# Ewens Sampling and Invariable Generation

GERANDY  BRITO[1],  CHRISTOPHER  FOWLER[1],  MATTHEW  JUNGE[2]

and  AVI  LEVY[3]

[1] School of Mathematics, Georgia Institute of Technology, 686 Cherry St, Atlanta, GA 30309, USA
(e-mail: gbrito3@gatech.edu)
[2] Department of Mathematics, Duke University, Durham, NC 27708, USA
(e-mail: jungem@math.duke.edu)
[3] Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, USA
(e-mail: me@avilevy.org)

We study the number of random permutations needed to invariably generate the symmetric group $S_n$ when the distribution of cycle counts has the strong $\alpha$-logarithmic property. The canonical example is the Ewens sampling formula, for which the special case $\alpha = 1$ corresponds to uniformly random permutations.

For strong $\alpha$-logarithmic measures and almost every $\alpha$, we show that precisely $\lceil (1 - \alpha \log 2)^{-1} \rceil$ permutations are needed to invariably generate $S_n$ with asymptotically positive probability. A corollary is that for many other probability measures on $S_n$ no fixed number of permutations will invariably generate $S_n$ with positive probability. Along the way we generalize classic theorems of Erdős, Tehran, Pyber, Łuczak and Bovey to permutations obtained from the Ewens sampling formula.

2010 *Mathematics subject classification*:  Primary 60C05
Secondary 12Y05, 68W20, 68W30, 68W40

## 1. Introduction

In 1934, van der Waerden [32] raised the question of the minimal number of uniformly random permutations required to invariably generate $S_n$ with positive probability. Recent breakthroughs by Pemantle, Peres and Rivin [31] and Eberhard, Ford and Green [15] have resolved van der Waerden's question, showing that precisely four permutations are required. We consider the same question in the context of more general permutation measures related to the cycle structure. The primary example is the Ewens sampling formula (1) introduced in [19]. It aligns a permutation's cycle structure with conditioned independent Poisson random variables and is defined in (1.1).

Elements $g_1, \ldots, g_m$ of a group $G$ *generate* the group if the smallest subgroup containing them is $G$ itself. The *symmetric group* $S_n$ is the group of permutations of the set $[n] := \{1, \ldots, n\}$. We

say a set of permutations $\pi_1, \ldots, \pi_m$ *generate* $S_n$ *invariably* if $\pi'_1, \ldots, \pi'_m$ generate $S_n$ whenever $\pi'_i$ is conjugate to $\pi_i$ for all $i$ [13]. Invariable generation is conjugacy invariant, and conjugacy classes of the symmetric group are precisely permutations of the same *cycle type*, defined to be the multiset of cycle lengths in the cycle decomposition of a permutation. A permutation fixes a set of size $k$ if some combination of the permutation cycle lengths sums up to $k$. If a collection of permutations $\pi_1, \ldots, \pi_m$ all fix sets of size $k$, then we can choose a set of conjugates $\pi'_1, \ldots, \pi'_m$ that all fix the same set of size $k$, and thus do not generate all of $S_n$. For this reason, invariable generation concerns only the cycle structure of permutations, rather than the action on $[n]$.

Recent developments show that cycle counts converge to independent Poisson random variables for many measures on $S_n$ as $n$ tends to infinity. Arratia, Barbour and Tavaré have shown that strong $\alpha$-logarithmic measures have this property and numerous further results [2, 3, 4, 5]. Measures with general cycle weights are considered in [8, 16] and related permutons are introduced in [25]. Following this work, Mukherjee used Stein's method to deduce that a variety of permutation measures have Poisson limiting cycle counts [29]. Furthermore, the structure of the short cycles in the Mallows measure was characterized in [21].

In addition to its intrinsic interest, the minimal number of permutations required for invariable generation has applications to computational Galois theory. One may use this quantity to estimate the run time of a Monte Carlo polynomial factorization algorithm [12, 24, 30, 31]. This establishes a connection between the irreducible factors of polynomials and the cycle structure of permutations. The cycle structure of a permutation is also compared with prime factorization of integers in the survey [22]. An analogy between divisors of large integers and invariable generation of permutations is then mentioned in the introduction of [15], postulating a link between the threshold for invariable generation and small divisors of a set of random integers.

Despite the age of van der Waerden's question, it was only recently shown that the number of uniformly random permutations required to invariably generate $S_n$ is bounded. A bound of $O(\sqrt{\log n})$ was first established in [13]. This was improved to $O(1)$ in [26], but the constant was large ($\approx 2^{100}$). Pemantle, Peres and Rivin reduced the bound to four in [31], which was then shown to be sharp by Eberhard, Ford and Green in [15].

We extend the results for uniform permutations to the Ewens sampling formula and beyond. Along the way, we extend results in [9, 17, 26], replacing counting arguments with a tool known as the Feller coupling (3.2) introduced in [20, p. 815].

## 1.1. Ewens sampling and the logarithmic property

The Ewens sampling formula appears throughout mathematics, statistics and the sciences. In the words of Harry Crane, it 'exemplifies the harmony of mathematical theory, statistical application, and scientific discovery'. This is stated in his survey [10]. It and its follow-up [11] give a nice tour of the formula's universal character, describing applications to evolutionary molecular genetics, the neutral theory of biodiversity, Bayesian non-parametrics, combinatorial stochastic processes and inductive inference, to name a few. The sampling formula also underpins foundational mathematics in number theory, stochastic processes and algebra.

We start by formally defining the *Ewens sampling formula*, as well as the strong $\alpha$-logarithmic property. Let $\mathrm{Poi}(\lambda)$ denote the law of a Poisson random variable with mean $\lambda$. Fix $\alpha > 0$ and let $X_1, \ldots, X_n$ be independent random variables in which $X_k$ has law $\mathrm{Poi}(\alpha/k)$. For a permutation $\pi \in S_n$ let $C_k$ denote the number of $k$-cycles in $\pi$. More precisely, $C_k$ is the number of disjoint

$k$-element subsets of $[n]$ on which $\pi$ restricts to a cyclic permutation. We say that the vector $\mathbf{C} = (C_1, \ldots, C_n)$ has distribution $\mathrm{ESF}(\alpha, n)$ if it satisfies the joint distribution

$$\mathbb{P}[C_1 = x_1, \ldots, C_n = x_n] = \mathbb{P}\left[X_1 = x_1, \ldots, X_n = x_n \,\Big|\, \sum_1^n k X_k = n\right]. \quad (1.1)$$

As observed in [19], this can be written explicitly as

$$\mathbb{P}[C_1 = x_1, \ldots, C_n = x_n] = \frac{n!}{\alpha_{(n)}} \prod_{j=1}^n \left(\frac{\alpha}{j}\right)^x \frac{1}{j} \frac{1}{x_j!} \mathbf{1}\{\textstyle\sum_{\ell=1}^n \ell \cdot x_\ell = n\},$$

where $\alpha_{(n)} = \alpha(\alpha+1)\cdots(\alpha+n-1)$. The formulation in (1.1) will be most useful for our purposes.

The tuple $(C_1, \ldots, C_n)$ uniquely specifies the cycle type of a permutation. Furthermore, it is easy to see that as $\pi$ ranges over $S_n$, the tuple $(C_1, \ldots, C_n)$ associated with $\pi$ ranges over the set

$$\left\{(C_1, \ldots, C_n) \in \{0, \ldots, n\}^n : \sum_{k=1}^n k C_k = n\right\}.$$

Thus, (1.1) specifies a measure on the conjugacy classes of $S_n$. We extend this to a measure on $S_n$ (also denoted by $\mathrm{ESF}(\alpha, n)$) by weighting all elements of the same class uniformly. Note that $\mathrm{ESF}(1, n)$ is the uniform measure, and therefore the results we establish regarding $\mathrm{ESF}(\alpha, n)$ will generalize the corresponding results about uniform permutations appearing in [15, 31].

A more general family that includes $\mathrm{ESF}(\alpha, n)$ is obtained by replacing the $X_k$ in (1.1) with an arbitrary sequence of independent non-negative integer valued random variables $Z_1, Z_2, \ldots$. To avoid intractable distortions when conditioning on $\sum k Z_k = n$, logarithmic growth of $\sum Z_k$ is traditionally assumed (see [3]). For our purposes, we require the $Z_k$ satisfy the *strong $\alpha$-logarithmic condition*:

$$\begin{aligned} |i\mathbb{P}[Z_i = 1] - \alpha| &< e(i)c_1, \\ i\mathbb{P}[Z_i = \ell] &\leqslant e(i)c_\ell, \quad l \geqslant 2, \end{aligned} \quad (1.2)$$

where $\sum i^{-1} e(i)$ and $\sum \ell c_\ell$ are both finite.

Besides Poisson, the most commonly used distributions for the $Z_i$ are the negative-binomial and binomial distributions. These are called assemblies, multisets and selections, respectively (see [3] for more discussion). The proof that all three of these different distributional choices for the $Z_i$ satisfy the strong $\alpha$-logarithmic property is in [3, Proposition 1.1].

It is no trouble to work at this level of generality, because the limiting cycle structure of any family satisfying the strong $\alpha$-logarithmic property is the same as that for an $\mathrm{ESF}(\alpha, n)$ permutation. We will describe this in more detail in Section 1.4, but in short, the number of $\ell$-cycles converges to an independent Poisson with mean $\alpha/\ell$.

## 1.2. Statement of Theorem 1.1

We state our result in terms of random variables $Z_1, Z_2, \ldots$ satisfying the strong $\alpha$-logarithmic property in (1.2). The example to keep in mind though is the measure $\mathrm{ESF}(\alpha, n)$, since all of these have the same limiting cycle counts.

Given such a collection let $\mu_n(\alpha, Z_1, Z_2, \ldots, Z_n)$ be the measure induced on $S_n$ via the relation in (1.1) with $X_k$ replaced by $Z_k$. We define $m_\alpha = m_\alpha(Z_1, Z_2, \ldots)$ to be the minimum number
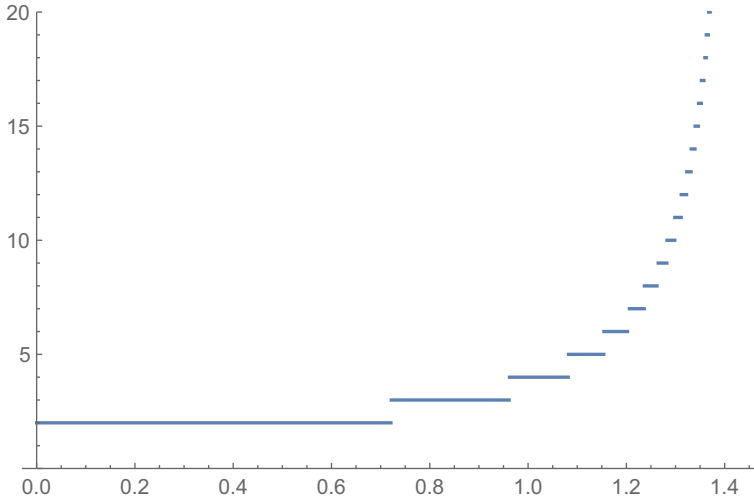
*Figure 1.* The graph $h(\alpha)$ defined in (1.4). We show that $m_\alpha = h(\alpha) = s_\alpha$, but are not sure about the value at the discontinuities of $h$. See Question 1 for more discussion.

of permutations sampled according to $\mu_n$ to invariably generate $S_n$ with positive probability as $n \to \infty$:

$$m_\alpha = \inf_{m \geqslant 2} \left\{ m : \inf_{n \geqslant 2} \mathbb{P}[\{\pi_1, \ldots, \pi_m\} \subset S_n, \pi_i \sim \mu_n, \text{ invariably generate } S_n] > 0 \right\}. \tag{1.3}$$

Our theorem gives a closed formula for $m_\alpha = h(\alpha)$, save for a countable exceptional set. Because we will reference it several times, we write the expression here:

$$h(\alpha) = \begin{cases} \lceil (1 - \alpha \log 2)^{-1} \rceil & 0 < \alpha < 1/\log 2, \\ \infty & \alpha \geqslant 1/\log 2. \end{cases} \tag{1.4}$$

See Figure 1.

**Theorem 1.1.** *Let $Z_1, Z_2, \ldots$ be a collection of random variables satisfying the strong $\alpha$-log-arithmic condition in (1.2). Let $m_\alpha = m_\alpha(Z_1, Z_2, \ldots)$, as in (1.3), be the minimum number of permutations to invariably generate $S_n$ with positive probability. Let $h(\alpha)$ be as in (1.4). For points of continuity of $h$ it holds that $m_\alpha = h(\alpha)$. At points of discontinuity we have $h(\alpha) \leqslant m_\alpha \leqslant h(\alpha) + 1$.*

This has broader implications beyond measures that have the strong $\alpha$-logarithmic property. We use it to deduce that any finite collection of random permutations with cycle counts asymptotically dominating a $\text{Poi}(1/j \log 2)$ random variable will fail to generate a transitive subgroup of $S_n$.

**Corollary 1.2.** *Suppose that $c_j \geqslant 1/\log 2$ and $X_k \sim \text{Poi}(c_j/j)$. Let $\kappa_n$ be a sequence of probability measures on $S_n$ such that for random permutations $\pi \sim \kappa_n$ the cycle structure, $\mathbf{C}$ satisfies*

$$d((C_1, \ldots, C_{k_0}), (X_1, \ldots, X_{k_0}))_{TV} \to 0$$

*for all fixed $k_0$. Then, the probability that any fixed number of permutations sampled according to $\kappa_n$ invariably generate a transitive subgroup of $S_n$ goes to 0 as $n \to \infty$.*

This corollary is particularly relevant to the results in [8] and [16]. They describe measures on $S_n$ that are formed by weighting cycle lengths by parameters $c_j$. The limiting cycle structure has $C_j \sim \mathrm{Poi}(c_j/j)$. Thus, we can create any limiting Poisson cycle structure we like – in particular, those satisfying Corollary 1.2.

Another source of alternative measures with limiting Poisson cycle counts comes from [29]. This fits into the larger body of work on permutons initiated in [25]. Of particular interest is Mallow's measure (introduced in [28]). This measure specifies a parameter $q_n > 0$ that biases towards more or less inversions in a permutation. A permutation $\pi$ has probability proportional to $q_n^{\mathrm{inv}(\pi)}$, where $\mathrm{inv}(\pi) := |\{(s,t)|s < t \text{ and } \pi s > \pi t\}|$.

The recent article [21] begins to characterize the cycle structure of Mallow's measure. They find that for $(1-q_n)^{-2} \ll n$ all cycles are of the order of $o(n)$. The high density of small cycles, along with [29, Theorem 1.4], which proves a limiting Poisson cycle profile, suggests that no finite collection of permutations sampled according to Mallow's measure in this regime will invariable generate $S_n$. Note that in the regime $(1-q_n)^{-2} \gg n$ the cycle counts converge to those of a uniformly random permutation. This is a new and exciting area. We are hopeful our result will find more applications as these objects become better understood.

### 1.3. Generalizations for the Ewens sampling formula

In proving Theorem 1.1 we connect an approximation sumset model (described in Section 1.4) back to fixed-set sizes in random permutations. This also takes place in [31] and [15]. These two articles use a small-cycle limit theorem for uniformly random permutations ([5, Theorem 1]). Namely, the number of $\ell$-cycles in a uniformly random permutation converges to a Poisson random variable with mean $1/\ell$ that is independent of the other small cycles.

Permutations sampled from $\mu_n$ have an analogous limit theorem ([3, Theorem 3.2]), except that now the $\ell$-cycles behave like Poisson random variables with mean $\alpha/\ell$. This lets us use similar ideas to analyse the corresponding sumset model. However, our approach diverges significantly when we connect back to random permutations, especially in obtaining an upper bound on $m_\alpha$. The difference is that both [31] and [15] have access to a large canon of results for uniformly random permutations. We do not. This requires a deep excavation where we extend classical results to $\pi \sim \mathrm{ESF}(\alpha,n)$. Along the way we prove many new estimates (see Lemmas 3.1, 3.2 and 3.3, Proposition 3.4 and Lemmas 3.9 and 3.11) for larger cycle lengths in $\mathrm{ESF}(\alpha,n)$ permutations.

We begin by showing that with high probability the only transitive subgroups containing a permutation $\pi$ sampled with distribution $\mathrm{ESF}(\alpha,n)$ are $A_n$ and $S_n$. This is the analogue of what [26, Theorem 1] proves for uniformly random permutations. This was recently studied in more detail by Eberhard, Ford and Green in [14]. We plan to explore this vein for $\mathrm{ESF}(\alpha,n)$ permutations in future work. For our current purposes, the following result suffices.

