

Asymptotic normality of pattern counts in conjugacy classes*

Valentin Féray[†]

Mohamed Slim Kammoun[‡]

Abstract

We prove, under mild conditions on fixed points and 2-cycles, the asymptotic normality of vincular pattern counts for a permutation chosen uniformly at random in a conjugacy class. Additionally, we prove that the limiting variance is always non-degenerate for classical pattern counts. The proof uses weighted dependency graphs.

Keywords: permutations; patterns; asymptotic normality; dependency graphs; cumulants.

MSC2020 subject classifications: 60C05; 05A05.

Submitted to EJP on December 19, 2023, final version accepted on March 14, 2024.

Supersedes arXiv:2312.08756.

1 Introduction

1.1 Background

A natural question in the context of random combinatorial structures is the asymptotic behaviour of the number of substructures of a given type. In general, we expect such numbers to be asymptotically normal or Poisson distributed, but proving it might be difficult. For permutations, a good notion of substructures is that of *permutation patterns*. Unsurprisingly, the asymptotic normality of patterns counts (either classical, consecutive or vincular patterns) in *uniform random permutations* has been extensively studied. We refer for example to [2, 4, 10, 12, 13] for such results.

More recently, there has been some interest in extending these results to other models of random permutations, and in particular to uniform random permutations in some conjugacy classes. In this spirit, asymptotic normality results have been obtained for specific patterns: see [7, 16] for descent counts and [8] for the number of peaks (the latter is a sum of vincular pattern counts). These results use minimal assumptions on the conjugacy class under consideration. In another direction, the second author [15] and

*Mohamed Slim Kammoun is supported by ERC Project LDRAM: ERC-2019-ADG Project 884584. Valentin Féray is partially supported by a *Future Leader* grant from the initiative *Lorraine Université d'Excellence* (LUE).

[†]Université de Lorraine, CNRS, IECL, F-54000 Nancy, France. E-mail: valentin.feray@univ-lorraine.fr

[‡]Université de Lyon, CNRS, France. E-mail: slim.kammoun@ens-lyon.fr

independently Hamaker and Rhoades, [9, Theorem 8.8] have both proved asymptotic normality results for general patterns, but with some restrictions on the chosen conjugacy classes. In the latter reference, the problem was raised of proving the asymptotic normality of pattern counts assuming only the convergence of the proportion of fixed points and 2-cycles in our permutations [9, Problem 9.9]. We solve this problem in the present paper; see Theorem 1.2 below. Additionally, we prove that the limiting Gaussian distribution is always non-degenerate for classical pattern counts (Theorem 1.3).

Note: In parallel to this work, the asymptotic normality of *classical* pattern counts in random permutations in conjugacy classes has been obtained by Dubach [3] by a different method. His method only applies to classical patterns (and not to the general case of vincular patterns that we treat here), but can provide stronger results in this case (speed of convergence and large deviation estimates). It can also be used to analyze other types of statistics on random permutations in conjugacy classes, such as longest monotone subsequences, the Robinson–Schensted shape, and the number of records; all of these are out of reach with the method of weighted dependency graphs used in this paper.

1.2 Vincular patterns

We use the notation $[n]$ to represent the set $\{1, 2, \dots, n\}$. Additionally, we utilize the one-line notation for permutations $\sigma = \sigma_1\sigma_2 \dots \sigma_n$.

Definition 1.1. *Let π be a permutation of size k and A a subset of $[k - 1]$. Then an occurrence of the vincular pattern (π, A) in a permutation σ is a tuple (i_1, \dots, i_k) with $i_1 < \dots < i_k$ such that:*

- (i_1, \dots, i_k) is an occurrence of π in σ , i.e.

$$\sigma_{i_{\pi_1-1}} < \dots < \sigma_{i_{\pi_k-1}}.$$

- for all s in A , we have $i_{s+1} = i_s + 1$.

In words, the subsequence $(\sigma_{i_1}, \dots, \sigma_{i_k})$ is in the same relative order as (π_1, \dots, π_k) and A indicates entries of the subsequence which must be consecutive in σ . For example, let us consider the permutation $\sigma = 2173456$ and the triple $(i_1, i_2, i_3) = (2, 3, 7)$ (the corresponding subsequence $(\sigma_{i_1}, \sigma_{i_2}, \sigma_{i_3}) = (1, 7, 6)$ is in blue in σ).

- The triple $(i_1, i_2, i_3) = (2, 3, 7)$ is an occurrence of $(\pi = 132, \emptyset)$ since $\pi = 132$ is the permutation with the same relative order as the subsequence $(\sigma_{i_1}, \sigma_{i_2}, \sigma_{i_3}) = (1, 7, 6)$. Equivalently,

$$\sigma_{i_{\pi_1-1}} = \sigma_{i_1} = 1 < \sigma_{i_{\pi_2-1}} = \sigma_{i_3} = 6 < \sigma_{i_{\pi_3-1}} = \sigma_{i_2} = 7.$$

- The triple $(2, 3, 7)$ is also an occurrence of $(132, \{1\})$. Indeed, it is an occurrence of 132 as explained before, and it satisfies $i_2 = i_1 + 1$ (i.e. the two first elements in the corresponding subsequence $(1, 7, 6)$ are adjacent in σ).
- The triple $(2, 3, 7)$ is however not an occurrence of $(132, \{2\})$ since $i_3 \neq i_2 + 1$; said otherwise, the elements 7 and 6 in the corresponding subsequence $(1, 7, 6)$ are not adjacent in σ .

Vincular patterns have been introduced in [1]. They encompass the notions of classical and consecutive patterns. Indeed, when $A = \emptyset$, there are no constraints on indices being consecutive, and we recover the notion of classical patterns. On the other hand, when $A = [k - 1]$, all indices must be consecutive, and we recover consecutive patterns. Here are some examples of occurrences of vincular patterns in the literature.

- An occurrence of $(21, \emptyset)$ is a classical occurrence of 21, i.e. an inversion.
- An occurrence of $(21, \{1\})$ is a consecutive occurrence of 21, i.e. a descent.
- An occurrence of $(123 \dots k, \emptyset)$ is an increasing subsequence of length k .
- A peak is an occurrence of $(132, \{1, 2\})$ or $(231, \{1, 2\})$. Symmetrically, a valley is an occurrence of $(213, \{1, 2\})$ or $(312, \{1, 2\})$.
- The longest alternating subsequence of a permutation equals, up to an additive term equal to 1 or 2, the number of valleys, plus the number of peaks of the permutation. In other terms, it is almost a linear combination of vincular patterns (in that case, consecutive patterns). This result and some equivalent variants were obtained separately by different authors. We refer for example to Houdré and Restrepo [11] and Romik [17].
- The generating function of the vincular pattern $(312, \{1\})$ (together with other statistics) occurs as partition function of a model of statistical physics called PASEP (partially asymmetric simple exclusion process); see [14].

1.3 Our main result: asymptotic normality of vincular pattern counts for uniform random permutation in conjugacy classes

Throughout this article, we fix a vincular pattern (π, A) such that π is of size k . We let $X^{(\pi, A)}(\sigma)$ be the number of the occurrences of (π, A) in σ and we define m by $m = k - |A|$, which is the number of blocks in (π, A) . For an (integer) partition λ of n , we let \mathcal{C}_λ be the conjugacy class indexed by λ in the symmetric group S_n , and σ_λ be a uniform random permutation in \mathcal{C}_λ . Our main result is the following statement, where $m_i(\lambda^n)$ denotes the number of parts of size i in λ^n .

Theorem 1.2. *For any vincular pattern (π, A) , there exist two polynomial functions f and g such that if $\frac{m_1(\lambda^n)}{n} \rightarrow p_1$ and $\frac{m_2(\lambda^n)}{n} \rightarrow p_2$, then*

$$\frac{X^{(\pi, A)}(\sigma_{\lambda^n}) - \mathbb{E}(X^{(\pi, A)}(\sigma_{\lambda^n}))}{n^{m - \frac{1}{2}}} \xrightarrow[n \rightarrow \infty]{d} \mathcal{N}(0, f(p_1) + p_2g(p_1)). \tag{1.1}$$

Moreover, we have convergence of all moments.

We prove Theorem 1.2 in Section 4. A difficulty, compared to the case of uniform random permutations in the whole symmetric group S_n , is that, when working in a given conjugacy class, occurrences of patterns at disjoint sets of positions are no longer independent events. Our proof thus relies then on a technique known as weighted dependency graphs introduced in [5] in order to control higher cumulants for sums of random dependent variables. The main ingredient is Theorem 3.2 (proved in Section 3), which states the existence of a weighted dependency graph structure for the random permutation σ_{λ^n} . Both the proofs of Theorem 3.2 and of Theorem 1.2, starting from Theorem 3.2, require delicate combinatorial arguments, to bound complicated multi-indexed sums.

We believe that Theorem 3.2 is interesting in itself. It can be used to prove that other classes of permutation statistics are asymptotically normal: for example, it is immediate to adapt the proof of Theorem 1.2 to the number of excedances¹, peaks and/or valleys (and thus to the length of the longest alternating subsequence). Also, the proof is readily adapted to establish joint convergence in Theorem 1.2, i.e. the limiting joint distribution of

$$\left(\frac{X^{(\pi^1, A^1)}(\sigma_{\lambda^n}) - \mathbb{E}(X^{(\pi^1, A^1)}(\sigma_{\lambda^n}))}{n^{m_1 - \frac{1}{2}}}, \dots, \frac{X^{(\pi^r, A^r)}(\sigma_{\lambda^n}) - \mathbb{E}(X^{(\pi^r, A^r)}(\sigma_{\lambda^n}))}{n^{m_r - \frac{1}{2}}} \right)$$

is a multivariate Gaussian distribution for any family of patterns.

¹An excedance is a position i such that $\sigma(i) > i$.

1.4 Non-degeneracy of the limiting Gaussian law for classical patterns

A natural problem following Theorem 1.2 is to determine whether the limiting law is degenerate or not, i.e. whether the limiting variance $f(p_1) + p_2g(p_1)$ is positive or equal to 0. In the case of a uniform permutation in S_n , which corresponds² to $p_1 = p_2 = 0$, it is known that the limiting variance, $f(0)$ in this case, is always positive; see [10, 13]. Since f and g are polynomial functions, this implies that $f(p_1) + p_2g(p_1)$ can only vanish when (p_1, p_2) lives in a subvariety of codimension 1 of $\{(p_1, p_2) \in \mathbb{R}_+^2 : p_1 + 2p_2 \leq 1\}$. In other terms, we know that the limiting variance is generically positive. It would nevertheless be desirable to have a complete characterization of the degeneracy cases.

We could solve the question in the case of classical patterns, i.e. the case $A = \emptyset$, for which we prove that the limiting Gaussian law is always non-degenerate (except in the trivial case $p_1 = 1$, where most points of the permutation are fixed). For simplicity, we write π instead of (π, \emptyset) . Our second main result is the following.

Theorem 1.3. *For any **classical** pattern π and any $(p_1, p_2) \in \mathbb{R}_+^2$ with $p_1 + 2p_2 \leq 1$, except $(p_1, p_2) = (1, 0)$, we have $f(p_1) + p_2g(p_1) > 0$.*

Note that in the case $p_1 = 1, p_2 = 0$, we can choose $\lambda^n = (1^n)$, i.e. σ_{λ^n} is a.s. the identity permutation. Thus, in this case, the limiting variance $f(1)$ is indeed 0.

