

# FOUR RANDOM PERMUTATIONS CONJUGATED BY AN ADVERSARY GENERATE $\mathcal{S}_n$ WITH HIGH PROBABILITY

ROBIN PEMANTLE, YUVAL PERES, AND IGOR RIVIN

ABSTRACT. We prove a conjecture dating back to a 1978 paper of D.R. Musser [Mus78], namely that four random permutations in the symmetric group  $\mathcal{S}_n$  generate a transitive subgroup with probability  $p_n > \varepsilon$  for some  $\varepsilon > 0$  independent of  $n$ , even when an adversary is allowed to conjugate each of the four by a possibly different element of  $\mathcal{S}_n$  (in other words, the cycle types already guarantee generation of  $\mathcal{S}_n$ ). This is closely related to the following random set model. A random set  $M \subseteq \mathbb{Z}^+$  is generated by including each  $n \geq 1$  independently with probability  $1/n$ . The sumset  $\text{sumset}(M)$  is formed. Then at most four independent copies of  $\text{sumset}(M)$  are needed before their mutual intersection is no longer infinite.

## 1. INTRODUCTION

**1.1. Background and motivation.** The roots of this work are in computational algebra. It is a result going back to van der Waerden [vdW34] that most polynomials  $p(x) \in \mathbb{Z}[x]$  of degree  $n$  have Galois group  $\mathcal{S}_n$ . Computing the Galois group is a central problem in computational number theory and is a fundamental building block for the solution of seemingly unrelated problems (see [Riv13] for an extensive discussion). Therefore, one cannot take for granted being in the “generic” case and one would like an effective and speedy algorithm for determining whether the Galois group of  $p(x)$  is the full symmetric group.

There are deterministic polynomial time algorithms to answer this. The first is due to S. Landau; a simpler and more efficient algorithm was proposed by the third author (see [Riv13]). These algorithms, however, are of purely theoretical interest due to their very long run times (their complexity is of the order of  $O(n^{40})$ , where  $n$  is the degree of the polynomial). The best algorithms in practice are Monte Carlo algorithms. To discuss Monte Carlo testing for full Galois group, one begins with two classical results<sup>1</sup>.

**Theorem 1.1** (Dedekind). *If  $p(x)$  is square-free modulo a prime  $q$  and the factorization of  $p(x)$  modulo  $q$  into irreducible factors yields degrees  $d_1, d_2, \dots, d_k$ , then the Galois group of  $G$   $p(x)$  has an element whose cycle decomposition has lengths precisely  $\{d_1, \dots, d_k\}$ .*

**Theorem 1.2** (Frobenius Density Theorem). *The density of prime numbers  $q$  for which  $p(x) \pmod q$  has factors whose degrees are  $d_1, \dots, d_k$  is equal to the density in the Galois group  $G \subseteq \mathcal{S}_n$  for  $p(x)$  of elements of  $\mathcal{S}_n$  with cycle type  $d_1, \dots, d_k$ .*

---

1991 *Mathematics Subject Classification.* 60C05;12Y05; 68W20; 68W30; 68W40.

*Key words and phrases.* sumset, cycle, Poisson, dimension, Galois group.

Igor Rivin would like to thank the Brown University Mathematics Department and ICERM for their hospitality and financial support during the preparation of this paper. Robin Pemantle was supported in part by NSF grant # DMS-1209117.

<sup>1</sup>In the literature, the much harder Chebotarev Density Theorem is often used in place of the Frobenius Density Theorem.

*Remark 1.* Theorem 1.2 is useless without effective convergence bounds. The first step in this direction was made by the J. Lagarias and A. Odlyzko [LO77] – they proved *conditional* (on the Riemann hypothesis for certain L-functions) results with “effectively computable” (but quite hard to compute) constants. A couple of years later, Oesterlé [Oes79] claimed a computation of the constants, but his computation has not been published in the intervening 35 years (despite being used by J.-P. Serre in [Ser81]). Finally, the problem was put to rest by B. Winckler in [Win13]) at the end of 2013(!) – Winckler shows both unconditional and conditional results (with somewhat worse constants in the latter case than those claimed by Oesterlé).

Together, these two results tell us that without yet knowing  $G$  we can uniformly sample cycle decompositions  $V_i = \{d_{i,1}, \dots, d_{i,k(i)}\}$  of elements of  $G$  by sampling integers  $q_i$  at random and setting  $V_i$  equal to the set of degrees of the irreducible factors of  $p(x)$  modulo  $q_i$  (it should be noted that factoring modulo a prime can be done quite efficiently using variants of Berlekamp’s algorithm). A result of C. Jordan allows us to turn this into a probabilistic test for  $G = \mathcal{S}_n$  with certain acceptance and possible false rejection.

