|\sigma|} = \frac{1}{d!}5^{-(d-\lambda)/2}\phi^{3d-2\lambda-|\sigma|}. \tag{3.3}$$

Then, for some $\gamma_\sigma^2 \geqslant 0$,

$$\frac{\operatorname{occ}_\sigma(\boldsymbol{\pi}_n) - \widetilde{\mu}_\sigma n^d}{n^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0, \gamma_\sigma^2\bigr), \tag{3.4}$$

with convergence of all moments.

Furthermore, $\gamma_\sigma^2 > 0$ except in the case $\sigma = 1 \cdots d$ (the identity permutation with every $|\sigma_i| = 1$).

Again, the asymptotic variance γ_σ^2 can in principle be found from our proof, but we do not know any explicit formula; see Remark 9.1 and Example 9.2.

In the exceptional case $\sigma = 1 \cdots d$, the limit in (3.4) is 0, and a different normalization is required.

Theorem 3.6. *Let ι_d be the identity permutation $1 \cdots d$ for some $d \geqslant 2$. Then, for some $\gamma^2_{\iota_d} > 0$,*

$$\frac{\operatorname{occ}_{\iota_d}(\boldsymbol{\pi}_n) - \binom{n}{d} + \frac{5+\sqrt{5}}{10(d-2)!} n^{d-1}}{n^{d-3/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\big(0, \gamma^2_{\iota_d}\big), \tag{3.5}$$

with convergence of all moments.

Remark 3.7. If we consider several patterns, (3.1), (3.2), (3.4) and (3.5) extend to joint convergence to a multi-variate normal limit. This follows by the same proof, using Remarks 5.7 and 5.13. We omit the details. □

4. Preliminaries on tree and forest permutations

We recall some facts from (mainly) [1] (in our notation); for completeness we sometimes sketch the arguments, but we refer to [1] for further details.

Note first that a permutation is determined by its (labelled) permutation graph, in other words, the mapping $\pi \mapsto G_\pi$ is injective. Furthermore, the induced subgraphs of G_π are the inversion graphs of the patterns occuring in π, up to obvious relabelling.

In particular, it is easily seen that the only induced cycles in a permutation graph are C_3 and C_4 (as unlabelled graphs); these are the permutation graphs of 321 and 3412 (and no other permutations), which proves Proposition 1.2.

Moreover, G_π is a forest if and only if its component are trees, and thus π is a forest permutation if and only its blocks are tree permutations. In other words,

$$\pi \in \mathfrak{F}_* \iff \pi = \tau_1 \oplus \cdots \oplus \tau_m \tag{4.1}$$

for some (unique) sequence $\tau_1, \dots, \tau_m$ of tree permutations. (We will find the asymptotic distribution of the number of blocks in a random forest permutation in Theorem 9.3.)

Let $t_n := |\mathfrak{T}_n|$ be the number of tree permutations of length n. It is shown in [1] that

$$t_n = \begin{cases} 1, & n = 1, \\ 2^{n-2}, & n \geqslant 2, \end{cases} \tag{4.2}$$

and thus the corresponding generating fuction $T(z)$ is

$$T(z) := \sum_{n=1}^{\infty} t_n z^n = z + \frac{z^2}{1-2z} = \frac{z-z^2}{1-2z}, \qquad |z| < 1/2. \tag{4.3}$$

As a consequence of (4.3) and (4.1), if f_n is the number of forest permutations of length n (with $f_0 := 1$), then the corresponding generating function is

$$F(z) := \sum_{n=0}^{\infty} f_n z^n = \frac{1}{1-T(z)} = \frac{1-2z}{1-3z+z^2}. \tag{4.4}$$

The sequence (f_n) is A001519 in [13] (where many other interpretations are given).

In a permutation π, label the left-to-right maxima by L, and the right-to-left minima by R. Thus, i is labelled L if $\pi(j) < \pi(i)$ for every $j < i$, i.e., if there are no inversions (j, i) with $j < i$. In other words,

$$i \text{ is labelled } \mathsf{L} \iff i \text{ is the left endpoint of every adjacent edge in } G_\pi. \tag{4.5}$$

Similarly, i is labelled R if there are no inversions (i,j) with $j>i$, and

$$i \text{ is labelled } \mathsf{R} \iff i \text{ is the right endpoint of every adjacent edge in } G_\pi. \tag{4.6}$$

Now, let π be a forest permutation. If $i<j<k$, then ij and jk cannot both be edges in G_π, since otherwise, $\pi(i)>\pi(j)>\pi(k)$, so (i,k) would also be an inversion, and thus G_π would contain a cycle ijk. If follows that every $j\in[n]$ is labelled either L or R, or possibly both.

Moreover, (4.5)–(4.6) imply that i is labelled both L and R if and only if i is isolated in G_π. In a tree permutation π with $|\pi|\geqslant 2$, this is impossible. Thus, if π is a tree permutation with $|\pi|\geqslant 2$, then every $i\in[n]$ is labelled L or R, but not both. Each tree permutation τ with $|\tau|=n\geqslant 2$ may thus be represented by a string Ω_τ of n symbols L or R. (The notation in [1] is different: there W_1 $[W_0]$ denotes the set of i labelled L $[\mathsf{R}]$ here.) The first symbol in Ω_τ is always L and the last is R. We let $\Sigma_n := \bigl\{\mathsf{L}\{\mathsf{L},\mathsf{R}\}^{n-2}\mathsf{R}\bigr\}$ be the set of such strings, so $\Omega_\tau\in\Sigma_n$. It is shown in [1] that the map $\tau\mapsto\Omega_\tau$ is a bijection between $\mathfrak{T}_n$ and Σ_n, for every $n\geqslant 2$. (Note that $|\mathfrak{T}_n|=2^{n-2}=|\Sigma_n|$ by (4.2).) In other words, for $n\geqslant 2$, the tree permutations in $\mathfrak{T}_n$ can be encoded by the strings in Σ_n.

We follow [1] and define the *blocks* $B_1,\dots,B_{2m}$ of Ω_τ as the successive runs of L and R in Ω_τ. Note that since Ω_τ begins with L and ends with R, there is always an even number of blocks; an odd-numbered block B_{2l-1} is a run of L and an even-numbered block B_{2l} is a run of R. (Note that we also use 'block' in a different sense for the block decomposition of a permutation into its blocks (components) in Section 2.3; there should be no risk of confusion since the two different meanings of 'block' appear in different contexts, and we will not use both at the same time.) [1, Lemma 8] shows how the edges and vertex degrees in G_τ can be found explicitly from the code Ω_τ and the blocks B_i. We summarize this as follows.

Lemma 4.1 ([1]). *Let τ be a tree permutation with $|\tau|\geqslant 2$. Then the pairs of symbols in Ω_τ that correspond to edges in G_τ (and thus to inversions in τ) are:*

- (e1) *each L and the nearest following R;*
- (e2) *each R and the nearest preceding L;*
- (e3) *The last L in a block B_{2k-1} and the first R in B_{2k+2}.*

The symbols in Ω_τ that correspond to leaves in G_τ are the following:

- (l1) *every L that is not the last L in its block;*
- (l2) *the last but one symbol, if that is L;*
- (l3) *every R that is not the first R in its block;*
- (l4) *the second symbol, if that is R.*

Proof. As said above, this is [1, Lemma 8], in different notation. (The four cases (l1)–(l4) correspond to parts (a),(c),(d),(f) in that lemma.) ☐

If σ is a tree permutation with $|\sigma|\geqslant 2$ such that its code Ω_σ has $2m$ blocks, we define $b(\sigma):=m$; in other words the code of σ has $b(\sigma)$ L-blocks and $b(\sigma)$ R-blocks. If $|\sigma|=1$, we do not define any code Ω_σ, but we define (for later convenience) $b(\sigma):=1$.

5. Preliminaries on U-statistics

A *U-statistic* is a random variable of the form

$$U_n=U_n(f)=\sum_{1\leqslant i_1<\dots<i_d\leqslant n} f\bigl(X_{i_1},\dots,X_{i_d}\bigr), \qquad n\geqslant 0, \tag{5.1}$$

where $X_1, X_2, \dots$ is an i.i.d. sequence of random variables with values in some measurable space $\mathcal{S}$, and $f : \mathcal{S}^d \to \mathbb{R}$ is a given measurable function of $d \geqslant 1$ variables. (It is often assumed that f is a symmetric function; we do not assume this.) U-statistics were introduced by Hoeffding [6]; we will use versions and results from [9] and [11], see also [10] for similar applications to pattern occurences in some other pattern classes.

The fundamental central limit theorem for U-statistics, due to Hoeffding [6] in the symmetric case, can in the general (asymmetric) case be stated as follow, see [7, Theorem 11.20] and [9, Corollary 3.5 and (moment convergence) Theorem 3.15]. Assume that the random variables X_i are i.i.d., let X denote a generic X_i, and define (for a given f)

$$\mu := \mathbb{E}\, f(X_1, \dots, X_d), \tag{5.2}$$

$$f_i(x) := \mathbb{E}\big[f(X_1, \dots, X_d) \mid X_i = x\big], \tag{5.3}$$

$$\sigma_{ij} := \operatorname{Cov}\big[f_i(X), f_j(X)\big], \tag{5.4}$$

$$\sigma^2 := \sum_{i,j=1}^{d} \frac{(i+j-2)!\,(2d-i-j)!}{(i-1)!\,(j-1)!\,(d-i)!\,(d-j)!\,(2d-1)!}\sigma_{ij}. \tag{5.5}$$

Note that $f_i(x)$ in [9; 11] is $f_i(x) - \mu$ in the present notation.

Proposition 5.1 ([7; 9]). *Suppose that $(X_i)_1^\infty$ are i.i.d. random variables, and that $\mathbb{E}\,|f(X_1, \dots, X_d)|^2 < \infty$. Then, with the notation in* (5.2)–(5.5)*, as $n \to \infty$,*

$$\frac{U_n - \binom{n}{d}\mu}{n^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\big(0, \sigma^2\big). \tag{5.6}$$

Furthermore, $\sigma^2 > 0$ unless $f_i(X) = \mu$ a.s. for $i = 1, \dots, d$.

Moreover, if $\mathbb{E}\,|f(X_1, \dots, X_d)|^p < \infty$ for some $p \geqslant 2$, the (5.6) *holds with convergence of all moments of order $\leqslant p$.* □

We will need a renewal theory version of Proposition 5.1. In addition to a sequence $(X_i)_1^\infty$ and a function f as above, let $h : \mathcal{S} \to \mathbb{R}$ be another measurable function, and assume (for simplicity) that $h(X_i) \geqslant 0$ a.s. Define

$$\nu := \mathbb{E}\, h(X_i), \tag{5.7}$$

$$S_n = S_n(h) := \sum_{i=1}^{n} h(X_i), \tag{5.8}$$

and let for each $x > 0$

$$N(x) := \inf\{N : S_N \geqslant x\}. \tag{5.9}$$

Remark 5.2. The definition (5.9) agrees with $N_+(x)$ in [10] but differs slightly from $N_+(x)$ and $N_-(x)$ in [9] and [11]; this does not affect the asymptotic results used here, see [11, Remark 3.19]. (For integer valued h and integer x, as in our application, $N(x) = N_+(x-1)$.) We will use results from [9] and [11]; note that the event $\{S_k = n \text{ for some } k \geqslant 0\}$ equals $\{S_{N(n)} = n\}$ in the present notation, and $\{U_{N_-(n)} = n\}$ in the notation of [9] and [11]. (When we condition on this event in propositions below, we tacitly consider only n such that the event has positive probability.) □

The following results are special cases of [9, Theorems 3.11, 3.13(iii) and 3.18] (with somewhat different notation).