The proof of this non-degeneracy statement is rather involved. It uses a method introduced in [10]; see also [6]. The basic idea is to use a recursive construction of the random objects, say σ_n , under consideration. Typically, such a recursive construction uses a random object of smaller size, say σ_{n-1} , and an extra source of randomness, say I . Then one can compute the variance of the quantity of interest, $X(\sigma_n)$, by conditioning on I and using the law of total variance. If the recursive construction is well-chosen, then some terms in the law of total variance involve the variance of $X(\sigma_{n-1})$, i.e., one can compute, or at least bound from below, the variance recursively. This is the basic principle, the details being quite subtle; see Section 5.

2 Preliminaries: weighted dependency graphs

In this section, we present the notion of weighted dependency graphs, which is the main tool in the proof of Theorem 1.2. The results presented in this section are taken from [5].

2.1 Definition

We start by recalling the notion of mixed cumulants. The mixed cumulant of a family of random variables X_1, X_2, \dots, X_r defined on the same probability space and having finite moments is defined as

$$\kappa(X_1, X_2, \dots, X_r) := [t_1 \cdots t_r] \log \mathbb{E}[\exp(t_1 X_1 + \cdots + t_r X_r)]. \quad (2.1)$$

If $X_i = X$ for all i , we abbreviate $\kappa(X_1, X_2, \dots, X_r)$ as $\kappa_r(X)$. This is the standard cumulant of a random variable.

An important property of cumulants is the following: if $\{X_1, X_2, \dots, X_r\}$ can be written as a disjoint union of two mutually independent sets of random variables, then $\kappa(X_1, X_2, \dots, X_r) = 0$. Hence, cumulants can be seen as a kind of measure of dependency, and the definition of weighted dependency graphs is based on these heuristics. Another important and elementary property of cumulants is that X follows the normal distribution if and only if $\kappa_r(X) = 0$ for every $r \geq 3$.

²Uniform random permutations in the whole symmetric group S_n are of course not a special case of random permutations in conjugacy classes. However, it follows easily from the results of [9, Section 8] that the limiting variance for uniform random permutations is the same as for any conjugacy class with proportions of fixed points and 2 cycles tending to 0.

In the sequel, a weighted graph is a graph with weights on its edges, belonging to $(0, 1]$. Non-edges can be interpreted as edges of weight 0, so that a weighted graph can be equivalently seen as an assignment of weights in $[0, 1]$ to the edges of the complete graph. All our definitions are compatible with this convention.

For a weighted graph H , we define $\mathcal{M}(H)$ to be the *maximal weight of a spanning tree* of H , the weight of a spanning tree being the product of the weights of its edges (if H is disconnected, there is no spanning tree and as a consequence of the above convention, we have $\mathcal{M}(H) = 0$). For example, for $\varepsilon < 1$, the weighted graph H of Fig. 1 satisfies $\mathcal{M}(H) = \varepsilon^4$. The following definition was introduced in [5].

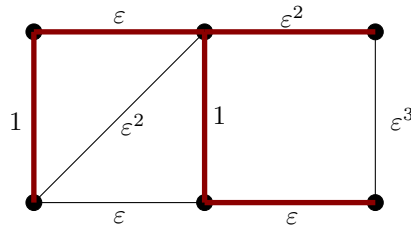


Figure 1: Example of a weighted graph H with a marked spanning tree of maximum weight in red.

Definition 2.1. Let V be a finite set and let $\{Y_v, v \in V\}$ be a family of random variables with finite moments, defined on the same probability space. We fix a sequence $\mathbf{C} = (C_1, C_2, \dots)$ of positive real numbers and a real-valued function Ψ defined on multisets of elements of V .

A weighted graph L with vertex set V is a (Ψ, \mathbf{C}) weighted dependency graph for $\{Y_v, v \in V\}$ if, for any multiset $U = \{v_1, \dots, v_r\}$ of elements of V , one has

$$\left| \kappa(Y_v; v \in U) \right| \leq C_r \Psi(U) \mathcal{M}(L[U]). \tag{2.2}$$

Here, $L[U]$ is the weighted graph with vertex set U and the same weights as L . In examples of weighted dependency graphs, Ψ and \mathbf{C} are simple or universal quantities, so that the meaningful term is $\mathcal{M}(L[U])$. Note that the smaller the weights on edges are, the smaller $\mathcal{M}(L[U])$ is, i.e. the closer to independence are the corresponding random variables $\{Y_v, v \in U\}$. Hence, the edge weights in a weighted dependency graph quantify in some sense the dependency between the random variables $\{Y_v, v \in V\}$.

2.2 A useful example: uniform random permutations in S_n

Let $n \geq 1$ be an integer, and τ be a uniform random permutation in the symmetric group S_n . For (i, j) in $[n]^2$, we let $A_{i,j} = \mathbf{1}[\tau_i = j]$. Note that variables $A_{i,j}$ and $A_{k,\ell}$ cannot be both equal to 1 if $(i, j) \neq (k, \ell)$ but either $i = k$ or $j = \ell$. If $i \neq k$ and $j \neq \ell$, we have

$$\mathbb{P}[A_{i,j} A_{k,\ell}] = \frac{1}{n(n-1)},$$

which is close but different from $\mathbb{P}[A_{i,j}] \cdot \mathbb{P}[A_{k,\ell}] = 1/n^2$. Hence, variables $A_{i,j}$ and $A_{k,\ell}$ are weakly dependent. It turns out that this family of random variables admits a nice weighted dependency graph. This was shown in [5, Proposition 8.1], which we copy here. (We use the symbol $\#$ for the number of distinct elements in a set or multiset.)

Proposition 2.2. Consider the weighted graph L on vertex set $[n]^2$ defined as follows:

- if two pairs $v_1 = (i_1, j_1)$ and $v_2 = (i_2, j_2)$ satisfy either $i_1 = i_2$ or $j_1 = j_2$, then they are linked in L by an edge of weight 1.
- otherwise, they are linked in L by an edge of weight $1/n$.

Then L is a (Ψ, \mathbf{C}) weighted dependency graph, for the family $\{Y_{i,j}, (i, j) \in [n]^2\}$, where

- $\Psi(U) = n^{-\#(U)}$ for any multiset U of elements of $[n]^2$;
- $\mathbf{C} = (C_r)_{r \geq 1}$ is a sequence that does not depend on n .

2.3 Cumulant bounds

Weighted dependency graphs are useful to bound cumulants and prove asymptotic normality results. In particular, in this paper, we will use the following lemma [5, Lemma 4.10]

Lemma 2.3. *Let L be a $(\Psi, \mathbf{C} = (C_r)_{r \geq 1})$ weighted dependency graph for a family of variables $\{Y_v, v \in V\}$. Define R and T_ℓ (for $\ell \geq 1$) as follows:*

$$R = \sum_{v \in V} \Psi(\{v\}); \tag{2.3}$$

$$T_\ell = \max_{v_1, \dots, v_\ell \in V} \left[\sum_{u \in V} \left(\max_{i \in [\ell]} w(u, v_i) \right) \cdot \frac{\Psi(\{v_1, \dots, v_\ell, u\})}{\Psi(\{v_1, \dots, v_\ell\})} \right], \tag{2.4}$$

where $w(u, v)$ is the weight of the edge (u, v) in L . Then, for $r \geq 1$, we have

$$\left| \kappa_r \left(\sum_{v \in V} Y_v \right) \right| \leq C_r r! R T_1 \cdots T_{r-1}.$$

2.4 Power of weighted dependency graphs

An important property of weighted dependency graphs is the following stability property. Consider a family of random variables $\{Y_v, v \in V\}$ with a (Ψ, \mathbf{C}) weighted dependency graph L and fix some integer $d \geq 1$. We are interested in monomials $\mathbf{Y}_I := \prod_{v \in I} Y_v$ of degree at most d , i.e. where $I \in V^d$ is a list of elements of V of length d (possibly with repetitions). This defines a new family of random variables $\{\mathbf{Y}_I, I \in V^d\}$. It turns out that this family admits a natural weighted dependency graph inherited from that of $\{Y_v, v \in V\}$.

To state this formally, we need to introduce some more notation. Let L be a weighted graph with vertex set V . The d -th power L^d of L is a weighted graph with vertex set V^d . If I and J are in V^d , the weight between vertices I and J is

$$W(I, J) := \max_{u \in I, v \in J} w(u, v),$$

where $w(u, v)$ is the weight between u and v in L .

Besides, we say that a function Ψ defined on multisets of elements of a set V is super-multiplicative if $\Psi(U_1 \uplus U_2) \geq \Psi(U_1)\Psi(U_2)$ for all multisets U_1 and U_2 (here and throughout the paper, \uplus is the *disjoint union* of multisets, i.e. if an element belongs to both U_1 and U_2 with multiplicity m_1 and m_2 , it belongs to $U_1 \uplus U_2$ with multiplicity $m_1 + m_2$). Finally, a function Ψ defined on lists of V induces a function on lists of multisets of elements of V by setting

$$\Psi((I_1, \dots, I_r)) = \Psi(I_1 \uplus \dots \uplus I_r). \tag{2.5}$$

The following proposition was proved in [5, Proposition 5.11].

Proposition 2.4. Consider random variables $\{Y_v, v \in V\}$ with a (Ψ, \mathbf{C}) weighted dependency graph L and assume that Ψ is super-multiplicative.

We fix some integer $d \geq 1$. Then L^d is a (Ψ, \mathbf{D}_d) weighted dependency graph for the family $\{Y_I, I \in V^d\}$, where the constants $\mathbf{D}_d = (D_{d,r})_{r \geq 1}$ depend only on d, r and \mathbf{C} .

3 A weighted dependency graph structure for uniform random permutation in conjugacy classes

Fix an integer partition λ^n of size n , for (i, j) in $[n]^2$, we introduce the random variables

$$B_{i,j} = \mathbf{1}[\sigma_{\lambda^n}(i) = j].$$

Note that the law of $B_{i,j}$ depends on λ^n . Given a subset $\alpha = \{(i_t, j_t), t \leq T\}$ of $[n]^2$, we write $S(\alpha) = \{i_1, \dots, i_T, j_1, \dots, j_T\}$ for the support of α . Furthermore, we denote by $cc(V, E)$ the number of connected components of the graph (V, E) . We also write $CC(\alpha)$ instead of $cc(S(\alpha), \alpha)$; this is the number of components of the graph with edge set α and no isolated vertices. We now let

$$\Psi(\alpha) = n^{CC(\alpha) - \#\alpha}. \tag{3.1}$$

We define similarly Ψ on lists of elements of $[n]^2$, and it is insensitive to the order and repetitions of arguments.

The choice of Ψ may seem surprising, but in fact, it is quite natural. Indeed, the joint moments $M_\alpha := \mathbb{E}(\prod_{(i,j) \in \alpha} B_{i,j})$ are nonzero if and only if $(S(\alpha), \alpha)$ is a disjoint union of directed paths and cycles. In this case, M_α is $O(\Psi(\alpha))$, and this is optimal (in the sense that $M_\alpha = \Theta(\Psi(\alpha))$ if $m_k(\lambda) = \Theta(n)$ for any fixed k). Furthermore, as proved in the next lemma, Ψ has been chosen to be super-multiplicative, so that we can use Proposition 2.4.

Lemma 3.1. The function Ψ defined by (3.1) is super-multiplicative.