**Theorem 1.3.** *Let $\pi \sim \mathrm{ESF}(\alpha,n)$. Then with probability $1-o(1)$ the only transitive subgroups containing $\pi$ are $A_n$ and $S_n$.*

Pyber and Łuczak rely on an important theorem of Bovey regarding primitive subgroups. We need the analogue of it for $\text{ESF}(\alpha,n)$ permutations. The *minimal degree of* $\langle\pi\rangle = \{\pi^k : k \in \mathbb{Z}\}$ is the minimum number of elements of $\{1,\ldots,n\}$ displaced by some power of $\pi$ that is not the identity.

**Example 1.4.** *Consider* $\pi = (1234)(567)(89)$ *and* $\tau = (1234567)(89)$ *in* $S_9$. *The order of* $\pi$ *displacing the fewest elements of* $\{1,\ldots 9\}$ *is* $\pi^4 = (567)$, *so the minimal degree of* $\langle\pi\rangle$ *is* 3. *Meanwhile, the order of* $\tau$ *displacing the fewest elements is* $\tau^7 = (89)$, *so the minimal degree of* $\langle\tau\rangle$ *is* 2.

Theorem 1 of [9] says that if $\pi$ is a uniformly random permutation, then for each $\varepsilon > 0$ and $0 < \beta < 1$ we have

$$\mathbb{P}[\text{minimal degree of } \langle\pi\rangle > n^\beta] < C_{\varepsilon,\beta} n^{\varepsilon-\beta}.$$

We establish a weaker analogue of this result.

**Theorem 1.5.** *Let* $\pi \sim \text{ESF}(\alpha,n)$. *For each* $0 < \beta < 1$ *and any* $\alpha$ *it holds that*

$$\mathbb{P}[\text{minimal degree of } \langle\pi\rangle > n^\beta] = o(1).$$

Establishing Theorem 1.3 also requires a generalization of a classical theorem of Erdős and Turán (see [17, Theorem V]).

**Theorem 1.6.** *Let* $\pi \sim \text{ESF}(\alpha,n)$ *and* $\omega(n)$ *be an arbitrary function with* $\omega(n) \to \infty$. *The probability that the largest prime which divides* $\prod_{\ell=1}^n \ell^{C_\ell}$ *is larger than* $n\exp(-\omega(n)\sqrt{\log n})$ *is* $1 - o(1)$.

The proof of Theorem 1.3 uses both Theorems 1.5 and 1.6. For all three generalizations we follow a somewhat similar blueprint to their predecessors. However, the previous work often uses special features of uniformly random permutations. Counting arguments are heavily employed. In general, these techniques do not translate to $\text{ESF}(\alpha,n)$ permutations. Our approach is to use the Feller coupling (see Section 3.1) to directly relate the cycle structure to independent Poisson random variables. In some places this yields more elegant proofs, in others it becomes a bit technical. Keeping in mind that $\text{ESF}(1,n)$ is the uniform measure, this is a nice high-level approach to extending these results.

## 1.4. Overview of proof and result for sumsets

Because a fixed set's size is a sum of cycle lengths of a permutation, invariable generation is related to sumsets formed from random multisets. The link is developed over a collaborative arc of Arratia, Barbour and Tavaré [2, 3, 4, 6]. They study the relationship between $\alpha$-logarithmic structures and Poisson random variables. Most relevant for our purposes are the descriptions of the cycle counts $\mathbf{C} = (C_1,\ldots,C_n)$ for permutations sampled according to $\mu_n$.

In Theorems 3.1 and 3.2 of [3] they prove that for permutations induced by the strong $\alpha$-logarithmic property, the small cycle counts evolve to be independent Poisson random

variables, and the large cycle counts behave according to the Ewens sampling formula. Furthermore, Lemma 1 of [2] (also discussed in [4]) shows that the Ewens sampling formula can be cleanly related via the Feller coupling to independent $\text{Poi}(\alpha/k)$ random variables. The pay-off is that we can model the sizes of fixed sets in a permutation with independent Poisson random variables.

Let $\mathbf{X}(\alpha) = (X_1, X_2, \ldots)$ be a sequence of Poisson random variables where $X_j$ has mean $\alpha/j$. The coupling discussed in the previous paragraph allows us to model the sizes of fixed sets in a permutation with the random sumset

$$\mathcal{L}(\mathbf{X}(\alpha)) = \left\{ \sum_{j \geqslant 1} j x_j : 0 \leqslant x_j \leqslant X_j \right\}.$$

Returning to our opening question, we can bound the required number of permutations from above if no fixed set size is common to all permutations. And we can bound the number from below if there is guaranteed to be a common fixed set size. Consider independent realizations $\mathbf{X}^{(1)}(\alpha), \ldots, \mathbf{X}^{(m)}(\alpha)$. The upper and lower bounds correspond to finding extremal values of $m$ such that

(i) $\bigcap_{i=1}^{m} \mathcal{L}(\mathbf{X}^{(i)}(\alpha)) = \{0\}$ with positive probability, and

(ii) $|\bigcap_{i=1}^{m} \mathcal{L}(\mathbf{X}^{(i)}(\alpha))| = \infty$ almost surely.

Now the cycle counts of these permutations converge to independent Poisson random variables [2, 3, 4, 5]. Thus common elements among independent sumsets correspond to common fixed-set sizes of independent permutations. So invariable generation is analogous to a trivial sumset intersection. On the other hand, an infinite intersection means that common fixed-sets persist among the permutations. For the case $\alpha = 1$, the results of [31, Theorem 1.6] and [15, Corollary 2.5] each imply that four sets suffice in (i). That (ii) holds with three sets is due to [15, Corollary 3.9]. Here is the analogue for general $\alpha > 0$.

**Theorem 1.7.** *Let* $s_\alpha := \inf\{m : \mathbb{P}[\cap_1^m \mathcal{L}(\mathbf{X}^{(i)}(\alpha)) = \{0\}] > 0\}$ *be the smallest number of i.i.d. sumset intersections so that the resulting set is trivial with positive probability. Let* $h(\alpha)$ *be as in* (1.4)*, and* $m_\alpha$ *as in Theorem 1.1. At points where $h$ is continuous we have $s_\alpha = h(\alpha)$, and at the discontinuities* $h(\alpha) \leqslant s_\alpha \leqslant h(\alpha) + 1$.

To establish this result for sumsets, we build upon the ideas in [31] and [15] to prove matching lower and upper bounds for $s_\alpha$. These bounds can be found in Propositions 4.2 and 5.1. Parts of the argument are as simple as swapping 1s for $\alpha$s, but in others, particularly the lower bound, some care is required.

A relevant quantity for establishing the upper bound is $p_k = p_k(\alpha)$, the probability that the element $k$ appears in $\mathcal{L}(\mathbf{X})$. Estimating $p_k(1)$ is the major probabilistic hurdle in [31]. There are two difficulties. The first is called a lottery effect: certain unlikely events greatly skew $p_k$. This is circumvented by using *quenched* probabilities $\tilde{p}_k$, which ignore an $o(1)$ portion of the probability space (that $\sum X_j$ and $\sum j X_j$ are uncharacteristically large). With this restriction, it is shown in [31, Lemma 2.3] that $\tilde{p}_k(1) \leqslant k^{\log(2)-1}$. Notice that $4(\log 2 - 1) < -1$ and a Borel–Cantelli-type argument implies finiteness of the intersection of the $m$-independent $\mathcal{L}(\mathbf{X}^{(i)}(\alpha))$. Establishing the quenched formula is the second hurdle. It requires a clever counting and partitioning of

$\mathscr{L}(\mathbf{X})$. We generalize these ideas in Lemma 4.1 to obtain the upper bound

$$\tilde{p}_k(\alpha) \leqslant C_\varepsilon k^{-1+\alpha\log 2+\varepsilon}$$

for any $\varepsilon > 0$. Taking into account the lottery effect, and ignoring a vanishing portion of the probability space, we can write $\mathbb{P}[k \in \cap_1^m \mathscr{L}(\mathbf{X}^{(i)})] = \prod_1^m \tilde{p}_k$. When $m \geqslant \lceil (1 - \alpha\log 2 + \varepsilon)^{-1} \rceil$, the product is smaller than $n^{-1}$, and the probabilities are summable. With this, the Borel–Cantelli lemma implies the intersection is almost surely finite, and thus it is trivial with positive probability.

The lower bound in Proposition 5.1 requires more care. We generalize the framework in [15, Section 3] to higher dimensions to incorporate more than three sumsets as well as different values of $\alpha$. Given $I \subseteq \mathbb{N}$ and $m$ sumsets we consider the set of differences

$$S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}) := \{(n_1 - n_m, n_2 - n_m, \ldots, n_{m-1} - n_m) : n_i \in \mathscr{L}(\mathbf{X}^{(i)}) \cap I\}.$$

We use Fourier analysis (as in [15] and [27]) to study a smoothed version of the indicator function of this set. In particular, we will work with its Fourier transform $F : \mathbb{T}^{m-1} \to \mathbb{C}$. After obtaining a pointwise bound on $F$, we integrate around that torus to obtain, for sufficiently large $k$ and $I = (k^{1-\beta}, k]$, that $|S(I, M_1, \ldots, M_m)| \gg k^{m-1}$ and lies in the cube $[-ck, ck]^{m-1}$ for some $c > 0$. The density of these sets is sufficiently high that if we look in a larger interval, with positive probability we can find $j_1, \ldots, j_m \in (k, Dk]$ such that $X_{j_1}^{(1)}, \ldots, X_{j_m}^{(m)} > 0$ and

$$(j_m - j_1, \ldots, j_m - j_{m-1}) \in S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}).$$

Combining this with the corresponding point in the set of differences then produces a number in $\mathscr{L}(\mathbf{X}^{(1)}) \cap \cdots \cap \mathscr{L}(X^{(m)})$ made solely from values in $(k^{1-\beta}, Dk]$ for some constants $\beta, D > 0$. By partitioning $\mathbb{N}$ into infinitely many disjoint sets of this form and repeating this argument for each of them, we conclude that the set of intersections is almost surely infinite.

## 1.5. Further questions

There is still much to be done for both components of this paper – random permutations and random sumsets. Recall that Theorems 1.1 and 1.7 do not pin down the exact behaviour at points of discontinuity of $h$. A glaring question is to characterize $m_\alpha$ and $s_\alpha$ at these points. We are doubtful our approach generalizes. Characterizing the behaviour at the critical values is likely difficult, and will require a new idea.

**Question 1.** *What are $m_\alpha$ and $s_\alpha$ at discontinuity points of $h$?*

For random permutations we would like a more complete characterization than in Corollary 1.2, in particular an upper bound.

**Question 2.** *Show that if a permutation has cycle counts that are asymptotically dominated by Poisson random variables with mean $h^{-1}(2)/k$, then two will invariably generate $S_n$ with positive probability.*

Corollary 1.2 shows that for many permutation measures no finite number will invariably generate $S_n$. Possibly if we let the number of permutations grow with $n$ we will see different behaviour.

**Question 3.** *Let $\alpha > 1/\log 2$. Prove that there are constants $\beta < \beta'$ such that the probability $\beta \log n$ Ewens-$\alpha$ permutations generate $S_n$ is bounded away from 0, while $\beta' \log n$ fail to generate $S_n$ with high probability?*

We believe that $\log n$ is the correct order by the following approximation model. When $\alpha > 1/\log 2$, the fixed set sizes in a random permutation are dense in $[n]$. So, we can model the occurrence of each fixed size by independent Bernoulli random variables with some parameter $p_\alpha > 0$. If one performs $N$ independent thinnings of $[n]$ by including each integer $k$ with probability $p_\alpha$, then the probability that $k$ belongs to all $N$ subsets is $p_\alpha^N$. Hence, the probability that there are no common elements in the $N$-fold intersection is $(1 - p_\alpha^N)^n$. If we choose $N = c \log n$ this will converge to either 0 or a positive number as we increase $c$. The difficulty in answering the question for permutations is that we do not have independence for fixing different fixed set sizes.

We also have a question in the simpler setting of generating $S_n$. Dixon's theorem [13] implies that two uniformly random permutations will generate either $A_n$ or $S_n$ with high probability. Does the analogue hold for $\mathrm{ESF}(\alpha, n)$ permutations? Because generation relies on more than just cycle structure, this may be a hard question.

**Question 4.** *How many $\mathrm{ESF}(\alpha, n)$ permutations are needed to generate one of $A_n$ or $S_n$ with high probability?*

### 1.6. Notation

Note that $\alpha > 0$ is fixed throughout. As already discussed, we denote cycle counts of a permutation by $\mathbf{C} = (C_1, \ldots, C_n)$ with $C_j$ the number of $j$-cycles. Unless otherwise noted, our permutations and cycle vectors come from an $\mathrm{ESF}(\alpha, n)$ distribution. We will let $\mathbf{Y}(\alpha)$ denote a vector of independent $\mathrm{Poi}(\alpha/k)$ random variables that are obtained from the Feller coupling in (3.2). We will also take $\mathbf{X}(\alpha)$ to be a vector of independent $\mathrm{Poi}(\alpha/k)$ random variables. This is done to distinguish the permutation and random sumset settings. Henceforth we will suppress the $\alpha$ dependence unless there is reason to call attention to it. For an infinite vector, such as $\mathbf{Y}$, we will use the notation $\mathbf{Y}[i, j]$ to denote the sub-vector $(Y_i, Y_{i+1}, \ldots, Y_j)$.

For $I$ an interval in $\mathbb{N}$ define the sumset

$$\mathscr{L}(I, \mathbf{X}) := \left\{ \sum_{j \in I} j x_j : 0 \leqslant x_j \leqslant X_j \right\},$$

and set $\mathscr{L}(\mathbf{X}) := \mathscr{L}(\mathbb{N}, \mathbf{X})$. Given $\mathbf{X}$ define

$$f_{\ell,k}(\mathbf{X}) = \sum_{\ell < j \leqslant k} X_j, \quad g_{\ell,k}(\mathbf{X}) = \sum_{\ell < j \leqslant k} j X_j. \tag{1.5}$$

It will be convenient to abbreviate $f_k := f_{1,k}(\mathbf{X})$ and $g_k := g_{1,k}(\mathbf{X})$. Thus, $f_k$ is the number of available summands smaller than $k$, and $g_k$ is the largest sum attainable with them.

We will use the notation $f \ll g$ and $f(n) = O(g(n))$ interchangeably for the existence of $c$ such that for all large enough $n$ it holds that $f(n) \leqslant cg(n)$. When the $c$ depends on our choice of $\varepsilon$ we will use $f \ll_\varepsilon g$, similarly $f \ll_{m,\varepsilon}$ denotes dependence of the constant on both $\varepsilon$ and the number of intersections $m$. We use the little-$o$ notation $f(n) = o(g(n))$ to mean that $f(n)/g(n) \to 0$. Also, $f(n) = \tilde{O}(g(n))$ means there is $k > 0$ such that $f(n) = O(\log^k(n)g(n))$. The symbol $f(n) \approx g(n)$ means that $f(n)/g(n) \to c$ for some $c > 0$. A sequence of events $E_n$ occurs *with high probability* if $\liminf \mathbb{P}[E_n] = 1$. Also $i \mid j$ means that $i$ divides $j$ (*i.e.* there is an integer $k$ with $ik = j$.)

## 1.7. Outline of paper

We will lead off with a proof of Theorem 1.1. This comes in two parts: an upper bound from Theorem 1.7, and a matching lower bound in Proposition 2.3. The upper bound in Theorem 1.7 has two distinct components. We connect random sumsets to $\mathrm{ESF}(\alpha, n)$ permutations in Section 3.1. This section includes the proofs of Theorems 1.3, 1.5 and 1.6. The second component is analysing the random sumset model. We do this in Section 4 and provide an upper bound on $s_\alpha$ at Proposition 4.2. The lower bound is Proposition 2.3 and is proved in Section 5.

## 2. Proof of Theorem 1.1

First we show that $h(\alpha) + 1$ is an upper bound for the number of $\mathrm{ESF}(\alpha, n)$ permutations to invariably generate $S_n$. We will use a few lemmas and propositions that come later in the paper, but prove this now to give the reader a sense of the important bounds needed.

**Proposition 2.1.** *Suppose $\alpha$ is a continuity point of h. A collection of $h(\alpha) < \infty$ independent permutations sampled according to an $\mathrm{ESF}(\alpha, n)$ invariably generates $S_n$ with positive probability as $n \to \infty$.*

**Proof.** Suppose $m = h(\alpha)$. Call the permutations $\pi^{(1)}, \ldots, \pi^{(m)}$. Condition $\pi^{(1)}$ to be odd. Lemma 3.11 guarantees this happens with probability bounded away from zero. Define the events

$$E_{n,k} = \{\pi^{(1)}, \ldots, \pi^{(m)} \text{ fix an } \ell\text{-set for some } k \leqslant \ell \leqslant n/2\},$$
$$F_{n,k} = \{\pi^{(1)}, \ldots, \pi^{(m)} \text{ fix an } \ell\text{-set for some } 1 \leqslant \ell \leqslant k\}.$$

The probability of a common fixed set size amongst all of the permutations is bounded above by

$$\mathbb{P}[E_{n,k} \cup F_{n,k}] \leqslant \mathbb{P}[E_{n,k}] + \mathbb{P}[F_{n,k}]. \tag{2.1}$$

Therefore, the $\pi^{(i)}$ will invariably generate a transitive subgroup of $S_n$ if this quantity is less than 1. By Theorem 1.3 we know this transitive subgroup is with high probability either $A_n$ or $S_n$. Because $\pi^{(1)}$ is odd, it must be $S_n$.