Proposition 5.3 ([9]). *Suppose that (X_i) are i.i.d., $\mathbb{E}|f(X_1,\dots,X_d)|^2 < \infty$, and $h(X) \geqslant 0$ a.s., with $\nu := \mathbb{E}\,h(X) > 0$ and $\mathbb{E}\,h(X)^2 < \infty$. Then, with notations as above, as $x \to \infty$,*

$$\frac{U_{N(x)} - \mu\nu^{-d}d!^{-1}x^d}{x^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\big(0,\gamma^2\big), \tag{5.10}$$

where, with σ^2 given by (5.5),

$$\gamma^2 := \nu^{1-2d}\sigma^2 - 2\frac{\mu\nu^{-2d}}{(d-1)!\,d!}\sum_{i=1}^{d}\operatorname{Cov}\big[f_i(X),h(X)\big] + \frac{\mu^2\nu^{-2d-1}}{(d-1)!^2}\operatorname{Var}\big[h(X)\big]. \tag{5.11}$$

Moreover, $\gamma^2 > 0$ unless $f_i(X) = \frac{\mu}{\nu}h(X)$ a.s. for $i = 1,\dots,d$. □

Proposition 5.4 ([9]). *Suppose in addition to the hypotheses in Proposition 5.3 that $h(X)$ is integer-valued. Then* (5.10) *holds also conditioned on $S_{N(x)} = x$ (cf. Remark 5.2) for integers $x \to \infty$.* □

Proposition 5.5 ([9]). *Suppose in addition to the hypotheses in Proposition 5.3 or 5.4 that $\mathbb{E}|f(X_1,\dots,X_d)|^p < \infty$ and $\mathbb{E}|h(X)|^p < \infty$ for every $p < \infty$. Then the conclusion* (5.10) *holds with convergence of all moments.* □

Remark 5.6. In the special case $d = 1$, when the U-statistic (5.1) is a standard single sum, (5.2)–(5.5) and (5.11) simplify to $f_1 = f$, $\sigma^2 = \sigma_{11} = \operatorname{Var} f(X)$, and

$$\begin{aligned}\gamma^2 &= \frac{1}{\nu}\sigma^2 - 2\frac{\mu}{\nu^2}\operatorname{Cov}\big[f(X),h(X)\big] + \frac{\mu^2}{\nu^3}\operatorname{Var}h(X)\\ &= \frac{1}{\nu}\operatorname{Var}\Big[f(X) - \frac{\mu}{\nu}h(X)\Big].\end{aligned} \tag{5.12}$$

This special case is classical, see e.g. [5, Theorem 4.2.3]. □

Remark 5.7. The results in Propositions 5.1–5.5 hold jointly for several f (possibly with different d). This is not stated explicitly in [9] (except for (5.6)), but it follows by the same proofs as in [9] (perhaps, for convenience, using the Skorohod coupling theorem [12, Theorem 4.30] and a.s. convergence in the proofs). See also [11]. □

5.1. Constrained U-statistics.

In this subsection we extend some of the results above to constrained U-statistics, defined as follows. We consider here only a case relevant for the application in the present paper; for more general definitions and results, see [11] (with somewhat different notation).

Let, as above, $(X_i)_1^\infty$ be an i.i.d. sequence of random variables in some measurable space $\mathcal{S}$.

Let $d \geqslant 1$ and let $b_1,\dots,b_d$ be given non-negative integers. (These are regarded as fixed in this subsection.) Let

$$b'_j := b_j - 1, \tag{5.13}$$

$$D_j := \sum_1^j b_i, \qquad 0 \leqslant j \leqslant d, \tag{5.14}$$

$$D'_j := \sum_1^j b'_i = D_j - j, \qquad 0 \leqslant j \leqslant d, \tag{5.15}$$

$$D := D_d = \sum_{i=1}^d b_i = D'_d + d. \tag{5.16}$$

Suppose that $f : \mathcal{S}^D \to \mathbb{R}$ is a measurable function, and define the *constrained U-statistic*

$$\widehat{U}_n = \widehat{U}_n(f) := \sum_{i_1,\dots,i_d} f\bigl((X_{i_1+k})_{k=0}^{b'_1}, (X_{i_2+k})_{k=0}^{b'_2}, \dots, (X_{i_d+k})_{k=0}^{b'_d}\bigr) \tag{5.17}$$

summing over all $i_1,\dots,i_d$ such that $i_1 \geqslant 1$, $i_1+b'_1 < i_2$, $i_2+b'_2 < i_3$, …, $i_{d-1}+b'_{d-1} < i_d$, and $i_d+b'_d \leqslant n$. (We have grouped the arguments of f in (5.17), using an obvious notation.) In other words, $\widehat{U}_n$ is defined as U_n in (5.1), with d replaced by D, but only summing over $i_1,\dots,i_D$ such that the b_1 first indices are consecutive, as well as the next b_2, and so on. In particular, in the special case $b_1 = \dots = b_d = 1$, $\widehat{U}_n$ equals the unconstrained U-statistic U_n in (5.1).

By replacing i_j by $i_j - D'_{j-1}$ in (5.17), we obtain the alternative formula

$$\widehat{U}_n := \sum_{1\leqslant i_1<i_2<\dots<i_d\leqslant n-D'_d} f\bigl((X_{i_1+k})_{k=0}^{D'_1}, (X_{i_2+k})_{k=D'_1}^{D'_2}, \dots, (X_{i_d+k})_{k=D'_{d-1}}^{D'_d}\bigr). \tag{5.18}$$

Define, as in (5.2),

$$\mu = \mu_f := \mathbb{E}\, f(X_1,\dots,X_D). \tag{5.19}$$

By (5.18), the mean of $\widehat{U}_n$ is

$$\mathbb{E}\,\widehat{U}_n = \binom{n-D'_d}{d}\mu. \tag{5.20}$$

Proposition 5.1 extends to constrained U-statistics as follows.

Proposition 5.8 ([11]). *Let $\widehat{U}_n = \widehat{U}_n(f)$ be a constrained U-statistic defined as above, with $(X_i)_1^\infty$ i.i.d., and assume $\mathbb{E}\,|f(X_1,\dots,X_D)|^2 < \infty$. Then, with $\mu = \mu_f$ given by* (5.19) *and some $\sigma^2 = \sigma_f^2 \geqslant 0$,*

$$\frac{\widehat{U}_n - \binom{n}{d}\mu}{n^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0,\sigma^2\bigr). \tag{5.21}$$

Moreover, if $\mathbb{E}\,|f(X_1,\dots,X_d)|^p < \infty$ for some $p \geqslant 2$, the (5.21) *holds with convergence of all moments of order $\leqslant p$.*

It does not matter whether we subtract $\mathbb{E}\,\widehat{U}_n$ or $\binom{n}{d}\mu$ in (5.21), since the difference is $O\bigl(n^{d-1}\bigr) = o\bigl(n^{d-1/2}\bigr)$ by (5.20).

Proof. This is a special case of [11, Theorems 3.9 and 3.15]. □

The variance σ^2 in (5.21) can be calculated explicitly, see [11, Remark 6.2], but the formulas are a bit complicated, and we omit them. Instead, we give a criterion that often can be used in applications to show that $\sigma^2 > 0$. We define, in analogy with (5.3),

$$f_j\bigl(x_1,\dots,x_{b_j}\bigr) := \mathbb{E}\bigl[f\bigl(X_1,\dots,X_D\bigr) \mid (X_{D_{j-1}+1},\dots,X_{D_j}) = (x_1,\dots,x_{b_j})\bigr]. \tag{5.22}$$

We extend the definition (5.8) to functions $g : \mathcal{S}^b \to \mathbb{R}$ for any $b \geqslant 1$ by defining, for such g,

$$S_n(g) := \sum_{i=1}^n g\bigl(X_i,\dots,X_{i+b-1}\bigr). \tag{5.23}$$

Proposition 5.9. *In Proposition 5.8, the asymptotic variance $\sigma_f^2 = 0$ if and only if for every $j \in [d]$, there exists a function $\psi_j : \mathcal{S}^{b_j-1} \to \mathbb{R}$ such that a.s.*

$$f_j\bigl(X_1, \dots, X_{b_j}\bigr) - \mu = \psi_j\bigl(X_2, \dots, X_{b_j}\bigr) - \psi_j\bigl(X_1, \dots, X_{b_j-1}\bigr), \tag{5.24}$$

and thus a.s., for every $n \geqslant 1$,

$$S_n(f_j - \mu) = \psi_j\bigl(X_{n+1}, \dots, X_{n+b_j-1}\bigr) - \psi_j\bigl(X_1, \dots, X_{b_j-1}\bigr). \tag{5.25}$$

Consequently, if $\sigma_f^2 = 0$, then $S_n(f_j)$ is independent of $X_{b_j}, \dots, X_n$ for every $j \in [d]$ and $n \geqslant b_j$.

Proof. This is essentially a special case of [11, Theorem 8.4]; the difference is mainly notational. The function g_j in [11, Theorem 8.4 and Remark 6.2] is, in our case, given by

$$g_j\bigl(x_1, \dots, x_{D'_d+1}\bigr) = f_j\bigl(x_{D'_{j-1}+1}, \dots, x_{D'_{j-1}+b_j}\bigr) - \mu; \tag{5.26}$$

thus g_j is essentially the same as $f_j - \mu$ but contains some redundant variables. [11, Theorem 8.4] says that $\sigma_f^2 = 0$ if and only if there exists a function $\varphi_j : \mathcal{S}^{D'_d} \to \mathbb{R}$ such that a.s.

$$g_j\bigl(X_1, \dots, X_{D'_d+1}\bigr) = \varphi_j\bigl(X_2, \dots, X_{D'_d+1}\bigr) - \varphi_j\bigl(X_1, \dots, X_{D'_d}\bigr). \tag{5.27}$$

This is (5.24), except that we have redundant variables. These may be eliminated one by one. For example, if $D'_{j-1} > 0$, and thus g_j does not depend on x_1 by (5.26), then (5.27) implies that for a.e. fixed $x_1 \in \mathcal{S}$, we have $\varphi_j\bigl(X_1, \dots, X_{D'_d}\bigr) = \varphi_j\bigl(x_1, X_2, \dots, X_{D'_d}\bigr)$ a.s., and thus a.s.

$$\varphi_j\bigl(X_1, \dots, X_{D'_d}\bigr) = \varphi'_j\bigl(X_2, \dots, X_{D'_d}\bigr) \tag{5.28}$$

for some function $\varphi'_j : \mathcal{S}^{D'_d-1} \to \mathbb{R}$. Continuing in this way, from both ends, we see that a.s.

$$\varphi_j\bigl(X_1, \dots, X_{D'_d}\bigr) = \psi_j\bigl(X_{D'_{j-1}+1}, \dots, X_{D'_{j-1}+b_j-1}\bigr) \tag{5.29}$$

for some function ψ_j, and thus (5.27) reduces to (5.24). (Alternatively, one might note that (5.27) implies $\operatorname{Var}\bigl[S_n(f_j - \mu)\bigr] = \operatorname{Var} S_n(g_j) = O(1)$, and then [8, Theorem 2] yields (5.24) – this essentially repeats part of the argument in [11] yielding (5.27).) Conversely, (5.24) trivially yields (5.27) for a suitable φ_j. $\square$

We will use a renewal theory version of constrained U-statistics. We assume again that $h : \mathcal{S} \to \mathbb{R}$ with $h(X_i) \geqslant 0$ a.s., and use the notation (5.7)–(5.9). The following results are special cases of [11, Theorems 3.20, 8.7, 3.21, and 3.23].

Proposition 5.10 ([11])**.** *Let $\widehat{U}_n = \widehat{U}_n(f)$ be a constrained U-statistic defined as above, with $(X_i)_1^\infty$ i.i.d. Suppose that $\mathbb{E}\,|f(X_1, \dots, X_D)|^2 < \infty$, and that $h(X) \geqslant 0$ a.s., with $\nu := \mathbb{E}\, h(X) > 0$ and $\mathbb{E}\, h(X)^2 < \infty$. Then, with notations as above, as $x \to \infty$,*

$$\frac{\widehat{U}_{N(x)} - \mu\nu^{-d} d!^{-1} x^d}{x^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0, \gamma^2\bigr), \tag{5.30}$$

for some $\gamma^2 \geqslant 0$. Moreover, $\gamma^2 > 0$ unless, for each $j = 1, \dots, d$, the conditions in Proposition 5.9 hold with $f - \mu$ replaced by the function $f_j(X_1, \dots, X_{b_j}) - \frac{\mu}{\nu} h(X_1)$.

Proof. The limit (5.30) is a special case of [11, Theorem 3.20]. The only detail that requires a comment is that [11, Theorem 8.7] says that if $\gamma^2 = 0$, then a.s.

$$g_j\bigl(X_1,\dots,X_{D'_d+1}\bigr) + \mu - \frac{\mu}{\nu}h(X_1) = \varphi_j\bigl(X_2,\dots,X_{D'_d+1}\bigr) - \varphi_j\bigl(X_1,\dots,X_{D'_d}\bigr) \tag{5.31}$$

for some function φ, where as above g_j is given by (5.26). If we use (5.26) and define

$$\overline{\varphi}_j\bigl(x_1,\dots,x_{D'_d}\bigr) := \varphi_j\bigl(x_1,\dots,x_{D'_d}\bigr) - \sum_{i=1}^{D'_{j-1}} \frac{\mu}{\nu}h(x_i), \tag{5.32}$$

then (5.31) is equivalent to

$$f_j\bigl(X_{D'_{j-1}+1},\dots,X_{D'_{j-1}+b_j}\bigr) - \frac{\mu}{\nu}h\bigl(X_{D'_{j-1}+1}\bigr) = \overline{\varphi}_j\bigl(X_2,\dots,X_{D'_d+1}\bigr) - \overline{\varphi}_j\bigl(X_1,\dots,X_{D'_d}\bigr). \tag{5.33}$$

The result follows by eliminating redundant variables as in the proof of Proposition 5.9. □

Proposition 5.11 ([11])**.** *Suppose in addition to the hypotheses in Proposition 5.10 that $h(X)$ is integer-valued. Then* (5.30) *holds also conditioned on $S_{N(x)} = x$ for integers $x \to \infty$.* □

Proposition 5.12 ([11])**.** *Suppose in addition to the hypotheses in Proposition 5.10 or 5.11 that $\mathbb{E}\,|f(X_1,\dots,X_D)|^p < \infty$ and $\mathbb{E}\,|h(X)|^p < \infty$ for every $p < \infty$. Then the conclusion* (5.30) *holds with convergence of all moments.* □

Remark 5.13. Again, the results in Propositions 5.8 and 5.10–5.12 hold jointly for several f (possibly with different d and $b_1,\dots,b_d$), see [11]. □

6. Patterns and codes of tree permutations

Consider an occurrence of a tree permutation $\sigma \in \mathfrak{T}_\ell$ in another tree permutation $\tau \in \mathfrak{T}_n$. The occurrence is defined by a subset $I = \{i_1,\dots,i_\ell\}$ of the index set $[n]$. We colour each symbol in the code Ω_τ *red* if its index belongs to I, and *black* otherwise. We use also the same colours for the corresponding vertices in G_τ. (All colourings in this paper are in red and black. We may regard the red symbols or vertices as marked.)