**Theorem 1.3** (C. Jordan). *Suppose a subgroup  $H$  of  $\mathcal{S}_n$  ( $n > 12$ ) acts transitively on  $[n]$ . If it contains at least one cycle of prime length between  $n/2 + 1$  and  $n - 5$ , then it is either  $\mathcal{S}_n$  or the alternating group  $\mathcal{A}_n$ .*

Certification that  $G$  is not alternating and contains at least one long prime cycle is trivial: we just check that at least one of the lists  $V_1, \dots, V_r$  corresponds to an odd permutation class and at least one contains a prime value in  $[n/2 + 1, n - 5]$  – some power of the corresponding permutation will be a long prime cycle. In a uniform random permutation, the cycle containing a given element, say 1, has length exactly uniform on  $[n]$ . The Prime Number Theorem guarantees that the number of primes in  $[n/2 + 1, n - 5]$  is asymptotic to  $n/(2 \log n)$ . It follows that if  $G$  is truly  $\mathcal{S}_n$ , then each  $V_i$  contains a large prime with probability at least  $(1 + o(1))/(2 \log n)$ . Also, each  $V_i$  corresponds to an odd class with probability  $1/2$ . Therefore, if  $G$  is truly  $\mathcal{S}_n$ , we will quickly discover that the hypotheses of Theorem 1.3 other than transitivity are satisfied.

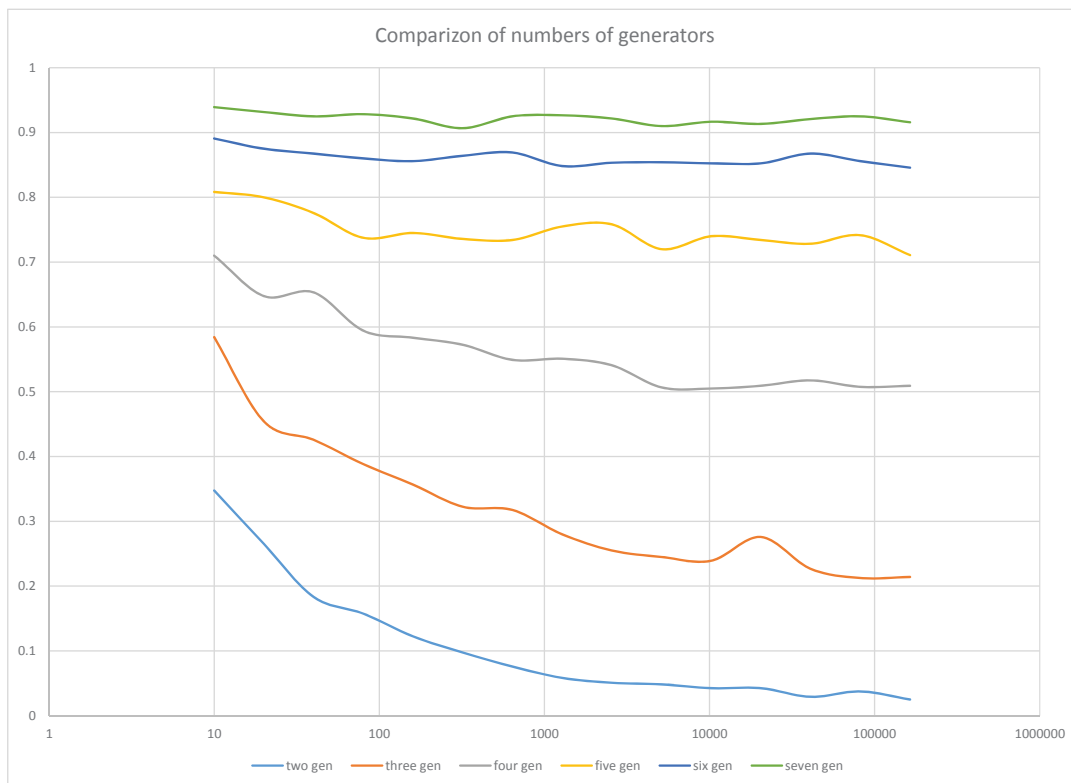
Establishing transitivity of  $G$  when we know only  $V_1, \dots, V_r$  must involve showing that *any* set of cycles in these respective conjugacy classes generates a transitive subgroup of  $\mathcal{S}_n$ . Let us say in this case that classes  $V_1, \dots, V_r$  **invariably generate** a transitive group. If the action of  $G$  leaves a subset  $I$  of  $[n]$  invariant, then  $|I|$  will appear as a sum of cycle sizes of every element of  $G$ . The converse holds as well: if the sumsets of  $V_1, \dots, V_r$  have no common intersection then the corresponding permutations invariably generate a transitive group. This leads to the following test:

**Algorithm:** Sample some random primes  $\{q_1, q_2, \dots, q_r\}$ , compute the degree sets  $V_i := \{d_{i,1}, \dots, d_{i,k(i)}\}$  of the factors of  $p$  modulo  $q_i$ , and the sumsets  $S_i := \text{sumset}(V_i)$ . If the sets  $S_i$  have some element in common other than 0 and  $n$ , or if none of the sets  $V_i$  contains a prime greater than  $n/2$ , or if all  $r$  conjugacy classes are even, then output NEGATIVE, otherwise output POSITIVE.