Proof. Let $\alpha = \{(i_t, j_t), t \leq T\}$ and $\alpha' = \{(i'_t, j'_t), t \leq T'\}$. We want to prove that

$$\begin{aligned} CC(\alpha \cup \alpha') + \#\{i_1, \dots, i_T, j_1, \dots, j_T\} + \#\{i'_1, \dots, i'_{T'}, j'_1, \dots, j'_{T'}\} \\ \geq CC(\alpha) + CC(\alpha') + \#\{i_1, \dots, i_T, j_1, \dots, j_T, i'_1, \dots, i'_{T'}, j'_1, \dots, j'_{T'}\}. \end{aligned} \tag{3.2}$$

When α contains a cycle C of length at least 2, removing an edge of C from α does not change any quantity of (3.2). By symmetry, we may assume that neither α nor α' contain a cycle, except possibly loops. Then, writing $\text{loops}(\alpha)$ for the number of loops in α (counted with multiplicity), one has

$$CC(\alpha) + T = \#\{i_1, \dots, i_T, j_1, \dots, j_T\} + \text{loops}(\alpha)$$

and

$$CC(\alpha') + T' = \#\{i'_1, \dots, i'_{T'}, j'_1, \dots, j'_{T'}\} + \text{loops}(\alpha').$$

Consequently, (3.2) is equivalent to

$$CC(\alpha \cup \alpha') + T + T' \geq \#\{i_1, \dots, i_T, j_1, \dots, j_T, i'_1, \dots, i'_{T'}, j'_1, \dots, j'_{T'}\} + \text{loops}(\alpha) + \text{loops}(\alpha').$$

The latter holds true since for any multigraph $G = (V, E)$, one has

$$cc(G) + |E| \geq |V| + \text{loops}(E). \quad \square$$

The goal of this section is to prove the following proposition.

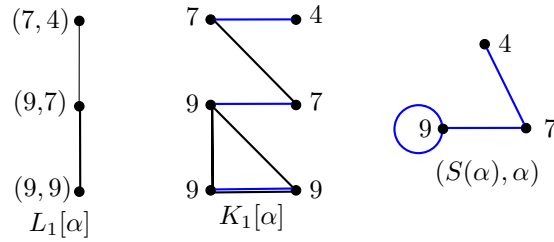


Figure 2: From left to right: the three graphs $L_1[\alpha]$, $K_1[\alpha]$ and $(S(\alpha), \alpha)$ associated with the $r = 3$, $i_1 = j_2 = 7$, $j_1 = 4$, $i_2 = i_3 = j_3 = 9$. Edges of the first kind in $K_1[\alpha]$ are plotted in blue. In this case, all three graphs are connected.

Theorem 3.2. *There is a universal sequence $C = (C_r)_{r \geq 1}$ such that the following holds. For any $n \geq 1$ and any $\lambda \vdash n$, the complete graph L on $[n]^2$ with weights*

$$w((i, j), (k, \ell)) = \begin{cases} \frac{1}{n} & \text{if } \{i, j\} \text{ and } \{k, \ell\} \text{ are disjoint;} \\ 1 & \text{otherwise,} \end{cases} \tag{3.3}$$

is a (Ψ, C) weighted dependency graph for the family $\{B_{i,j}, (i, j) \in [n]^2\}$.

Fix $r \geq 1$ and let $\alpha = \{(i_t, j_t), 1 \leq t \leq r\}$, be a multiset of size r of elements of $[n]^2$. We want to show that

$$|\kappa(B_{i_1, j_1}, \dots, B_{i_r, j_r})| \leq C_r \Psi(\alpha) \mathcal{M}(L[\alpha]), \tag{3.4}$$

where $\mathcal{M}(L[\alpha])$ is defined in Section 2.1. We start with a lemma.

Lemma 3.3. *For any $\alpha = \{(i_1, j_1), \dots, (i_r, j_r)\}$, we have $\mathcal{M}(L[\alpha]) = n^{-\text{CC}(\alpha)+1}$.*

Proof. By definition, the weighted graph L , and therefore also $L[\alpha]$, have only edges with weight 1 or n^{-1} . To form a spanning tree of maximal weight, we need to use as many edges of weight 1 as possible. It follows that $\mathcal{M}(L[\alpha]) = n^{-c+1}$, where c is the number of connected components of the subgraph $L_1[\alpha]$ of $L[\alpha]$ obtained by keeping only edges of weight 1.

It remains to relate connected components of $L_1[\alpha]$ to those of $(S(\alpha), \alpha)$. By construction, $L_1[\alpha]$ has r vertices decorated with pairs (i_t, j_t) and edges between pairs with a non-empty intersection. Consider a variant, denoted $K_1[\alpha]$, where we have $2r$ vertices with decorations i_t and j_t (t running from 1 to r) and with two kinds of edges

- for each t , the vertex decorated with i_t is connected to that decorated with j_t ;
- vertices with equal decorations are connected.

See Fig. 2 for an example. Contracting the first type of edges in $K_1[\alpha]$ gives $L_1[\alpha]$. Contracting the second type of edges instead gives the graph $(S(\alpha), \alpha)$ considered above. Hence, both have the same number of connected components and the lemma follows. \square

Using Lemma 3.3 and the definition of Ψ (Eq. (3.1)), Eq. (3.4) reduces to

$$|\kappa(B_{i_1, j_1}, \dots, B_{i_r, j_r})| \leq C_r n^{-\#\{i_1, \dots, i_r, j_1, \dots, j_r\}+1}. \tag{3.5}$$

We now prove this inequality. The idea is to represent σ_λ as $\tau^{-1} \circ \rho \circ \tau$, where ρ is a fixed permutation in \mathcal{C}_λ and τ a uniform random permutation in S_n . As in Section 2.2, we write $A_{i,j} = \mathbf{1}[\tau_i = j]$ and we have

$$B_{i,j} = \sum_{k=1}^n A_{i,k} A_{j, \rho_k}.$$

By multilinearity of cumulants, one has

$$\kappa(B_{i_1, j_1}, \dots, B_{i_r, j_r}) = \sum_{1 \leq k_1, \dots, k_r \leq n} \kappa(A_{i_1, k_1} A_{j_1, \rho_{k_1}}, \dots, A_{i_r, k_r} A_{j_r, \rho_{k_r}}). \quad (3.6)$$

Combining Proposition 2.2 together with Proposition 2.4 for $m = 2$, we know that L^2 is a (Ψ, \mathbf{D}) weighted dependency graph for the family $\{A_{i, k} A_{j, k'}, (i, k, j, k') \in [n]^4\}$, where

- L^2 is the second power of the graph L defined in Proposition 2.2; namely, in the graph L^2 , the vertex associated with $A_{i_s, k_s} A_{j_s, k'_s}$ is linked to that of $A_{i_t, k_t} A_{j_t, k'_t}$ by an edge of weight 1 if $\{i_s, j_s\} \cap \{i_t, j_t\} \neq \emptyset$ or $\{k_s, k'_s\} \cap \{k_t, k'_t\} \neq \emptyset$, and by an edge of weight $1/n$ otherwise.
- $\Psi((i_1, k_1, j_1, k'_1), \dots, (i_r, k_r, j_r, k'_r)) = n^{-\#\{(i_t, k_t), t \leq r\} \cup \{(j_t, k'_t), t \leq r\}}$;
- $\mathbf{D} = (D_r)_{r \geq 1}$ is a universal sequence.

In particular, for any (i_1, \dots, i_r) , (j_1, \dots, j_r) and (k_1, \dots, k_r) , we have

$$\left| \kappa(A_{i_1, k_1} A_{j_1, \rho_{k_1}}, \dots, A_{i_r, k_r} A_{j_r, \rho_{k_r}}) \right| \leq D_r \cdot n^{-\#\{(i_t, k_t), t \leq r\} \cup \{(j_t, \rho_{k_t}), t \leq r\}} \cdot \mathcal{M}(L^2[(i_1, k_1, j_1, \rho_{k_1}), \dots, (i_r, k_r, j_r, \rho_{k_r})]). \quad (3.7)$$

The weighted graph L^2 has only edges with weight 1 and $1/n$. Therefore for any subgraph $L^2[U]$, we have $\mathcal{M}(L^2[U]) = n^{-\text{cc}(L^2[U])+1}$, where $L^2_1[U]$ is the subgraph of $L^2[U]$ formed by its edges of weight 1.

In the case we are interested in, we get

$$\mathcal{M}(L^2[(i_1, k_1, j_1, \rho_{k_1}), \dots, (i_r, k_r, j_r, \rho_{k_r})]) = n^{-\text{cc}(G_{k_1, \dots, k_r}^A)+1}, \quad (3.8)$$

where G_{k_1, \dots, k_r}^A is the graph with vertex set $[r]$ and, for each s, t in $[r]$, it has an edge between s and t if and only if $\{i_s, j_s\} \cap \{i_t, j_t\} \neq \emptyset$ or $\{k_s, \rho_{k_s}\} \cap \{k_t, \rho_{k_t}\} \neq \emptyset$. We note that G_{k_1, \dots, k_r}^A depends on $i_1, \dots, i_r, j_1, \dots, j_r$, but since these indices are fixed throughout the proof, we keep this dependence implicit in the notation. Only the dependence in k_1, \dots, k_r is made explicit.

Let us illustrate this definition with an example. We take $\lambda = (6, 2, 1)$ and choose ρ to be $(1, 2, 3, 4, 5, 6)(7, 8)(9)$ (in the product of cycle notation). As in Fig. 2, we let $r = 3$, $i_1 = j_2 = 7$, $j_1 = 4$, $i_2 = i_3 = j_3 = 9$. We consider the term indexed by $(k_1, k_2, k_3) = (9, 5, 4)$, yielding $(\rho_{k_1}, \rho_{k_2}, \rho_{k_3}) = (9, 6, 5)$. Then 1 is linked to 2 in G_{k_1, \dots, k_r}^A because $\{i_1, j_1\} \cap \{i_2, j_2\} = \{7\} \neq \emptyset$. Likewise $\{i_2, j_2\} \cap \{i_3, j_3\} = \{9\} \neq \emptyset$ and $\{k_2, \rho_{k_2}\} \cap \{k_3, \rho_{k_3}\} = \{5\} \neq \emptyset$, implying that 2 is linked to 3 in G_{k_1, \dots, k_r}^A (in fact, only one of these intersections being nonempty would suffice for 2 to be linked to 3). One can check however that 1 is not linked to 3 since $\{i_1, j_1\} \cap \{i_3, j_3\} = \{4, 7\} \cap \{9\} = \emptyset$ and $\{k_1, \rho_{k_1}\} \cap \{k_3, \rho_{k_3}\} = \{9\} \cap \{4, 5\} = \emptyset$. Hence, G_{k_1, \dots, k_r}^A is the path $1 \text{---} 2 \text{---} 3$.

To help us in our analysis, we now introduce, for each r -tuple k_1, \dots, k_r , four (loop-free) graphs denoted $G_{k_1, \dots, k_r}^{(1)}$, $G_{k_1, \dots, k_r}^{(2)}$, G_{k_1, \dots, k_r}^\vee and $G_{k_1, \dots, k_r}^\wedge$ all on vertex set $[r] \times \{1, 2\}$. For all these graphs, the idea is to label the vertex $(s, 1)$ with the pair (i_s, k_s) and $(s, 2)$ with the pair (j_s, ρ_{k_s}) . Then we put an edge between two vertices labelled by (x, y) and (x', y')

- in $G_{k_1, \dots, k_r}^{(1)}$ if $x = x'$, i.e. if the first coordinates in the labels coincide;
- in $G_{k_1, \dots, k_r}^{(2)}$ if $y = y'$, i.e. if the second coordinates in the labels coincide;
- in $G_{k_1, \dots, k_r}^\wedge$ if $(x, y) = (x', y')$, i.e. if the labels coincide;
- in G_{k_1, \dots, k_r}^\vee if either $x = x'$ or $y = y'$, i.e. if one coordinate in the labels coincide.