Note that in the resulting coloured copy of Ω_τ, the red symbols form the code Ω_σ of σ; this is a consequence of (4.5)–(4.6) and the fact that the corresponding (red) induced subgraph of G_τ equals G_σ up to an order-preserving relabelling. However, not every subset of ℓ symbols in the right order corresponds to an occurrence of σ. There is a 1–1 correspondence between

(1) (nonempty) subsets of $[n]$,
(2) (nonempty) subsequences of Ω_τ,
(3) occurences of some permutation υ in τ,
(4) (nonempty) labelled subgraphs of the permutation graph G_τ.

However, the subgraph in (4) is not necessarily a tree, and thus, the permutation υ in (3) is not necessarily a tree permutation.

We may characterize the subsets of symbols in Ω_τ that yield occurences of σ as follows.

Lemma 6.1. *Let τ and σ be tree permutations with $|\tau| \geqslant |\sigma| \geqslant 2$. A colouring of the code Ω_τ corresponds to an occurrence of σ in τ if and only if we may the delete the black symbols one by one in some order according to the following rules (always interpreted for the current string) until only red symbols remain, and these form the code Ω_σ. The allowed deletions are (in any order, and possibly repeated):*

(A1) *a black* L *that is immediately followed by another* L*;*
(A2) *a black* L *in the last but one position;*
(A3) *a black* R *that is immediately preceded by another* R*;*
(A4) *a black* R *in position* 2*.*

Proof. Consider first the case of deleting one vertex $i \in [n]$ from the tree G_τ, i.e., restricting the permutation τ to $[n] \setminus \{i\}$ and then relabelling to get a permutation τ_1 in $\mathfrak{S}_{n-1}$. The permutation graph G_{τ_1} is an induced subgraph of G_τ, and is thus always a forest; it is a tree if and only if it is connected, which is the case exactly when i is leaf in G_τ. By Lemma 4.1, the black vertices that may be deleted leaving a tree correspond precisely to the symbols listed in (A1)–(A4).

Thus, to repeatedly remove black symbols according to the rules in the lemma, is equivalent to repeatedly removing black leaves of G_τ, leaving a red subtree; if the resulting red code is Ω_σ, then this yields an occurence of σ.

Conversely, if the colouring of Ω_τ corresponds to an occurrence of σ in τ, then the red vertices form a red subtree in G_τ, and we may remove the black vertices of G_τ is some order such that we always remove a black leaf of the current tree; this means that we may remove the black symbols in some order such that the rules (A1)–(A4) are followed. $\square$

We may invert the deletions in Lemma 6.1, and instead insert black symbols into Ω_σ.

Lemma 6.2. *Let τ and σ be tree permutations with $|\tau| \geqslant |\sigma| \geqslant 2$. A colouring of the code Ω_τ corresponds to an occurrence of σ in τ if and only if we may obtain it by from a red code Ω_σ by inserting black symbols one by one according to the following rules (always interpreted for the current string). The allowed insertions are (in any order, and possibly repeated):*

(B1) *a black* L *immediately to the left of any* L*;*
(B2) *a black* L *immediately to the left of the last symbol;*
(B3) *a black* R *immediately to the right of any* R*;*
(B4) *a black* R *immediately to the right of the first symbol.*

Proof. Immediate from Lemma 6.1. $\square$

We have so far considered deleting or inserting one symbol at a time. Since only the end result matters, the following version is more convenient for our purposes. (Recall that the first red symbol always is L, and the last is R.)

Lemma 6.3. *Let σ be a tree permutation with $|\sigma| \geqslant 2$. A coloured code Ω corresponds to a marked (red) occurrence of σ in some tree permutation τ if and only if we may obtain Ω from a red code Ω_σ by inserting black symbols as follows (the strings may be empty):*

(C1) *a string of black* L *immediately to the left of each red* L *except the first;*
(C2) *a string of black* L *immediately to the left of the last red* R*;*
(C3) *a string of black* R *immediately to the right of each red* R *except the last;*

(C4) *a string of black* R *immediately to the right of the first red* L*;*
(C5) *any black string that is empty or begins with* L *before the first red symbol;*
(C6) *any black string that is empty or ends with* R *after the last red symbol.*

Proof. It is easily seen that if we take any coloured code obtained by these rules, and insert another black symbol according to the rules in Lemma 6.2, then the result is also described by (C1)–(C6). Hence, by induction, all possible coloured codes are given by the insertions (C1)–(C6).

Conversely, suppose that Ω is obtained from a red Ω_σ by (C1)–(C6); we have to show that it also can be obtained by repeating (B1)–(B4) in some order. Evidently, (C1)–(C4) can be obtained by repeating (B1)–(B4), so it remains only to show that we may add an arbitrary black string beginning with L before the first red symbol, and an arbitrary black string ending with R after the last red symbol. To see this, note that we may first add a black L to the left by (B1). Then, when the code begins with a black L, we may by either add a black R as the second symbol by (B4), or a black L as the first symbol by (B1), but the latter gives the same result as adding a black L as the second symbol. Hence, we may add an arbitrary black symbol immediately after the first one, and by repeating this we may obtain any black string beginning with L, verifying (C5). The argument for the right side is symmetric. ☐

Lemma 6.4. *Fix a tree permutation σ with $|\sigma| \geqslant 2$. For every n, let $a_{n;\sigma}$ be the number of pairs (τ, σ') of a tree permutation τ of length $|\tau| = n$ together with a marked occurence σ' of the pattern σ. Define also the generating function*

$$A_\sigma(z) := \sum_{n \geqslant |\sigma|} a_{n;\sigma} z^n. \tag{6.1}$$

Then,

$$A_\sigma(z) = \frac{z^{|\sigma|}}{(1-z)^{|\sigma|-2}(1-2z)^2}. \tag{6.2}$$

Proof. By Lemma 6.3, $a_{n;\sigma}$ equals the number of coloured codes of length n that can be obtained from a red Ω_σ by the rules (C1)–(C6). These insertions are independent of each other, so they correspond to multiplying factors in the generating function $A_\sigma(z)$.

Each possible application of (C1)–(C4) yields a factor $\sum_{k=0}^{\infty} z^k = (1-z)^{-1}$. There is one possible such application for each symbol in Ω_σ, by (C1) or (C4) for each L, and by (C2) or (C3) for each R. Hence, the total contribution of (C1)–(C4) is $(1-z)^{-|\sigma|}$.

By (C5), we may to the left add a black prefix that is either empty or is an arbitrary sting beginning with L, which gives 2^{k-1} possible prefixes of length k for every $k \geqslant 1$ (and 1 prefix of length 0). This contributes to $A_\sigma(z)$ a factor

$$1 + \sum_{k=1}^{\infty} 2^{k-1} z^k = 1 + \frac{z}{1-2z} = \frac{1-z}{1-2z}. \tag{6.3}$$

Black suffixes by (C6) contribute the same factor. These factors all multiply the term corresponding to the original red symbols Ω_σ, which is $z^{|\sigma|}$. Hence, we obtain

$$A_\sigma(z) = z^{|\sigma|}(1-z)^{-|\sigma|}\Bigl(\frac{1-z}{1-2z}\Bigr)^2, \tag{6.4}$$

which yields (6.2). ☐

This yields an exact formula for the expected number of occurences of σ; note that the result depends only on $|\sigma|$ and n.

Theorem 6.5. *Fix a tree permutation σ with $|\sigma| \geqslant 2$. Then, for $n \geqslant |\sigma|$,*

$$\begin{aligned}\mathbb{E}\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) &= [z^n]\bigl(z^{|\sigma|}(2-z)^{2-|\sigma|}(1-z)^{-2}\bigr) = [z^{n-|\sigma|}]\bigl((2-z)^{2-|\sigma|}(1-z)^{-2}\bigr)\\ &= n+3-2|\sigma| + 2^{-n}\sum_{i=0}^{|\sigma|-3}(|\sigma|-2-i)2^{|\sigma|-i-1}\binom{n-|\sigma|+i}{i}. \end{aligned} \tag{6.5}$$

Proof. The total number of occurences of σ in tree permutations of length n is $a_{n;\sigma}$, and the number of such tree permutations is $t_n = 2^{n-2}$ by (4.2). Hence, by (6.1)–(6.2),

$$\begin{aligned}\mathbb{E}\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) &= \frac{a_{n;\sigma}}{2^{n-2}} = [z^n]\bigl(2^{2-n}A_n(z)\bigr) = [z^n]\bigl(4A_n(z/2)\bigr)\\ &= [z^n]\frac{z^{|\sigma|}}{(2-z)^{|\sigma|-2}(1-z)^2}, \end{aligned} \tag{6.6}$$

which gives the first two expressions in (6.5); the explicit formula then follows from the partial fraction expansion, with $m = |\sigma|-2 \geqslant 0$,

$$\frac{1}{(2-z)^m(1-z)^2} = \frac{1}{(1-z)^2} - \frac{m}{1-z} + \sum_{j=1}^{m}\frac{m-j+1}{(2-z)^j}. \tag{6.7}$$

□

7. A random tree permutation of random length

Recall that $T(z)$ is the generating function in (4.3), and let, throughout the paper, p be the (unique) positive root of

$$T(p) = 1. \tag{7.1}$$

By (4.3), this yields $0 < p < 1/2$ and $p - p^2 = 1 - 2p$, or $p^2 - 3p + 1 = 0$, and thus

$$p = \frac{3-\sqrt{5}}{2} = 0.381966\ldots. \tag{7.2}$$

Recalling the golden ration ϕ in (2.2), we thus have

$$p = \phi^{-2} = 2 - \phi. \tag{7.3}$$

We note also

$$1 - p = \phi - 1 = \phi^{-1}, \qquad\qquad 1 - 2p = p(1-p) = \phi^{-3}. \tag{7.4}$$

We now define a random tree permutation $\widetilde{\boldsymbol{\tau}}$ to be a random element of $\mathfrak{T}_*$ with the distribution

$$\mathbb{P}(\widetilde{\boldsymbol{\tau}} = \tau) = p^{|\tau|}, \qquad \tau \in \mathfrak{T}_*. \tag{7.5}$$

Note that the sum over all $\tau \in \mathfrak{T}_*$ of the probabilities in (7.5) equals $\sum_n t_n p^n = T(p) = 1$, and thus (7.5) really defines a probability distribution.

The random tree permutation $\widetilde{\boldsymbol{\tau}}$ thus has random length. It follows from (7.5) that the probability generating function of $|\widetilde{\boldsymbol{\tau}}|$ is

$$G_{|\widetilde{\boldsymbol{\tau}}|}(z) := \sum_{n=1}^{\infty} t_n p^n z^n = T(pz). \tag{7.6}$$

Lemma 7.1. *We have*

$$\mathbb{E}\,|\widetilde{\boldsymbol{\tau}}| = \phi + 2 = \frac{5+\sqrt{5}}{2} = \sqrt{5}\,\phi \doteq 3.618, \tag{7.7}$$

$$\mathbb{E}\,|\widetilde{\boldsymbol{\tau}}|^2 = 11\phi + 8 = \frac{27+11\sqrt{5}}{2} \doteq 25.798, \tag{7.8}$$

$$\operatorname{Var}|\widetilde{\boldsymbol{\tau}}| = 6\phi + 3 = 3\phi^3 = 6 + 3\sqrt{5} \doteq 12.708, \tag{7.9}$$

$$\mathbb{E}\,|\widetilde{\boldsymbol{\tau}}|^k < \infty, \qquad \forall k < \infty. \tag{7.10}$$

Proof. By (7.6) and straightforward calculations using (7.3)–(7.4),

$$\mathbb{E}\,|\widetilde{\boldsymbol{\tau}}| = G'_{|\widetilde{\boldsymbol{\tau}}|}(1) = pT'(p) = \frac{p(1-2p+2p^2)}{(1-2p)^2} = \phi^4(\phi^{-3} + 2\phi^{-4}) = \phi + 2. \tag{7.11}$$

Similarly,

$$\mathbb{E}\big[|\widetilde{\boldsymbol{\tau}}|(|\widetilde{\boldsymbol{\tau}}|-1)\big] = G''_{|\widetilde{\boldsymbol{\tau}}|}(1) = p^2T''(p) = \frac{2p^2}{(1-2p)^3} = 2\phi^5 = 10\phi + 6 \tag{7.12}$$

and thus, combining (7.11) and (7.12),

$$\mathbb{E}\,|\widetilde{\boldsymbol{\tau}}|^2 = 11\phi + 8 \tag{7.13}$$

and

$$\operatorname{Var}|\widetilde{\boldsymbol{\tau}}| = \big(11\phi + 8\big) - (\phi+2)^2 = 6\phi + 3. \tag{7.14}$$

This shows (7.7)–(7.9).