Now $\alpha$ is a continuity point of $h$, so

$$m = h(\alpha) = \lceil (1 - \alpha \log 2)^{-1}) \rceil > (1 - \alpha \log 2)^{-1}.$$

It follows that there exists some $\varepsilon > 0$ such that

$$m = (1 - (\alpha + \varepsilon) \log 2)^{-1}.$$

Fix $\varepsilon'$ such that $0 < \varepsilon' < \varepsilon$. By Lemma 3.10 there exists a $c$ such that for any $k$ and $n$ and $i = 2, \ldots, m$,

$$\mathbb{P}[\pi^{(i)} \text{ fixes a } k\text{-set}] \leqslant ck^{-\delta},$$

where

$$\delta = 1 - (\alpha + \varepsilon') \log 2.$$

Similarly, by Corollary 3.13 for any $k$,

$$\mathbb{P}[\pi^{(1)} \text{ fixes a } k\text{-set}] \leqslant ck^{-\delta} \cdot \max\left\{ \frac{\alpha+1}{\alpha}, \frac{\alpha+1}{1} \right\}.$$

By independence of the $\pi^{(i)}$, we conclude

$$\mathbb{P}[\pi^{(1)}, \ldots, \pi^{(m)} \text{all fix a } k\text{-set}] \leqslant (ck^{-\delta})^m \cdot \max\left\{ \frac{\alpha+1}{\alpha}, \frac{\alpha+1}{1} \right\}.$$

By construction we have that $\delta m > 1$. A union bound for $k_0 \leqslant k \leqslant n/2$ ensures that

$$\mathbb{P}[E_{n,k_0}] \leqslant Ck_0^{1-\delta m}. \tag{2.2}$$

As $\delta m > 1$ it follows from (2.2) that we can make $\mathbb{P}[E_{n,k}]$ as small as we like by fixing a large $k$.

Now we turn our attention to bounding the probability of $F_{n,k}$. We will show that there exists $\beta > 0$ such that

$$\inf_{k \geqslant 1} \liminf_{n \to \infty} \mathbb{P}[F_{n,k}] < 1 - \beta. \tag{2.3}$$

Once we have (2.3) we apply (2.2) with $k_0$ large enough so that $\mathbb{P}[E_{n,k_0}] < \beta/2$. This then gives (2.1) is strictly less than 1 as $n \to \infty$, as desired.

It remains to establish (2.3). Fix $k$ and let $\pi$ be an ESF$(n, \alpha)$. The Feller coupling in (3.2) ensures that each cycle count $C_\ell \leqslant Y_\ell + \mathbf{1}\{J_n = 0\}$ for all $1 \leqslant \ell \leqslant n$ for a certain random variable $J_n$ (defined just above (3.2)). Thus, with high probability we have

$$C_\ell \leqslant Y_\ell \text{ for all } 1 \leqslant \ell \leqslant k.$$

Since this dominance relation is with high probability, it also applies to $\pi^{(1)}$ despite being conditioned to be odd. This is because $\{\pi^{(1)} \text{ is odd}\}$ occurs with positive probability by Lemma 3.11.

The event $F_{n,k}^c$ can be coupled so that it occurs when $m$ independent Poisson sumsets have empty intersection. Theorem 1.7 ensures that this occurs with some positive probability which is independent of $k$ (since we are considering infinite sets). This probability provides the claimed $\beta$ in (2.3) and completes the argument. $\qquad\square$

**Remark 2.2.** The argument above implies that $m_\alpha \leqslant m_{\alpha+\eta}$ for all $\eta > 0$, since any appropriate choice of $\varepsilon'$ in the latter case will hold for the former. It follows that $h(\alpha) + 1$ permutations will invariably generate $S_n$ with positive probability at discontinuity points of $h$.

Now we prove a matching lower bound.

**Proposition 2.3.** *A collection of $m < h(\alpha)$ permutations sampled independently according to an $\mathrm{ESF}(\alpha, n)$ with high probability do not invariably generate $S_n$, or any transitive subgroup of $S_n$.*

**Proof.** Let $m < h(\alpha)$, and fix some $\varepsilon > 0$. By Proposition 5.1 the intersection of the i.i.d. sumsets with parameter $\alpha$, $\bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)})$, is almost surely infinite. It follows that there is some $k_0 = k_0(\varepsilon)$ such that if $k \geqslant k_0$ then

$$\bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)}) \cap [1, k] \neq \varnothing$$

with probability at least $1 - \varepsilon/2$.

Fix some $k \geqslant k_0$ and call the random permutations $\pi^{(1)}, \ldots, \pi^{(m)}$. By [3, Theorem 3.1] the cycle counts $\mathbf{C}^{(i)}[1, k]$ of $\pi^{(i)}$ converge in total variation to $\mathbf{X}^{(i)} \cap [1, k]$. Then there exists some $n_0 = n_0(\varepsilon, k)$ such that if $n \geqslant n_0$, with probability greater than $1 - \varepsilon/2$ we have $\mathbf{C}^{(i)}[1, k] = \mathbf{X}^{(i)} \cap [1, k]$ for all $i = 1, \ldots, m$. Since

$$\bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)}) \cap [1, k] \subset \bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)} \cap [1, k]),$$

with probability at least $1 - \varepsilon$ we have that $\pi^{(1)}, \ldots, \pi^{(m)}$ each fix a set of size $\ell$ for some $\ell \leqslant k$. As $\varepsilon$ was arbitrary, this completes our proof. $\square$

Theorem 1.1 is a straightforward consequence of the last two propositions.

**Proof of Theorem 1.1.** Together, Theorems 3.1 and 3.2 of [3] guarantee that the cycle counts of any measure satisfying the strong $\alpha$-logarithmic property converge in total variation to the same distribution as an $\mathrm{ESF}(\alpha, n)$ permutation. Hence it suffices to consider $\mathrm{ESF}(\alpha, n)$ permutations.

By Proposition 2.1 we have $m_\alpha \leqslant h(\alpha)$ at point of continuity of $h$. At points of discontinuity, we use Remark 2.2 to deduce that $m_\alpha \leqslant m_{\alpha+\eta} \leqslant h(\alpha) + 1$. Then, by Proposition 2.3 we have that $h(\alpha) \leqslant m_\alpha$ for all $\alpha$, completing our proof. $\square$

## 3. The upper bound for $m_\alpha$

The main goal of this section is to prove that $m_\alpha \leqslant s_\alpha$. Our primary tool for relating Poisson sumsets to permutations is the Feller coupling. This is described at the onset of Section 3.1. We then use it in that section to prove several estimates on the cycle structure of $\mathrm{ESF}(\alpha, n)$ permutations. With these we prove Theorems 1.3, 1.5 and 1.6 in Section 3.2. We put all of this together and establish Proposition 2.1 in Section 3.3.

### 3.1. The Feller Coupling

The Ewens sampling formula can be obtained via an elegant coupling attributed to Feller [20, p. 815] for uniform permutations. The articles [2, 4] have nice descriptions. We start with a sequence $\xi_1, \xi_2, \ldots$ of mixed Bernoulli random variables with $\mathbb{P}[\xi_i = 1] = \alpha/(\alpha + i - 1)$ and $\mathbb{P}[\xi_i = 0] = (i - 1)/(\alpha + i - 1)$. Define an $\ell$-*spacing* to be any occurrence of $\ell - 1$ zeros enclosed

by a 1 on the left and right (*e.g.* $1,0,0,0,1$ is a 4 spacing). Remarkably, the total number of $\ell$-spacings, $Y_\ell$, is distributed as $\text{Poi}(\alpha/\ell)$. Moreover, the collection $Y_1, Y_2, \ldots$ are independent (see [2, 4]).

The counts of spacings in $\xi_1, \ldots, \xi_n, 1$ generate a partition of $n$, which can be filled in uniformly randomly to form a permutation with $C_\ell$ cycles of size $\ell$, where $C_\ell$ are sampled according to an $\text{ESF}(\alpha, n)$ distribution. Explicitly,

$$C_\ell = \text{the number of } \ell\text{-spacings in } \xi_1, \ldots, \xi_n, 1$$

$$= \xi_{n-\ell+1}(1-\xi_{n-\ell+2})\cdots(1-\xi_n) + \sum_{k=1}^{n-\ell} \xi_k(1-\xi_{k+1})\cdots(1-\xi_{k+\ell-1})\xi_{k+\ell}. \qquad (3.1)$$

This gives an intuition for why $\text{ESF}(1, n)$ corresponds to a uniformly random permutation. Indeed, we can inductively construct such a permutation by letting $\xi_i$ indicate the decision to complete a cycle, when there is an $i$-way choice for the next element.

Letting $R_n$ be the index of the rightmost 1 in $\xi_1, \ldots, \xi_n$ and $J_n = n + 1 - R_n$, it follows from the previous discussion that

$$C_\ell \leqslant Y_\ell + \mathbf{1}\{J_n = \ell\}. \qquad (3.2)$$

This says that the cycle counts can be obtained from independent Poisson random variables through a random number of deletions, and possibly one insertion. We require a few estimates on the behaviour of these perturbations.

**Lemma 3.1.** *Let $J_n$ be as defined above, and let $D_n = \sum_{\ell=1}^n Y_\ell - \mathbf{1}\{J_n = \ell\} - C_\ell$ be the number of deletions. The following four inequalities hold:*

$$\mathbb{P}[J_n = \ell] \leqslant \frac{\alpha}{n-\ell}, \qquad \ell < n. \qquad (3.3)$$

$$\mathbb{P}[J_n = \ell] \leqslant cn^{-\alpha(1-\gamma)}, \qquad n - n^\gamma < \ell < n, \quad 0 < \gamma < 1, \quad \textit{for some } c = c(\gamma) > 0. \qquad (3.4)$$

$$\mathbb{P}[J_n = \ell] = O(n^{-\alpha/(1+\alpha)}), \quad \ell < n. \qquad (3.5)$$

$$\mathbb{E}D_n = O(1). \qquad (3.6)$$

**Proof.** We can write the density of $J_n$ explicitly for all $\ell < n$ as

$$\mathbb{P}[J_n = \ell] = \frac{\alpha}{\alpha+n-\ell} \prod_{n-\ell+2}^n \frac{i-1}{\alpha+i-1} \leqslant \prod_{n-\ell+1}^n \frac{i-1}{\alpha+i-1}. \qquad (3.7)$$

As $\alpha > 0$ we can use the leading term $\alpha/(\alpha+n-\ell) < \alpha/(n-\ell)$ to obtain (3.3).

Fix some $\gamma \in (0,1)$. For the case $\ell > n - n^\gamma$ in (3.4) we use that when $J_n = \ell$ we must have $\xi_{n-\ell+1} = 1$ and $\xi_k = 0$ for all larger index terms up to $\xi_n$. The probability of this can be computed explicitly as

$$\mathbb{P}[J_n = \ell] = \frac{\alpha}{\alpha+n-\ell} \prod_{n-\ell+2}^n \frac{i-1}{\alpha+i-1} \leqslant \prod_{n-\ell+1}^n \frac{i-1}{\alpha+i-1}.$$

We estimate the product by converting it to the sum

$$\prod_{n-\ell+1}^n \frac{i-1}{\alpha+i-1} = \exp\left(-\sum_{n-\ell+1}^n \log \frac{\alpha+i-1}{i-1}\right) \leqslant \exp\left(-\int_{n-\ell+1}^n \log \frac{\alpha+x-1}{x-1} dx\right).$$

The last inequality follows from the fact that $\log((\alpha+x-1)/(x-1))$ is a decreasing function, and the sum can be viewed as a right-sided Riemann (over-)approximation. This integral has the closed form

$$-\int_{n-\ell+1}^{n} \log\left(\frac{\alpha+x-1}{x-1}\right) dx = -\log\left(\frac{(\alpha+n-1)^{\alpha+n-1}}{(n-1)^{n-1}}\right) + \log\left(\frac{(\alpha+n-\ell)^{\alpha+n-\ell}}{(n-\ell)^{n-\ell}}\right).$$

Hence, exponentiating the above line gives

$$\prod_{n-\ell+1}^{n} \frac{i-1}{\alpha+i-1} \leqslant \frac{(n-1)^{n-1}}{(\alpha+n-1)^{\alpha+n-1}} \times \frac{(\alpha+n-\ell)^{\alpha+n-\ell}}{(n-\ell-1)^{n-\ell}}. \tag{3.8}$$

The leftmost product is less than $n^{-\alpha}$. The rightmost product can be written as

$$(\alpha+n-\ell)^{\alpha}\left(\frac{n-\ell+\alpha}{n-\ell-1}\right)^{n-\ell} = (\alpha+n-\ell)^{\alpha}\left(\frac{1+\alpha/(n-\ell)}{1-1/(n-\ell)}\right)^{n-\ell}. \tag{3.9}$$

The rightmost product in (3.9) is asymptotic to $e^{\alpha}/e^{-1}$ as $n-\ell \to \infty$, and thus may be bounded by some constant for all values of $n-\ell$. Using the hypothesis that $n-\ell \leqslant n^{\gamma}$, it follows that (3.9) is universally bounded by $cn^{\alpha\gamma}$ for some constant $c$. We conclude that (3.8) is less than $cn^{-\alpha+\alpha\gamma} = cn^{-\alpha(1-\gamma)}$. This establishes (3.4).

To obtain the universal bound in (3.5), observe when $\ell \leqslant n-n^{\alpha/(1+\alpha)}$ we can input this into (3.3) to obtain

$$\mathbb{P}[J_n = \ell] \leqslant \frac{\alpha}{n-(n-n^{\alpha/(1+\alpha)})} = O(n^{-\alpha/(1+\alpha)}).$$

And, for $\ell > n-n^{\alpha}$ we substitute $\gamma = \alpha/(1+\alpha)$ into (3.4) to obtain the same asymptotic inequality: $O(n^{-\alpha/(1+\alpha)})$.

Lastly, (3.6) is proved in [2, Theorem 2]. $\qquad\square$

We will also require a bound on the joint probability that two cycles occur simultaneously. There will be some poly-log terms that, besides being a minor nuisance, do not change the tack of our proof. Recall the $\tilde{O}$ notation that ignores logarithmic contribution. It is described carefully in Section 1.6.

**Lemma 3.2.** *For any $i < j \leqslant n$ it holds that*

$$\mathbb{P}[C_i C_j > 0] = \tilde{O}(i^{-1}j^{-\alpha/(1+\alpha)}).$$

**Proof.** We start by using the formula in (3.1) to bound $C_i$ with the $\xi_k$:

$$C_i \leqslant \xi_{n-i+1}(1-\xi_{n-i+2})\cdots(1-\xi_n) + \sum_{k=1}^{n-i} \xi_k \xi_{k+i}.$$

Thus,

$$
C_i C_j \leqslant \left( \xi_{n-i+1}(1-\xi_{n-i+2})\cdots(1-\xi_n) + \sum_{k=1}^{n-i} \xi_k \xi_{k+i} \right)
$$
$$
\times \left( \xi_{n-j+1}(1-\xi_{n-j+2})\cdots(1-\xi_n) + \sum_{\ell=1}^{n-j} \xi_\ell \xi_{\ell+j} \right)
$$
$$
= (x+y)(z+w).
$$

The last line is just labelling the terms in the product immediately above in the natural way (*i.e.* $y = \sum_{k=1}^{n-i} \xi_k \xi_{k+i}$). This is so we can bound $\mathbb{E}[xz + xw + yw + yz]$ in an organized fashion, one term at a time.