Pattern counts in conjugacy classes

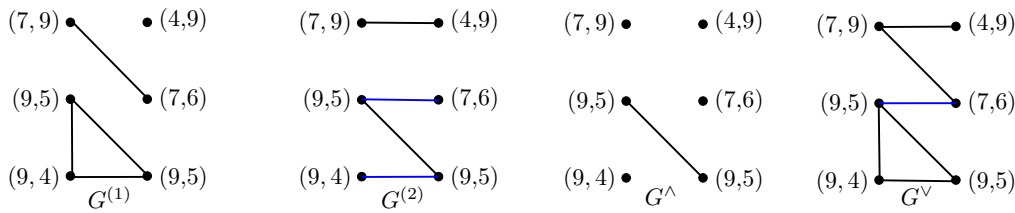


Figure 3: From left to right: the four graphs $G_{k_1, \dots, k_r}^{(1)}$, $G_{k_1, \dots, k_r}^{(2)}$, $G_{k_1, \dots, k_r}^{\vee}$ and $G_{k_1, \dots, k_r}^{\wedge}$ associated with the particular values of $i_1, \dots, i_r, j_1, \dots, j_r, k_1, \dots, k_r$ given on page 9.

Additionally, in $G_{k_1, \dots, k_r}^{(2)}$ and $G_{k_1, \dots, k_r}^{\vee}$, we add edges between $(s, 1)$ and $(s, 2)$ for any $s \leq r$ (for clarity, these extra edges will be drawn in blue). We continue our example above by representing the four associated graphs $G_{k_1, \dots, k_r}^{(1)}$, $G_{k_1, \dots, k_r}^{(2)}$, $G_{k_1, \dots, k_r}^{\vee}$ and $G_{k_1, \dots, k_r}^{\wedge}$ in Fig. 3. We note that a coincidence between a value in $\{i_1, \dots, i_r, j_1, \dots, j_r\}$ and a value in the set $\{k_1, \dots, k_r, \rho_{k_1}, \dots, \rho_{k_r}\}$ has no influence on any of the graphs considered here.

Comparing their definitions, we immediately see that $G_{k_1, \dots, k_r}^{\wedge}$ can be obtained from $G_{k_1, \dots, k_r}^{\vee}$ by merging, for each $s \leq r$, the vertices $(s, 1)$ and $(s, 2)$ into a single vertex s . Since $G_{k_1, \dots, k_r}^{\vee}$ has edges between $(s, 1)$ and $(s, 2)$ for any $s \leq r$, this merge operation does not change the number of connected components. Besides, again directly from definition, we see that connected components of $G_{k_1, \dots, k_r}^{\wedge}$ correspond to distinct pairs in the set $\{(i_t, k_t), t \leq r\} \cup \{(j_t, \rho_{k_t}), t \leq r\}$.

With these observations, Eqs. (3.7) and (3.8) imply

$$\left| \kappa(A_{i_1, k_1} A_{j_1, \rho_{k_1}}, \dots, A_{i_r, k_r} A_{j_r, \rho_{k_r}}) \right| \leq D_r n^{-cc(G_{k_1, \dots, k_r}^{\wedge}) - cc(G_{k_1, \dots, k_r}^{\vee}) + 1}. \quad (3.9)$$

We now need the following easy, possibly well-known, graph theoretical lemma.

Lemma 3.4. *Let V be a vertex set and E_1, E_2 be two sets of edges on V . Then*

$$cc(V, E_1 \cup E_2) + cc(V, E_1 \cap E_2) \geq cc(V, E_1) + cc(V, E_2). \quad (3.10)$$

Proof. Assume E_1 contains a cycle C . If E_2 contains C as well, we can remove any edge of C from E_1 without changing any of the quantities appearing in (3.10). If E_2 does not contain C , removing from E_1 an edge of C which is not in E_2 does not change any of the quantities appearing in (3.10) either. Hence, it is enough to prove (3.10) when E_1 does not contain any cycle. With the same argument, we can assume that E_2 does not contain any cycle as well. In this case, we have

$$cc(V, E_1) = |V| - |E_1|, \quad cc(V, E_2) = |V| - |E_2|, \quad cc(V, E_1 \cap E_2) = |V| - |E_1 \cap E_2|.$$

The union $E_1 \cup E_2$ may however contain cycles, so we only have an inequality

$$cc(V, E_1 \cup E_2) \geq |V| - |E_1 \cup E_2|.$$

Together with the standard inclusion-exclusion principle $|E_1 \cup E_2| + |E_1 \cap E_2| = |E_1| + |E_2|$, this proves the lemma. \square

By construction, the graph $cc(G_{k_1, \dots, k_r}^{\vee})$ is the union of $G_{k_1, \dots, k_r}^{(1)}$ and $G_{k_1, \dots, k_r}^{(2)}$. On the other hand, the edges of $G_{k_1, \dots, k_r}^{\wedge}$ are included in that of the intersection of $G_{k_1, \dots, k_r}^{(1)}$ and $G_{k_1, \dots, k_r}^{(2)}$ (since we have added edges $\{(s, 1), (s, 2)\}$ in $G_{k_1, \dots, k_r}^{(2)}$, it might happen that such an edge is in both $G_{k_1, \dots, k_r}^{(1)}$ and $G_{k_1, \dots, k_r}^{(2)}$, but not in $G_{k_1, \dots, k_r}^{\wedge}$; this is for example the case

of the bottom edge in Fig. 3). Consequently, $G_{k_1, \dots, k_r}^\wedge$ has more connected components than this intersection. Combining these observations with Lemma 3.4, for any k_1, \dots, k_r , we have

$$\begin{aligned} \text{cc}(G_{k_1, \dots, k_r}^\wedge) + \text{cc}(G_{k_1, \dots, k_r}^\vee) &\geq \text{cc}(G_{k_1, \dots, k_r}^{(1)} \cap G_{k_1, \dots, k_r}^{(2)}) + \text{cc}(G_{k_1, \dots, k_r}^{(1)} \cup G_{k_1, \dots, k_r}^{(2)}) \\ &\geq \text{cc}(G_{k_1, \dots, k_r}^{(1)}) + \text{cc}(G_{k_1, \dots, k_r}^{(2)}). \end{aligned}$$

Back to (3.9), we get that, for any k_1, \dots, k_r ,

$$\left| \kappa(A_{i_1, k_1} A_{j_1, \rho_{k_1}}, \dots, A_{i_r, k_r} A_{j_r, \rho_{k_r}}) \right| \leq D_r n^{-\text{cc}(G_{k_1, \dots, k_r}^{(1)}) - \text{cc}(G_{k_1, \dots, k_r}^{(2)}) + 1}. \tag{3.11}$$

At this stage, we observe the graph $G_{k_1, \dots, k_r}^{(1)}$ is independent of k_1, \dots, k_r (it compares the first coordinates of the labels, which are either i 's or j 's). Its number of connected components is the number of distinct elements in $\{i_1, \dots, i_r, j_1, \dots, j_r\}$. Hence, summing (3.11) over k_1, \dots, k_r and recalling (3.6), we obtain

$$\left| \kappa(B_{i_1, j_1}, \dots, B_{i_r, j_r}) \right| \leq D_r \cdot n^{-\#\{i_1, \dots, i_r, j_1, \dots, j_r\} + 1} \cdot \sum_{1 \leq k_1, \dots, k_r \leq n} n^{-\text{cc}(G_{k_1, \dots, k_r}^{(2)})}. \tag{3.12}$$

We need a final lemma.

Lemma 3.5. Fix ρ in S_n . For any graph G on vertex set $[r] \times \{1, 2\}$, the number of lists (k_1, \dots, k_r) in $[n]^r$ with $G_{k_1, \dots, k_r}^{(2)} = G$ is at most $n^{\text{cc}(G)}$.

Proof. Observe that if $G_{k_1, \dots, k_r}^{(2)} = G$ and if (t, h) is linked to (s, h') in G , the value of k_s is determined by that of k_t ; they should either be equal if $h = h'$, or they should satisfy $k_s = \rho_{k_t}$ if $h = 1$ and $h' = 2$, or $k_t = \rho_{k_s}$ if $h = 2$ and $h' = 1$. Hence, a list (k_1, \dots, k_r) in $[n]^r$ with $G_{k_1, \dots, k_r}^{(2)} = G$ is determined by the value of one k_t in each component of G . This proves the lemma. \square

Consequently, for any graph G on vertex set $[r] \times \{1, 2\}$, the lists (k_1, \dots, k_r) in $[n]^r$ with $G_{k_1, \dots, k_r}^{(2)} = G$ contributes at most 1 to the sum $\sum_{1 \leq k_1, \dots, k_r \leq n} n^{-\text{cc}(G_{k_1, \dots, k_r}^{(2)})}$. Thus the latter sum is bounded by a constant depending only on r , namely $2^{\binom{2r}{2}}$. Inserting this inequality in (3.12), this proves (3.5) and thus Proposition 3.2.

4 Proof of asymptotic normality of vincular pattern counts

By the method of moments, convergence (1.1), together with moment convergence, hold if

- $\text{Var}(X^\pi(\sigma_{\lambda^n})) = n^{2m-1}(f(p_1) + p_2 g(p_1))(1 + o(1))$;
- and higher cumulants tend to 0 after normalization, i.e. $\kappa_r(X^\pi(\sigma_{\lambda^n})) = o(n^{r(m-\frac{1}{2})})$ for $r \geq 3$.

The existence of polynomial functions f and g such that the first item holds is proven in [9] (see Proposition 7.2 and Theorem 8.14 there). We will bound the cumulants using the weighted dependency graph structure of Section 3.

Note that we have

$$X^{(\pi, A)}(\sigma_{\lambda^n}) = \sum_{\substack{i_1 < \dots < i_k \\ i_{s+1} = i_s + 1 \text{ for } s \in A}} \sum_{\substack{j_1, \dots, j_k \\ j_{\pi^{-1}(1)} < \dots < j_{\pi^{-1}(k)}}} B_{i_1, j_1} \cdots B_{i_k, j_k}.$$

We denote by $\mathcal{I}^{(\pi, A)}$ the set of tuples $(i_1, \dots, i_k, j_1, \dots, j_k)$ as in the above sum. To bound the cumulant of $X^{(\pi, A)}(\sigma_{\lambda^n})$, we need a weighted dependency graph for the family of

products $B_{i_1, j_1} \cdots B_{i_k, j_k}$. This is given by applying the stability property of weighted dependency graphs of Proposition 2.4 to the specific dependency graph of Proposition 3.2, as we now explain.