Finally, (7.10) follows because $G_{|\widetilde{\boldsymbol{\tau}}|}(z)$ has radius of convergence greater than 1. (Or directly from (4.2) and (7.5).) ☐

7.1. **From random trees to random forests.** Recall that forest permutations are sums of tree permutations (4.1). Let $\widetilde{\boldsymbol{\tau}}_1, \widetilde{\boldsymbol{\tau}}_2, \dots$ be an infinite sequence of independent random tree permutations with the distribution (7.5), and let

$$S_m := \sum_{i=1}^m |\widetilde{\boldsymbol{\tau}}_i|, \tag{7.15}$$

the total length of the m first of these tree permutations. Thus, for any $m \geqslant 1$, $\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_m$ is a forest permutation of length S_m, having m blocks.

Suppose that π is a forest permutation with m blocks $\tau_1, \dots, \tau_m$. Then, by (7.5),

$$\mathbb{P}\big(\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_m = \pi\big) = \mathbb{P}\big(\widetilde{\boldsymbol{\tau}}_i = \tau_i, \forall i \leqslant m\big) = \prod_{i=1}^m \mathbb{P}\big(\widetilde{\boldsymbol{\tau}}_i = \tau_i\big) = \prod_{i=1}^m p^{|\tau_i|} = p^{|\pi|}. \tag{7.16}$$

Note that this depends only on $|\pi|$.

In order to obtain arbitrary forest permutations, we have to consider a random number of blocks. We use a renewal theoretic approach. For any $n \geqslant 1$, let, as in (5.9),

$$N(n) := \min\{m \geqslant 1 : S_m \geqslant n\}. \tag{7.17}$$

Then, $S_{N(n)} \geqslant n$. Moreover, if $\pi \in \mathfrak{F}_n$ has m blocks $\pi_1, \dots, \pi_m$, then $\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_m = \pi$ entails $S_m = |\pi| = n$, and thus $N(n) = m$. Hence, using also (7.16),

$$\mathbb{P}\big(\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_{N(n)} = \pi\big) = \mathbb{P}\big(N(n) = m \;\&\; \widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_m = \pi\big)$$

$$= \mathbb{P}\bigl(\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_m = \pi\bigr) = p^{|\pi|} = p^n. \tag{7.18}$$

This probability is thus the same for all $\pi \in \mathfrak{F}_n$. Consequently, conditioned on $S_{N(n)} = n$, so that $\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_{N(n)} \in \mathfrak{F}_n$, (7.18) implies that $\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_{N(n)}$ has the uniform distribution in $\mathfrak{F}_n$, and thus

$$\boldsymbol{\pi}_n \overset{\mathrm{d}}{=} \bigl(\widetilde{\boldsymbol{\tau}}_1 \oplus \cdots \oplus \widetilde{\boldsymbol{\tau}}_{N(n)} \mid S_{N(n)} = n\bigr). \tag{7.19}$$

In words, we can construct a uniformly random $\boldsymbol{\pi}_n \in \mathfrak{F}_n$ from the infinite sequence $(\widetilde{\boldsymbol{\tau}}_i)$ by composing $\widetilde{\boldsymbol{\tau}}_1, \widetilde{\boldsymbol{\tau}}_2, \dots$ until their total length is at least n, and then condition on the total length being exactly n.

8. Trees in a random tree permutation $\widetilde{\boldsymbol{\tau}}$

The construction (7.19) suggests that it is useful to study the random variable $\operatorname{occ}_\sigma(\widetilde{\boldsymbol{\tau}})$, for a given permutation σ. We do this first for a tree permutation σ.

Lemma 8.1. *Let σ be a tree permutation, and let $\widetilde{\boldsymbol{\tau}}$ be random with the distribution (7.5). Then,*

$$\mu_\sigma := \mathbb{E}[\operatorname{occ}_\sigma(\widetilde{\boldsymbol{\tau}})] = \begin{cases} \mathbb{E}\,|\widetilde{\boldsymbol{\tau}}| = \phi + 2, & |\sigma| = 1, \\ p^{|\sigma|}(1-p)^{-|\sigma|}\bigl(\frac{1-p}{1-2p}\bigr)^2 = p^{|\sigma|/2-2} = \phi^{4-|\sigma|}, & |\sigma| \geqslant 2. \end{cases} \tag{8.1}$$

$$\mathbb{E}[\operatorname{occ}_\sigma(\widetilde{\boldsymbol{\tau}})^k] < \infty, \qquad \forall k \geqslant 1. \tag{8.2}$$

Proof. First, if $|\sigma| = 1$, i.e., $\sigma = 1$, then trivially $\operatorname{occ}_\sigma(\tau) = |\tau|$ for any permutation τ, and thus this case of (8.1) follows from Lemma 7.1.

Assume now $|\sigma| \geqslant 2$, and let $a_{n;\sigma}$ and $A_\sigma(z)$ be as in Lemma 6.4. Then,

$$\sum_{\tau \in \mathfrak{T}_n} \operatorname{occ}_\sigma(\tau) = a_{n;\sigma}, \tag{8.3}$$

and thus it follows from (7.5) that

$$\mathbb{E}\operatorname{occ}_\sigma(\widetilde{\boldsymbol{\tau}}) = \sum_{\tau \in \mathfrak{T}_*} \operatorname{occ}_\sigma(\tau) p^{|\tau|} = \sum_{n \geqslant |\sigma|} p^n \sum_{\tau \in \mathfrak{T}_n} \operatorname{occ}_\sigma(\tau) = \sum_{n \geqslant |\sigma|} p^n a_{n;\sigma} = A_\sigma(p). \tag{8.4}$$

Consequently, (8.1) follows from (8.4) och (6.4), using (7.4).

Finally, (8.2) follows from (7.10), since $\operatorname{occ}_\sigma(\tau) \leqslant |\tau|$ for any σ. ☐

Example 8.2. The only tree permutation σ with $|\sigma| = 2$ is 21, and $\operatorname{occ}_{21}(\tau)$ counts the number of inversions in τ, i.e., the number of edges in G_τ. If τ is a tree permutation, we thus have $\operatorname{occ}_{21}(\tau) = |\tau| - 1$. Indeed, Lemma 8.1 yields $\mathbb{E}\operatorname{occ}_{21}(\widetilde{\boldsymbol{\tau}}) = \phi^2$, which equals $\mathbb{E}\bigl[|\widetilde{\boldsymbol{\tau}}| - 1\bigr] = \mathbb{E}\,|\widetilde{\boldsymbol{\tau}}| - 1 = \phi + 1$ given by Lemma 7.1. ☐

9. Patterns in a random forest permutation

We are now prepared to prove Theorems 3.5 and 3.6 on patterns in π_n.

Proof of Theorem 3.5. Let $\pi \in \mathfrak{S}_n$ have block decomposition $\pi = \pi_1 \oplus \cdots \oplus \pi_N$. If $\sigma = \sigma_1 \oplus \cdots \oplus \sigma_d$ occurs as a pattern in π, then each block σ_j is mapped into some block π_{i_j}, but it is possible that several blocks of σ fit in the same block of π. Let $\operatorname{occ}'_\sigma(\pi)$ be the number of occurrences of σ such that the blocks are mapped to different blocks in π, i.e., where the function $j \mapsto i_j$ is injective, and let $\operatorname{occ}''_\sigma(\pi)$ denote the number of the remaining occurrences.

Let us first consider $\operatorname{occ}'_\sigma$, which will be the main term. We have

$$\operatorname{occ}'_\sigma(\pi) = \sum_{1\leqslant i_1<\cdots<i_d\leqslant N} \prod_{j=1}^d \operatorname{occ}_{\sigma_j}(\pi_{i_j}). \tag{9.1}$$

Thus, by (7.19),

$$\operatorname{occ}'_\sigma(\boldsymbol{\pi}_n) \overset{\mathrm{d}}{=} \Bigl(\sum_{1\leqslant i_1<\cdots<i_d\leqslant N(n)} \prod_{j=1}^d \operatorname{occ}_{\sigma_j}(\widetilde{\boldsymbol{\tau}}_{i_j}) \Bigm| S_{N(n)} = n\Bigr). \tag{9.2}$$

This is a conditioned U-statistic of the type in Proposition 5.4, based on the i.i.d. sequence $X_i := \widetilde{\boldsymbol{\tau}}_i$, with $\mathcal{S} = \mathfrak{S}_*$, the (discrete) space of all permutations, and $h(\tau) := |\tau|$; more precisely, we then have $\operatorname{occ}'_\sigma(\boldsymbol{\pi}_n) \overset{\mathrm{d}}{=} \bigl(U_{N(n)}(f) \mid S_{N(n)} = n\bigr)$ with

$$f\bigl(\tau_1,\dots,\tau_d\bigr) := \prod_{j=1}^d \operatorname{occ}_{\sigma_j}(\tau_j). \tag{9.3}$$

Note that (8.2) and Hölder's inequality imply that $\mathbb{E}\bigl[\bigl|f(\widetilde{\boldsymbol{\tau}}_1,\dots,\widetilde{\boldsymbol{\tau}}_d)\bigr|^p\bigr] < \infty$ for every $p<\infty$. Similarly, $\mathbb{E}\bigl[h(\widetilde{\boldsymbol{\tau}}_1)^p\bigr] < \infty$ by (7.10).

It follows from Proposition 5.4 that (3.4) holds for $\operatorname{occ}'_\sigma$, with some $\widetilde{\mu}_\sigma$ and γ_σ^2; note that in the notation of Section 5, by Lemma 7.1,

$$\nu := \mathbb{E}\, h(\widetilde{\boldsymbol{\tau}}) = \mathbb{E}\,|\widetilde{\boldsymbol{\tau}}| = \phi + 2, \tag{9.4}$$

and by (5.2), (9.3), the independence of $\widetilde{\boldsymbol{\tau}}_i$, and (8.1) in Lemma 8.1,

$$\mu = \mu_\sigma := \prod_{j=1}^d \mathbb{E}\bigl[\operatorname{occ}_{\sigma_j}(\widetilde{\boldsymbol{\tau}}_j)\bigr] = \prod_{j=1}^d \mu_{\sigma_j} = (\phi+2)^\lambda \phi^{4(d-\lambda)-(|\sigma|-\lambda)}. \tag{9.5}$$

Thus, by (5.10), $\widetilde{\mu}_\sigma$ in (3.4) (so far for $\operatorname{occ}'_\sigma$) is given by

$$\widetilde{\mu}_\sigma = \frac{\mu_\sigma}{\nu^d d!} = \frac{\mu_\sigma}{(\phi+2)^d d!} = \frac{1}{d!}(\phi+2)^{\lambda-d}\phi^{4d-3\lambda-|\sigma|}, \tag{9.6}$$

which yields (3.3).

Similarly, by (5.3),

$$f_i(\tau) = \operatorname{occ}_{\sigma_i}(\tau)\prod_{j\neq i}\mathbb{E}\operatorname{occ}_{\sigma_j}(\widetilde{\boldsymbol{\tau}}_j) = \prod_{j\neq i}\mu_{\sigma_j}\cdot\operatorname{occ}_{\sigma_i}(\tau) = \frac{\mu_\sigma}{\mu_{\sigma_i}}\operatorname{occ}_{\sigma_i}(\tau). \tag{9.7}$$

Suppose that $|\sigma_i| > 1$. We may have, with positive probabilities,

(1) $|\widetilde{\boldsymbol{\tau}}| = 1$, and then $\operatorname{occ}_{\sigma_i}(\widetilde{\boldsymbol{\tau}}) = 0$,
(2) $\widetilde{\boldsymbol{\tau}} = \sigma_i$, and then $\operatorname{occ}_{\sigma_i}(\widetilde{\boldsymbol{\tau}}) = 1 > 0$.

Thus it is impossible to have $f_i(\widetilde{\boldsymbol{\tau}}) = c|\widetilde{\boldsymbol{\tau}}|$ a.s., for any real c. Consequently, Proposition 5.3 yields $\gamma_\sigma^2 > 0$ if any block σ_i with $|\sigma_i| > 1$ exists.