**Term** *xz*. First, consider

$$
\xi_{n-i+1}(1-\xi_{n-i+2})\cdots(1-\xi_n)\xi_{n-j+1}(1-\xi_{n-j+2})\cdots(1-\xi_n).
$$

When $i < j$ this product contains the term $\xi_{n-i+1}(1-\xi_{n-i+1}) \equiv 0$, thus it is always zero. We conclude that

$$
\mathbb{E}[xz] \equiv 0. \tag{3.10}
$$

**Term** *xw*. Next, we consider the cross-term

$$
\xi_{n-i+1}(1-\xi_{n-i+2})\cdots(1-\xi_n)\sum_{\ell=1}^{n-j} \xi_\ell \xi_{\ell+j}.
$$

Many of the summands are zero. In fact, the overlapping terms with $n-j-i+1 < \ell \leqslant n-j$ are identically zero. So, for the sake of obtaining disjoint terms, we consider the smaller sum

$$
\xi_{n-i+1}(1-\xi_{n-i+2})\cdots(1-\xi_n)\sum_{\ell=1}^{n-j-i+1} \xi_\ell \xi_{\ell+j}.
$$

Aside from the $\ell = n-j-i+1$ term, all the indices are all disjoint. Ignoring this term for now, the expected value is

$$
\frac{\alpha}{\alpha+n-i}\frac{n-i+1}{\alpha+n-i+1}\cdots\frac{n-1}{\alpha+n-1}\sum_{\ell=1}^{n-j-i}\frac{\alpha^2}{(\alpha+\ell-1)(\alpha+\ell-1+j)}. \tag{3.11}
$$

Now $i$ and $j$ correspond to cycle lengths, so we must have $i+j \leqslant n$ to have a non-zero expectation at all. Since $i < j$ we conclude that $i < n/2$. A simple bound on the product outside the sum is to plug this value into the first term and ignore the rest. This gives the bound

$$
\frac{\alpha}{\alpha+n-i}\frac{n-i+1}{\alpha+n-i+1}\cdots\frac{n-1}{\alpha+n-1} < \frac{\alpha}{n-(n/2)+1} < \frac{2\alpha}{n}.
$$

Ignoring the product outside of it, the sum can be compared to the integral

$$
\int_1^n \frac{1}{x(x+j)}dx
$$

to prove it is $O(\log(j)/j)$. Thus the expression in (3.11) is $\tilde{O}(1/nj)$.

Rather than using the rough bound from above, the full $\ell = n - j - i + 1$ term is

$$\xi_{n-j-i+1}(1 - \xi_{n-j-i+2})\cdots(1 - \xi_{n-i})\xi_{n-i+1}^2(1 - \xi_{n-i+2})\cdots(1 - \xi_n).$$

The expectation of this term is

$$\frac{\alpha}{\alpha + n - j - i}\left(\prod_{k=n-j-i+2}^{n-i}\frac{k-1}{\alpha + k - 1}\right)\frac{\alpha}{\alpha + n - i}\left(\prod_{k=n-i+2}^{n}\frac{k-1}{\alpha + k - 1}\right).$$

But this is just

$$\frac{\alpha}{\alpha + n - i}\frac{\alpha + n - i}{n - i}\mathbb{P}[J_n = j + i] = \frac{\alpha}{n - i}\mathbb{P}[J_n = j + i]$$
$$\leqslant \frac{2\alpha}{n}\mathbb{P}[J_n = j + i].$$

Then by (3.5) of Lemma 3.1 we have the bound

$$\mathbb{P}[J_n = j + i] = O(n^{-\alpha/(1+\alpha)}).$$

Thus this term is $O(i^{-1}j^{-\alpha/(1+\alpha)})$, and

$$\mathbb{E}[xw] = \tilde{O}(i^{-1}j^{-\alpha/(1+\alpha)}). \tag{3.12}$$

**Term** $yw$. Because $i < j$, the term

$$\xi_{n-j+1}(1 - \xi_{n-j+2})\cdots(1 - \xi_n)\sum_{\ell=1}^{n-i}\xi_\ell\xi_{\ell+i}$$

behaves differently from the previous case. In fact, it is the largest-order term amongst the four. The first step is the same as before, though. A few of the terms are identically zero. So we can write $yw$ as

$$\xi_{n-j+1}(1 - \xi_{n-j+2})\cdots(1 - \xi_n)\sum_{\ell=1}^{n-i-j+1}\xi_\ell\xi_{\ell+i}.$$

Except for when $\ell = n - i - j + 1$, all of the indices are now disjoint. When we take expectation of the terms with disjoint indices, we obtain

$$\frac{\alpha}{\alpha + n - j}\frac{n - j + 1}{\alpha + n - j + 1}\cdots\frac{n - 1}{\alpha + n - 1}\sum_{\ell=1}^{n-j-i}\frac{\alpha^2}{(\alpha + \ell - 1)(\alpha + \ell + i - 1)}.$$

Notice that $j$ could be quite large. The best bound we have for the product outside the sum comes from (3.5). This implies it is $O(n^{-\alpha/(1+\alpha)})$. As with (3.11), the sum term is $O(\log i/i)$. This yields that the whole expression is $\tilde{O}(i^{-1}j^{-\alpha/(1+\alpha)})$.

As before, the full $\ell = n - i - j + 1$ term is

$$\xi_{n-i-j+1}(1 - \xi_{n-j-i+2})\cdots(1 - \xi_{n-j})\xi_{n-j+1}^2(1 - \xi_{n-i+2})\cdots(1 - \xi_n).$$

The expectation of this term is

$$\frac{\alpha}{\alpha + n - j - i}\left(\prod_{k=n-j-i+2}^{n-j}\frac{k-1}{\alpha + k - 1}\right)\frac{\alpha}{\alpha + n - i}\left(\prod_{k=n-j+2}^{n}\frac{k-1}{\alpha + k - 1}\right).$$

This simplifies to

$$\frac{\alpha}{\alpha+n-j}\frac{\alpha+n-j}{n-j}\mathbb{P}[J_n = j+i] = \frac{\alpha}{n-j}\mathbb{P}[J_n = j+i].$$

By construction $i \leqslant n - j$, and again by (3.5) of Lemma 3.1, we have that $\mathbb{P}[J_n = j+i] = O(n^{-\alpha/(1+\alpha)})$. Thus this term has expectation that is $O(i^{-1}j^{-\alpha/(1+\alpha)})$. It follows that

$$\mathbb{E}[yw] = \tilde{O}(i^{-1}n^{-\alpha/(1+\alpha)}). \tag{3.13}$$

**Term $yz$.** The last step is to bound the product

$$\left(\sum_{k=1}^{n-i}\xi_k\xi_{k+i}\right)\left(\sum_{\ell=1}^{n-j}\xi_\ell\xi_{\ell+j}\right) = \sum_{k,\ell=1}^{n-i,n-j}\xi_k\xi_{k+i}\xi_\ell\xi_{\ell+j}. \tag{3.14}$$

Call the index set

$$\Lambda = \{(k,\ell) : k \in [1,\ldots,n-i], \ell \in [1,\ldots,n-j]\}.$$

As $i < j$, we can divide $\Lambda$ into disjoint sets

$$\begin{aligned}
\Lambda_0 &= \{k, k+i, \ell, \text{ and } \ell+j \text{ are all distinct}\},\\
\Lambda_1 &= \{k = \ell\},\\
\Lambda_2 &= \{k = \ell+j\},\\
\Lambda_3 &= \{k+i = \ell\},\\
\Lambda_4 &= \{k+i = \ell+j\}.
\end{aligned}$$

This yields the following upper bound on (3.14):

$$\sum_{\Lambda_0}\xi_k\xi_{k+i}\xi_\ell\xi_{\ell+j} + \sum_{\Lambda_1}\xi_k\xi_{k+i}\xi_{\ell+j} + \sum_{\Lambda_2}\xi_k\xi_{k+i}\xi_\ell + \sum_{\Lambda_3}\xi_k\xi_{k+i}\xi_{\ell+j} + \sum_{\Lambda_4}\xi_k\xi_{k+i}\xi_\ell.$$

The pay-off of this partition is that each product of $\xi$ above is of distinct, independent terms. This lets us tidily compute the expectation of each. For $(k,\ell) \in \Lambda_0$ we have the $k, k+i, \ell, \ell+j$ are all distinct. This gives

$$\mathbb{E}[\xi_k\xi_{k+i}\xi_\ell\xi_{\ell+j}] \leqslant \frac{A}{k(k+i)\ell(\ell+j)}$$

for some constant $A$ depending only on $\alpha$. We can take the expectation of the entire sum:

$$\mathbb{E}\sum_{\Lambda_0}\xi_k\xi_{k+i}\xi_\ell\xi_{\ell+j} \leqslant \sum_{k=1,\ell=1}^{n}\mathbb{E}[\xi_k\xi_{k+i}\xi_{k+j}] \leqslant \sum_{k=1,\ell=1}^{n}\frac{A}{k(k+i)\ell(\ell+j)} = \tilde{O}(1/ij).$$

We can parametrize $\Lambda_1$ by $k$ to write it as $\Lambda_1 = \{(k,k) : k = 1,\ldots,n-j\}$. This gives the bound

$$\mathbb{E}\sum_{\Lambda_1}\xi_k\xi_{k+i}\xi_{\ell+j} \leqslant \sum_{k=1}^{n}\mathbb{E}[\xi_k\xi_{k+i}\xi_{k+j}] \leqslant \sum_{k=1}^{n}\frac{A}{k(k+i)(k+j)} = \tilde{O}(1/ij).$$

We can similarly parametrize the sums over $\Lambda_2, \Lambda_3$ and $\Lambda_4$ to obtain

$$\mathbb{E}[yz] = \tilde{O}(1/ij). \tag{3.15}$$

All of the expressions in (3.10), (3.12), (3.13) and (3.15) are $\tilde{O}(i^{-1}j^{-\alpha/(1+\alpha)})$. Markov's inequality yields

$$\mathbb{P}[C_iC_j > 0] = \tilde{O}(i^{-1}j^{-\alpha/(1+\alpha)}). \tag{3.16}$$

This proves the claimed inequality. $\qquad\square$

We also need a bound on the probability that cycles of the same size occur.

**Lemma 3.3.** *For any $i \leqslant n$ it holds that*

$$\mathbb{P}[C_i \geqslant 2] = \tilde{O}(i^{-2} + n^{-1-\alpha/(1+\alpha)}).$$

**Proof.** We start by bounding $\mathbb{P}[C_i > 2]$. By (3.2) we have

$$\mathbb{P}[Y_i + \mathbf{1}\{J_n = i\} > 2] \leqslant \mathbb{P}[Y_i > 1] = O(1/i^2). \tag{3.17}$$

It remains to bound $\mathbb{P}[C_i = 2]$. Conditioning on the value of $J_n$, we obtain the decomposition

$$\mathbb{P}[C_i = 2] = \mathbb{P}[C_i = 2 \mid J_n = i]\mathbb{P}[J_n = i] + \mathbb{P}[C_i = 2 \mid J_n \neq i]\mathbb{P}[J_n \neq i]. \tag{3.18}$$

First we estimate the first summand on the right-hand side of (3.18). Because two cycles larger than $n/2$ cannot occur, we must have $i \leqslant n/2$. Then by (3.3) of Lemma 3.1, $\mathbb{P}[J_n = i] \leqslant 2\alpha/n$. When $J_n = i$, it ensures that the sequence $\xi_1, \ldots, \xi_n$ has $\xi_{n-i+1} = 1$ and the larger index $\xi_k$ are zero. The smaller indices $\xi_i$ are independent of this event. The value of $C_i$ conditioned on $J_n = i$ is thus one more than the number of $i$-spacings in the sequence

$$\xi_1, \ldots, \xi_{n-i}, 1,$$

with $\xi_1, \ldots, \xi_{n-i}$ independent and distributed according the Feller coupling of an $\mathrm{ESF}(\alpha, n-i)$ distribution. Call the number of $i$-spacings in the above sequence $C_i^-$. So, conditional on $J_n = i$ we have $C_i = C_i^- + 1$. By writing explicitly what it means for an $i$-spacing to arise, we can bound $C_i^-$ in terms of the $\xi_k$:

$$C_i^- \leqslant \xi_{n-2i+1}(1 - \xi_{n-2i+2}) \cdots (1 - \xi_{n-i}) + \sum_{k=1}^{n-2i} \xi_k \xi_{k+i}.$$

Looking to apply Markov's inequality, we take the expectation of the above line and obtain

$$\mathbb{E}C_i^- \leqslant \frac{\alpha}{\alpha + n - 2i} \frac{n - 2i + 1}{\alpha + n - 2i + 1} \cdots \frac{n - i - 1}{\alpha + n - i - 1} + \sum_{k=1}^{n-2i} \frac{\alpha^2}{(\alpha + k - 1)(\alpha + k - 1 + i)}.$$

The product above is equivalent to $\mathbb{P}[J_{n-i} = i]$, then we use (3.5) of Lemma 3.1 and the fact that $i < n/2$ to bound it by $\tilde{O}((n-i)^{-\alpha/(1+\alpha)}) = \tilde{O}(n^{-\alpha/(1+\alpha)})$. The sum, as we have seen from (3.11), has order $\tilde{O}(1/i)$. Recalling that (3.3) implies $\mathbb{P}[J_n = i] = O(1/n)$, it follows that

$$\mathbb{P}[C_i = 2 \mid J_n = i]\mathbb{P}[J_n = i] = \tilde{O}(n^{-1-\alpha/(1+\alpha)} + 1/in) = \tilde{O}(i^{-2} + n^{-1-\alpha/(1+\alpha)}). \tag{3.19}$$

Now we estimate the second summand on the right-hand side of (3.18). Under the event $\{J_n \neq i\}$, for $C_i$ to equal 2, it is necessary that two $i$-spacings occur in the infinite sequence $\xi_1, \xi_2, \ldots$.

The count of $i$-spacings is distributed as $Y_i$ from (3.2). Therefore,

$$\mathbb{P}[C_i = 2 \text{ and } J_n \neq i] \leqslant \mathbb{P}[Y_i \geqslant 2] = O(1/i^2). \tag{3.20}$$

Then (3.17), (3.19) and (3.20) imply that

$$\mathbb{P}[C_i \geqslant 2] = \tilde{O}(i^{-2} + n^{-1-\alpha/(1+\alpha)}), \tag{3.21}$$

as desired. $\qquad\square$

Using the previous two lemmas we may prove the following.

**Proposition 3.4.** *Let $\pi \sim \mathrm{ESF}(\alpha, n)$ and $a = 1 - \alpha/(4\alpha + 3)$. The probability that $\pi$ has two cycles whose lengths have a common divisor larger than $n^a$ is $o(1)$.*

**Proof.** Call $E_a$ the event that $\pi$ has two cycles with common divisor larger than $n^a$. We utilize a union bound, then apply Lemmas 3.2 and 3.3 to write

$$\mathbb{P}[E_a] \leqslant \sum_{d=\lceil n^a \rceil}^{n} \left( \sum_{1 \leqslant i < j}^{\lfloor n^{1-a} \rfloor} \mathbb{P}[C_{id} C_{jd} > 0] + \sum_{i=1}^{\lfloor n^{1-a} \rfloor} \mathbb{P}[C_{id} \geqslant 2] \right)$$

$$= \tilde{O}\left( \sum_{d=\lceil n^a \rceil}^{n} \sum_{1 \leqslant i < j}^{\lfloor n^{1-a} \rfloor} (id)^{-1} (jd)^{-\alpha/(1+\alpha)} + \sum_{d=\lceil n^a \rceil}^{n} \sum_{i=1}^{\lfloor n^{1-a} \rfloor} \left( \frac{1}{(id)^2} + n^{-1-\alpha/(1+\alpha)} \right) \right).$$

Since $d \geqslant n^a$ the above line is

$$\tilde{O}(n^1 n^{2(1-a)} n^{-a-a\alpha/(1+\alpha)} + n^1 n^{1-a} n^{-2a} + n^{1-a-\alpha/(1+\alpha)}).$$

We can simplify the exponents in each term to

$$3 - 3a - a\alpha/(1+\alpha), \quad 2 - 3a, \quad \text{and} \quad 1 - a - \alpha/(1+\alpha),$$

respectively. Some algebra shows that all are negative as long as $a > 1 - \alpha/(4\alpha + 3)$. For such $a$ we have $\mathbb{P}[E_a] = o(1)$. $\qquad\square$

## 3.2. Proofs of Theorems 1.3, 1.5 and 1.6
We now have enough to establish our bound on the minimal degree of an $\mathrm{ESF}(\alpha, n)$ permutation.