In the algorithm above, we have implicitly defined *sumset*:

**Definition 1.4.** *The sumset of a (multi)set  $S = \{k_1, \dots, k_l\}$  is the (multi)set of all sums of subsets of  $S$ .*

FIGURE 1. Experimental results



What needs to be checked next is that that we can choose  $r(\varepsilon)$  not too large so that if  $p(x)$  does have full Galois group then a NEGATIVE output has probability less than  $\varepsilon$ . For this we need to answer the question: given  $\varepsilon > 0$ , how many uniformly random permutations in  $\mathcal{S}_n$  do we have to choose before their cycle length sumsets have no common value in  $\{1, \dots, n-1\}$ ? If there is a number  $m_0$  such that this probability is at least  $\delta$  for  $m_0$  random permutations, then it is at least  $1 - (1 - \delta)^j$  for  $j m_0$  permutations. Therefore we may begin by asking about the value of  $m_0$ : how many IID uniform permutations are needed so that their cycle length sumsets have no nontrivial common value with probability that remains bounded away from zero as  $n \rightarrow \infty$ ?

It turns out that this question was first raised by D. R. Musser [Mus78], for reasons similar to ours. Musser did some experiments (where he was hindered both by the performance of the hardware of the time and by using an algorithm considerably inferior to the one we describe below), and observed that 5 elements should be sufficient; see also [DS00]. More modern experimental evidence (See Figure 1.1) is as follows. Each curve represents the probability that some number of random elements of  $S_n$  invariably generates a transitive subgroup, where the  $x$ -axis measures  $n$  logarithmically. The goal is to prove that one of these curves does not go to zero as  $n \rightarrow \infty$ . Evidently, even the lowest of these curves does

not seem to go to zero very fast (the horizontal axis is logarithmic), thus we might believe the question to be delicate.

This question (again, for Galois-theoretic reasons) was considered by J. Dixon in his 1992 paper [Dix92], and he succeeded in showing that  $O(\sqrt{\log n})$  elements are sufficient for fixed  $\varepsilon$ . Pictorially, to get above  $\varepsilon$  on the graph, it would suffice to go to the curve numbered  $C\sqrt{\log n}$  from the bottom. He conjectured, as did Musser, that his bound was not sharp, and  $O(1)$  elements should suffice. He proved that if that is, indeed, true, then to check that the Galois group is all of  $S_n$  we need to factor modulo  $O(\log \log n)$  primes. Dixon's  $O(1)$  conjecture was proved by T. Luczak and L. Pyber in 1993 ([LP93]), however the implied constant was absurdly high: on the order of  $2^{100}$  (and it can be shown that their method cannot be improved to yield a qualitatively better result).

**1.2. Main results.** Our main result is that  $m_0 \leq 4$ . We do not settle whether  $m_0$  could be 2 or 3, though we discuss why very likely  $m_0 = 4$  (though experimental evidence is inconclusive) and why proving this via analyses such as ours would require significantly more work.

Let  $\mathbb{P}_N$  denote the uniform measure on the symmetric group,  $\mathcal{S}_N$ . For a permutation  $\sigma \in \mathcal{S}_n$ , let  $\mathcal{I}(\sigma)$  denote the set of sizes of invariant sets of  $\sigma$ , that is,

$$\mathcal{I}(\sigma) := \{|I| : I \text{ is a proper subset of } [N] \text{ and } \sigma[I] = I\}.$$

In other words,  $\mathcal{I}(\sigma) = \text{sumset}(V(\sigma))$  when  $V$  is the multiset of cycle lengths of  $\sigma$ . Trivially, the set  $\mathcal{I}(\sigma)$  is symmetric about  $N/2$ , meaning that it is closed under  $k \mapsto N - k$ . As usual,  $\mathbb{P}_N^j$  denotes the  $j$ -fold product of uniform measures on  $\mathcal{S}_N$ .

**Theorem 1.5** (Main result). *There is a positive number  $b_0$  such that for all  $N$ ,*

$$\mathbb{P}_N^4 \left\{ (\sigma_1, \sigma_2, \sigma_3, \sigma_4) : \bigcap_{j=1}^4 \mathcal{I}(\sigma_j) = \emptyset \right\} \geq b_0.$$

The ideas behind the proof of this are more evident when we take  $N$  to infinity, resulting in the following Poisson model. Let  $\mathbb{P}$  denote the probability measure on  $(\mathbb{Z}^+)^{\infty}$  making the coordinates  $X_j(\omega) := \omega_j$  into independent Poisson variables with  $\mathbb{E}X_j = 1/j$ . Let  $M = M(\omega)$  be the multiset having  $X_k$  copies of the positive integer  $k$ . Let  $S = S(\omega) = \text{sumset}(M(\omega))$  be the sumset; we may define this formally by

$$S = \left\{ \sum_k a_k \cdot k : a_k \leq X_k \text{ for all } k \right\}.$$

This is the analogue in the Poisson model of the set  $\mathcal{I}(\sigma)$  of sums of cycle lengths in the group theoretic model.