Let us recall the notation from Section 2.4, adapted to the present situation. If each of $\alpha_1, \dots, \alpha_\ell$ is a list of k elements (or k -list for short) of $[n]^2$, we define

$$\Psi(\alpha_1, \dots, \alpha_\ell) = \Psi(\alpha_1 \cup \dots \cup \alpha_\ell),$$

where in the right-hand side, lists are seen as sets, and we recall that Ψ is defined in (3.1). We also consider the k -th power L^k of the weighted graph L from Proposition 3.2. Concretely, L^k has vertices indexed by k -lists of $[n]^2$ and its edge weights are given by

$$w_k(\alpha, \beta) := \max_{(i, \ell) \in \alpha, (j, \ell') \in \beta} w((i, \ell), (j, \ell')) = \begin{cases} \frac{1}{n} & \text{if } S(\alpha) \cap S(\beta) = \emptyset; \\ 1 & \text{otherwise.} \end{cases}$$

Using Proposition 2.4, there is a universal sequence $\mathbf{D}_k = (C_r^k)_{r \geq 1}$ such that L^k is a (Ψ, \mathbf{D}_k) weighted dependency graph for the set of random variables $\prod_{t=1}^k B_{i_t, j_t}$, indexed by k -lists $\alpha = ((i_1, j_1), \dots, (i_k, j_k))$ of elements of $[n]^2$. We shall consider the restriction of this weighted graph to k -lists in $\mathcal{I}^{(\pi, A)}$: this is a (Ψ, \mathbf{D}_k) weighted dependency graph for the set of random variables $\prod_{t=1}^k B_{i_t, j_t}$, indexed by $\mathcal{I}^{(\pi, A)}$.

To use Lemma 2.3, we want to bound the quantities R and $(T_\ell)_{\ell \geq 1}$, whose definitions are now recalled for this specific weighted dependency graph.

$$R = \sum_{\alpha} \Psi(\alpha) = \sum_{(i_1, \dots, i_k, j_1, \dots, j_k) \in \mathcal{I}^{(\pi, A)}} n^{\text{CC}((i_1, j_1), \dots, (i_k, j_k)) - \#\{i_1, \dots, i_k, j_1, \dots, j_k\}}, \quad (4.1)$$

and

$$T_\ell = \max_{\alpha_1, \dots, \alpha_\ell} \left[\sum_{\beta} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \right], \quad (4.2)$$

where, in the indices, each of $\alpha_1, \dots, \alpha_\ell, \beta$ is a k -list in $\mathcal{I}^{(\pi, A)}$, and where

$$W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) = \max_i w_k(\beta, \alpha_i) = \begin{cases} \frac{1}{n} & \text{if } S(\beta) \cap (S(\alpha_1) \cup \dots \cup S(\alpha_\ell)) = \emptyset; \\ 1 & \text{otherwise.} \end{cases} \quad (4.3)$$

We start by bounding R .

Lemma 4.1. $R \leq 2^{\binom{2k}{2}} n^m$.

Proof. As in Section 3, we introduce two graphs $G^{(I)}$ and $G^{(A)}$ associated with a sequence $(i_1, \dots, i_k, j_1, \dots, j_k)$ (the dependence in $(i_1, \dots, i_k, j_1, \dots, j_k)$ is implicit in the notation). Both have vertex set $[k] \times \{1, 2\}$, where the vertex $(s, 1)$, resp. $(s, 2)$, is labelled with i_s , resp. j_s . Edges are chosen as follows:

- in $G^{(I)}$, we connect $(s, 1)$ and $(s, 2)$ for all $s \leq k$;
- in $G^{(A)}$, we connect $(s, 1)$ and $(s + 1, 1)$ for $s \in A$;
- in both graphs, we additionally connect vertices with equal labels.

With the same arguments as Lemma 3.3, we see that connected components of $G^{(I)}$ correspond to that of $(S((i_1, j_1), \dots, (i_k, j_k)), \{(i_1, j_1), \dots, (i_k, j_k)\})$, implying

$$\text{cc}(G^{(I)}) = \text{CC}((i_1, j_1), \dots, (i_k, j_k)).$$

Furthermore, let $G^{(\cap)}$ be the intersection of $G^{(I)}$ and $G^{(A)}$; it only has edges connecting vertices with equal labels, and hence its number of components is the number of distinct values in $\{i_1, \dots, i_k, j_1, \dots, j_k\}$. Thus (4.1) can be rewritten as

$$R = \sum_{(i_1, \dots, i_k, j_1, \dots, j_k) \in \mathcal{I}(\pi, A)} n^{\text{cc}(G^{(I)}) - \text{cc}(G^{(\cap)})}. \tag{4.4}$$

We now introduce the union $G^{(\cup)}$ of $G^{(I)}$ and $G^{(A)}$ (union in terms of edge-set). Since it contains the edges $\{(s, 1), (s, 2)\}, s \leq k \cup \{(s, 1), (s + 1, 1)\}, s \in A$, it has at most $m = k - \#A$ connected components. Using Eq. (4.4) and Lemma 3.4, we get

$$R \leq \sum_{(i_1, \dots, i_k, j_1, \dots, j_k) \in \mathcal{I}(\pi, A)} n^{\text{cc}(G^{(\cup)}) - \text{cc}(G^{(A)})} \leq n^m \sum_{(i_1, \dots, i_k, j_1, \dots, j_k) \in \mathcal{I}(\pi, A)} n^{-\text{cc}(G^{(A)})}.$$

But, for a given graph G , there are at most $n^{\text{cc}(G)}$ lists $(i_1, \dots, i_k, j_1, \dots, j_k)$ in $\mathcal{I}(\pi, A)$ such that $G^{(A)} = G$. Indeed, choosing a value i_s or j_s corresponding to a vertex of G forces the values of variables i_t or j_t in the same component. We finally get $R \leq 2^{\binom{2k}{2}} n^m$. \square

The next step is to bound T_ℓ for all $\ell \geq 1$.

Lemma 4.2. *For any $\ell \geq 1$, there exists $C''_{k,\ell}$ (depending on the size k of the pattern (π, A) and on ℓ , but not on n) such that, the following holds true*

$$T_\ell \leq C''_{k,\ell} n^{m-1}.$$

To prove this lemma, we need the following simple combinatorial fact to control the various terms in (4.2).

Lemma 4.3. *For any k -lists $\alpha_1, \dots, \alpha_\ell$ and $\beta = ((i_1, j_1), \dots, (i_k, j_k))$ of elements of $[n]^2$, one has*

$$\text{CC}(\alpha_1 \cup \dots \cup \alpha_\ell \cup \beta) - \text{CC}(\alpha_1 \cup \dots \cup \alpha_\ell) \leq k - I(\beta),$$

where $I(\beta)$ is the number of pairs (i_s, j_s) in β which intersect $S(\alpha_1 \cup \dots \cup \alpha_\ell)$.

Proof. We can see the graph with edge set $\alpha_1 \cup \dots \cup \alpha_\ell \cup \beta$ (and no isolated vertex) as obtained from that with edge set $\alpha_1 \cup \dots \cup \alpha_\ell$ adding successively the edges $(i_1, j_1), \dots, (i_k, j_k)$ (with their extremity as new vertices if needed). When adding the edge (i_s, j_s) , if (i_s, j_s) intersects $S(\alpha_1 \cup \dots \cup \alpha_\ell)$, the number of components does not increase, otherwise it might increase by 1. This proves the lemma. \square

Proof of Lemma 4.2. We first fix $\alpha_1, \dots, \alpha_\ell$ and write $S(\alpha) = S(\alpha_1) \cup \dots \cup S(\alpha_\ell)$. We split the sum in (4.2) according to the intersection $S(\beta) \cap S(\alpha)$ being empty or not. When $S(\beta) \cap S(\alpha) = \emptyset$, we have $W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) = \frac{1}{n}$ (Equation (4.3)) and $I(\beta) = 0$. Thus Lemma 4.3 gives

$$\frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} = n^{\text{CC}(\alpha_1 \cup \dots \cup \alpha_\ell \cup \beta) - \text{CC}(\alpha_1 \cup \dots \cup \alpha_\ell) - (\#S(\alpha_1 \cup \dots \cup \alpha_\ell, \beta) - \#S(\alpha_1 \cup \dots \cup \alpha_\ell))} \leq n^{k - \#S(\beta)}.$$

This yields the following bound:

$$\sum_{\beta: S(\beta) \cap S(\alpha) = \emptyset} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \leq \sum_{i=1}^{2k} \sum_{\beta: \#S(\beta)=i} n^{k-i-1}.$$

But the number of k -lists β in $\mathcal{I}(\pi, A)$ with $\#S(\beta) = i$ is bounded by $C_{k,i}n^{i-|A|}$ for some constant $C_{k,i}$ (to construct such a β , one chooses $i - |A|$ values, the other being forced by the adjacencies conditions $i_{s+1} = i_s + 1$ for $s \in A$). This yields, recalling that $m = k - \#A$,

$$\sum_{\beta: S(\beta) \cap S(\alpha) = \emptyset} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \leq \left(\sum_{i=1}^k C_{k,i} \right) n^{m-1}.$$

Let us now consider the summands in (4.2) indexed by terms β in $\mathcal{I}(\pi, A)$ such that $S(\beta) \cap S(\alpha) \neq \emptyset$. In this case, we have $W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) = 1$ (see Equation (4.3)) and the analysis is more subtle.

We write as usual $\beta = ((i_1, j_1), \dots, (i_k, j_k))$. First we recall that the constraints $i_{s+1} = i_s + 1$ for $s \in A$ splits the sequence (i_1, \dots, i_k) into m blocks of necessary consecutive values. This allows to define $F_\alpha(\beta)$ (F stands for *forced*) as the number of indices s such that i_s is not in $S(\alpha)$, but either i_{s-1} is in the same block as i_s (i.e. $i_s - 1 \in A$), or one i_t in the same block is in $S(\alpha)$. Then, for fixed α, i, j , the number of lists β in $\mathcal{I}(\pi, A)$ with $F_\alpha(\beta) = j$ and $\#(S(\beta) \setminus S(\alpha_1 \cup \dots \cup \alpha_\ell)) = i$ is bounded by $C'_{k,\ell,i,j}n^{i-j}$ for some constant $C'_{k,\ell,i,j}$. Indeed, to construct β , one needs i new values, but j of them are forced by the constraints $i_{s+1} = i_s$ for $s \in A$.

On the other hand, in each block of indices, all but possibly one i_s in the block are in $F_\alpha(\beta) \cup S(\alpha)$. Therefore, $|F_\alpha(\beta)| + |S(\alpha) \cap \{i_1, \dots, i_k\}| \geq k - m$. Equality occurs if and only if $|S(\alpha) \cap \{i_1, \dots, i_k\}| = 0$. Looking at the definition of $I(\beta)$ in Lemma 4.3, we trivially have $|I(\beta)| \geq |S(\alpha) \cap \{i_1, \dots, i_k\}|$. Putting both inequalities together, we have $|F_\alpha(\beta)| + |I(\beta)| \geq k - m$, with equality if and only if $|I(\beta)| = 0$. But $|I(\beta)| = 0$ implies $S(\beta) \cap S(\alpha) = \emptyset$. Since we are considering the case $S(\beta) \cap S(\alpha) \neq \emptyset$, the inequality above must be strict and we have in fact $|F_\alpha(\beta)| + |I(\beta)| \geq k - m + 1$.