**Lemma 3.5.** *Let $\pi \sim \mathrm{ESF}(\alpha, n)$. For each $0 < \beta < 1$, there exists $\lambda \in (0, 1)$ such that if we define the set*

$$Q_n = \{d < n^\beta : \text{there exists a prime } p > n^{\lambda\beta} \text{ with } p \mid d\},$$

*then*

$$\mathbb{P}[\exists C_d = 1 \text{ with } d \in Q_n] = 1 - o(1). \tag{3.22}$$

**Proof.** We start with [3, Theorem 3.1], which implies

$$\sum_{d \in Q_n} C_d \xrightarrow{TV} \mathrm{Poi}(\kappa_n), \tag{3.23}$$

where

$$\kappa_n = \sum_{d \in Q_n} \frac{\alpha}{d}.$$

Borrowing an argument from [9, p. 50], we have

$$\sum_{d \in Q_n} \frac{1}{d} = \sum_{d < n^\beta} \frac{1}{d} - \sum_{\substack{d < n^\beta \\ p \mid d \Rightarrow p \leqslant n^{\lambda\beta}}} \frac{1}{d}$$

$$\geqslant \sum_{d < n^\beta} \frac{1}{d} - \prod_{p \leqslant n^{\lambda\beta}} (1 - 1/p)^{-1}$$

$$= \beta \log n - e^{\gamma^*} \lambda \beta \log n + O(1)$$

where $\gamma^*$ is Euler's constant. When $\lambda$ is such that $e^{\gamma^*} \lambda \beta < \beta/2$ we have $\kappa_n > (\beta/2) \log n$ for all $n$. It follows from (3.23) and standard estimates on the concentration of a Poisson random variable that

$$\mathbb{P}\left[\sum_{d \in Q_n} C_d > (\beta/4) \log n\right] \to \mathbb{P}[\mathrm{Poi}(\kappa_n) > (\beta/4) \log n] = 1 - o(1).$$

To obtain (3.22), observe that the convergence statement in [3, Theorem 3.1] implies that $\mathbb{P}[C_d > 1] \to \mathbb{P}[Y_d > 1] = O(1/d^2)$. Here $Y_d$ is a Poisson random variable as in (3.2). The Borel–Cantelli lemma implies that only finitely many $Y_d$ are larger than 1. Hence $\sum_{Y_i > 1} Y_i$ is almost surely bounded. We have seen that, for $n$ large, the sum of $C_d$ with $d \in Q_n$ is at least $\mathrm{Poi}((\beta/4) \log n)$ with probability $1 - o(1)$. This diverges, and so one or more of the positive $C_d$ must be equal to one. Our claim follows. $\qquad\square$

For a random permutation $\pi \sim \mathrm{ESF}(\alpha, n)$, we define $\Phi(\pi) = \prod_{\ell=1}^n \ell^{C_\ell}$. To prove Theorem 1.5, it suffices to show that with high probability $\pi$ has a cycle of length $d$ such that

$$d \leqslant n^\beta, \tag{3.24}$$

$$d^2 \nmid \Phi(\pi). \tag{3.25}$$

First we will need an estimate on $d^2$ dividing $\Pi(\pi_n)$ conditional on $C_d = 1$.

**Lemma 3.6.** *Let $\pi$, $\beta$, $\lambda$, and $Q_n$ be as in Lemma 3.5, then for any $d \in Q_n$*

$$\mathbb{P}[d^2 \text{ divides } \Phi(\pi_n) \mid C_d = 1] = o(n^{-\lambda\beta+\varepsilon})$$

*for all $\varepsilon > 0$.*

**Proof.** It follows from [5, Theorem 3] that the cycle counts $C_i$ with $i \neq d$ are with high probability distributed like $\mathrm{ESF}(\alpha, n - d)$. Accordingly let $n' = n - d$ and $\pi_{n'}$ be a permutation sampled from this distribution. Note that from our construction of $Q_n$, there exists a prime $p > n^{\lambda\beta}$ such that

$$\{d \mid \Phi(\pi_{n'})\} \subset \{p \mid \Phi(\pi_{n'})\}.$$

For such prime number $p$, the event on the right-hand side above satisfies

$$\{p \mid \Phi(\pi_{n'})\} \subset \left\{ \sum_{i=1}^{n'/p} C_{ip} > 0 \right\}$$

Note we are omitting $i = d/p$ from the above summation, and every subsequent one.

Call the last event $G_{n'}$. The Feller coupling in (3.2) allows us to bound the probability by

$$\mathbb{P}[G_{n'}] \leqslant \mathbb{P}\left[ \sum_{i=1}^{\lceil n'/p \rceil} Y_{ip} + \mathbf{1}\{J_{n'} = ip\} > 0 \right]. \tag{3.26}$$

This, in turn, is bounded by

$$\mathbb{P}\left[ \text{Poi}\left( \sum_{i=1}^{\lceil n'/p \rceil} \frac{\alpha}{ip} \right) > 0 \right] + \mathbb{P}\left[ \sum_{i=1}^{\lceil n'/p \rceil} \mathbf{1}\{J_{n'} = ip\} > 0 \right].$$

The mean of the above Poisson random variable is asymptotically bounded by $\alpha \log(n/p)/p$. Since $p \geqslant n^{\lambda\beta}$, this ensures that the left summand is $o(n^{-\lambda\beta+\varepsilon})$. To bound the right summand we write

$$\mathbb{P}\left[ \sum_{i=1}^{\lceil n/p \rceil} \mathbf{1}\{J_n = ip\} > 0 \right] \leqslant \sum_{ip \leqslant n - n^\gamma} \mathbb{P}[J_n = ip] + \sum_{ip > n - n^\gamma} \mathbb{P}[J_n = ip].$$

Here $\gamma$ is a constant we will determine in a moment. We begin by bounding the first summand of the right-hand side. By (3.3) of Lemma 3.1,

$$\begin{aligned}
\sum_{ip \leqslant n' - n^\gamma} \mathbb{P}[J_{n'} = ip] &\leqslant \sum_{i=1}^{\lfloor (n' - n^\gamma)/p \rfloor} \frac{\alpha}{n' - ip} \\
&\leqslant \int_1^{(n - n^\gamma)/p + 1} \frac{\alpha}{n - d - px} dx \\
&= O\left( \frac{\log n}{p} \right) \\
&= o(n^{-\lambda\beta+\varepsilon}).
\end{aligned}$$

The last line follows since $p \geqslant n^{\lambda\beta}$. For the second summand, from (3.4) of Lemma 3.1 there exists some constant $c > 0$ such that

$$\begin{aligned}
\sum_{ip > n' - n^\gamma} \mathbb{P}[J_n = ip] &\leqslant \sum_{ip > n - n^\gamma} cn^{-\alpha(1-\gamma)} \\
&\leqslant cn^{\gamma - \lambda\beta} n^{-\alpha(1-\gamma)}.
\end{aligned}$$

The last line follows as there are at most $n^{\gamma - \lambda\beta}$ different values of $i$ such that $n' - n^\gamma < ip \leqslant n'$.

We therefore have that

$$\mathbb{P}[G_{n'}] = O(n^{\gamma - \lambda\beta - \alpha(1-\gamma)}).$$

An easy calculation confirms that $\gamma$ satisfying

$$\gamma < \frac{\alpha + \varepsilon}{1 + \alpha}$$

is sufficient to ensure

$$\mathbb{P}[G_{n'}] = o(n^{-\lambda\beta+\varepsilon}).$$

This completes our proof. $\qquad\square$

**Lemma 3.7.** *Let $\pi$, $\beta$, $\lambda$, and $Q_n$ be as in Lemma 3.5. Then*

$$\mathbb{P}[\exists d \in Q_n \text{ such that } d^2 \nmid \Phi(\pi) \text{ and } C_d = 1] = 1 - o(1). \tag{3.27}$$

**Proof.** We will show the stronger result that

$$\mathbb{P}[d^2 \mid \Phi(\pi) \text{ for some } d \in Q_n \text{ such that } C_d = 1] = o(1).$$

Denote the above event by $E_n$ so that

$$E_n = \bigcup_{d \in Q_n} \{C_d = 1 \text{ and } d^2 \mid \Phi(\pi)\}.$$

A union bound gives

$$
\begin{aligned}
\mathbb{P}[E_n] &\leqslant \sum_{d \in Q_n} \mathbb{P}[C_d = 1 \text{ and } d^2 \mid \Phi(\pi)] \\
&= \sum_{d \in Q_n} \mathbb{P}[d^2 \mid \Phi(\pi) \mid C_d = 1]\mathbb{P}[C_d = 1] \\
&\leqslant \max_{d \in Q_n}(\mathbb{P}[d^2 \mid \Phi(\pi) \mid C_d = 1])\left(\sum_{d \in Q_n} \mathbb{P}[C_d = 1]\right).
\end{aligned}
$$

By Lemma 3.6, the left probability is $o(n^{-\lambda\beta+\varepsilon})$ for arbitrarily small $\varepsilon > 0$. The Feller coupling ensures that the right-hand sum can be bounded by $\alpha\beta\log n$. It follows that $\mathbb{P}[E_n] = o(1)$, completing our proof. $\qquad\square$

**Proof of Theorem 1.5.** By Lemma 3.7, with probability $1 - o(1)$ there exists a cycle length of length $d$ satisfying (3.24) and (3.25). When these conditions are met, for $K = \Phi(\pi)/d$, we have $\pi^K$ displaces $d \leqslant n^\beta$ elements. Thus, with high probability $\pi$ has minimal degree no more than $n^\beta$. This completes our proof. $\qquad\square$

**Proof of Theorem 1.6.** The theorem statement is such that if it holds for $\omega(n)$, then it also holds for any $\omega'(n) \geqslant \omega(n)$. So, it suffices to prove the statement for all $\omega(n) \leqslant \sqrt[4]{\log n}$.

Fix such an $\omega(n)$ and set $b_n = n\exp(-\omega(n)\sqrt{\log n})$. Let $P$ be the set of all primes and define $P_n = P \cap [b_n, n]$. Given $p \in P_n$, let $I_p = [1, 2, \ldots, \lfloor n/p \rfloor]$ be the set of all $i$ for which $pi$ is a multiple of $p$ between $p$ and $n$. We also introduce the entire collection of such multiples smaller than $n$:

$$W_n = \bigcup_{p \in P_n} \{(p, i) : i \in I_p\}.$$

It suffices to show that the event

$$E_n = \{\exists(p, i) \in W_n : ip \mid \Phi(\pi)\}$$

occurs with high probability. We can write $E_n$ as a union of events:

$$E_n = \bigcup_{(p,i)\in W_n} \{C_{ip}^{(n)} > 0\}.$$

Notice that any two numbers in the set of products

$$\{ip : (p,i) \in W_n\}$$

are different. To see this we argue by contradiction: suppose that $ip = i'p'$ with $p < p'$. Because of primality, this could only happen if $p' \mid i$. Thus $i \geqslant p'$. Also note that our assumption $\omega(n) \leqslant \sqrt[4]{\log n}$ ensures $b_n > \sqrt{n}$. Thus $p, p' > \sqrt{n}$. These two inequalities give

$$ip \geqslant p'p > \sqrt{n}\sqrt{n} = n.$$

But $ip > n$ is a contradiction, since our construction of $I_p$ ensures that $ip \leqslant n$.

We have now shown that the occurrence of $E_n$ is equivalent to the sum of cycle counts being positive: $E_n = \{\sum_{W_n} C_{ip} > 0\}$. The Feller coupling (3.2) ensures that

$$\sum_{W_n} C_{ip} \geqslant -D_n + \sum_{W_n} Y_{ip}, \tag{3.28}$$

with $D_n$ the total number of deletions. It follows that

$$\mathbb{P}[E_n] \geqslant \mathbb{P}\left[-D_n + \sum_{W_n} Y_{ip} > 0\right]. \tag{3.29}$$

Using additivity of independent Poisson random variables, the sum on the right of (3.28) is distributed as a Poisson with mean

$$\mu_{n,\alpha} = \sum_{p\in P_n} \sum_{i=1}^{\lfloor n/(p)\rfloor} \alpha/ip \approx \alpha \sum_{p\in P_n} \frac{\log(n/(p))}{p}$$

$$= \alpha\left(\log n \sum_{p\in P_n} \frac{1}{p} - \sum_{p\in P_n} \frac{\log p}{p}\right). \tag{3.30}$$

Lemma 3.8 shows that $\mu_{n,\alpha} \to \infty$. To establish that (3.29) occurs with high probability, we will show that the number of deletions, $D_n$, in the Feller coupling is unlikely to remove every successful $Y_{ip}$ in (3.28).

As $\mu_{n,\alpha} \to \infty$, we take any $h(n) = o(\mu_{n,\alpha})$ with $h(n) \to \infty$. By (3.6) of Lemma 3.1 we know that $\mathbb{E}D_n \leqslant c$ for some $c = c(\alpha)$. It follows from Markov's inequality that

$$\mathbb{P}[D_n \geqslant h(n)] \leqslant \frac{c}{h(n)} = o(1).$$

Standard estimates on a Poisson random variable tell us that

$$\mathbb{P}[\mathrm{Poi}(\mu_{n,\alpha}) \leqslant h(n)] = 1 - o(1).$$

So $D_n$ is with high probability smaller than $\sum_{W_n} Y_{ip}$, and combined with (3.28) we have established that the right side of (3.29) occurs with high probability. $\qquad\square$

**Lemma 3.8.** *Let $\mu_{n,\alpha}$ be as in the proof of Theorem 1.6. It holds that $\mu_{n,\alpha} \to \infty$.*

**Proof.** It suffices to show that (3.30) tends to infinity. Two classical theorems attributed to Euler (see [18]) and Chebyshev (see [1]), respectively, state that for sums of primes smaller than $n$

$$\sum_{p \leqslant n} \frac{1}{p} = \log\log n + O(1) \quad \text{and} \quad \sum_{p \leqslant n} \frac{\log p}{p} = \log n + O(1).$$

Recalling that $P_n = P \cap [b_n, n]$, it follows that (3.30) is asymptotic to

$$\log n \log \frac{\log n}{\log b_n} - \log \frac{n}{b_n}.$$

We can express the second term precisely:

$$\log \frac{n}{b_n} = \log \exp(\omega(n)\sqrt{\log n}) = \omega(n)\sqrt{\log n}. \tag{3.31}$$

Similarly we compute the order of the first term. Notice that

$$\log \frac{\log n}{\log b_n} = \log \left( \frac{\log n}{\log n - \omega(n)\sqrt{\log n}} \right)$$

$$= \log \left( 1 + \frac{\omega(n)\sqrt{\log n}}{\log n - \omega(n)\sqrt{\log n}} \right). \tag{3.32}$$

As $\omega(n) \leqslant \sqrt[4]{\log n}$ the above term is $\log(1 - o(1))$. Using the fact that $\log(1-x) \approx -x$ as $x \to 0$ we have that (3.32) is asymptotic to

$$\frac{\omega(n)\sqrt{\log n}}{\log n - \omega(n)\sqrt{\log n}}.$$

Reintroducing the $\log n$ factor, we now have

$$\log n \sum_{P_n} \frac{1}{p} \approx \log n \frac{\omega(n)\sqrt{\log n}}{\log n - \omega(n)\sqrt{\log n}}. \tag{3.33}$$

Combining (3.31) and (3.33) gives

$$\mu_{n,\alpha} \approx \log n \frac{\omega(n)\sqrt{\log n}}{\log n - \omega(n)\sqrt{\log n}} - \omega(n)\sqrt{\log n}$$

$$= \omega(n)^2 \frac{\log n}{\log n - \omega(n)\sqrt{\log n}}$$

$$= \omega(n)^2 (1 - o(1)).$$

Since $\omega(n) \to \infty$, we conclude that $\mu_{n,\alpha} \to \infty$. $\qquad \square$

Now we set our sights on proving Theorem 1.3. This will require Theorems 1.5 and 1.6, as well as a few additional lemmas.