Let  $\mathbb{P}^4$  denote the fourfold product of  $\mathbb{P}$  on  $((\mathbb{Z}^+)^{\infty})^4$  and for a 4-sequence  $(\omega^1, \omega^2, \omega^3, \omega^4) \in ((\mathbb{Z}^+)^{\infty})^4$ , let  $X_{r,k}$  denote the  $k^{\text{th}}$  coordinate of  $\omega^r$ . Let  $S(\omega_r)$  denote the set of sumsets of  $\omega^r$ . Our main result on the Poisson model is:

**Theorem 1.6** (Poisson result).

$$\mathbb{P}^4 \left( \bigcap_{r=1}^4 S(\omega^r) = \emptyset \right) > 0.$$

We require a number of estimates of probabilities associated with the random sumset  $S$ . The most straightforward quantity to define, though, as it turns out, not the most useful, is the marginal probability  $p_n := \mathbb{P}(n \in S)$  of finding a number  $n$  in the random sumset. This quantity is estimated as follows.

**Theorem 1.7** (marginal probabilities). *Let  $\eta = \frac{1 - \log 2 - \log(1/\log 2)}{\log 2} \approx -0.08607\dots$*

*Then  $p_n = n^{\eta+o(1)}$ .*

We remark that the exponent  $\eta$  is familiar from number theoretic contexts. For example, the asymptotic density of integers  $m$  having a divisor in the interval  $[N, 2N]$ , is a quantity  $g(N)$  known to satisfy  $g(N) \sim (\log N)^{\eta+o(1)}$  as  $N \rightarrow \infty$  (see, e.g., [HT88]).

**1.3. Discussion.** The analysis relies on the following lemma of Arratia and Tavaré, to the effect that the joint distribution of number of cycles of lengths up to  $m = o(N)$  of a random permutation of  $\mathcal{S}_N$  look like independent Poissons (see also [Gra06] for further refinements).

**Lemma 1.8** ([AT92, Theorem 2]). *Let  $Q_{N,m}$  be the joint distribution, for  $1 \leq k \leq m$ , of the number of  $k$ -cycles in a uniform random permutation in  $\mathcal{S}_N$ . Let  $\nu_m := \prod_{j=1}^m \mathcal{P}(1/j)$  denote the product of Poisson laws with respective means  $1/j$ . Then there is a constant  $C > 0$  such that the **total variation distance** between these two distributions is bounded above by*

$$\|Q_{N,m} - \nu_m\|_{TV} \leq \exp(-C(N/m) \log(N/m)).$$

*In particular,  $\|Q_{N,m} - \nu_m\|_{TV} \rightarrow 0$  as  $N/m \rightarrow \infty$ .*

Our main result is proved by showing that the random set  $\mathcal{I}(\sigma)$  behaves roughly like a set of dimension  $\ln 2$ , that is, it typically has density  $n^{\ln 2 - 1 + o(1)}$  near  $n$ . It follows that intersecting four of these yields a co-dimension greater than 1, which is characteristic of a random set which is almost surely finite and possibly empty.

Given the relatively clean Poisson approximation, one might wonder why there is any difficulty at all in proving such a result. The reason for the difficulty is that the averages of certain quantities are dominated by exceptionally large contributions from sets of small probability and therefore do not represent the typical values. For example, let  $q_{n,k}$  be the probability that there is an invariant set of size  $k$  and let  $e_{n,k}$  be the expected number of invariant sets of size  $k$ . Because  $q_{n,k} \ll 1$ , one might expect that  $e_{n,k} \approx q_{n,k}$ , but it turns out

that  $e_{n,k} = 1$  precisely, for all  $n$  and  $k$  (simply check that each  $k$ -set has probability  $\binom{n}{k}^{-1}$  of being an invariant set). Thus  $\mathbb{P}_N(k \in \mathcal{I}(\sigma))$  is much smaller than the expected number of representations of  $k$  as a sum of cycle lengths. A similar phenomenon holds for the Poisson model. The expected number of ways that the integer  $n$  is the sum of elements of the random

multiset  $M$  is the  $z^n$  coefficient in the generating function  $\prod_{k=1}^{\infty} \exp(z^k/k)$ , which simplifies to precisely 1. We see that  $p_n$  is much smaller than this expectation.

What is more subtle is that even  $p_n$  does not give the right estimate. The right estimate is what is known in the statistical physics as the **quenched** estimate. This is the estimate obtained when a  $o(1)$  portion of the probability space is excluded which contributes non-negligibly to the quantity in question, in this case  $p_n$ . Holding key parameters at their typical values produces a “correct”, quenched estimate,  $\tilde{p}_n$ . It may sound strange to ask what is the probability that  $n$  is in the sumset under typical behavior because  $p_n$  is already a probability,

meaning it is averaged over all behaviors. To be clear, to obtain the quenched estimate  $\tilde{p}_n$ , we exclude a set of arbitrarily small probability (but a single set for all  $n$ ), such that off of this set the probability  $\tilde{p}_n$  of finding  $n \in S$  is much smaller than  $n^{-\eta}$ , decaying instead like  $n^{\log 2 - 1}$ .