It remains to show that $\operatorname{occ}''_\sigma(\boldsymbol{\pi}_n)$ is negligible. By grouping the blocks of σ that are mapped into the same block of π, we see that $\operatorname{occ}''_\sigma(\pi)$ can be written as a sum over all decompositions $\sigma = \tilde{\sigma}_1 \oplus\cdots\oplus\tilde{\sigma}_k$ with $k<d$, of the number of occurrences with each $\tilde{\sigma}_i$ mapped into a block of π, with these blocks distinct. (Here $\tilde{\sigma}_i$ are necessarily forest permutations.) It follows, using again (7.19), and $N(n)\leqslant n$, that

$$\mathbb{E}\operatorname{occ}''_\sigma(\boldsymbol{\pi}_n) \leqslant \frac{1}{\mathbb{P}(S_{N(n)=n})}\,\mathbb{E}\operatorname{occ}''_\sigma\bigl(\widetilde{\boldsymbol{\tau}}_1\oplus\cdots\oplus\widetilde{\boldsymbol{\tau}}_{N(n)}\bigr) \leqslant C\,\mathbb{E}\operatorname{occ}''_\sigma\bigl(\widetilde{\boldsymbol{\tau}}_1\oplus\cdots\oplus\widetilde{\boldsymbol{\tau}}_n\bigr)$$

$$= C\sum_{k=1}^{d-1}\sum_{\tilde\sigma_1,\dots,\tilde\sigma_k}\sum_{1\leqslant i_1<\dots<i_k\leqslant n}\mathbb{E}\prod_{j=1}^{k}\operatorname{occ}_{\tilde\sigma_j}(\widetilde{\boldsymbol{\tau}}_{i_j}). \tag{9.8}$$

The number of terms in the multiple sum is $O\bigl(n^{d-1}\bigr)$, and each term is $O(1)$, using independence, the trivial $\operatorname{occ}_{\tilde\sigma_j}(\widetilde{\boldsymbol{\tau}})\leqslant|\widetilde{\boldsymbol{\tau}}|^{|\sigma_j|}$, and (7.10). Hence, $\mathbb{E}\operatorname{occ}''_\sigma(\boldsymbol{\pi}_n)=O\bigl(n^{d-1}\bigr)$, and (3.4) follows from the result for $\operatorname{occ}'_\sigma(\boldsymbol{\pi}_n)$.

Moment convergence follows in the same way, using Proposition 5.5 and Minkowski's inequality; we omit the details. □

Proof of Theorem 3.6. We have the trivial identity

$$\sum_{\sigma\in\mathfrak{S}_d}\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)=\binom{n}{d}. \tag{9.9}$$

Furthermore, we only have to consider forest permutations $\sigma\in\mathfrak{F}_d$ in (9.9), since otherwise $\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)=0$.

Let $\sigma\in\mathfrak{F}_d$, and let d' be its number of blocks. If $\sigma\neq\iota_d$, then $d'<d$. If $d'\leqslant d-2$, then (3.4) implies that $\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)/n^{d-3/2}\xrightarrow{\mathrm{p}}0$, so such terms can be ignored.

The remaining terms in (9.9) have $d'=d-1$, and thus 1 block of length 2 and $d-2$ blocks of length 1. There are $d-1$ such permutations; for example, if $d=4$, they are 2134, 1324 and 1243. For each such σ, we have by (3.3)

$$\widetilde{\mu}_\sigma=\frac{1}{(d-1)!}(\phi+2)^{-1}\phi^{4(d-1)-3(d-2)-d}=\frac{1}{(d-1)!}(\phi+2)^{-1}\phi^2, \tag{9.10}$$

where, see (7.7),

$$(\phi+2)^{-1}\phi^2=\frac{\phi^2}{\sqrt5\phi}=\frac{\phi}{\sqrt5}=\frac{5+\sqrt5}{10}. \tag{9.11}$$

Hence, Theorem 3.5 yields

$$\frac{\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)-\frac{5+\sqrt5}{10}(d-1)!^{-1}n^{d-1}}{n^{d-3/2}}\xrightarrow{\mathrm{d}}\mathsf{N}\bigl(0,\gamma_\sigma^2\bigr), \tag{9.12}$$

Moreover, the proof of Theorem 3.5 applies also to the sum $\sum'_\sigma\operatorname{occ}_\sigma$ over these $d-1$ permutations σ. (Consider the sum of the corresponding functions (9.3). See also Remark 3.7.) Thus,

$$\frac{\sum'_\sigma\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)-\frac{5+\sqrt5}{10}(d-2)!^{-1}n^{d-1}}{n^{d-3/2}}\xrightarrow{\mathrm{d}}\mathsf{N}\bigl(0,\gamma^2\bigr), \tag{9.13}$$

where $\gamma^2>0$ by the argument in the proof of Theorem 3.5.

As said above, we may add all $\sigma\in\mathfrak{F}_k$ with less than $d-1$ blocks to the sum in (9.13) without changing the limit. The resulting sum is, by (9.9),

$$\sum_{\sigma\in\mathfrak{F}_d\setminus\{\iota_d\}}\operatorname{occ}_\sigma(\boldsymbol{\pi}_n)=\binom{n}{d}-\operatorname{occ}_{\iota_d}(\boldsymbol{\pi}_n), \tag{9.14}$$

and thus (3.5) follows, with $\gamma_{\iota_d}^2=\gamma^2$ in (9.13),

Moment convergence follows by the same argument. □

Remark 9.1. The asymptotic variance γ_σ^2 can by (5.11) and (5.5) be computed from variances and covariances of the $\operatorname{occ}_{\sigma_i}(\widetilde{\boldsymbol{\tau}})$ and $|\widetilde{\boldsymbol{\tau}}|$. (See also Remark 5.6 when σ is a tree permutation, so $d=1$.) We do not know any general formula, but at least for a specific σ, it should be possible to calculate these using methods similar to those in the proof of Lemmas 8.1 and 6.4. $\square$

Example 9.2. Consider the simplest example $\sigma = 21$, where we count the number of inversions in a random forest permutation $\boldsymbol{\pi}_n$. In this case, σ is indecomposable, so $d = 1$. Furthermore, by (9.3) and Example 8.2,

$$f(\tau) = \operatorname{occ}_{21}(\tau) = |\tau| - 1 = h(\tau) - 1, \tag{9.15}$$

and thus, using also (9.4),

$$\mu_{21} = \mathbb{E} f(\widetilde{\boldsymbol{\tau}}) = \nu - 1 = \phi + 1 = \phi^2, \tag{9.16}$$

in agreement with (9.5). Hence, by (9.6) (or (3.3)) and (9.11),

$$\widetilde{\mu}_{21} = \frac{\mu_{21}}{\nu} = \frac{\phi^2}{\phi + 2} = \frac{5 + \sqrt{5}}{10}. \tag{9.17}$$

Moreover, (5.12) yields, using also (7.9) and (7.7),

$$\gamma_{21}^2 = \frac{1}{\nu} \operatorname{Var}\Bigl[|\widetilde{\boldsymbol{\tau}}| - 1 - \frac{\nu - 1}{\nu}|\widetilde{\boldsymbol{\tau}}|\Bigr] = \nu^{-3} \operatorname{Var}|\widetilde{\boldsymbol{\tau}}| = \frac{3\phi^3}{(\sqrt{5}\phi)^3} = 3 \cdot 5^{-3/2} \doteq 0.268. \tag{9.18}$$

Consequently, Theorem 3.5 yields

$$\frac{\operatorname{occ}_{21}(\boldsymbol{\pi}_n) - \frac{5+\sqrt{5}}{10}n}{n^{1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0, 3 \cdot 5^{-3/2}\bigr). \tag{9.19}$$

This implies also that for the case $d = 2$ of Theorem 3.6, we have $\gamma_{12}^2 = \gamma_{21}^2 = 3 \cdot 5^{-3/2}$. $\square$

Note that $\operatorname{occ}_{21}(\boldsymbol{\pi}_n)$ equals the number of edges in the forest $G_{\boldsymbol{\pi}_n}$, and thus $n - \operatorname{occ}_{21}(\boldsymbol{\pi}_n)$ is the number of components of $G_{\boldsymbol{\pi}_n}$, which equals the number of blocks in $\boldsymbol{\pi}_n$. Hence, Example 9.2 implies a central limit theorem for the number of blocks in a random forest permutation:

Theorem 9.3. *Let $T(\boldsymbol{\pi}_n)$ be number of blocks in a random forest permutation $\boldsymbol{\pi}_n$, i.e., the number of tree permutations in a decomposition* (4.1) *of $\boldsymbol{\pi}_n$. Then*

$$\frac{T(\boldsymbol{\pi}_n) - \frac{5-\sqrt{5}}{10}n}{n^{1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\bigl(0, 3 \cdot 5^{-3/2}\bigr), \tag{9.20}$$

with convergence of all moments. $\square$

10. Random tree permutations from random blocks

In the remaining sections, we study patterns in a random tree permutation $\boldsymbol{\tau}_n$. In analogy with the construction of $\boldsymbol{\pi}_n$ from random tree permutations $\widetilde{\boldsymbol{\tau}}_i$ in Section 7, we may construct the random tree permutation $\boldsymbol{\tau}_n$ with given length from a code with blocks of random lengths. There is only one L-block or R-block of each length, and therefore (cf. (7.5)) we simply let $(L_i)_1^\infty$ and $(R_i)_1^\infty$ be two infinite sequences of random variables, all i.i.d., with the geometric distribution

$$\mathbb{P}(L_i = \ell) = \mathbb{P}(R_i = \ell) = 2^{-\ell}, \qquad \ell \geqslant 1. \tag{10.1}$$

We also define the random vector

$$X_i := (L_i, R_i), \tag{10.2}$$

and, for a vector $x = (\ell, r)$,

$$h(x) := \ell + r. \tag{10.3}$$

We use the notation of Section 5; in particular,

$$S_m := \sum_{i=1}^{m} h(X_i) = \sum_{i=1}^{m} (L_i + R_i). \tag{10.4}$$

For $m \geqslant 1$, let $\boldsymbol{\tau}^*_m$ be the random tree permution that has a code with $2m$ blocks of lengths $L_1, R_1, \dots, L_m, R_m$, and thus (random) length S_m. Then, for every tree permutation τ having a code Ω_τ with $2m$ blocks with lengths $\ell_1, r_1, \dots, \ell_m, r_m$, by independence and (10.1),

$$\begin{aligned} \mathbb{P}\bigl(\boldsymbol{\tau}^*_m = \tau\bigr) &= \mathbb{P}\bigl(L_1 = \ell_1, R_1 = r_1, \dots, L_m = \ell_m, R_m = r_m\bigr) \\ &= \prod_{i=1}^{m} \mathbb{P}(L_1 = \ell_i)\, \mathbb{P}(R_i = r_i) = \prod_{i=1}^{m} 2^{-\ell_i} 2^{-r_i} = 2^{-|\tau|}. \end{aligned} \tag{10.5}$$

It follows as in Section 7.1, cf. (7.18)–(7.19), that if τ is a tree permutation of length n that has $2m$ blocks in its code, then

$$\mathbb{P}\bigl(\boldsymbol{\tau}^*_{N(n)} = \tau\bigr) = \mathbb{P}\bigl(N(n) = m \;\&\; \boldsymbol{\tau}^*_m = \tau\bigr) = \mathbb{P}\bigl(\boldsymbol{\tau}^*_m = \tau\bigr) = 2^{-|\tau|}, \tag{10.6}$$

which is the same for all $\tau \in \mathfrak{T}_n$, and thus

$$\boldsymbol{\tau}_n \overset{\mathrm{d}}{=} \bigl(\boldsymbol{\tau}^*_{N(n)} \mid S_{N(n)} = n\bigr). \tag{10.7}$$

Note that $X_{N(n)}$ does not have the same distribution as X_i for a fixed i, see e.g. [5, Section 2.6]. We will use a simple (coarse) estimate (valid for much more general X_i and $h(X_i)$). Define for convenience $h(X_i) := 0$ for $i \leqslant 0$.