**Lemma 3.9.** *Let $0 < \gamma < 1$ and $\psi(n) \to \infty$ such that $\psi(n) = o(\log n)$. The probability that simultaneously for all $2 \leqslant r \leqslant \exp(\psi(n))$ there exists $C_{j_r} > 0$ such that $j_r > n^\gamma$ and $r \nmid j_r$ is $1 - o(1)$.*

**Proof.** Let $r_0 = \exp(\psi(n))$ and consider the complement of the event in question:

$$E = \{\exists r < r_0 : r \mid j \text{ for all } n^\gamma < j \text{ such that } C_j > 0\}.$$

It suffices to show that $\mathbb{P}[E] = o(1)$. In order to model the values $C_j$ with independent Poisson random variables we will work with a larger event. For any $\gamma < \gamma' < 1$, we consider

$$E_{\gamma'} = \{\exists r < r_0 : r \mid j \text{ for all } n^\gamma < j < n^{\gamma'} \text{ such that } C_j > 0\}.$$

A union bound leads to

$$\mathbb{P}[E_{\gamma'}] \leqslant \sum_{r=2}^{r_0} \mathbb{P}[r \mid j \text{ for all } n^\gamma < j < n^{\gamma'} \text{ such that } C_j > 0]$$

$$= \sum_{r=2}^{r_0} \mathbb{P}[C_j = 0 \text{ for all } n^\gamma < j < n^{\gamma'} \text{ such that } r \nmid j]. \tag{3.34}$$

Let $A_r$ be the set of integers in $(n^\gamma, n^{\gamma'}]$ that $r$ fails to divide. Formally,

$$A_r = [n^\gamma, n^{\gamma'}] \cap ([n] \setminus \{ir : i \leqslant \lfloor n/r \rfloor\}).$$

As $n^{\gamma'} = o(n)$ we can apply [3, Theorem 3.1] to conclude that $C[n^\gamma, n^{\gamma'}] \xrightarrow{TV} Y[n^\gamma, n^{\gamma'}]$. Thus, the probabilities in (3.34) converge to

$$\sum_{r=2}^{r_0} \prod_{j \in A_r} \mathbb{P}[Y_j = 0] = \sum_{r=2}^{r_0} \exp\left(-\sum_{j \in A_r} \frac{\alpha}{j}\right). \tag{3.35}$$

The quantity $\sum_{j \in A_r}(\alpha/j)$ is minimized at $r = 2$, and in this case $\sum_{j \in A_2}(1/j) \gg \log n$. We can apply this to (3.35) to obtain

$$\mathbb{P}[E_{\gamma'}] \leqslant \sum_{r=2}^{r_0} n^{-\delta} \leqslant r_0 n^{-\delta} = o(1),$$

for some $\delta > 0$. As $E \subseteq E_{\gamma'}$, this completes the proof. $\qquad\square$

**Lemma 3.10.** *Let $\delta = \delta(\varepsilon, \alpha) = 1 - (\alpha + \varepsilon)\log 2$. For any $\varepsilon > 0$, there exists a constant $c = c(\varepsilon, \alpha)$ such that the probability of $\pi$ fixing a set of size $k$ is bounded by $ck^{-\delta}$ uniformly for all $1 \leqslant k \leqslant n/2$.*

**Proof.** In Lemma 4.1 we show that $\mathbb{P}[k \in \mathcal{L}(\mathbf{X}(\alpha))] \leqslant Ck^{-\delta}$ for some $C$ and all $k \leqslant n$. The Feller coupling lets us relate the probability of a fixing a set size to this quantity, the only complication is a possible contribution from $\mathbf{1}\{J_n = \ell\}$. However, the location of this extra cycle is unconcentrated enough to not alter the order of these probabilities.

Inequality (3.3) of Lemma 3.1 ensures that $\mathbb{P}[J_n = \ell] \leqslant \alpha/(n - \ell)$. We make use of this as we condition on the value of $J_n$:

$$\mathbb{P}[\pi \text{ has a } k\text{-cycle}] \leqslant \mathbb{P}[k \in \mathcal{L}(\mathbf{X}(\alpha))] + \sum_{\ell \leqslant k} \mathbb{P}[k - \ell \in \mathcal{L}(\mathbf{X}(\alpha))]\mathbb{P}[J_n = \ell] \tag{3.36}$$

$$\leqslant Ck^{-\delta} + \sum_{\ell \leqslant k} C(k - \ell)^{-\delta} \frac{\alpha}{n - \ell} \tag{3.37}$$

$$\leqslant Ck^{-\delta} + \frac{2\alpha}{n} \sum_{\ell \leqslant k} C\ell^{-\delta} \tag{3.38}$$

$$\leqslant Ck^{-\delta} + C'n^{-\delta},$$

where in (3.37) we use our bound on $\mathbb{P}[k \in \mathcal{L}(\mathbf{X}(\alpha))]$ from Lemma 4.1 and bound $\mathbb{P}[J_n = \ell]$ with (3.3) of Lemma 3.1. The next line, (3.38), reindexes the sum and uses that $\ell \leqslant n/2$ to bound $\alpha/(n-\ell) \leqslant 2\alpha/n$. Our claim then follows by bounding the leftmost sum by $n^{-\delta}$ times some constant $C'$, and the fact that $k^{-\delta} > n^{-\delta}$. $\qquad\square$

We now prove a result analogous to [26, Theorem 1]. Namely, if a permutation sampled according to the Ewens sampling formula is in a transitive subgroup, this subgroup is likely to be either all of $S_n$ or the alternating group, $A_n$. Although the lemmas leading up to this used rather different arguments, our proof from here closely follows that of [26]. Recall that a *primitive subgroup* is a transitive subgroup that does not fix any partition of $[n]$.

**Proof of Theorem 1.3.** Let $\pi \in G$ for a transitive subgroup $G$. According to [23], it follows from results of [7] that the minimal degree of a primitive subgroup not containing $A_n$ is at least $(\sqrt{n}-1)/2$. On the other hand, by Theorem 1.5, with high probability the minimal degree of $\langle\pi\rangle$ is less than $n^{0.4}$. Thus, the probability that $G$ is primitive is $o(1)$.

Assume $G$ is a non-primitive, transitive subgroup. Then it must fix some partition $\{B_i\}_{i=1}^r$ of $[n]$. By transitivity every block in such a partition will be mapped to every other, thus the partition must be into blocks of equal size, that is,

$$|B_i| = \frac{n}{r}.$$

Note that any cycle of a permutation that preserves this partition must have the same number of elements in each block it acts on. We will use this to establish that any $\pi \in G$ with high probability cannot fix this partition for any $r$.

We consider three cases for the possibly amount, $r$, of blocks.

**Case 1.** $2 \leqslant r \leqslant r_0 = \exp(\log\log n\sqrt{\log n})$. By Lemma 3.9 with $\psi(n) = \log\log n\sqrt{\log n}$, with high probability for every $r$ there exists $j_r > n^{.99}$ such that $\pi$ contains a cycle of length $j_r$ and $r \nmid j_r$. Fix some $r$ and $j_r$, and let $B$ be the union of all blocks this cycle acts upon. Because $r$ does not divide $j_r$, $B$ is a proper invariant subset of $[n]$ with size $sn/r$, where $s$ is the number of blocks in $B$. By Lemma 3.10 this occurs with probability bounded by

$$\sum_{r=2}^{r_0}\sum_{s=1}^r c(sn/r)^{-\delta} \leqslant cr_0^2(n/r_0)^{-\delta} = o(1).$$

**Case 2.** $r_0 \leqslant r \leqslant n/r_0$. By Theorem 1.6, with high probability there exists a prime $q > n/r_0$ that divides $\Omega(\pi)$. Thus there exists a cycle of length $j$ such that $q \mid j$. Let $t \leqslant r$ be the number of blocks this cycle intersects. Because of the bounds on $r$ we have

$$\frac{j}{t} \leqslant |B_i| = \frac{n}{r} \leqslant \frac{n}{r_0} < q.$$

Since $q$ divides $j = t(j/t)$ and $j/t < q$ by the above inequality, we must have that $q \mid t$. However,

the total number of blocks $r$ satisfies

$$r < \frac{n}{r_0} < q,$$

so this is impossible.

**Case 3.** $n/r_0 \leqslant r < n$. Let $a$ be as in Proposition 3.4 and $\varepsilon > 0$ such that $a + \varepsilon < 1$. Using Lemma 3.9 with $\psi(n) = \log\log n \sqrt{\log n}$ one more time, we now find $j > n^{a+\varepsilon}$ such that $\pi$ contains a cycle of length $j$ and the block size, $s = n/r$ does not divide $j$. Again, let $B$ be the union of blocks which intersect this cycle. Because the size of the blocks does not divide $j$, this intersection is a proper subset of $B$. We conclude that $\pi$ must contain another cycle of length $j'$ that intersects each block in $B$. Hence both cycle lengths $j$ and $j'$ are divisible by the number of blocks in $B$. We bound the number of blocks:

$$\frac{|B|}{s} > \frac{j}{s} > \frac{n^{a+\varepsilon}}{s} > n^a.$$

However, by Proposition 3.4, with probability $1 - o(1)$ no two cycle lengths have a common divisor greater than $n^a$. This completes our proof.

$\square$

### 3.3. Proving Proposition 2.1

To rule out generating $A_n$ we need an estimate that shows an $\mathrm{ESF}(\alpha, n)$ permutation is odd with probability bounded away from zero. Of course, this probability ought to converge to $1/2$, but we were unable to find a proof. We make do with an absolute lower bound.

**Lemma 3.11.** *Let $\pi \sim \mathrm{ESF}(\alpha, n)$. Then*

$$\inf_{n \geqslant 2} \min\{\mathbb{P}[\pi \text{ is odd}], \mathbb{P}[\pi \text{ is even}]\} \geqslant \min\left\{\frac{1}{\alpha + 1}, \frac{\alpha}{\alpha + 1}\right\}.$$

**Proof.** We will proceed by induction. It is straightforward to check via the Feller coupling that an $\mathrm{ESF}(\alpha, 2)$ permutation is odd with probability $\alpha/(\alpha + 1)$, and it is even with probability $1 - \alpha/(\alpha + 1) = 1/(\alpha + 1)$.

Suppose that for $2 \leqslant m \leqslant n - 1$. If $\pi_m \sim \mathrm{ESF}(\alpha, m)$, then

$$\min\{\mathbb{P}[\pi_m \text{ is odd}], \mathbb{P}[\pi_m \text{ is even}]\} \geqslant \varepsilon,$$

for some yet to be determined $\varepsilon > 0$. Now, it suffices to prove that

$$\min\{\mathbb{P}[\pi \text{ is odd}], \mathbb{P}[\pi \text{ is even}]\} \geqslant \varepsilon. \tag{3.39}$$

Consider the sequence $\xi_1, \xi_2, \ldots$ of Bernoulli random variables for the Feller coupling. Recall that $J_n = (n+1) - R_n$, with $R_n$ the index of the rightmost 1 in $\xi_1, \ldots, \xi_n$. We will condition on the value of $J_n$, but also need to take into account the parity of $n$.

**$n$ is odd.** We condition on the value of $J_n$. For convenience let $q_\ell = \mathbb{P}[J_n = \ell]$. If $J_n = \ell$, then $\xi_{n+1-\ell} = 1$ and we can view the beginning sequence $\xi_1, \ldots, \xi_{n-\ell}, 1$ as corresponding to a random

permutation sampled with distribution ESF$(\alpha, n - \ell)$. We denote such a permutation by $\pi_{n-\ell}$. Noting that when $J_n = n$, $\pi$ is a single $n$-cycle and therefore even, it follows that

$$\mathbb{P}[\pi \text{ is odd}] = \sum_{\ell \text{ odd}}^{n-3} \mathbb{P}[\pi_{n-\ell} \text{ is even}]q_\ell + \sum_{\ell \text{ even}}^{n-2} \mathbb{P}[\pi_{n-\ell} \text{ is odd}]q_\ell + q_{n-1}.$$

By the induction hypothesis,

$$\mathbb{P}[\pi \text{ is odd}] \geqslant \sum_{\ell=1}^{n-2} \varepsilon q_\ell + q_{n-1}.$$

Since $\sum_1^n q_\ell = \sum_1^n \mathbb{P}[J_n = \ell] = 1$, it suffices to prove that $q_{n-1} \geqslant \varepsilon(q_{n-1} + q_n)$. Equivalently, this requires that $(1 - \varepsilon)q_{n-1} \geqslant \varepsilon q_n$. Notice that

$$q_{n-1} = \mathbb{P}[\xi_2 = 1] \prod_{i=3}^{n} \mathbb{P}[\xi_i = 0] \quad \text{and} \quad q_n = \prod_{i=2}^{n} \mathbb{P}[\xi_i = 0].$$

Much of this cancels when we solve $(1 - \varepsilon)q_{n-1} \geqslant \varepsilon q_n$, so we arrive at the requirement

$$(1 - \varepsilon)\frac{\mathbb{P}[\xi_2 = 1]}{\mathbb{P}[\xi_2 = 0]} \geqslant \varepsilon.$$

The ratio $\mathbb{P}[\xi_2 = 1]/\mathbb{P}[\xi_2 = 0] = \alpha$, and thus we require

$$(1 - \varepsilon)\alpha \geqslant \varepsilon. \tag{3.40}$$

*n* **is even.**   The argument is similar. The key difference is that we instead need $q_n \geqslant \varepsilon(q_{n-1} + q_n)$. Rewriting as before, this is the same as

$$\varepsilon\alpha \leqslant (1 - \varepsilon). \tag{3.41}$$

After solving for $\varepsilon$ in (3.40) and (3.41), we see that (3.39) holds whenever

$$0 < \varepsilon \leqslant \min\left\{\frac{1}{\alpha + 1}, \frac{\alpha}{\alpha + 1}\right\}. \qquad \square$$

**Remark 3.12.** When $\alpha = 1$, this confirms the intuition that even and odd permutations occur with probability $1/2$ for all $n \geqslant 2$.

**Corollary 3.13.** *Let* $\delta = \delta(\alpha, \varepsilon)$ *and* $c = c(\alpha, \varepsilon)$ *be as in Lemma 3.10. It holds for all* $k \leqslant n/2$ *that*

$$\mathbb{P}[\pi \text{ fixes a } k\text{-set} \mid \pi \text{ is odd}] \leqslant ck^{-\delta} \cdot \max\left\{\frac{\alpha + 1}{\alpha}, \frac{\alpha + 1}{1}\right\}.$$

**Proof.**   We start with the reformulation

$$\mathbb{P}[\pi \text{ fixes a } k\text{-set} \mid \pi \text{ is odd}] = \frac{\mathbb{P}[\pi \text{ fixes a } k\text{-set and } \pi \text{ is odd}]}{\mathbb{P}[\pi \text{ is odd}]}.$$

The numerator is bounded by $\mathbb{P}[\pi$ fixes a $k$-set$]$, which is less than $ck^{-\delta}$ by Lemma 3.10. We bound the denominator away from zero by Lemma 3.11, completing our proof. $\qquad\square$

## 4. The upper bound for $s_\alpha$

We carry out the plan described in Section 1.4. First we bound the quenched probability $\tilde{p}_k$ defined formally below. The proof is similar to [31, Lemma 3.1]. We repeat a selection of the details for clarity, and point out where the argument differs with general $\alpha$. Subsequently, we mirror [31, Theorem 3.2] as we apply the Borel–Cantelli lemma.

Recall the functions $f_k$ and $g_k$ from (1.5), namely that $f_k$ is the number of available summands smaller than $k$, and $g_k$ is the largest sum attainable with them. Also, define

$$\tau_\varepsilon := \sup\{k : f_k \geqslant (\alpha + \varepsilon)\log k\},$$
$$\tau := \sup\{n : g_{\lambda_k} \geqslant k\},$$

where $\varepsilon > 0$ is yet to be chosen, but small, and $\lambda_k := \lfloor k/(\alpha \log k)\rfloor$. The proofs of [31, Lemmas 2.1, 2.2] generalize in a straightforward way, so that $T = T(\varepsilon, \delta) := \max\{\tau_\varepsilon, \tau\}$ is almost surely finite. As $\mathbb{P}[T > \lambda_k] \to 0$, we define the quenched probabilities to avoid this diminishing sequence of events:

$$\tilde{p}_k := \mathbb{P}[T < \lambda_k \text{ and } k \in \mathscr{L}(\mathbf{X})].$$

**Lemma 4.1.** *Suppose that $\alpha < 1/\log 2$. For each $\varepsilon > 0$ there exists $C = C(\varepsilon)$ such that for all $k \geqslant 1$,*

$$\tilde{p}_k(\alpha) \leqslant Ck^{-1+(\alpha+\varepsilon)\log 2}.$$

**Proof.** Fix $\varepsilon > 0$. Let $G_k$ be the event that $T < \lambda_k$ while also $k \in \mathscr{L}(\mathbf{X})$. Call a sequence $\mathbf{y} = (y_1, y_2, \ldots)$ admissible if it is coordinate-wise less than or equal to $\mathbf{X}$. When $G_k$ occurs, it is not possible for $n$ to be a sum of $\sum_j jy_j$ for an admissible $\mathbf{y}$ with $y_j$ vanishing for $j > \lambda_k$. Indeed, on the event $G_k$, we have $\sum_{j \leqslant \lambda_k} jX_j < k$. The event satisfying the following three conditions contains $G_k$:

(i) $f_{\lambda_k} \leqslant (\alpha + \varepsilon)\log \lambda_k$,
(ii) $g_{\lambda_k} < k$,
(iii) there is some $k$ with $k = \sum_j (y'_j + y''_j)$ with $\mathbf{y}'$ supported on $[0, \lambda_k]$ and $\mathbf{y}''$ non-zero and supported on $[\lambda_k + 1, k]$, with both $\mathbf{y}'$ and $\mathbf{y}''$ admissible.