Using the definition of Ψ (Eq. (3.1)), Lemma 4.3 and the inequality above, we have

$$\frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \leq n^{k-|I(\beta)|-\#S(\beta) \setminus S(\alpha_1 \cup \dots \cup \alpha_\ell)} \leq n^{|F_\alpha(\beta)|+m-1-\#S(\beta) \setminus S(\alpha_1 \cup \dots \cup \alpha_\ell)}.$$

This yields the following bound:

$$\begin{aligned} \sum_{\beta: S(\beta) \cap S(\alpha) \neq \emptyset} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} &\leq \sum_{\substack{i \leq 2m-1 \\ j \leq |A|}} \sum_{\substack{\beta: \#(S(\beta) \setminus S(\alpha_1 \cup \dots \cup \alpha_\ell)) = i \\ |F_\alpha(\beta)| = j}} n^{j+m-1-i}. \end{aligned}$$

Recalling that there are at most $C'_{k,\ell,i,j}n^{i-j}$ terms in the latter sum, we conclude that

$$\sum_{\beta: S(\beta) \cap S(\alpha) \neq \emptyset} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \leq \left(\sum_{i \leq 2m-1, j \leq |A|}^m C'_{k,\ell,i,j} \right) n^{m-1}.$$

Bringing everything together, we get that

$$\sum_{\beta} W(\{\beta\}, \{\alpha_1, \dots, \alpha_\ell\}) \frac{\Psi(\alpha_1, \dots, \alpha_\ell, \beta)}{\Psi(\alpha_1, \dots, \alpha_\ell)} \leq C''_{k,\ell} n^{m-1}.$$

Since this holds for any tuple $(\alpha_1, \dots, \alpha_\ell)$, we have $T_\ell \leq C''_{k,\ell} n^{m-1}$, as wanted. \square

Back to the proof of Theorem 1.2, Lemma 2.3 gives us

$$\left| \kappa_r(X^{(\pi,A)}(\sigma_{\lambda^n})) \right| \leq D_r R \prod_{i=1}^{r-1} T_i \leq C'''_{k,r} n^{r(m-1)+1},$$

for some constants $C'''_{k,r}$. Consequently, if $r > 2$, we get:

$$\kappa_r \left(\frac{X^{(\pi,A)}(\sigma_{\lambda^n}) - \mathbb{E}(X^{(\pi,A)}(\sigma_{\lambda^n}))}{n^{m-\frac{1}{2}}} \right) = O(n^{1-\frac{r}{2}}) \rightarrow 0.$$

This concludes the proof of Theorem 1.2.

5 Non-degeneracy of the limiting law

In this section, we prove Theorem 1.3. Recall that this theorem only considers classical patterns, i.e. the case $A = \emptyset$, and that we write π instead of (π, \emptyset) for simplicity. Note that in this case, the parameter m in Theorem 1.2 is simply equal to the size k of the pattern.

Let us define $p_3 = \frac{1}{3}(1 - p_1 - 2p_2) \geq 0$. For $n \geq 0$, we set $m_i = \lfloor np_i \rfloor$ for $i \in \{2, 3\}$ and $m_1 = n - 2m_2 - 3m_3$, so that $(1^{m_1}, 2^{m_2}, 3^{m_3})$ is a partition of n , which we denote by λ^n . From [9, Theorem 8.14], it holds that

$$\text{Var}(X^\pi(\sigma_{\lambda^n})) = (f(p_1) + p_2g(p_1)) n^{2k-1} + O(n^{2k-2}). \tag{5.1}$$

In fact, in [9, Theorem 8.14], the error term is $o(n^{2k-1})$ and this is best possible assuming only $m_i/n \rightarrow p_i$, but a quick inspection of the proof reveals that if $m_i/n = p_i + O(1/n)$ for $i \in \{1, 2\}$, then the error term is $O(n^{2k-2})$, as claimed.

We shall prove that $\text{Var}(X^\pi(\sigma_{\lambda^n})) \geq K_1 n^{2k-3/2}$, for some constant $K_1 = K_1(p_1, p_2) > 0$. Comparing with Eq. (5.1), this forces $f(p_1) + p_2g(p_1) > 0$. The method of proof is inspired from [10, Section 4] and [6, Section 3.4]. The idea is to establish a recursive inequality on $\text{Var}(X^\pi(\sigma_{\lambda^n}))$, using a recursive construction of σ_{λ^n} and the law of total variance. We first consider the case $p_2 > 0$ (in Sections 5.1 to 5.4) and indicate later (in Section 5.5) the appropriate modifications in the case $p_2 = 0, p_1 < 1$.

5.1 Preliminary: one point conditional pattern densities

In this section, we introduce some quantity which will be useful later. Let $P_i = (x_i, y_i)$, for $i \leq k$, be points in the unit square with distinct coordinates, i.e. the x_i 's are distinct and the y_i 's are distinct, but we allow here some x_i to be equal to some y_j . We reorder the points $(P_i)_{i \leq k}$ as $P_{(i)} = (x_{(i)}, y_{(i)})$ such that $x_{(1)} < \dots < x_{(k)}$. Then there exists a unique permutation π such that $y_{(\pi^{-1}(1))} < \dots < y_{(\pi^{-1}(k))}$, and we set $\text{Perm}(P_1, \dots, P_k) := \pi$. For example, if $P_1 = (0.1, 0.5)$, $P_2 = (0.7, 0.2)$ and $P_3 = (0, 0.33)$ then $P_{(1)} = P_3 = (0, 0.333)$, $P_{(2)} = P_1 = (0.1, 0.5)$, $P_{(3)} = P_2 = (0.7, 0.2)$ and $\text{Perm}(P_1, P_2, P_3) = 231$.

Let us introduce some randomness. We denote by Λ the Lebesgue measure on $[0, 1]^2$ and by Δ that on the diagonal $\{(x, x), x \in [0, 1]\}$ with total mass 1. In particular, for any $p_1 \in [0, 1]$, $p_1\Delta + (1 - p_1)\Lambda$ has uniform marginals. We then let P_1, \dots, P_{k-1} be i.i.d. random points with distribution $p_1\Delta + (1 - p_1)\Lambda$ and set

$$\varphi_{p_1}^\pi(x, y) = \mathbb{P}[\text{Perm}(P_1, \dots, P_{k-1}, (x, y)) = \pi]. \tag{5.2}$$

Lemma 5.1. *Assume π has size at least 2 and $p_1 < 1$. Then the function*

$$(x, y) \mapsto \varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, x)$$

is non constant.

Proof. Let $\tilde{\pi}$ be the pattern obtained from π by removing the first entry, and standardizing. We look at $\varphi_{p_1}^\pi((0, \frac{1}{2}))$, i.e the probability that random points P_1, \dots, P_{k-1} together with $(0, \frac{1}{2})$ induce the pattern π . This event is equivalent to the fact that P_1, \dots, P_{k-1} induce $\tilde{\pi}$ and that exactly $\pi_1 - 1$ points among P_1, \dots, P_{k-1} have a y -coordinate smaller than $\frac{1}{2}$. Hence, we have

$$\begin{aligned} \varphi_{p_1}^\pi(0, \frac{1}{2}) &= \mathbb{P}[\text{Perm}(P_1, \dots, P_{k-1}) = \tilde{\pi} \wedge \#\{i \leq k-1 : y_i < \frac{1}{2}\} = \pi_1 - 1] \\ &\geq \mathbb{P}[\text{Perm}(P_1, \dots, P_{k-1}) = \tilde{\pi}, \wedge \#\{i \leq k-1 : y_i < \frac{1}{2}\} = \pi_1 - 1 \wedge \forall i, x_i \neq y_i]. \end{aligned}$$

The latter event “ $\forall i, x_i \neq y_i$ ” is a.s. equivalent to the P_i ’s being sampled according to Λ and hence has probability $(1 - p_1)^{k-1}$ (recall that P_1, \dots, P_{k-1} have distribution $p_1\Delta + (1 - p_1)\Lambda$). Conditionally to “ $\forall i, x_i \neq y_i$ ”, the events

$$\text{Perm}(P_1, \dots, P_{k-1}) = \tilde{\pi} \text{ and } \#\{i \leq k-1 : y_i < \frac{1}{2}\} = \pi_1 - 1$$

are independent (in general, the pattern induced by uniform random points is independent of the set of its y -coordinates). These events occur with probability $\frac{1}{(k-1)!}$ and $\binom{k-1}{\pi_1-1} 2^{-k+1}$ respectively. Thus we get

$$\varphi_{p_1}^\pi(0, \frac{1}{2}) \geq \frac{(1 - p_1)^{k-1}}{(\pi_1 - 1)! (k - \pi_1)! 2^{k-1}}.$$

In particular, when $p_1 < 1$, we have $\varphi_{p_1}^\pi(0, \frac{1}{2}) > 0$. On the other hand,

- If $\pi_1 \neq 1$, then $\varphi_{p_1}^\pi(0, 0) = 0$;
- If $\pi_1 = 1$ and if π has size at least 2, then $\varphi_{p_1}^\pi(0, 1) = \varphi_{p_1}^\pi(1, 0) = 0$.

In both cases, we can conclude that the function $\varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, x)$ is not constant. \square

We also note that $|\varphi_{p_1}^\pi(x, y) - \varphi_{p_1}^\pi(x', y')|$ is bounded from above by the probability that at least one of P_1, \dots, P_{k-1} has its x -coordinate between x and x' or its y -coordinate between y and y' . Using a union bound, this implies

$$|\varphi_{p_1}^\pi(x, y) - \varphi_{p_1}^\pi(x', y')| \leq (k - 1)(|x - x'| + |y - y'|),$$

i.e. $\varphi_{p_1}^\pi$ is $(k - 1)$ -Lipschitz when using the L_1 norm on $[0, 1]$.

5.2 A recursive inequality on the variance

Let λ be a partition with at least one part equal to 2 and denote by λ' the partition obtained by removing a part of length 2 to λ . Then a uniform random permutation $\sigma = \sigma_\lambda$ in \mathcal{C}_λ can be constructed as follows.

- Choose two distinct indices i and j uniformly at random between 1 and n , and set $\sigma_i = j$ and $\sigma_j = i$
- Take a uniform random permutation σ' of $[n] \setminus \{i, j\}$ of cycle type λ' , independent of $\{i, j\}$, and set $\sigma_k = \sigma'_k$ for $k \notin \{i, j\}$.

Using the law of total variance, we have

$$v(\lambda) := \text{Var}(X^\pi(\sigma)) = \mathbb{E}[\text{Var}(X^\pi(\sigma)|i, j)] + \text{Var}[\mathbb{E}(X^\pi(\sigma)|i, j)].$$

Moreover, the quantity $X^\pi(\sigma)$ can be decomposed as $X^\pi(\sigma) = X^\pi(\sigma') + C$, where C is the number of occurrences of π in σ , using position i or j (or both). Note that $X^\pi(\sigma')$ is

independent of i, j , so that $\mathbb{E}(X^\pi(\sigma')|i, j)$ is the *constant* random variable a.s. equal to $\mathbb{E}(X^\pi(\sigma'))$. This implies that

$$\text{Var} [\mathbb{E}(X^\pi(\sigma)|i, j)] = \text{Var} [\mathbb{E}(X^\pi(\sigma')) + \mathbb{E}(C|i, j)] = \text{Var} [\mathbb{E}(C|i, j)].$$

In particular, we get an initial bound

$$v(\lambda) \geq \text{Var} [\mathbb{E}(C|i, j)]. \tag{5.3}$$

On the other hand, from Cauchy-Schwarz and Jensen inequalities, since $X^\pi(\sigma')$ and i, j are independent, we have

$$\begin{aligned} \left| \mathbb{E} [\text{Cov}(X^\pi(\sigma'), C|i, j)] \right| &\leq \mathbb{E} [\sqrt{\text{Var}(X^\pi(\sigma'))} \sqrt{\text{Var}(C|i, j)}] \\ &\leq \sqrt{\text{Var}(X^\pi(\sigma'))} \sqrt{\mathbb{E}[\text{Var}(C|i, j)]}. \end{aligned}$$

Expanding $\text{Var}(X^\pi(\sigma)|i, j) = \text{Var}(X^\pi(\sigma') + C|i, j)$ by bilinearity, we find that

$$\begin{aligned} \mathbb{E} [\text{Var}(X^\pi(\sigma)|i, j)] &= \text{Var}(X^\pi(\sigma')) + 2\mathbb{E} [\text{Cov}(X^\pi(\sigma'), C|i, j)] + \mathbb{E} [\text{Var}(C|i, j)] \\ &\geq \text{Var}(X^\pi(\sigma')) - 2\sqrt{\text{Var}(X^\pi(\sigma'))} \sqrt{\mathbb{E}[\text{Var}(C|i, j)]}. \end{aligned}$$

Note that $\text{Var}(X^\pi(\sigma')) = v(\lambda')$, since σ' is a uniform random permutation of cycle-type λ' . Summing up, we get the following recursive inequality on $v(\lambda)$:

$$v(\lambda) \geq v(\lambda') \left(1 - 2\sqrt{\frac{\mathbb{E}[\text{Var}(C|i, j)]}{v(\lambda')}} \right) + \text{Var} [\mathbb{E}(C|i, j)]. \tag{5.4}$$

5.3 Analysing the initial bound (5.3)

Recall that C counts the number of occurrences of π in σ that use position i or j (or both). We first consider the number C_1 of those using i but not j .