Lemma 10.1. *For any $j \geqslant 0$, $k \geqslant 1$ and $n \geqslant 1$,*

$$\mathbb{P}\bigl[h(X_{N(n)-j}) = k\bigr] \leqslant \bigl(k + j\,\mathbb{E}\,h(X_1)\bigr)\,\mathbb{P}\bigl(h(X_1) = k\bigr). \tag{10.8}$$

Hence, for any $q > 0$,

$$\mathbb{E}\bigl[h(X_{N(n)-j})^q\bigr] \leqslant \mathbb{E}\bigl[h(X_1)^{q+1}\bigr] + j\,\mathbb{E}\,h(X_1)\,\mathbb{E}\bigl[h(X_1)^q\bigr]. \tag{10.9}$$

Proof. Write $Y_i := h(X_i)$ and $Z_i := \sum_{s=1}^{j} Y_{i+s}$. If $Y_{N(n)-j} = k$, then there exists some $m \geqslant 0$ (viz. $N(n)-j-1$) such that $S_m < n$, $Y_{m+1} = k$, and $S_m + Y_{m+1} + Z_{m+1} \geqslant n$. For a given m, S_m, Y_{m+1} and Z_{m+1} are independent, and thus

$$\begin{aligned} \mathbb{P}\bigl(Y_{N(n)-j} = k\bigr) &\leqslant \sum_{m=0}^{\infty} \sum_{i=0}^{n-1} \mathbb{P}\bigl(S_m = i,\, Y_{m+1} = k,\, k + Z_{m+1} \geqslant n - i\bigr) \\ &= \sum_{m=0}^{\infty} \sum_{i=0}^{n-1} \mathbb{P}\bigl(S_m = i\bigr)\,\mathbb{P}\bigl(Y_{m+1} = k\bigr)\,\mathbb{P}\bigl(k + Z_{m+1} \geqslant n - i\bigr) \\ &= \mathbb{P}\bigl(Y_1 = k\bigr) \sum_{i=0}^{n-1} \sum_{m=0}^{\infty} \mathbb{P}\bigl(S_m = i\bigr)\,\mathbb{P}\bigl(k + Z_1 \geqslant n - i\bigr) \\ &\leqslant \mathbb{P}\bigl(Y_1 = k\bigr) \sum_{i=0}^{n-1} \mathbb{P}\bigl(k + Z_1 \geqslant n - i\bigr) \end{aligned}$$

$$\begin{aligned}&\leqslant \mathbb{P}\bigl(Y_1=k\bigr)\sum_{s=1}^{\infty}\mathbb{P}\bigl(k+Z_1\geqslant s\bigr)=\mathbb{P}\bigl(Y_1=k\bigr)\,\mathbb{E}\bigl(k+Z_1\bigr)\\&=(k+j\,\mathbb{E}\,Y_1)\,\mathbb{P}\bigl(Y_1=k\bigr).\end{aligned}\tag{10.10}$$

This proves (10.8). We obtain (10.9) by multiplying (10.8) by k^q and summing over k. ☐

We record a simple fact.

Lemma 10.2. *We have $\mathbb{E}\,L_i=\mathbb{E}\,R_i=2$, and thus*

$$\nu:=\mathbb{E}\,h(X_i)=4.\tag{10.11}$$

Proof. By definition, $L_i\overset{\mathrm{d}}{=}R_i\sim\mathrm{Ge}(1/2)$, and thus, as is well known, $\mathbb{E}\,L_i=\mathbb{E}\,R_i=2$. (See also (11.16) below.) Hence, (10.11) follows. ☐

11. Trees in a given tree permutation

We next express the number of occurences of a pattern σ in a tree permutation using codes and block lengths. We consider here only the case when σ is a tree permutation.

Lemma 11.1. *Let σ be a tree permutation with $|\sigma|\geqslant 3$ having a code with $2b$ blocks of lengths $\ell_1,r_1,\dots,\ell_b,r_b$, and let τ be a tree permutation with $|\tau|\geqslant 3$ having a code with $2m$ blocks of lengths $\ell'_1,r'_1,\dots,\ell'_m,r'_m$. Then*

$$\operatorname{occ}_\sigma(\tau)=\sum_{s=0}^{m-b}\prod_{i=1}^{b}\alpha_{\mathsf{L},i}\bigl(\tilde\ell'_{i+s}\bigr)\alpha_{\mathsf{R},i}\bigl(\tilde r'_{i+s}\bigr)\tag{11.1}$$

where

$$\tilde\ell'_k:=\ell'_k-\mathbf{1}\{k=1,\ell_1>1\},\tag{11.2}$$

$$\tilde r'_k:=r'_k-\mathbf{1}\{k=m,r_b>1\},\tag{11.3}$$

and

$$\alpha_{\mathsf{L},i}(\ell'):=\binom{\ell'-1+\mathbf{1}\{i=1,\ell_1>1\}+\mathbf{1}\{i=b,r_b=1\}}{\ell_i-1+\mathbf{1}\{i=b,r_b=1\}},\tag{11.4}$$

$$\alpha_{\mathsf{R},i}(r'):=\binom{r'-1+\mathbf{1}\{i=b,r_b>1\}+\mathbf{1}\{i=1,\ell_1=1\}}{r_i-1+\mathbf{1}\{i=1,\ell_1=1\}}.\tag{11.5}$$

Proof. The occurrences of σ in τ are described by colourings of Ω_τ that can be obtained as in Lemma 6.3. Consider one such colouring, $\widehat\Omega_\tau$ say. We find some properties of it.

(i): Consider first the red symbols in $\widehat\Omega_\tau$ that correspond to a single block B_j in Ω_σ. These red symbols have the same type (L of R), and there are no other red symbols between them. It follows from Lemma 6.3 that they have to belong to the same block, B'_k say, in τ, except for the first and last blocks B_1 and B_{2b}. If $|B_1|\geqslant 2$, it is also possible that the first L in B_1 corresponds to the last in B'_{k-2}, while all others correspond to red L in B'_k (for some odd $k\geqslant 3$). We have a symmetric situation for the last block B_{2b} if $|B_{2b}|\geqslant 2$. Write $k=k(j)$ for the index of the block B'_k in $\widehat\Omega_\tau$ that corresponds to B_j. (To be precise in all cases, $B'_{k(j)}$ contains the last red L in B_j if j is odd, and the first red R in B_j if j is even.)

(ii): Furthermore, for an L-block B_{2i-1} in Ω_σ, the last L in the corresponding block $B'_{k(2i-1)}$ in $\widehat{\Omega}_\tau$ has to be red, except in the case of the last L-block B_{2b-1} if $|B_{2b}|=1$; in that exceptional case there is no restriction on the red subset of $B'_{k(2b-1)}$ (except it having the size ℓ_b of B_{2b-1}). For an R-block B_{2i} there is a symmetric condition, unless $i=1$ and $|B_1|=1$.

(iii): In all cases, $k=k(j)\equiv j \pmod 2$. Moreover, no completely black blocks can be inserted between the red symbols in two consecutive blocks of Ω_σ. Hence, $k(j+1)=k(j)+1$ for every $j<2b$, and thus there exists $s\in[0,m-b]$ such that $k(j)=j+2s$ for all i.

Conversely, any choice of red symbols satisfying (i)–(iii) for some $s\in[0,m-b]$ gives a colouring of the code Ω_τ that can be constructed as in Lemma 6.3, and thus corresponds to an occurrence of σ in τ.

For each choice of s, the choices of red symbols permitted by (i)–(iii) for an L-block B_{2i-1} is $\alpha_{\mathsf{L},i}(\tilde{\ell}'_{s+i})$; note that for $1<i<b$, this is just $\binom{\ell'_{i+s}-1}{\ell_i-1}$, while for $i=1$ and b there are (possibly) some adjustments that are taken care of by the indicator functions in (11.2) and (11.4). Similarly, the choices of red symbols for an R-block B_{2i} is $\alpha_{\mathsf{R},i}(\tilde{r}'_{s+i})$. Hence, still for a fixed s, the total number of choices of red symbols in Ω_τ is given by the product in (11.1), because the choices for the different blocks $B_1,\dots,B_{2b}$ can be made independently of each other. Consequently, (11.1) holds. □

Remark 11.2. The condition $|\sigma|\geqslant 3$ in Lemma 11.1 excludes the two cases $\sigma=1$ and $\sigma=21$. Recall that both these cases are trivial, with $\mathrm{occ}_1(\tau)=|\tau|$ and $\mathrm{occ}_{21}(\tau)=|\tau|-1$ for any tree permutation τ. (The latter because the number of inversions in τ equals the number of edges in the tree Ω_τ.) Note that 21 has the code LR, so in the notation above, it has $b=1$ and $\ell_1=r_1=1$; however, (11.1) is not valid in this case. □

Recall that b in Lemma 11.1 is denoted $b(\sigma)$, see Section 4, and that we also have defined $b(1):=1$ for the case $\sigma=1$. For any tree permutation σ we define, for $b=b(\sigma)$ vectors $x_j=(\ell'_j,r'_j)$,

$$\overline{f}_\sigma(x_1,\dots,x_b):=\prod_{i=1}^{b}\alpha_{\mathsf{L},i}(\ell'_i)\alpha_{\mathsf{R},i}(r'_i),\qquad \text{if } |\sigma|\geqslant 3, \tag{11.6}$$

with $\alpha_{\mathsf{L},i}$ and $\alpha_{\mathsf{R},i}$ given by (11.4)–(11.5), and

$$\overline{f}_\sigma(x_1):=\ell'_1+r'_1=h(x_1)\qquad \text{if } |\sigma|\leqslant 2. \tag{11.7}$$

(In the exceptional cases 1 and 21 where (11.7) applies, we have $b(\sigma)=1$.)

We compute also some expectations needed later.

Lemma 11.3. *Let σ be as in Lemma 11.1 and let $\alpha_{\mathsf{L},i}$ and $\alpha_{\mathsf{R},i}$ be given by* (11.4)–(11.5)*. Let L_i and R_i have the geometric distribution in* (10.1)*. Then,*

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i)=\bigl(1+\mathbf{1}\{i=1,\ell_1>1\}\bigr)\bigl(1+\mathbf{1}\{i=b,r_b=1\}\bigr), \tag{11.8}$$

$$\mathbb{E}\,\alpha_{\mathsf{R},i}(R_i)=\bigl(1+\mathbf{1}\{i=b,r_b>1\}\bigr)\bigl(1+\mathbf{1}\{i=1,\ell_1=1\}\bigr). \tag{11.9}$$

Proof. In the definition (11.4), there are two special cases: (I) $i=1$ and $\ell_1>1$; (II) $i=b$ and $r_b=1$. Note that both may occur together, if $b=1$; thus there are four possible combinations.

Case 1: Neither (I) nor (II). In this case, (11.4) is simply $\binom{\ell'-1}{\ell_i-1}$, and thus

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = \mathbb{E}\binom{L_i-1}{\ell_i-1}. \tag{11.10}$$

To compute this binomial moment, we note that the probability generating function of $L_i - 1$ is, by (10.1),

$$g_{L-1}(z) := \sum_{\ell=1}^{\infty} z^{\ell-1}2^{-\ell} = \frac{1/2}{1-z/2} = \frac{1}{2-z}, \tag{11.11}$$

and thus,

$$\mathbb{E}\binom{L_i-1}{k} = \frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k} g_{L-1}(1) = 1, \qquad k \geqslant 0. \tag{11.12}$$

(Alternatively, compute $[z^k]g_{L-1}(1+z)$.) Hence, in this case,

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = 1. \tag{11.13}$$

Case 2: (I) but not (II). Then, $\ell_1 \geqslant 2$ and (11.4) yields

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = \mathbb{E}\binom{L_i}{\ell_i-1}. \tag{11.14}$$

The probability generating function of L_i is, by (11.11),

$$g_L(z) = zg_{L-1}(z) = \frac{z}{2-z} = \frac{2}{2-z} - 1, \tag{11.15}$$

and thus,

$$\mathbb{E}\binom{L_i}{k} = \frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k} g_L(1) = 2, \qquad k \geqslant 1. \tag{11.16}$$

(Alternatively, use (11.12) and $\binom{L_i}{k} = \binom{L_i-1}{k} + \binom{L_i-1}{k-1}$.) Hence, (11.14) yields

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = 2. \tag{11.17}$$

Case 3: (II) but not (I). Then, (11.4) yields, using (11.16),

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = \mathbb{E}\binom{L_i}{\ell_i} = 2. \tag{11.18}$$

Case 4: Both (I) and (II). Then, $b = i = 1$, $\ell_1 \geqslant 2$, and (11.4) yields

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = \mathbb{E}\binom{L_i+1}{\ell_i}. \tag{11.19}$$

The probability generating function of $L_i + 1$ is, by (11.15),

$$g_{L+1}(z) = zg_L(z) = \frac{2z}{2-z} - z = \frac{4}{2-z} - 2 - z, \tag{11.20}$$

and thus,

$$\mathbb{E}\binom{L_i+1}{k} = \frac{1}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k} g_L(1) = 4, \qquad k \geqslant 2. \tag{11.21}$$

(Alternatively, use (11.16) and $\binom{L_i+1}{k} = \binom{L_i}{k} + \binom{L_i}{k-1}$.) Hence, by (11.19),

$$\mathbb{E}\,\alpha_{\mathsf{L},i}(L_i) = 4. \tag{11.22}$$

We may summarize the four cases (11.13), (11.17), (11.18) and (11.22) as (11.8). Similarly, by only notational changes, (11.5) yields (11.9). □

Lemma 11.4. *Let σ be any tree permutation, let $b := b(\sigma)$, and let $(X_i)_i$ be the i.i.d. random vectors defined in* (10.2)*. Then*

$$\mathbb{E}\,\overline{f}_\sigma\bigl(X_1,\dots,X_b\bigr) = 4. \tag{11.23}$$

Proof. The case $|\sigma| \leqslant 2$ is immediate by (11.7) and Lemma 10.2.