This is important because  $|\eta| = 0.08607$  is a bit larger than  $1/12$ , whereas  $1 - \log 2$  is a little larger than  $1/4$ . Showing that a set has co-dimension  $|\eta|$  indicates that one should intersect twelve independent copies in order to arrive at the empty set. When the random sets have co-dimension  $1 - \log 2$ , however, only four should be required. Interestingly, it is no easier to prove that 12 suffice than that 4 suffice, because the estimate of  $p_n$  is as hard as the estimate of  $\tilde{p}_n$ .

Finally, we note that this is in some sense the “easy” direction. To show that the fourfold intersection is finite in the Poisson model and often empty in the permutation model requires only an upper bound on the marginals  $p_k$ . To show that a threefold intersection does not suffice would require an upper bound on the probability of  $j$  and  $k$  both being in  $\mathcal{I}(\sigma)$ . This appears more difficult.

## 2. ESTIMATES FOR THE POISSON MODEL

Throughout this section we work on the probability space  $(\Omega, \mathcal{F}, \mathbb{P})$  where  $(\omega, \mathcal{F}) = (\mathbb{Z}^+, 2^{\mathbb{Z}^+})^\infty$  and  $\mathbb{P}$  is the probability measure making the coordinates independent Poissons, the  $n^{\text{th}}$  having mean  $1/n$ . Our notation includes the coordinate variables  $\{X_n\}$ , the random multiset  $M$  and its sumset  $S$ . We also define partial sums

$$\begin{aligned} Z_n &:= \sum_{k=1}^n X_k; \\ W_n &:= \sum_{k=1}^n kX_k. \end{aligned}$$

Thus  $Z_n$  counts the cardinality of  $M \cap [n]$  and  $W_n$  is the sum of all elements of  $M \cap [n]$ , which is the greatest element of  $\text{sumset}(M \cap [n])$ . We will need estimates for the right tail of  $Z_n$  and  $W_n$ , which are obtained in a straightforward way from their moment generating functions.

Let  $\phi_{Z,n}(\lambda) := \mathbb{E}e^{\lambda Z_n}$  denote the moment generating function for  $Z_n$  and let  $\psi_{Z,n}(\lambda)$  denote  $\log \phi_{Z,n}(\lambda)$ . Let  $\phi_{W,n}$  and  $\psi_{W,n}$  denote the corresponding functions for  $W_n$  in place of  $Z_n$ . Let  $H_n := \sum_{j=1}^n 1/j$  denote the  $n^{\text{th}}$  harmonic number. Using  $\mathbb{E}e^{\lambda X_j} = \exp[(e^{\lambda j} - 1)/j]$  and summing over  $j$  leads immediately to

$$(2.1) \quad \psi_{Z,n}(\lambda) = H_n \cdot (e^\lambda - 1) .$$

Similarly,

$$(2.2) \quad \psi_{W,n}(\lambda) = \sum_{j=1}^n \frac{e^{j\lambda} - 1}{j} .$$

Markov’s inequality implies an upper bound

$$(2.3) \quad \log \mathbb{P}(Z_n \geq a) \leq \psi_{Z,n}(\lambda) - a\lambda$$

for any  $a > \mathbb{E}Z_n = H_n$ . Similarly

$$(2.4) \quad \log \mathbb{P}(W_n \geq a) \leq \psi_{W,n}(\lambda) - a\lambda$$

for any  $a > \mathbb{E}W_n = n$ .

**Lemma 2.1.**

(i) There is a function  $\beta(\varepsilon) \sim \varepsilon^2/2$  as  $\varepsilon \downarrow 0$  such that

$$\mathbb{P}(Z_n \geq (1 + \varepsilon) \log n) \leq e n^{-\beta(\varepsilon)}.$$

(ii) For  $\varepsilon > 0$ , let  $\tau_\varepsilon := \sup\{n : Z_n \geq (1 + \varepsilon) \log n\}$ . Then  $\tau_\varepsilon < \infty$  almost surely.

PROOF OF LEMMA 2.1 For a one-sided bound one does not need to optimize (2.3) in  $\lambda$  but may take the near optimal  $\lambda = \log(1 + \varepsilon)$ . Set  $a = (1 + \varepsilon) \log n$  to obtain

$$(2.5) \quad \log \mathbb{P}(Z_n \geq (1 + \varepsilon) \log n) \leq H_n \varepsilon - (1 + \varepsilon) \log(1 + \varepsilon) \log n.$$

Letting  $\beta(\varepsilon) := (1 + \varepsilon) \log(1 + \varepsilon) - \varepsilon \sim \varepsilon^2/2$  and observing that  $\sup_j H_j - \log j = 1$  gives

$$\log \mathbb{P}(Z_n \geq (1 + \varepsilon)n) \leq -\frac{\varepsilon^2}{2} \log n + O(\varepsilon + \varepsilon^3 \log n)$$

which proves (i).