Lemma 5.2. *Let x, y be in $[0, 1]$ and $i = i(n)$ and $j = j(n)$ be two sequences chosen such that $i = xn + o(n)$ and $j = yn + o(n)$. Then we have*

$$\mathbb{E}(C_1|i, j) = \frac{1}{(k-1)!} n^{k-1} \varphi_{p_1}^\pi(x, y) + o(n^{k-1}),$$

where $\varphi_{p_1}^\pi$ is defined in Eq. (5.2) and the error term is uniform on i and j .

We start with some estimates on the distribution of finitely tuples $(\sigma(i_1), \dots, \sigma(i_m))$, where σ is a uniform random permutation of cycle type λ .

Lemma 5.3. 1. *Fix $m \geq 1$ and let i_1, \dots, i_m be distinct indices in $\{1, \dots, n\}$. Then*

$$\mathbb{P}(\exists s \neq t : \sigma_{i_s} = i_t) = O(n^{-1}).$$

2. *Let j_1, \dots, j_m also be distinct indices in $\{1, \dots, n\}$. We assume that there is no $s \neq t$ such that $i_s = j_t$ and we let $r = |\{t : i_t = j_t\}|$. Then*

$$\mathbb{P}(\forall s \leq m, \sigma_{i_s} = j_s) = (p_1)^r \left(\frac{1-p_1}{n} \right)^{m-r} (1 + O(n^{-1})).$$

In both estimates, the error terms depend on m , but are uniform in $(i_1, \dots, i_m, j_1, \dots, j_m)$ and in λ . Moreover, the same estimates hold conditionally to i, j in the construction of Section 5.2, provided that the tuple $(i_1, \dots, i_m, j_1, \dots, j_m)$ does not contain i or j .

Proof. For the first item, let us note that a given index i_s is a fixed point of σ with probability $m_1(\lambda)/n$. Conditionally of i_s not being a fixed point, by conjugation invariance, its image σ_{i_s} is uniformly distributed on $[n] \setminus \{i_s\}$, and thus for fixed $s \neq t$, we have

$$\mathbb{P}(\sigma_{i_s} = i_t) = \frac{n - m_1(\lambda)}{n} \cdot \frac{1}{n - 1} \leq \frac{1}{n - 1}.$$

The first item in the lemma follows by a simple union bound.

Let us consider the second item. By symmetry, we can assume that $i_s = j_s$ for $s \leq r$ and $i_s \neq j_s$ for $s > r$. Recalling that the set $F(\sigma)$ of fixed points of σ is a uniform random subset of $[n]$ of size $m_1(\lambda)$, we have

$$\mathbb{P}[\forall s \leq r, \sigma(i_s) = i_s] = \mathbb{P}[\{i_1, \dots, i_r\} \subset F(\sigma)] = \frac{\binom{n-r}{m_1-r}}{\binom{n}{m_1}} = p_1^r + O(n^{-1}).$$

Now, conditionally on i_1, \dots, i_r being fixed points, the probability that i_{r+1}, \dots, i_m are **not** fixed points is

$$\frac{\binom{n-m}{m_1-r}}{\binom{n-r}{m_1-r}} = (1 - p_1)^{m-r} + O(n^{-1}).$$

Combining this computation with item i), the probability of i_1, \dots, i_r being fixed points, of i_{r+1}, \dots, i_m being not fixed points and of $(\sigma_{i_{r+1}}, \dots, \sigma_{i_m})$ being disjoint from (i_{r+1}, \dots, i_m) is

$$p_1^r (1 - p_1)^{m-r} + O(n^{-1}).$$

But under this event, by conjugation invariance, the tuple $(\sigma_{i_{r+1}}, \dots, \sigma_{i_m})$ is a uniform random tuple of distinct values in $[n] \setminus \{i_1, \dots, i_m\}$. Hence, the conditional probability that is equal to (j_{r+1}, \dots, j_m) is $1/n^{m-r}(1 + O(n^{-1}))$. We finally get

$$\mathbb{P}(\forall s \leq m, \sigma(i_s) = j_s) = \frac{p_1^r (1 - p_1)^{m-r} + O(n^{-1})}{n^{m-r}(1 + O(n^{-1}))},$$

which proves the second item.

The conditional statement in the lemma follows easily from the unconditional one since for $h \notin \{i, j\}$, we have $\sigma_h = \sigma'_h$, where σ' is a uniform random permutation of cycle-type λ' of the set $[n] \setminus \{i, j\}$. □

Proof of Lemma 5.2. Recall that C_1 is the number of occurrences of π in σ using position i but not j . Denoting by i_1, \dots, i_{k-1} the other positions of such an occurrence of π (in any order, hence the symmetry factor in the next formula) and j_1, \dots, j_{k-1} their images by σ , we can write C_1 as follows:

$$C_1 = \frac{1}{(k-1)!} \sum_{\substack{i_1, \dots, i_{k-1} \in [n] \setminus \{i, j\} \\ \text{distinct}}} \sum_{\substack{j_1, \dots, j_{k-1} \in [n] \setminus \{i, j\} \\ \text{distinct}}} \left(\mathbb{1} \left[\text{Perm} \left(\left(\frac{i_1}{n}, \frac{j_1}{n} \right), \dots, \left(\frac{i_{k-1}}{n}, \frac{j_{k-1}}{n} \right), \left(\frac{i}{n}, \frac{j}{n} \right) \right) = \pi \right] \cdot \mathbb{1} [\sigma(i_\ell) = j_\ell, \forall 1 \leq \ell \leq k-1] \right). \quad (5.5)$$

By Lemma 5.3, item i), for each i_1, \dots, i_{k-1} , the total contribution of terms with some $j_t \neq i_t$, but $j_t \in \{i_1, \dots, i_{k-1}, i\}$, is $O(n^{-1})$ in expectation. Hence, the total contribution of such terms to the double sum in (5.5) is $O(n^{k-2})$ in expectation.

Furthermore, from Lemma 5.3, item ii), if (j_1, \dots, j_{k-1}) is such that for each t , either $j_t = i_t$ or $j_t \notin \{i_1, \dots, i_{k-1}, i\}$, we have, conditionally to i, j ,

$$\mathbb{P}[\sigma(i_\ell) = j_\ell, \forall 1 \leq \ell \leq k-1 \mid i, j] = (p_1)^{\#\{\ell: i_\ell = j_\ell\}} \left(\frac{1-p_1}{n}\right)^{\#\{\ell: i_\ell \neq j_\ell\}} (1 + O(\frac{1}{n})).$$

This gives

$$\begin{aligned} \mathbb{E}(C_1 \mid i, j) &= \sum_{\substack{i_1, \dots, i_{k-1} \in [n] \setminus \{i, j\} \\ \text{distinct}}} \sum_{\substack{j_1, \dots, j_{k-1} \in [n] \setminus \{i, j\} \\ \text{distinct}}} \\ &\left(\mathbb{1} \left[\text{Perm} \left(\left(\frac{i_1}{n}, \frac{j_1}{n} \right), \dots, \left(\frac{i_{k-1}}{n}, \frac{j_{k-1}}{n} \right), \left(\frac{i}{n}, \frac{j}{n} \right) \right) = \pi \right] \cdot (p_1)^{\#\{\ell: i_\ell = j_\ell\}} \left(\frac{1-p_1}{n}\right)^{\#\{\ell: i_\ell \neq j_\ell\}} \right) \\ &+ O(n^{k-2}). \end{aligned} \tag{5.6}$$

We let $(P_i^{(n)})_{i \leq k-1}$ be a vector of independent random variables with the following distribution, conditioned not to share coordinates:

- with probability p_1 , we set $P_1^{(n)} = (h/n, h/n)$, where h is uniform in $[n] \setminus \{i, j\}$;
- with probability $1 - p_1$, we set $P_1^{(n)} = (h/n, h'/n)$, where h, h' are uniform in $[n] \setminus \{i, j\}$ conditioned to satisfy $h \neq h'$;

Using this, (5.6) rewrites as

$$\mathbb{E}(C_1 \mid i, j) = \frac{1}{(k-1)!} n^{k-1} \mathbb{P}[\text{Perm}(P_1^{(n)}, \dots, P_{k-1}^{(n)}, (\frac{i}{n}, \frac{j}{n})) = \pi] + O(n^{k-2}).$$

But the tuple $(P_1^{(n)}, \dots, P_{k-1}^{(n)}, (\frac{i}{n}, \frac{j}{n}))$ converges in distribution to $(P_1, \dots, P_{k-1}, (x, y))$ as in Eq. (5.2). Since $(P_1, \dots, P_{k-1}, (x, y))$ have distinct coordinates with probability 1, and since the map $\mathbb{1}[\text{Perm}(\cdot, \dots, \cdot) = \pi]$ is continuous on the set of k -tuples of points with distinct coordinates, we get

$$\mathbb{E}(C_1 \mid i, j) = \frac{1}{(k-1)!} n^{k-1} \mathbb{P}[\text{Perm}(P_1, \dots, P_{k-1}, (x, y)) = \pi] + o(n^{k-1}).$$

It is easy to check that the error term is uniform in i, j , concluding the proof of the lemma. □

Symmetrically, letting C_2 be the number of occurrences of π in σ that use the position j but not i , we have

$$\mathbb{E}(C_2 \mid i, j) = \frac{n^{k-1}}{(k-1)!} \varphi_{p_1}^\pi(y, x) + o(n^{k-1}).$$

Moreover, it is easy to see that there are less than n^{k-2} occurrences of π in σ using simultaneously i and j , so that $C = C_1 + C_2 + O(n^{k-2})$. Summing up, we get that, provided that $i = xn + O(1)$ and $j = yn + O(1)$,

$$\mathbb{E}(C \mid i, j) = \frac{n^{k-1}}{(k-1)!} (\varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, x)) + o(n^{k-1}). \tag{5.7}$$

Since i, j are a uniform random pair of distinct integers in $[n]$, one can couple them with i.i.d. random uniform variables (U, V) in $[0, 1]$ such that $i = Un + O(1)$ and $j = Vn + O(1)$. Hence,

$$\text{Var} [\mathbb{E}(C \mid i, j)] = \frac{n^{2k-2}}{((k-1)!)^2} \text{Var} (\varphi_{p_1}^\pi(U, V) + \varphi_{p_1}^\pi(V, U)) + o(n^{2k-2}).$$

Using Lemma 5.1 and the fact that $\varphi_{p_1}^\pi$ is $(k - 1)$ -Lipschitz, we know that

$$K_2 := \text{Var}(\varphi_{p_1}^\pi(V, U) + \varphi_{p_1}^\pi(U, V)) > 0.$$

Setting $K'_2 = K_2/(k!)^2$, Eq. (5.3) implies that for any n sufficiently large,

$$v(\lambda) \geq \text{Var} [\mathbb{E}(C|i, j)] \geq K'_2 n^{2k-2}. \tag{5.8}$$

5.4 Improving the initial lower bound

Suppose that $p_2 \neq 0$. The lower bound (5.8), is not sufficient to prove that the limiting variance does not vanish. We will use the recursive inequality Eq. (5.4) to improve it. To this end, we first need to analyse the term $\mathbb{E}[\text{Var}(C|i, j)]$.