Assume thus $|\sigma| \geqslant 3$. Then, by (11.6), independence, and Lemma 11.3,

$$\begin{aligned}
\mathbb{E}\,\overline{f}_\sigma(X_1,\dots,X_b) &= \mathbb{E}\prod_{i=1}^b \alpha_{\mathsf{L},i}(L_i)\alpha_{\mathsf{R},i}(R_i) = \prod_{i=1}^b \mathbb{E}\,\alpha_{\mathsf{L},i}(L_i)\prod_{i=1}^b \mathbb{E}\,\alpha_{\mathsf{R},i}(R_i)\\
&= \bigl(1+\mathbf{1}\{\ell_1>1\}\bigr)\bigl(1+\mathbf{1}\{r_b=1\}\bigr)\cdot\bigl(1+\mathbf{1}\{r_b>1\}\bigr)\bigl(1+\mathbf{1}\{\ell_1=1\}\bigr)\\
&= \bigl(1+\mathbf{1}\{\ell_1>1\}\bigr)\bigl(1+\mathbf{1}\{\ell_1=1\}\bigr)\cdot\bigl(1+\mathbf{1}\{r_b=1\}\bigr)\bigl(1+\mathbf{1}\{r_b>1\}\bigr)\\
&= (1+1)(1+1) = 4,
\end{aligned} \tag{11.24}$$

which completes the proof. □

We have no simple explanation for the, perhaps surprising, fact that the expectation (11.23) is the same for every tree permutation σ, cf. Problem 3.4.

12. Patterns in a random tree permutation of given length

We next consider the occurrences of a pattern σ in a random tree permutation $\boldsymbol{\tau}_n$. We use the construction and notation in Sections 10 and 5. In particular, X_n and S_m are defined by (10.1)–(10.4) and $N(n)$ by (5.9).

We first consider the case of a tree permutation σ. Recall $\overline{f}_\sigma$ defined by (11.6)–(11.7).

Lemma 12.1. *Let σ be a tree permutation and let $b := b(\sigma)$. Then*

$$\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) \overset{\mathrm{d}}{=} \Bigl(\sum_{s=0}^{N(n)-b} \overline{f}_\sigma\bigl(X_{s+1},\dots,X_{s+b}\bigr)\Bigm| S_{N(n)} = n\Bigr) + O_{L^*}(1). \tag{12.1}$$

Proof. Assume first $|\sigma| \geqslant 3$, so $\overline{f}_\sigma$ is given by (11.6). Recall $\boldsymbol{\tau}^*_m$ defined in Section 10, and note that $\boldsymbol{\tau}^*_{N(n)}$ is a tree permutation having a code with $2N(n)$ blocks of lengths $L_1,\dots,R_{N(n)}$. Lemma 11.1 thus shows that

$$\begin{aligned}
\operatorname{occ}_\sigma\bigl(\boldsymbol{\tau}^*_{N(n)}\bigr) = \sum_{s=0}^{N(n)-b}\prod_{i=1}^b &\alpha_{\mathsf{L},i}\bigl(L_{i+s}-\mathbf{1}\{s=0,i=1,\ell_1>1\}\bigr)\\
&\cdot\alpha_{\mathsf{R},i}\bigl(R_{i+s}-\mathbf{1}\{s=N(n)-b,i=b,r_b>1\}\bigr).
\end{aligned} \tag{12.2}$$

Except in the extreme cases $s=0$ and $s=N(n)-m$, the product in the sum in (12.2) is

$$\prod_{i=1}^b \alpha_{\mathsf{L},i}\bigl(L_{i+s}\bigr)\alpha_{\mathsf{R},i}\bigl(R_{i+s}\bigr) = \overline{f}_\sigma\bigl(X_{s+1},\dots,X_{s+b}\bigr). \tag{12.3}$$

In the cases $s = 0$ and $s = N(n) - b$, the product might be smaller, but is still $\geqslant 0$. Hence, (12.2) yields

$$\sum_{s=1}^{N(n)-b-1} \overline{f}_\sigma\bigl(X_{s+1},\dots,X_{s+b}\bigr) \leqslant \operatorname{occ}_\sigma\bigl(\boldsymbol{\tau}^*_{N(n)}\bigr) \leqslant \sum_{s=0}^{N(n)-b} \overline{f}_\sigma\bigl(X_{s+1},\dots,X_{s+b}\bigr). \tag{12.4}$$

We claim that

$$\overline{f}_\sigma\bigl(X_1,\dots,X_b\bigr),\ \overline{f}_\sigma\bigl(X_{N(n)-b+1},\dots,X_{N(n)}\bigr) = O_{L^*}(1). \tag{12.5}$$

This implies that the difference of the first and last sums in (12.4) is $O_{L^*}(1)$, and thus

$$\operatorname{occ}_\sigma\bigl(\boldsymbol{\tau}^*_{N(n)}\bigr) = \sum_{s=0}^{N(n)-b} \overline{f}_\sigma\bigl(X_{s+1},\dots,X_{s+b}\bigr) + O_{L^*}(1). \tag{12.6}$$

To show (12.5), note first that (11.6) and (11.4)–(11.5) imply that

$$\overline{f}_\sigma(X_{k+1},\dots,X_{k+b}) \leqslant \prod_{i=1}^{b}(L_{k+i}+R_{k+i})^c = \prod_{i=1}^{b} h(X_{k+i})^c, \tag{12.7}$$

for some $c < \infty$ depending on σ only. Hence, using Hölder's inequality, (12.5) follows if we show that for every $q < \infty$ and every $j \in [1, b]$,

$$\mathbb{E}\, h(X_j)^q = O(1), \qquad\qquad \mathbb{E}\, h(X_{N(n)-b+j})^q = O(1), \tag{12.8}$$

The first part is trivial, since for any fixed j, we have $\mathbb{E}\, h(X_j)^q = \mathbb{E}\, h(X_1)^q < \infty$. The second part follows from Lemma 10.1.

Hence, (12.6) holds, and (12.1) follows by conditioning on $S_{N(n)} = n$, recalling (10.7). Note that the error term $O_{L^*}(1)$ survives this conditioning, because $\mathbb{P}(S_{N(n)} = n) \to 1/\,\mathbb{E}\, h(X_1) > 0$, see e.g. [5, Theorem 2.4.2], and thus for any $q < \infty$,

$$\mathbb{E}\bigl[|O_{L^*}(1)|^q \mid S_{N(n)} = n\bigr] \leqslant \frac{\mathbb{E}\bigl[|O_{L^*}(1)|^q\bigr]}{\mathbb{P}\bigl[S_{N(n)} = n\bigr]} = O(1). \tag{12.9}$$

Finally, if $|\sigma| \leqslant 2$, then $b = 1$ and

$$\sum_{s=0}^{N(n)-b} \overline{f}_\sigma(X_{s+1}) = \sum_{s=0}^{N(n)-1} h(X_{s+1}) = S_{N(n)}. \tag{12.10}$$

Furthermore, $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) = n$ or $n-1$, and thus (12.1) is trivial. ☐

The sum in (12.1) is a constrained U-statistic of the type in (5.17), with $d = 1$ and $b_1 = b(\sigma)$. We extend Lemma 12.1 to forest permutations σ.

Lemma 12.2. *Let σ be a forest permutation with block decomposition $\sigma = \sigma_1 \oplus \cdots \oplus \sigma_d$. Let $b_j := b(\sigma_j)$, and define*

$$\overline{f}_\sigma\bigl((x_{1i})_{i=1}^{b_1},\dots,(x_{di})_{i=1}^{b_d}\bigr) := \prod_{j=1}^{d} \overline{f}_{\sigma_j}\bigl(x_{j1},\dots,x_{jb_j}\bigr). \tag{12.11}$$

Then

$$\operatorname{occ}_\sigma(\boldsymbol{\tau}_n) \overset{\mathrm{d}}{=} \Bigl(\widehat{U}_{N(n)}(\overline{f}_\sigma) \Bigm| S_{N(n)} = n\Bigr) + O_{L^*}(n^{d-1}). \tag{12.12}$$

Proof. Recall again $\boldsymbol{\tau}^*_m$ from Section 10, and consider first $\operatorname{occ}_\sigma(\boldsymbol{\tau}^*_m)$, for some given m. By definition, $\boldsymbol{\tau}^*_m$ has $2m$ blocks, which we denote by $B'_1, \dots, B'_{2m}$.

As before, we mark an occurrence of σ in $\tau = \boldsymbol{\tau}^*_m$ by colouring the corresponding symbols in the code Ω_τ red (and the remaining ones black). Then each σ_j corresponds to a set of red symbols, A_j say; these sets A_j are subsets of $\{1, \dots, |\boldsymbol{\tau}_m|\}$.

As in (5.13), let $b'_j := b_j - 1$. For each σ_j with $|\sigma_j| \geqslant 3$, the red symbols A_j are as in the proof of Lemma 11.1, and they lie in some blocks $B'_{2i_j-1}, \dots, B'_{2(i_j+b'_j)}$, possibly also with a red symbol in B'_{2i_j-3} or $B'_{2(i_j+b'_j)+2}$.

If $|\sigma_j| = 2$, so $\sigma_j = 21$, then the red symbols in A_j are an L and an R forming an edge, and thus described by Lemma 4.1(e1)–(e3); we then define i_j so that the L belong to B'_{2i_j-1} (and thus the R to B'_{2i_j} or B'_{2i_j+2}).

Finally, if $\sigma_j = 1$, A_j is a single red symbol, which can be either L or R; we define i_j such that this symbol belongs to B'_{2i_j-1} or B'_{2i_j}.

The sets A_j follow each other in order, and thus we must have $1 \leqslant i_1 \leqslant i_2 \leqslant \dots i_d \leqslant m$. (Equality is possible, e.g. if $|\sigma_j| = 1$.) Moreover, for a given sequence $i_1, \dots, i_d$, if all gaps $i_{j+1} - i_j \geqslant 3$, then the sets A_j can be chosen independently, without interfering with each other (by colliding, having symbols in wrong order, or causing edges between two of them). If furthermore $i_1 > 1$ and $i_d + b'_d < m$, the number of choices for each σ_j with $|\sigma_j| \geqslant 3$ is $\overline{f}_{\sigma_j}\bigl(X_{i_j}, \dots, X_{i_j+b'_j}\bigr)$ by the proof of Lemma 11.1. The same holds for $|\sigma_j| \leqslant 2$ by the definition (11.7): if $\sigma_j = 1$, then A_j is one of the $L_{i_j} + R_{i_j}$ symbols in $B'_{2i_j-1} \cup B'_{2i_j}$; if $\sigma = 21$, then A_j consists of an L in B'_{2i_j-1} and an R in B'_{2i_j} or B'_{2i_j+2} chosen according to one of (e1)–(e3) in Lemma 4.1, and this too gives $L_{i_j} + R_{i_j}$ choices. (Note that (e1) and (e2) overlap in one possibility.) Hence, for such $i_1, \dots, i_d$ the number of possible choices of $A_1, \dots, A_d$ is

$$\prod_{j=1}^{d} \overline{f}_{\sigma_j}\bigl(X_{i_j}, \dots, X_{i_j+b'_j}\bigr) = \overline{f}_\sigma\bigl((X_i)_{i=i_1}^{i_1+b'_1}, \dots, (X_i)_{i=i_d}^{i_d+b'_d}\bigr). \tag{12.13}$$

If some gap $i_{j+1} - i_j \leqslant 2$, the number of possibilities may be smaller, but we may conclude that, recalling the definition (5.17),

$$\bigl|\operatorname{occ}_\sigma\bigl(\boldsymbol{\tau}^*_m\bigr) - \widehat{U}_m(\overline{f}_\sigma)\bigr| \leqslant \sum\nolimits^* \overline{f}_\sigma\bigl((X_i)_{i=i_1}^{i_1+b'_1}, \dots, (X_i)_{i=i_d}^{i_d+b'_d}\bigr), \tag{12.14}$$

where $\sum^*$ denotes the sum over $i_1, \dots, i_d \in [1, m - b'_d]$ such that either $i_1 = 1$, $i_d = m - b'_d$, or $i_j \leqslant i_{j+1} \leqslant i_j + 2$ for some j.

We now take $m = N(n)$, condition on $S_{N(n)} = n$ and use (10.7). It remains only to show that the sum in (12.14) (with $m = N(n)$) is $O_{L^*}(n^{d-1})$; this then survives the conditioning as in (12.9). To see this, consider first the terms with $i_1 = 1$ or $i_j \leqslant i_{j+1} \leqslant i_j + 2$ for some i. Since $m = N(n) \leqslant n$, we may extend the sum to all $i_1, \dots, i_d \in [1, n]$ satisfying one of these conditions. This is a sum of $O(n^{d-1})$ terms, and each term is $O_{L^*}(1)$ by (12.11), (12.7)–(12.8) and Hölder's inequality. Hence the sum of these terms is $O_{L^*}(n^{d-1})$ by Minkowski's inequality.