For (ii), apply (i) with  $e^n$  in place of  $n$  for  $n = 1, 2, 3, \dots$  to see that

$$\mathbb{P}(Z_{e^n} \geq (1 + \varepsilon/3)n) \leq \exp(1 - \beta(\varepsilon/3)n).$$

By Borel-Cantelli,  $Z_{e^n} \geq (1 + \varepsilon/3)n$  finitely often almost surely. For  $e^{n-1} < k < e^n$ , the inequality

$$\frac{Z_k}{\log k} \leq \frac{Z_{e^n}}{n-1} \leq \frac{n}{n-1} \frac{Z_{e^n}}{n}$$

implies that  $Z_k \leq (1 + \varepsilon) \log k$  as long as  $Z_{e^n} n \leq 1 + \varepsilon/3$  and  $n/(n-1) < 1 + \varepsilon/3$ . We have seen by Borel-Cantelli that these are both true for  $n$  sufficiently large, proving (ii).  $\square$

The upper tail of  $W_n$  may be estimated in a similar way. Throughout the paper from this point on we will use the notation

$$(2.6) \quad m(n) := \lfloor n / \log n \rfloor.$$

**Lemma 2.2.**

$$(2.7) \quad \log \mathbb{P}(W_{m(n)} \geq n) \leq -\log n (\log \log n - 1).$$

It follows by Borel-Cantelli that  $\tau := \sup\{n : W_{m(n)} \geq n\}$  is almost surely finite.

PROOF OF LEMMA 2.2 The near optimal choice of  $\lambda$  in (2.4) is a little more complicated than was the optimal choice in (2.3). We take  $\lambda := \log n \log \log n / n$  and find that

$$\log \mathbb{P}(W_{m(n)} \geq n) \leq \underbrace{\sum_{j=1}^{n/\log n} \frac{\exp(j \log n \log \log n / n) - 1}{j}}_{S_n} - \log n \log \log n.$$

A glance at its power series shows the function  $(e^{\beta x} - 1)/x$  to be increasing in  $x$  for positive  $\beta$ . Hence the sum  $S_n$  above may be bounded above if we replace each term with the last term. The number of terms is  $n/\log n$  so this yields

$$\log \mathbb{P}(W_{m(n)} \geq n) \leq \frac{n}{\log n} \left[ \frac{\log n - 1}{n/\log n} \right] - \log n \log \log n = \log n - 1 - \log n \log \log n.$$

Thus  $\mathbb{P}(W_{m(n)} \geq n) = O(n^{-\alpha})$  for any  $\alpha$ , and in particular is summable. This proves (2.7) and the summability of  $n^{1-\log \log n}$  (since  $1 - \log \log n \ll 2$ ) finishes the Borel-Cantelli argument.  $\square$

The estimates (2.3)–(2.4) are sharp in the limit when optimized over  $\lambda$ .

**Lemma 2.3.** *For fixed  $x > 1$ , as  $n \rightarrow \infty$ ,*

$$\frac{1}{\log n} \log \mathbb{P}(Z_n \geq x \log n) = x - 1 - x \log x + o(1).$$

**PROOF OF LEMMA 2.3:**

Upper bound: Again we optimize (2.3). The optimal value of  $\lambda$  is  $\log(a/H_n)$  but in fact we may use the simpler, near-optimal value  $\lambda = \log(a/\log n)$ . Setting  $a = x \log n$  and  $\lambda = \log x$  yields

$$\log \mathbb{P}(Z_n \geq x \log n) \leq H_n(\log x - 1) - x \log n \log x$$

and plugging in  $H_n = \log n + \gamma + o(1)$  yields

$$\begin{aligned} \log \mathbb{P}(Z_n \geq x \log n) &= (\log n + \gamma + o(1))(\log x - 1) - x \log x \log n \\ &= \log n(\log x - 1 - x \log x) + (\gamma + o(1))(\log x - 1). \end{aligned}$$

For  $1 < x \leq e$  this gives the exact upper bound

$$(2.8) \quad \log \mathbb{P}(Z_n \geq x \log n) \leq \log n(\log x - 1 - x \log x)$$

while for  $x > e$ , one has an asymptotically negligible remainder term on the right hand side of  $(\gamma + o(1))(\log x - 1)$ .