Define the probability $p'_k = \mathbb{P}[f_{\lambda_k} \leqslant (\alpha + \varepsilon)\log \lambda_k$ and $g_{\lambda_k} < n$ and $\ell = \sum_j jy'_j]$, where $\mathbf{y}'$ is admissible and supported on $[1, \lambda_n]$. Also, define $p''_\ell = \mathbb{P}[\ell = \sum_j jy''_j]$, where $\mathbf{y}''$ is admissible and supported on $[\lambda_k + 1, k]$. Using independence we can obtain an upper bound on $\tilde{p}_k$ by

decomposing into these two admissible vectors:

$$\tilde{p}_k \leqslant \sum_{k=\lambda_k+1}^{k} p'_{k-\ell} p''_\ell$$

$$\leqslant \left( \sum_{\ell=\lambda_k+1}^{k} p'_{k-\ell} \right) \max_{\lambda_k \leqslant \ell \leqslant k} p''_\ell. \qquad (4.1)$$

The first factor above is (by Fubini's theorem) equal to $\mathbb{E}|\mathscr{L}(\mathbf{X}) \cap [0, \lambda_k]|$. Using the trivial bound where we assume each sum from $\mathbf{X}$ is distinct and that we are working on the event $f_{\lambda_k} \leqslant (\alpha + \varepsilon) \log \lambda_k$, we have

$$|\mathscr{L}(\mathbf{X}) \cap [0, \lambda_k]| \leqslant 2^{Z_m} \leqslant \lambda_k^{(\alpha+\varepsilon)\log 2}. \qquad (4.2)$$

We now bound the second term. We will show that there exists a constant $C$ such that

$$p''_\ell \leqslant C \frac{\log^2 k}{k}, \quad \text{for all } \ell \in [\lambda_k + 1, k]. \qquad (4.3)$$

Let $H$ be the event that $X_j \geqslant 2$ for some $j \in [\lambda_k, k]$. As $\mathbb{P}[X_j \geqslant 2] \leqslant \alpha/j^2$, we can write

$$\mathbb{P}[H] \leqslant \sum_{j=\lambda_k}^{k} \alpha/j^2 \leqslant \alpha/\lambda_k \ll \frac{\log k}{k}. \qquad (4.4)$$

The event that $\ell = \sum_j j y''_j$ for an admissible $\ell$ supported on $[\lambda_k, k]$, but that $H$ does not occur, is contained in the union of events $E_j$ that $X_j = 1$ and $\ell = j = \sum_i i y''_i$ for some admissible $\mathbf{y}''$ supported on $[\lambda_k, k] \setminus \{j\}$. Using independence of $X_j$ from the other coordinates of $\mathbf{X}$, along with our description of what must happen if $H$ does not, we obtain

$$p''_\ell \leqslant \mathbb{P}[H] + \sum_{j=\lambda_k}^{k} \frac{1}{j} p''_{\ell-j}$$

$$\leqslant \mathbb{P}[H] + \frac{1}{\lambda_k} \sum_{j=\lambda_k}^{k} p''_{\ell-j}$$

$$\leqslant \frac{\log k}{k} \left( 1 + \sum_{j=\lambda_k}^{k} p''_{\ell-j} \right). \qquad (4.5)$$

It remains to bound the summation in the last factor. Here taking $\alpha \neq 1$ is a minor nuisance, but we circumvent any difficulties with the crude bound $\alpha \leqslant 2$. The idea is to use the generating function

$$F(z, \mathbf{X}) := \prod_{j \in [\lambda_k, k]: X_j = 1} (1 + z^j),$$

writing $[z^j]F$ for the $j$th coefficient of $F$. Observe that this corresponds to forming $j$ using the $j \in [\lambda_k, k]$ for which $X_j = 1$. It follows that

$$\sum_{j=\lambda_k}^{k} p''_{\ell-j} \leqslant \sum_{j=\lambda_k}^{k} \mathbb{E}([z^{\ell-j}]F(z, \mathbf{X})).$$

Using that $[z^j]F(z,\mathbf{X}) \leqslant F(1,\mathbf{X})$, we can bound the above line by $\mathbb{E}F(1,\mathbf{X})$. This is tractable using independence of the $X_j$. Carrying the calculation out we obtain

$$\mathbb{E}F(1,\mathbf{X}) = \prod_{j=\lambda_k}^{n} \mathbb{E}[1 + \mathbf{1}\{X_j = 1\}]$$

$$= \prod_{j=\lambda_k}^{k} \left(1 + \frac{\alpha}{j}\right)$$

$$\leqslant \prod_{j=\lambda_k}^{n} \left(1 + \frac{2}{j}\right).$$

The last term can be expanded and written as

$$\frac{k^2 + 3k + 2}{\lambda_k^2 + \lambda_k} \ll \frac{k^2}{\lambda_k^2} \ll \log^2 k.$$

Combine this with (4.5) and we have

$$p_\ell'' \ll \frac{\log k}{k}(1 + \log^2 k) \ll \frac{\log^3 k}{k}.$$

This yields (4.3). To conclude, we plug (4.2) and (4.3) into (4.1) and arrive at

$$\tilde{p}_k \leqslant C(k - \lambda_k)^{(\alpha+\varepsilon)\log 2} \frac{\log^3 k}{k}.$$

This is bounded above by a constant multiple of $k^{-1+(\alpha+2\varepsilon)\log 2}$, which, after swapping all $\varepsilon$ for $\varepsilon/2$, establishes the lemma. □

**Proposition 4.2.** *Let $s_\alpha$ be as in Theorem 1.7 and $h(\alpha)$ be as in (1.4). For continuity points of $h$ it holds that $s_\alpha \leqslant h(\alpha)$, and $s_\alpha \leqslant h(\alpha) + 1$ for all $\alpha$.*

**Proof of Proposition 4.2.** Fix a small $\varepsilon > 0$ and let $s_{\alpha,\varepsilon} = \lceil (1 - (\alpha - 2\varepsilon)\log 2)^{-1} \rceil$. Sample $s_{\alpha,\varepsilon}$ vectors $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(s_{\alpha,\varepsilon})}$ and let $T(\mathbf{X}^{(1)}), \ldots, T(\mathbf{X}^{(s_{\alpha,\varepsilon})})$ be the a.s. random times distributed as $T$. Take $T^* = \max\{T(\mathbf{X}^{(i)})\}_1^{s_{\alpha,\varepsilon}}$. By independence of the $\mathbf{X}^{(i)}$ and the bound in Lemma 4.1 we have the probability that $T^* < \lambda_k$ and $k \in \mathscr{L}(\mathbf{X}^{(i)})$ for all $1 \leqslant i \leqslant s_{\alpha,\varepsilon}$ is at most a constant multiple of $k^{s_{\alpha,\varepsilon}(-1+(\alpha+\varepsilon)\log 2)}$. Our choice of $s_{\alpha,\varepsilon}$ ensures that the exponent is less than $-1$. The series is thus summable. It follows that finitely many $k > T^*$ belong to the intersection of the $\mathscr{L}(\mathbf{X}^{(i)})$. As $T^*$ is almost surely finite (the proof of this can be found in [31, Section 2]), the result follows by letting $\varepsilon \to 0$ so that $s_{\alpha,\varepsilon} \to h(\alpha)$. □

## 5. The lower bound for $s_\alpha$

Here we establish the lower bounds on $m_\alpha$.

**Proposition 5.1.** *Let $s_\alpha$ be as in Theorem 1.7 and $h(\alpha)$ be as in (1.4). For $0 \leqslant \alpha < 1/\log 2$, it holds that $s_\alpha \geqslant h(\alpha)$. Moreover, $\bigcap_{i=1}^{h(\alpha)} \mathscr{L}(\mathbf{X}^{(i)}(\alpha)) = \infty$ a.s. Lastly, when $\alpha \geqslant 1/\log 2$ it holds that $s_\alpha = \infty$.*

The proof is at the end of this section. The idea is to use Fourier analysis like in [15]. We begin by considering continuity points $\alpha$ of $h$, and set $m = h(\alpha)$. If $\alpha \geqslant 1/\log 2$ so that $h(\alpha) = \infty$, we take $m$ to be a fixed arbitrarily large integer. We then fix $m$ i.i.d. Poisson vectors $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}$.

**Lemma 5.2.** *Let $I = (0,k]$ and $\varepsilon > 0$. Then the event*

$$A = \left\{ \bigcup_{i=1}^{m} \mathcal{L}(I, \mathbf{X}^{(i)}) \subset \left[ 0, \frac{mk}{\alpha \varepsilon} \right] \right\}$$

*holds with probability $\mathbb{P}[A] \geqslant 1 - \varepsilon$.*

**Proof.** This is a straightforward application of Markov's inequality. $\qquad\square$

Recall our notation that $\mathbf{X}^{(i)} = (X_j^{(i)})_{i=1}^{\infty}$.

**Lemma 5.3.** *Fix $\varepsilon \in (0, 1/2)$. For any $\delta > 0$, there is a constant $C = C(\varepsilon, \delta, m, \alpha)$ such that the event*

$$E = E(\varepsilon, \delta, m, \alpha) := \left\{ \min_i \sum_{\ell < j \leqslant k} X_j^{(i)} \geqslant (1 - \delta)\alpha \log(k/\ell) - C \quad \forall 1 \leqslant \ell \leqslant k \right\} \tag{5.1}$$

*holds with probability $\mathbb{P}[E] \geqslant 1 - \varepsilon$.*

We remark that this is a straightforward generalization of [15, Lemma 3.3].

**Proof.** In fact, we will take $C \gg (1/(\alpha\delta^2)) \log(m/\varepsilon)$ while ensuring that $C \geqslant 1$. Note that the claim holds vacuously when $\ell \geqslant e^{-C}k$, so we need only consider the case $\ell < e^{-C}k$.

Fix $1 \leqslant i \leqslant m$ and let $B$ denote the event that

$$\sum_{\ell < j \leqslant k} X_j^{(i)} < (1 - \delta)\alpha \log(k/\ell) - 1 \quad \text{for some } \ell \leqslant e^{-C}k.$$

Writing $\ell'$ for the smallest power of 2 with $\ell' > \ell$, we thus have

$$\sum_{\ell' < j \leqslant k} X_j^{(i)} \leqslant \alpha(1 - \delta) \log(k/\ell') \tag{5.2}$$

on $B$.

Let $\mathbb{D} = \{2^i\}_{i \geqslant 0}$ denote the set of non-negative powers of 2. From (5.2) it follows that

$$1_B \leqslant \sum_{\substack{\ell' \leqslant 2e^{-C}k, \\ \ell' \in \mathbb{D}}} (1 - \delta)^{\sum_{\ell' < j \leqslant \ell} X_j^{(i)} - \alpha(1-\delta)\log(k/\ell')}. \tag{5.3}$$

Each of the variables $X_j^{(i)}$ is independent and Poisson with mean $\mathbb{E}X_j^{(i)} = \alpha/j$. For a Poisson variable $P$ with mean $\lambda$, we have $\mathbb{E}a^P = e^{\lambda(a-1)}$. Moreover, the sum $\sum_{\ell' < j \leqslant k} X_j^{(i)}$ is Poisson with mean $\alpha \log(k/\ell') + O(1)$. Taking expectations on both sides of (5.3) therefore yields

$$\mathbb{P}[B] \ll \sum_{\substack{\ell' \leqslant 2e^{-C}k, \\ \ell' \in \mathbb{D}}} \exp[-\alpha\delta - \alpha(1-\delta)\log(1-\delta)] \log(k/\ell')$$

Near zero, $-\delta - (1-\delta)\log(1-\delta) = -(1+o(1))\delta^2$. Therefore

$$\mathbb{P}[B] \ll \sum_{\substack{\ell' \leqslant 2e^{-C}k, \\ \ell' \in \mathbb{D}}} (k/\ell')^{-\alpha\delta^2} \ll e^{-\alpha\delta^2 C},$$

since the sum of the geometric series is bounded by a constant multiple of its first term. Choosing $C \gg (1/(\alpha\delta^2))\log(m/\varepsilon)$ ensures that $\mathbb{P}[B] \leqslant \varepsilon/m$, and now the claim follows by the union of events bound. $\qquad\square$

An insight of Eberhard, Ford and Green [15] (which they credit to Maier and Tenenbaum [27]) is that Fourier analysis may be used to obtain a lower bound on the size of the set

$$S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}) := \{(n_i - n_m)_{1 \leqslant i < m} : n_i \in \mathscr{L}(I, \mathbf{X}^{(i)})\}. \tag{5.4}$$

We show that the technique extends to the present level of generality. Note that the analysis of [15] may be regarded as the special case $m = 3$ and $\alpha = 1$ of our arguments.

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the unit torus. For $\theta \in \mathbb{T}$, let $\|\theta\|$ denote the distance to $\mathbb{Z}$ and set $e(z) = e^{2\pi i z}$. Write $\theta = (\theta_1, \ldots, \theta_m) \in \mathbb{T}^m$. Define

$$\mathbb{T}_0^m = \{\theta \in \mathbb{T}^m : \theta_1 + \cdots + \theta_m = 0\}.$$

Fix an interval $I$. We define the function $F(\theta) = F_I(\theta, \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)})$ via the formula

$$F(\theta) := \prod_{j \in I} \prod_{i=1}^{m} \left( \frac{1 + e(j\theta_i)}{2} \right)^{X_j^{(i)}}. \tag{5.5}$$

Writing $\theta_m = -\theta_1 - \cdots - \theta_{m-1}$, we regard $F(\theta)$ as a function on $\mathbb{T}^{m-1}$. It is straightforward to verify that its Fourier transform is a function $\hat{F} : \mathbb{Z}^{m-1} \to \mathbb{C}$ supported on $S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)})$ defined above in (5.4).

The Cauchy–Schwarz inequality yields that

$$\sum_{a \in \mathbb{Z}^{m-1}} |\hat{F}(a)|^2 \sum_{a : \hat{F}(a) \neq 0} 1 \geqslant \left[ \sum_{a \in \mathbb{Z}^{m-1}} \hat{F}(a) \right]^2. \tag{5.6}$$

Recognizing that

$$S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}) \geqslant |\{a : \hat{F}(a) \neq 0\}|,$$

that $\sum_{a \in \mathbb{Z}^{m-1}} \hat{F}(a) = F(0) = 1$, and that by Parseval's identity

$$\sum_{a \in \mathbb{Z}^{m-1}} |\hat{F}(a)|^2 = \int_{\mathbb{T}_0^m} |F(\theta)|^2 \, d\theta,$$

it follows from (5.6) that

$$|S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)})| \geqslant \left( \int_{\mathbb{T}_0^m} |F(\theta)|^2 \, d\theta \right)^{-1}. \tag{5.7}$$

Our next task is to bound the integral appearing in (5.7). We will achieve this in two steps: first, by obtaining a pointwise bound with respect to $\theta$ in Lemma 5.4, and second, by integrating out the $\theta$ dependence in Corollary 5.5.

Let $\beta \in (0,1)$ be a real parameter satisfying

$$\beta \alpha \log 2 = 1 - m^{-1} + \delta_2,$$

where $\delta_2 > 0$ and $\delta_2 = o_\delta(1)$ is the parameter occurring in (5.1). (Note that this is possible since $\alpha$ is a continuity point of $h$, so $m < (1 - \alpha \log 2)^{-1}$.) Recall the event $E = E(\varepsilon, \delta, m, \alpha)$ from Lemma 5.3.

**Lemma 5.4.** *Fix $\varepsilon \in (0,1/2)$ and let $I$ denote the interval $I := (k^{1-\beta}, k]$. For a tuple $\theta = (\theta_1, \ldots, \theta_m)$, let $j(\theta)$ denote an index for which $\|\theta_j\|$ is maximal. Then*

$$\mathbb{E}_{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}} 1_E |F(\theta)|^2 \ll_{\varepsilon, m} \prod_{\substack{i:\ 1 \leqslant i \leqslant m, \\ i \neq j(\theta)}} (\|k \theta_i\| \vee 1)^{-1+\delta_2}. \tag{5.8}$$

**Proof.** For each $1 \leqslant i \leqslant m$, set $t_i := k\theta_i$ and define the cut-off parameter

$$k_i := \begin{cases} k^\beta & \|t_i\| \geqslant k^\beta, \\ \|t_i\| & 1 < \|t_i\| < k^\beta, \\ 1 & \|t_i\| \leqslant 1. \end{cases}$$

Note that $k_i \geqslant 1$ and $k_i \geqslant \|t_i\|^\beta$. For each $1 \leqslant i \leqslant m$, let $Y_i$ denote the expression

$$Y_i := \sum_{k_i < j \leqslant k} X_j^{(i)}.$$

By definition of the event $E$, when it occurs we have that

$$\sum_{i=1}^m Y_i \geqslant \alpha(1-\delta) \sum_{i=1}^m \log k_i - C(\varepsilon).$$

It will be convenient to rewrite this bound in the form

$$1_E \ll_\varepsilon \prod_{i=1}^m k_i^{-\alpha(1-\delta)\log 2} \prod_{i=1}^m 2^{Y_i}. \tag{5.9}$$

We define the quantity

$$R = R(\theta, \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}, \varepsilon, \delta, \alpha) := |F(\theta)|^2 \prod_{i=1}^m 2^{Y_i}. \tag{5.10}$$

To establish the bound (5.8), we will first show that $\mathbb{E}_{\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}} R \ll_{\varepsilon, m} 1$. Once this has been established, it will follow directly from (5.9) that

$$\mathbb{E} 1_E |F(\theta)|^2 \ll_{\varepsilon, m} \prod_{i=1}^m k_i^{-\alpha(1-\delta)\log 2}. \tag{5.11}$$

Obtaining (5.8) from (5.11) is a simple exercise in lower-bounding the maximum of a sequence by its geometric mean. Recalling the definition of $j(\theta)$ above, we have that

$$k_{j(\theta)} \geqslant \prod_{\substack{i:\ 1 \leqslant i \leqslant m, \\ i \neq j(\theta)}} k_i^{1/(m-1)}.$$

Consequently

$$\prod_{i=1}^{m} k_i^{-\alpha(1-\delta)\log 2} \leqslant \prod_{\substack{i:\, 1\leqslant i\leqslant m, \\ i\neq j(\theta)}} k_i^{-(m/(m-1))\alpha(1-\delta)\log 2}. \tag{5.12}$$

Substituting the lower bound $k_i \geqslant (\|k\,\theta_i\| \vee 1)^\beta$ into (5.12), we see that (5.11) implies

$$\mathbb{E}\,1_E |F(\theta)|^2 \ll_{\varepsilon,m} \prod_{\substack{i:\, 1\leqslant i\leqslant m, \\ i\neq j}} (\|k\,\theta_i\| \vee 1)^{-(m/(m-1))\alpha\beta(1-\delta)\log 2},$$

and since

$$\frac{m}{m-1}\alpha\beta(1-\delta)\log 2 = 1 + o_\delta(1)$$

by definition of $\beta$, the desired bound follows.