The remaining sum consists of terms with $i_d = m - b'_d = N(n) - b'_d$, and is thus

$$\leqslant \overline{f}_{\sigma_d}\bigl(X_{N(n)-b'_d}, \dots, X_{N(n)}\bigr) \sum_{i_1,\dots,i_{d-1}=1}^{n} \prod_{j=1}^{d-1} \overline{f}_{\sigma_j}\bigl(X_{i_j}, \dots, X_{i_j+b'_j}\bigr). \tag{12.15}$$

The first factor is $O_{L^*}(1)$ as shown in (12.7)–(12.8), and the sum is again a sum of $O(n^{d-1})$ terms that are $O_{L^*}(1)$, and thus this sum is $O_{L^*}(n^{d-1})$ by Minkowski's inequality. Hence, (12.15) is $O_{L^*}(n^{d-1})$ by Hölder's inequality, which completes the proof. □

Proof of Theorem 3.1. Lemma 12.2 and Proposition 5.11 show that

$$\frac{\mathrm{occ}_\sigma(\boldsymbol{\tau}_n) - \mu\nu^{-d}d!^{-1}n^d}{n^{d-1/2}} \xrightarrow{\mathrm{d}} \mathsf{N}\big(0,\gamma^2\big), \tag{12.16}$$

with convergence of all moments by Proposition 5.12; note that $\mathbb{E}\,|\overline{f}(X_1,\dots,X_D)|^p < \infty$ and $\mathbb{E}\,h(X_1)^p < \infty$ for every $p < \infty$ by (12.11), (12.5), and (12.8). Furthermore, (12.11) and Lemma 11.4 imply that

$$\mu = \mathbb{E}\,\overline{f}(X_1,\dots,X_D) = \prod_{j=1}^{d} \mathbb{E}\,\overline{f}_{\sigma_j}(X_1,\dots,X_{b_j}) = 4^d, \tag{12.17}$$

while $\nu = 4$ by Lemma 10.2. Hence, $\mu\nu^{-d} = 1$, and (3.1) follows from (12.16).

To see that $\gamma^2 > 0$ if some $|\sigma_j| \geqslant 3$, we use the criterion in Proposition 5.9. By (5.22) and (12.11),

$$\overline{f}_{\sigma,j}(x_1,\dots,x_{b_j}) = \overline{f}_{\sigma_j}(x_1,\dots,x_{b_j}) \prod_{i\neq j} \mathbb{E}\,\overline{f}_{\sigma_i}\big(X_1,\dots,X_{b_i}\big) = c\overline{f}_{\sigma_j}(x_1,\dots,x_{b_j}) \tag{12.18}$$

for some constant $c > 0$. Now suppose that $|\sigma_j| \geqslant 3$. Then, (12.18), (11.6) and (11.4)–(11.5) show that, with $x_i = (\ell_i', r_i')$, $\overline{f}_{\sigma,j}(x_1,\dots,x_{b_j})$ is a polynomial in $\{\ell_i', r_i'\}$ of total degree

$$\delta_j := \sum_{i=1}^{b_j} \big(\ell_i - 1 + r_i - 1\big) + \mathbf{1}\{r_{b_j} = 1\} + \mathbf{1}\{\ell_1 = 1\}. \tag{12.19}$$

We see also that the polynomial has only one term with this degree, and that this term has a positive coefficient. Note further that

$$\delta_j \geqslant \big(\ell_1 - 1 + \mathbf{1}\{\ell_1 = 1\}\big) + \big(r_{b_j} - 1 + \mathbf{1}\{r_{b_j} = 1\}\big) \geqslant 2. \tag{12.20}$$

In particular, if we take $x_1 = \dots = x_{b_j} = (s,s)$, then $\overline{f}_{\sigma,j}(x_1,\dots,x_{b_j})$ is a polynomial in s of degree $\delta_j \geqslant 2$. Hence, if we fix any $n > 2b_j$, and consider the event (which has positive probability)

$$X_i = (L_i, R_i) = \begin{cases} (s,s), & b_j < i \leqslant 2b_j, \\ (1,1), & \text{otherwise} \end{cases} \tag{12.21}$$

for an integer $s \geqslant 1$, we see that $S_n(\overline{f}_{\sigma,j})$ defined in (5.23) is a polynomial in s of degree $\delta_j \geqslant 2$. Furthermore, on the same event, $S_n(h)$ is a polynomial in s of degree 1, and thus, $S_n\big(\overline{f}_{\sigma,j} - \frac{\mu}{\nu}h\big)$ is a non-constant polynomial in s. Consequently, the condition in Proposition 5.9 cannot be satisfied for $\overline{f}_{\sigma,j} - \frac{\mu}{\nu}h$, and thus Proposition 5.10 shows that $\gamma^2 > 0$. □

We compute the asymptotic variance γ^2 only in a simple special case.

Example 12.3. Suppose that σ is a tree permutation with $b(\sigma)=1$; thus its code has only two blocks, of lengths $\ell_1=\ell$ and $r_1=r$. Then, the U-statistic $\widehat{U}_N(\overline{f}_\sigma)$ in (12.12) is simply a partial sum:

$$\widehat{U}_N(\overline{f}_\sigma)=\sum_{i=1}^{N}\overline{f}_\sigma(X_i)=S_N(\overline{f}_\sigma). \tag{12.22}$$

This is the special case $d=1$ of an unconstrained U-statistic discussed in Remark 5.6, and (5.12) yields, since $\mu=\nu=4$ by (12.17) and (10.11),

$$\begin{aligned}\gamma^2&=\frac{1}{4}\operatorname{Var}\bigl[\overline{f}_\sigma(X)-h(X)\bigr]\\&=\frac{1}{4}\Bigl(\operatorname{Var}\bigl[\overline{f}_\sigma(X)\bigr]-2\operatorname{Cov}\bigl[\overline{f}_\sigma(X),h(X)\bigr]+\operatorname{Var}\bigl[h(X)\bigr]\Bigr),\end{aligned} \tag{12.23}$$

where $X=(L,R)$ with independent $L,R\sim\mathrm{Ge}(1/2)$ as in (10.1)–(10.2). We recall that $\mathbb{E}\,L=\mathbb{E}\,R=2$. A simple calculation, for example using (11.16), yields $\operatorname{Var}L=\operatorname{Var}R=2$ and thus $\operatorname{Var}h(X)=\operatorname{Var}(L+R)=4$.

We consider several cases.

Case 1: $\ell=r=1$. This means $\Omega_\sigma=\mathsf{LR}$, and thus $\sigma=21$. As we have seen earlier, this case is trivial and $\operatorname{occ}_\sigma(\boldsymbol{\tau}_n)$ is deterministic. Indeed, we have $\overline{f}_\sigma=h$ and thus (12.23) yields $\gamma^2=0$.

Case 2: $\ell>1$, $r=1$. This means that σ is the permutation $23\cdots(\ell+1)1$.

By (11.6) and (11.4)–(11.5),

$$\overline{f}_\sigma(L,R)=\alpha_{\mathsf{L},1}(L)\alpha_{\mathsf{R},1}(R)=\binom{L+1}{\ell}\binom{R-1}{0}=\binom{L+1}{\ell}. \tag{12.24}$$

We have, using (11.21),

$$\mathbb{E}\left[L\binom{L+1}{\ell}\right]=(\ell-1)\,\mathbb{E}\binom{L+1}{\ell}+(\ell+1)\,\mathbb{E}\binom{L+1}{\ell+1}=8\ell,\qquad \ell\geqslant 2, \tag{12.25}$$

and thus (12.23) yields, using also $\mathbb{E}\,L^2=6$,

$$\begin{aligned}\gamma^2&=\frac{1}{4}\operatorname{Var}\left[\binom{L+1}{\ell}-(L+R)\right]=\frac{1}{4}\left(\operatorname{Var}\left[\binom{L+1}{\ell}-L\right]+2\right)\\&=\frac{1}{4}\left(\mathbb{E}\left[\left(\binom{L+1}{\ell}-L\right)^2\right]-2\right)=\frac{1}{4}\mathbb{E}\left[\binom{L+1}{\ell}^2\right]-4\ell+1.\end{aligned} \tag{12.26}$$

This can easily be evaluated for any $\ell\geqslant 2$, although we do not know a closed formula.

Case 3: $\ell=1$, $r>1$. This means that σ is the permutation $(r+1)1\cdots r\in\mathfrak{T}_{r+1}$. This case is the same as the preceding one, if we exchange $\ell\leftrightarrow r$ and $L\leftrightarrow R$.

Case 4: $\ell>1$, $r>1$. This means that $\sigma=2\cdots\ell(\ell+r)1(\ell+1)\cdots(\ell+r-1)$. By (11.6) and (11.4)–(11.5),

$$\overline{f}_\sigma(L,R)=\alpha_{\mathsf{L},1}(L)\alpha_{\mathsf{R},1}(R)=\binom{L}{\ell-1}\binom{R}{r-1}. \tag{12.27}$$

We have, using (11.16),

$$\mathbb{E}\left[L\binom{L}{\ell-1}\right]=(\ell-1)\,\mathbb{E}\binom{L}{\ell-1}+\ell\,\mathbb{E}\binom{L}{\ell}=4\ell-2,\qquad \ell\geqslant 2, \tag{12.28}$$

and thus (12.23) yields

$$\begin{aligned}\gamma^2 &= \frac{1}{4}\,\mathbb{E}\left[\left(\binom{L}{\ell-1}\binom{R}{r-1} - L - R\right)^2\right]\\ &= \frac{1}{4}\,\mathbb{E}\left[\binom{L}{\ell-1}^2\right]\mathbb{E}\left[\binom{R}{r-1}^2\right] - \frac{1}{2}\,\mathbb{E}\left[\binom{L}{\ell-1}\binom{R}{r-1}(L+R)\right] + \frac{1}{4}\,\mathbb{E}(L+R)^2\\ &= \frac{1}{4}\,\mathbb{E}\left[\binom{L}{\ell-1}^2\right]\mathbb{E}\left[\binom{R}{r-1}^2\right] - 4(\ell+r) + 9. \qquad (12.29)\end{aligned}$$

Again, this is easily evaluated for any $\ell, r \geqslant 2$.

Some numerical values for small ℓ and r are given in Table 1. These values are integers (but they do not seem to correspond to any integer sequence in [13]); we conjecture that $\gamma^2(\ell, r)$ is an integer for all $\ell, r \geqslant 1$, but we have no proof.

Note that $\gamma^2(1,3) \neq \gamma^2(2,2)$, which verifies our claim after Corollary 3.2 that γ_σ^2 can differ for different tree permutations σ, even if they have the same length. □

Problem 12.4. In Example 12.3, is γ^2 an integer for every $\ell, r \geqslant 1$?

Problem 12.5. Is γ_σ^2 an integer for every tree permutation σ? For every forest permutation σ?

$\ell \backslash r$	1	2	3	4	5
1	0	6	52	306	1664
2	6	2	28	174	944
3	52	28	154	800	4150
4	306	174	800	3946	20196
5	1664	944	4150	20196	103010

Table 1. Some numerical values of $\gamma^2 = \gamma^2(\ell, r)$ in Example 12.3.

References

[1] Hüseyin Acan & Paweł Hitczenko. On random trees obtained from permutation graphs. *Discrete Math.* **339** (2016), no. 12, 2871–2883. MR 3533335

[2] Frédérique Bassino, Mathilde Bouvel, Valentin Féray, Lucas Gerin, Mickaël Maazoun, Adeline Pierrot. Universal limits of substitution-closed permutation classes. *J. Eur. Math. Soc.* **22** (2020), no. 11, 3565–3639. MR 4167015

[3] Louis Comtet. *Advanced Combinatorics.* Reidel, Dordrecht, 1974. MR 0460128

[4] Scott Garrabrant & Igor Pak. Permutation patterns are hard to count. *Proceedings of the Twenty-Seventh Annual ACM–SIAM Symposium on Discrete Algorithms (SODA)*, 923–936, ACM, New York, 2016. MR 3478442

[5] Allan Gut. *Stopped Random Walks.* 2nd ed., Springer, New York, 2009. MR 2489436

[6] Wassily Hoeffding. A class of statistics with asymptotically normal distribution. *Ann. Math. Statistics* **19** (1948), 293–325. MR 0026294

[7] Svante Janson. *Gaussian Hilbert Spaces.* Cambridge Univ. Press, Cambridge, UK, 1997. MR 1474726

[8] Svante Janson. On degenerate sums of m-dependent variables. *J. Appl. Probab.* **52** (2015), no. 4, 1146–1155. MR 3439177

[9] Svante Janson. Renewal theory for asymmetric U-statistics. *Electron. J. Probab.* **23** (2018), Paper No. 129, 27 pp. MR 3896866

[10] Svante Janson. Patterns in random permutations avoiding some sets of multiple patterns. *Algorithmica* **82** (2020), no. 3, 616–641. MR 4058419

[11] Svante Janson. Asymptotic normality for m-dependent and constrained U¿-statistics, with applications to pattern matching in random strings and permutations. Preprint, 2021–2022. arXiv:2106.09401v3

[12] Olav Kallenberg. *Foundations of Modern Probability.* 2nd ed., Springer, New York, 2002. MR 1876169

[13] *The On-Line Encyclopedia of Integer Sequences*, N. J. A. Sloane ed. https://oeis.org.

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

Email address: svante.janson@math.uu.se

URL: http://www2.math.uu.se/∼svante/