Lower bound: We use a tilting argument. Let  $\mathbb{P}_x$  be the probability measure on  $(\Omega, \mathcal{F})$  making the coordinates independent Poissons with means  $x/n$ . Let  $G_n$  be the event that  $x \log n \leq Z_n \leq x \log n + (\log n)^{2/3}$ . The law under  $\mathbb{P}_x$  of  $Z_n$  is Poisson with mean  $xH_n = x(\log n + O(1))$ , from which it follows that  $\mathbb{P}_x(G_n) \rightarrow 1/2$  as  $n \rightarrow \infty$  for any fixed  $x > 1$ . The tilting argument is simply the inequality

$$\mathbb{P}(G_n) \geq \mathbb{P}_x(G_n) \inf_{\omega \in G_n} \frac{d\mathbb{P}}{d\mathbb{P}_x}(\omega)$$

On the  $\sigma$ -field  $\mathcal{F}_n$  generated by  $X_1, \dots, X_n$ , the Radon-Nikodym derivative is easily computed as

$$\begin{aligned} \frac{d\mathbb{P}}{d\mathbb{P}_x}(\omega) &= \prod_{k=1}^n e^{(x-1)/n} x^{-X_k} \\ (2.9) \quad &= \exp((x-1)H_n) x^{-Z_n}. \end{aligned}$$

On the event  $G_n$ , we have  $Z_n = x \log n + O(x)$ . Plugging this into (2.9) and using also  $H_n = \log n + O(1)$  shows that on  $G_n$ ,

$$\begin{aligned} \log \frac{d\mathbb{P}}{d\mathbb{P}_x} &= (x-1)H_n - Z_n \log x \\ (2.10) \quad &= \log n(x-1 - x \log x) + O(1), \end{aligned}$$

completing the proof.  $\square$



### 3. QUENCHED PROBABILITIES FOR THE POISSON MODEL AND THE PROOF OF THEOREM 1.6

The above lemmas are written for any  $\varepsilon > 0$  in case future work requires pushing  $\varepsilon$  arbitrarily close to zero. However, for our purposes,  $\varepsilon = 1/100$  will be fine. To simplify notation (and free up  $\varepsilon$  notationally for other uses) we set  $\varepsilon = 1/100$  in Lemma 2.1 and we set

$$T = \max\{\tau_{1/100}, \tau\}$$

where  $\tau$  is the supremum in Lemma 2.2. Define

$$(3.1) \quad p_n := \mathbb{P}[n \in S] ;$$

$$(3.2) \quad \tilde{p}_n := \mathbb{P}[T < m(n) \text{ and } n \in S] .$$

Thus  $\{\tilde{p}_n\}$  are the so-called quenched probabilities, with exceptional events  $\{T \geq m(n)\}$ . Although the exceptional events vary with  $n$ , they form a decreasing sequence, which allows us to assume without too much penalty that none of the exceptional events occurs. The following lemma encapsulates the dimension estimate in the Poisson model.

**Lemma 3.1** (dimension of  $S$ ). *There is a constant  $C$  such that for all  $n$ ,*

$$\tilde{p}_n \leq Cn^{-1+\ln 2+0.02} .$$

**PROOF OF LEMMA 3.1:** Let  $G_n$  be the event that  $T < m(n)$  while also  $n \in S$ ; this is the event whose probability we need to bound from above. Call a sequence  $\mathbf{y} = (y_1, y_2, \dots)$  admissible if it is coordinatewise less than or equal to  $\omega$ .

When  $G_n$  occurs, because  $\tau \leq m(n)$ , it is not possible for  $n$  to be a sum  $\sum_j jy_j$  for an admissible  $\mathbf{y}$  with  $y_j$  vanishing for  $j > m(n)$ : even setting  $y_j = \omega_j$  for  $j \leq m(n)$  does not give a big enough sum. Therefore, breaking any admissible vector into the part below  $m(n)$  and the part at  $m(n)$  or above, the event  $G_n$  is contained in the following event:

$$\begin{aligned} Z_{m(n)} \leq 1.01 \log m \text{ and } W_{m(n)} < n \text{ and there is some } k \text{ with } k = \sum_j j(y'_j + y''_j) \\ \text{with } \mathbf{y}' \text{ supported on } [1, m(n) - 1] \text{ and } \mathbf{y}'' \text{ nonzero and supported on } [m(n), n] \\ \text{and both } \mathbf{y}' \text{ and } \mathbf{y}'' \text{ admissible.} \end{aligned}$$

Let  $p'_k$  denote the probability that  $Z_{m(n)} \leq 1.01 \log m$  and  $W_{m(n)} \leq m$  and  $k = \sum_j jy'_j$  for an admissible  $\mathbf{y}'$  supported on  $[1, m(n) - 1]$  and let  $p''_k$  be the probability that  $k = \sum_j jy''_j$  for an admissible  $\mathbf{y}''$  supported on  $[m(n), n]$ . By independence of the coordinates  $\omega_k$  we see we have shown that

$$(3.3) \quad \begin{aligned} \tilde{p}_n &\leq \sum_{k=m(n)+1}^n p'_{n-k} p''_k \\ &\leq \left( \sum_{k=1}^{m(n)} p'_k \right) \cdot \max_{m(n)+1 \leq k \leq n} p''_k . \end{aligned}$$

By Fubini's theorem, the first of these factors is equal to the expected number of  $k \leq m(n)$  in  $S(\omega|_{m(n)})$  where  $\omega|_{m(n)}$  is the sequence  $\omega$  with all entries zeroed out above  $m(n)$ . Letting  $Z_m := \sum_{j=1}^{m(n)} \omega_j$  denote the size (with multiplicity) of  $\omega$  up to  $m(n)$ , it is immediate that the number of  $k \leq m(n)$  in  $S(\omega|_{m(n)})$  is at most  $2^{Z_m}$ . But on the event  $G_n$ , it always holds that