Finally we justify that $\mathbb{E}_{\mathbf{X}^{(1)},\ldots,\mathbf{X}^{(m)}} R \ll_{\varepsilon,m} 1$. Substituting (5.5) into (5.10) yields that

$$R = \prod_{i=1}^{m} \left[ 2^{Y_i} \prod_{j\in I} \left| \frac{1 + e(j\theta_i)}{2} \right|^{2X_j^{(i)}} \right].$$

Recall that $Y_i = \sum_{k_i < j \leqslant k} X_j^{(i)}$ and that $I = (k^{1-\beta}, k]$. Thus

$$R = \prod_{i=1}^{m} \left[ \prod_{k^{1-\beta} < j \leqslant k_i} \left| \frac{1 + e(j\theta_i)}{2} \right|^{2X_j^{(i)}} \prod_{k_i < j \leqslant k} \left| \frac{1 + e(j\theta_i)}{\sqrt{2}} \right|^{2X_j^{(i)}} \right].$$

To compute $\mathbb{E}R$, we recall that if $P$ is a Poisson random variable with mean $\lambda$, then we have $\log \mathbb{E}a^P = e^{\lambda(a-1)}$. Since the variables $X_j^{(i)}$ are independent Poisson variables with mean $\mathbb{E}X_j^{(i)} = \alpha/j$, it follows after some simplification that

$$\mathbb{E}R = \exp \sum_{i=1}^{m} \left[ \sum_{k^{1-\beta} < j \leqslant k_i} \frac{\alpha(\cos(2\pi j\theta_i) - 1)}{2j} + \sum_{k_i < j \leqslant k} \frac{\alpha\cos(2\pi j\theta_i)}{j} \right]. \tag{5.13}$$

Next we apply the following bound from [15, Lemma 3.4]:

$$\sum_{j\leqslant k} \frac{\cos(2\pi j\theta)}{j} = \log\min(k, \|\theta\|^{-1}) + O(1). \tag{5.14}$$

Substituting into (5.13) yields that

$$\mathbb{E}R = \exp \alpha \sum_{i=1}^{m} \left[ \frac{1}{2}\log\frac{k^\beta \vee \|k\,\theta_i\|}{k_i \vee \|k\,\theta_i\|} - \frac{1}{2}\log\frac{k^\beta}{k_i} + \log\frac{k_i \vee \|k\,\theta_i\|}{1 \vee \|k\,\theta_i\|} + O(1) \right].$$

By definition of $k_i$, each summand is $O(1)$. Therefore $\mathbb{E}R \ll_{\varepsilon,m} 1$, as desired. $\qquad\square$

Finally we integrate the bound (5.8) with respect to $\theta$ to obtain a bound on the integral appearing in (5.7).

**Corollary 5.5.** *Let* $I = (k^{1-\beta}, k]$ *and recall the function* $F(\theta) = F_I(\theta)$ *from (5.5). Then*

$$\int_{\mathbb{T}_0^m} \mathbb{E}\,1_E |F(\theta)|^2 \, d\theta \ll_{\varepsilon,m} k^{1-m}.$$

**Proof.** After the change of variables $t = k\theta$, the bound reduces to verifying that

$$\int_{k \cdot \mathbb{T}_0^m} \mathbb{E}1_E |F(\theta)|^2 \, dt \ll_{\varepsilon,m} 1.$$

Applying Lemma 5.4, we have

$$\int_{k \cdot \mathbb{T}_0^m} \mathbb{E}1_E |F(\theta)|^2 \, dt \ll_{\varepsilon,m} \int_{k \cdot \mathbb{T}_0^m} \prod_{\substack{i: \, 1 \leqslant i \leqslant m, \\ i \neq j(t)}} (\|t_i\| \vee 1)^{-(1+\delta_2)} \, dt. \tag{5.15}$$

By symmetry, we may upper-bound the integral (5.15) by

$$m \int_{k \cdot \mathbb{T}_0^m} \prod_{i=1}^{m-1} (\|t_i\| \vee 1)^{-(1+\delta_2)} \, dt = m \left( \int_{-k/2}^{k/2} (|t_i| \vee 1)^{-(1+\delta_2)} \, dt_i \right)^{m-1}. \tag{5.16}$$

The rightmost integral in (5.16) is bounded above by

$$m \left( 2 \int_1^\infty t_i^{-(1+\delta_2)} \, dt_i + 2 \right)^{m-1}.$$

The desired result now follows since the latter integral is finite and independent of $k$. $\qquad\square$

Substituting the result of Corollary 5.5 into (5.7) lets us show

$$|S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)})| \gg k^{m-1}$$

with high probability.

**Proposition 5.6.** *Let* $I = (k^{1-\beta}, k]$. *With probability greater than* $1/2$,

$$S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}) \subset [-ck, ck]^{m-1}$$

*for some constant* $c$ *and* $|S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)})| \gg k^{m-1}$.

**Proof.** By Lemma 5.2 with $\varepsilon = 1/3$, with probability greater than $2/3$ we have

$$\mathscr{L}(I, \mathbf{X}^{(i)}) \subset \left[ 0, \frac{3mk}{\alpha} \right] \subset [0, ck]$$

for some constant $c$ and all $i$. It follows that

$$S(I; \mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(m)}) \subset [-ck, ck]^{m-1}.$$

By Lemma 5.3 with $\varepsilon = 1/20$, the event $E(\varepsilon, \delta, m, \alpha)$ holds with probability greater than $19/20$, and by Corollary 5.5,

$$\int_{\mathbb{T}_0^m} \mathbb{E}1_E |F(\theta)|^2 \, d\theta \ll_{\varepsilon,m} k^{1-m}.$$

It follows by Markov's inequality that for some sufficiently large $L$,

$$\mathbb{P}\left[\left\{\int_{\mathbb{T}_0^m}|F(\theta)|^2\,d\theta \ll k^{1-m}\right\}^c\right] \leqslant \mathbb{P}[E^c] + \mathbb{P}\left[1_E\int_{\mathbb{T}_0^m}|F(\theta)|^2\,d\theta > Lk^{1-m}\right]$$

$$\leqslant \frac{1}{20} + \frac{\mathbb{E}1_E\int_{\mathbb{T}_0^m}|F(\theta)|^2\,d\theta}{Lk^{1-m}}$$

$$\leqslant \frac{1}{10}.$$

It then follows from (5.7) that with probability at least $9/10$,

$$|S(I;\mathbf{X}^{(1)},\dots\mathbf{X}^{(m)})| \geqslant \left(\int_{\mathbb{T}_0^m}|F(\theta)|^2\,d\theta\right)^{-1} \gg k^{m-1}.$$

Both events then occur simultaneously with probability at least $1/2$, completing our proof. $\square$

**Proposition 5.7.** *Let $D$ be a sufficiently large constant and let $I = (k^{1-\beta},Dk]$. With probability bounded away from zero there exists an array $(x_j^{(i)} : 1 \leqslant i \leqslant m, j \in I)$ for which*

$$0 \leqslant x_j^{(i)} \leqslant X_j^{(i)}$$

*for all $i$ and $j \in I$, and*

$$\sum_{j\in I} jx_j^{(1)} = \cdots = \sum_{j\in I} jx_j^{(m)}.$$

**Proof.** We define the interval $I' = (k^{1-\beta},k]$. By Proposition 5.6, with probability greater than $1/2$,

$$S(I';\mathbf{X}^{(1)},\dots,\mathbf{X}^{(m)}) \subset [-ck,ck]^{m-1}$$

and $|S(I';\mathbf{X}^{(1)},\dots,\mathbf{X}^{(m)})| \gg k^{m-1}$.

Let $d = 2(5/2)^{1/\alpha}$. Straightforward calculations tell us that independently with probability greater than $1/2$ there exists $j_m \in (2ck,cdk]$ such that $X_{j_m}^{(m)} > 0$. Given this $j_m$, there are $\gg k^{m-1}$ collections of constants $\{j_\ell\}_1^{m-1} \in (ck,(d+1)ck]^{m-1}$ such that

$$(j_m - j_1,\dots,j_m - j_{m-1}) \in S(I';\mathbf{X}^{(1)},\dots,\mathbf{X}^{(m)}).$$

Therefore, independently with probability bounded away from zero, there is one such collection $\{j_\ell\}_1^{m-1}$ such that $X_{j_\ell}^{(\ell)} > 0$ for all $\ell$.

By the definition of $S(I';\mathbf{X}^{(1)},\dots,\mathbf{X}^{(m)})$, we have a set of values $(n_i \in \mathscr{L}(I',\mathbf{X}^{(i)}))_{i=1}^m$ such that

$$(n_1 - n_m)_1^{m-1} = (j_m - j_\ell)_1^{m-1}.$$

Manipulating the terms coordinate-wise yields

$$j_1 + n_1 = \cdots = j_m + n_m.$$

Our claim follows by setting $D = (d+1)c$. $\square$

**Corollary 5.8.** $\bigcap_{i=1}^m \mathscr{L}(\mathbf{X}^{(i)})$ *is almost surely infinite.*

**Proof.**    We define $k_1$ to be sufficiently large for Proposition 5.7 to hold, then inductively define $k_{i+1} = (Dk_i)^{1/(1-\beta)}$. The intervals $I_i = (k_i^{1-\beta}, Dk_i]$ are pairwise disjoint and the set

$$\bigcap_{j=1}^{m} \mathscr{L}(I_i, \mathbf{X}^{(j)})$$

is non-empty with probability bounded away from zero. Since these events are independent for distinct values of $i$, the claim now follows from the second Borel–Cantelli lemma.    □

**Proof of Proposition 5.1.**    When $\alpha < 1/\log 2$, as we had fixed $m = h(\alpha)$, by Corollary 5.8 $s_\alpha \geqslant h(\alpha)$ holds for continuity points of $h$. For points of discontinuity, we fix $\alpha' < \alpha$ such that $h(\alpha') = h(\alpha) = m$. We couple two sets of independent realizations $\mathbf{X}^{(1)}(\alpha), \dots, \mathbf{X}^{(m)}(\alpha)$ and $\mathbf{X}^{(1)}(\alpha'), \dots, \mathbf{X}^{(m)}(\alpha')$ such that $X^{(i)}(\alpha') \subseteq X^{(i)}(\alpha)$ for all $i$. Now $\bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)}(\alpha'))$ is almost surely finite by Corollary 5.8. Our coupling ensures $\mathscr{L}(\mathbf{X}^{(i)}(\alpha') \subseteq \mathscr{L}(\mathbf{X}^{(i)}(\alpha))$ for all $i$, so

$$\bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)(\alpha')}) \subseteq \bigcap_{i=1}^{m} \mathscr{L}(\mathbf{X}^{(i)}(\alpha)).$$

Thus the second intersection is almost surely infinite, and our claim follows. When $\alpha \geqslant 1/\log 2$, all of the preliminary results in this section hold for arbitrarily large $m$, and thus every finite intersection is almost surely infinite.    □

**Proof of Theorem 1.7.**    This now follows immediately from Propositions 4.2 and 5.1.    □

## Acknowledgements

## References

[1] Apostol, T. M. (1976) *Introduction to Analytic Number Theory*, Undergraduate Texts in Mathematics, Springer.

[2] Arratia, R., Barbour, A. and Tavaré, S. (1992) Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2** 519–535.

[3] Arratia, R., Barbour, A. and Tavaré, S. (2000) Limits of logarithmic combinatorial structures. *Ann. Probab.* **28** 1620–1644.

[4] Arratia, R., Barbour, A. and Tavaré, S. (2016) Exploiting the Feller coupling for the Ewens sampling formula. *Statist. Sci.* **31** 27–29.

[5] Arratia, R. and Tavaré, S. (1992) The cycle structure of random permutations. *Ann. Probab.* **20** 1567–1591.

[6] Arratia, R. and Tavaré, S. (1994) Independent process approximations for random combinatorial structures. *Adv. Math.* **104** 90–154.

[7] Babai, L. (1981) On the order of uniprimitive permutation groups. *Ann. of Math.* **113** 553–568.

[8] Betz, V., Ueltschi, D. and Velenik, Y. (2011) Random permutations with cycle weights. *Ann. Appl. Probab.* **21** 312–331.

[9] Bovey, J. (1980) The probability that some power of a permutation has small degree. *Bull. London Math. Soc.* **12** 47–51.

[10] Crane, H. (2016) The ubiquitous Ewens sampling formula. *Statist. Sci.* **31** 1–19.

[11] Crane, H. (2016) Rejoinder: The ubiquitous Ewens sampling formula. *Statist. Sci.* **31** 37–39.

[12] Davenport, J. and Smith, G. (2000) Fast recognition of alternating and symmetric Galois groups. *J. Pure Appl. Algebra* **153** 17–25.

[13] Dixon, J. D. (1992) Random sets which invariably generate the symmetric group. *Discrete Math.* **105** 25–39.

[14] Eberhard, S., Ford, K. and Green, B. (2016) Permutations fixing a *k*-set. *Internat. Math. Res. Notices* **2016** 6713–6731.

[15] Eberhard, S., Ford, K. and Green, B. (2017) Invariable generation of the symmetric group. *Duke Math. J.* **166** 1573–1590.

[16] Ercolani, N. M. and Ueltschi, D. (2014) Cycle structure of random permutations with cycle weights. *Random Struct. Alg.* **44** 109–133.

[17] Erdős, P. and Turán, P. (1967) On some problems of a statistical group-theory, II. *Acta Math. Hungar.* **18** 151–163.

[18] Euler, L. (1737) Observationes circa series infinitas. *Commentarii Academiae Scientarum Petropolitanae* **9** 160–188.

[19] Ewens, W. (1972) The sampling theory of selectively neutral alleles. *Theoret. Popul. Biol.* **3** 87–112.

[20] Feller, W. (1945) The fundamental limit theorems in probability. *Bull. Amer. Math. Soc.* **51** 800–832.

[21] Gladkich, A. and Peled, R. (2018) On the cycle structure of Mallows permutations. *Ann. Probab.* **46** 1114–1169. doi: 10.1214/17-AOP1202.

[22] Granville, A. (2008) The anatomy of integers and permutations. Preprint.

[23] Guralnick, R. and Magaard, K. (1998) On the minimal degree of a primitive permutation group. *J. Algebra* **207** 127–145.

[24] Heintz, J. (1986) On polynomials with symmetric Galois group which are easy to compute. *Theoret. Comput. Sci.* **47** 99–105.

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

[26] Łuczak, T. and Pyber, L. (1993) On random generation of the symmetric group. *Combin. Probab. Comput.* **2** 505–512.

[27] Maier, H. and Tenenbaum, G. (1984) On the set of divisors of an integer. *Inventio. Math.* **76** 121–128.

[28] Mallows, C. L. (1957) Non-null ranking models, I. *Biometrika* **44** 114–130.

[29] Mukherjee, S. (2016) Fixed points and cycle structure of random permutations. *Electron. J. Probab.* **21** #40.

[30] Musser, D. R. (1978) On the efficiency of a polynomial irreducibility test. *J. Assoc. Comput. Mach.* **25** 271–282.

[31] Pemantle, R., Peres, Y. and Rivin, I. (2016) Four random permutations conjugated by an adversary generate $S_n$ with high probability. *Random Struct. Alg.* **49** 409–428.

[32] van der Waerden, B. (1934) Die Seltenheit der Gleichungen mit Affekt. *Math. Ann.* **109** 13–16.