Lemma 5.4. *There exist $K_3, K_4 > 0$ such that for any $n > K_3$,*

$$\mathbb{E}[\text{Var}(C|i, j)] \leq K_4 n^{2k-3}.$$

Proof. Conditionally on i, j , we have

$$\begin{aligned} C = & \sum_{r=1}^k \sum_{\substack{i_1 < \dots < i_k \in [n] \setminus \{j\} \\ i_r = i}} \sum_{\substack{j_{\pi-1(1)} < \dots < j_{\pi-1(k)} \in [n] \setminus \{i\} \\ j_r = j}} B_{i_r, j_r} \prod_{\ell \neq r} B_{i_\ell, j_\ell} \\ & + \sum_{r=1}^k \sum_{\substack{i_1 < \dots < i_k \in [n] \setminus \{i\} \\ i_r = j}} \sum_{\substack{j_{\pi-1(1)} < \dots < j_{\pi-1(k)} \in [n] \setminus \{j\} \\ j_r = i}} B_{i_r, j_r} \prod_{\ell \neq r} B_{i_\ell, j_\ell} \\ & + \sum_{r_1=1}^k \sum_{\substack{r_2=1 \\ r_2 \neq r_1}}^k \sum_{\substack{i_1 < \dots < i_k \in [n] \\ i_{r_1} = i, i_{r_2} = j}} \sum_{\substack{j_{\pi-1(1)} < \dots < j_{\pi-1(k)} \in [n] \\ j_{r_1} = j, j_{r_2} = i}} B_{i_{r_1}, j_{r_1}} B_{i_{r_2}, j_{r_2}} \prod_{\ell \notin \{r_1, r_2\}} B_{i_\ell, j_\ell}. \end{aligned}$$

Conditionally on i, j , we have, in the two first term $B_{i_r, j_r} = B_{i, j} = 1$, and in the third one $B_{i_{r_1}, j_{r_1}} = B_{i, j} = 1 = B_{j, i} = B_{i_{r_2}, j_{r_2}}$. The other B factors have indices different from i and j . Since σ' is a uniform permutation on the conjugacy class indexed by λ' we have (Ψ, \mathbf{D}_{k-1}) weighted dependency graph for the random variables $\prod_{t=1}^{k-1} B_{i_t, j_t}$.

One can easily adapt the proof of Section 4 to obtain that the quantities R and T_ℓ associated with the above dependency graph satisfy $R \leq 32^{\binom{k-1}{2}} (n - 2)^{k-1}$ and $T_\ell \leq K_{7,k} n^{k-2}$ (note that the product $\prod_{t=1}^{k-1} B_{i_t, j_t}$ under consideration have degree $k - 1$). From Lemma 2.3, this implies $\text{Var}(C|i, j) \leq K_4 n^{2k-3}$. One can conclude since K_4 is independent of the choice of i and j . \square

Plugging in Eq. (5.8) and Theorem 5.4 in Eq. (5.4), we get that, for some constant $K_5 > 0$,

$$v(\lambda) \geq v(\lambda')(1 - K_5 n^{-1/2}) + K'_2 n^{2k-2}. \tag{5.9}$$

For, $i \leq m_2(\lambda)$, we denote by $\lambda^{(i)} = ((\lambda') \dots)'$ the partition obtained by removing i blocks of size 2 from λ . For $i \leq K_6 n^{1/2}$ (where K_6 is a positive constant), the inequality (5.9) holds substituting λ by $\lambda^{(i)}$ and we have

$$v(\lambda^{(i)}) \geq v(\lambda^{(i+1)})(1 - K_5(n - 2i)^{-1/2}) + K'_2(n - 2i)^{2k-2}.$$

We start from the initial inequality (5.8), namely $v(\lambda^{(K_6 \sqrt{n})}) \geq K'_2(n - K_6 \sqrt{n})^{2k-2}$, and iterate the above recursive inequality: we get

$$v(\lambda) \geq \sum_{i=0}^{K_6 \sqrt{n}-1} K'_2(n - 2i)^{2k-2} (1 - K_5(n - 2i)^{-1/2}) \dots (1 - K_5 n^{-1/2})$$

For K_6 sufficiently small, e.g. $K_6 = (2K_5)^{-1}$, all products

$$(1 - K_5(n - 2i)^{-1/2}) \cdots (1 - K_5 n^{-1/2})$$

are bounded away from 0, so that each of the $\Theta(\sqrt{n})$ terms in the sum behaves as $\Theta(n^{2k-2})$. Therefore, there exists a constant $K_7 > 0$ such that

$$v(\lambda) \geq K_7 n^{2k-3/2}. \tag{5.10}$$

Comparing with (5.1), we conclude that $f(p_1) + p_2 g(p_1) > 0$.

5.5 The case without cycles of size 2

When $p_2 = 0$ and $p_1 < 1$, the proof can be easily adapted. In such cases, one has $p_3 = \frac{1}{3}(1 - p_1 - 2p_2) > 0$. In particular, for $n > \frac{3}{1-p_1}$, λ contains a block of size 3, and we let $\hat{\lambda}$ be the partition obtained by removing a block of size 3 from λ . One can construct σ_λ as follows:

- Choose i, j, h distinct uniformly at random in $[n]$ and set $\sigma_i = j$, $\sigma_j = h$ and $\sigma_h = i$.
- Take a uniform random permutation $\hat{\sigma}$ of $[n] \setminus \{i, j, h\}$ of cycle type $\hat{\lambda}$ and set $\sigma_\ell = \hat{\sigma}_\ell$ for any $\ell \in [n] \setminus \{i, j, h\}$

Let \hat{C} be the number of occurrences of π that contains i or j or h . The counterpart of (5.7) is the following.

Lemma 5.5. *Let x, y, z be in $[0, 1]$ and $i = i(n)$, $j = j(n)$ and $h = h(n)$ be three sequences of positive integers chosen such that $i = xn + o(n)$, $j = yn + o(n)$ and $h = zn + o(n)$. Then*

$$\mathbb{E}(\hat{C} | i, j, h) = \frac{n^{k-1}}{(k-1)!} (\varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, z) + \varphi_{p_1}^\pi(z, x)) + o(n^{k-1}),$$

where $\varphi_{p_1}^\pi$ is defined in Eq. (5.2).

The remainder of the proof is the same. One need only to check the following variant of Lemma 5.1.

Lemma 5.6. *Assume π has size at least 2 and $p_1 < 1$. Then the function*

$$f : (x, y, z) \mapsto \varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, z) + \varphi_{p_1}^\pi(z, x)$$

is not constant.

Proof. Assume, for the sake of contradiction, that f is constant. Then $f(x, y, y) = f(x, x, y)$ for any x and y . This implies that $\varphi_{p_1}^\pi(x, x) = \varphi_{p_1}^\pi(y, y)$ for any x and y , i.e. that the function $x \mapsto \varphi_{p_1}^\pi(x, x)$ is constant. As a consequence, the expression

$$(x, y) \mapsto f(x, x, y) - \varphi_{p_1}^\pi(x, x) = \varphi_{p_1}^\pi(x, y) + \varphi_{p_1}^\pi(y, x)$$

should also be constant, leading to a contradiction with Lemma 5.1. □

References

- [1] E. Babson and E. Steingrímsson, *Generalized permutation patterns and a classification of the Mahonian statistics*, *Sém. Lothar. Combin.* **44** (2000), 117–134. MR1758852
- [2] M. Bóna, *On three different notions of monotone subsequences*, *Permutation Patterns*, London Math. Soc. Lecture Note Series, vol. 376, Cambridge University Press, 2010, pp. 89–113. MR2732825
- [3] V. Dubach, *A geometric approach to conjugation-invariant permutations*, arXiv:2402.10116, 2024.

- [4] C. Even-Zohar, *Patterns in random permutations*, *Combinatorica* **40** (2020), no. 6, 775–804 (English). MR4208097
- [5] V. Féray, *Weighted dependency graphs*, *Electronic Journal of Probability* **23** (2018), paper no. 93, 65 pp. MR3858921
- [6] V. Féray, *Central limit theorems for patterns in multiset permutations and set partitions*, *The Annals of Applied Probability* **30** (2020), no. 1, 287–323. MR4068312
- [7] J. Fulman, *The distribution of descents in fixed conjugacy classes of the symmetric groups*, *J. Comb. Theory, Ser. A* **84** (1998), no. 2, 171–180 (English). MR1652841
- [8] J. Fulman, G. Kim, and S. Lee, *Central limit theorem for peaks of a random permutation in a fixed conjugacy class of S_n* , *Ann. Comb.* **26** (2022), no. 1, 97–123 (English). MR4407056
- [9] Z. Hamaker and B. Rhoades, *Characters of local and regular permutation statistics*, Preprint arXiv:2206.06567, 2022.
- [10] L. Hofer, *A central limit theorem for vincular permutation patterns*, *Discrete Mathematics & Theoretical Computer Science* **19** (2018), no. 2, #9, Special Issue for the conference Permutation Patterns 2016. MR3783365
- [11] C. Houdré and R. Restrepo, *A probabilistic approach to the asymptotics of the length of the longest alternating subsequence*, *Electron. J. Combin.* **17** (2010), no. 1, Research Paper 168, 19. MR2745721
- [12] S. Janson, B. Nakamura, and D. Zeilberger, *On the asymptotic statistics of the number of occurrences of multiple permutation patterns*, *Journal of Combinatorics* **6** (2015), no. 1, 117–143. MR3338847
- [13] S. Janson, *Asymptotic normality for m -dependent and constrained U -statistics, with applications to pattern matching in random strings and permutations*, *Advances in Applied Probability* **55** (2023), no. 3, 841–894. MR4624030
- [14] M. Josuat-Vergès, *Combinatorics of the three-parameter PASEP partition function*, *Electron. J. Comb.* **18** (2011), no. 1, research paper P22, 31 p. (English). MR2770127
- [15] M.S. Kammoun, *Universality for random permutations and some other groups*, *Stochastic Processes Appl.* **147** (2022), 76–106 (English). MR4375687
- [16] G. Kim and S. Lee, *Central limit theorem for descents in conjugacy classes of S_n* , *J. Combin. Th., Series A* **169** (2020), 105–123. MR3998822
- [17] D. Romik, *Local extrema in random permutations and the structure of longest alternating subsequences*, *Proceedings of the 23rd international conference on formal power series and algebraic combinatorics, FPSAC 2011, Reykjavik, Iceland, June 13–17, 2011, Nancy: The Association. Discrete Mathematics & Theoretical Computer Science (DMTCS), 2011, pp. 825–834 (English). MR2820763*

Acknowledgments. Both authors want to thank Zachary Hamaker and Victor Dubach for insightful discussions and an anonymous referee for their careful reading of the article and their insightful comments.