$Z_m \leq 1.01 \log m$  because  $\tau_{1/100} \leq T < m(n)$ . Therefore the first factor on the right-hand side of (3.3) is bounded above by

$$(3.4) \quad \sum_{k=1}^{m(n)} p'_k \leq 2^{1.01 \log m} = m^{1.01 \ln 2}.$$

Next we claim that there is a constant  $C$  such that

$$(3.5) \quad p''_k \leq C \frac{\log^2 n}{n} \text{ for all } k \in [m(n), n].$$

To prove this, start with the observation that  $\mathbb{P}(\omega_j \geq 2) \leq j^{-2}$  leading to

$$\mathbb{P}(H) \leq \sum_{j=m(n)}^n j^{-2} \leq m(n)^{-1} = \frac{\log n}{n}$$

where  $H$  is the event that  $\omega_j \geq 2$  for some  $j \in [m(n), n]$ . The event that  $k = \sum_j j y_j''$  for an admissible  $k$  supported on  $[m(n), n]$  but that  $H$  does not occur is contained in the union of events  $E_j$  that  $\omega_j = 1$  and  $k - j = \sum_i i y_i''$  for some admissible  $\mathbf{y}''$  supported on  $[m(n), n] \setminus \{j\}$ . Using independence of  $\omega_j$  from the other coordinates of  $\omega$ , along with our description of what must happen if  $H$  does not, we obtain

$$(3.6) \quad \begin{aligned} p''_k &\leq \mathbb{P}(H) + \sum_{j=m(n)}^n \frac{1}{j} p''_{k-j} \\ &\leq \mathbb{P}(H) + \frac{1}{m(n)} \sum_{j=m(n)}^n p''_{k-j} \\ &\leq \frac{\log n}{n} \left( 1 + \sum_{j=m(n)}^n p''_{k-j} \right). \end{aligned}$$

We may now employ a relatively easy upper bound on the summation in the last factor, namely we may use the expected number of ways of obtaining each number as a sum of large parts (recall that the expectation when not restricting the parts is too large to be useful). Accordingly, we define the generating function

$$F(z, \omega) := \prod_t (1 + z^t)$$

where the index  $t$  of the product ranges over values in  $[m(n), n]$  such that  $\omega_t = 1$  (recall we ruled out values of 2 or more). Then

$$\begin{aligned} \sum_{j=m(n)}^n p''_{k-j} &\leq \sum_{j=m(n)}^n \mathbb{E}[z_j] F(z, \omega) \\ &\leq \sum_{j=1}^{\infty} \mathbb{E} F(1, \omega) \\ &\leq \prod_{j=m(n)}^n \left(1 + \frac{1}{j}\right) \\ &= \frac{n+1}{m(n)}. \end{aligned}$$

Putting this together with (3.6) yields

$$p''_k \leq \frac{\log n}{n} \left(1 + \frac{n+1}{n} \log n\right) = O\left(\frac{\log^2 n}{n}\right)$$

proving (3.5).

Finally, plugging in (3.4) and (3.5) into (3.3) shows that

$$\tilde{p}_n \leq C m(n)^{1.01 \ln 2} \frac{\log^2 n}{n}$$

which is bounded about by a constant multiple of  $n^{\ln 2 + 0.02 - 1}$ , completing the proof of the lemma.  $\square$

It is now routine to establish something that is almost Theorem 1.6.

**Theorem 3.2.**

$$\mathbb{P}^4 \left( \bigcap_{r=1}^4 S(\omega^r) \text{ is finite} \right) = 1.$$

**PROOF:** Let  $T^1, \dots, T^4$  denote the quantities  $T(\omega)$  when  $\omega = \omega_1, \dots, \omega_4$  respectively. Let  $T^*$  denote the maximum of  $\{T^1, T^2, T^3, T^4\}$ . By Lemma 3.1 and the independence of the  $\omega^r$  for  $1 \leq j \leq 4$ , the probability that  $T^r < m(n)$  and  $n \in S(\omega^r)$  for all  $1 \leq r \leq 4$  is at most a constant multiple of  $n^{-3.94 + 4 \ln 2}$ . The exponent is less than  $-1$ , so the series is summable and we conclude that  $n$  is in the intersection of all four sets  $S(\omega^r)$  for finitely many  $n > T_*$  almost surely. Almost sure finiteness of  $T^*$  finishes the proof.  $\square$

**PROOF OF THEOREM 1.6:** By Theorem 3.2 we may choose  $L$  sufficiently large so that the  $\mathbb{P}(H_L) < 1/4$  where  $E_L$  is the event that  $[L, \infty) \cap \bigcap_{r=1}^4 S(\omega^r)$  is non-empty. The event  $E_L$  is an increasing function of the independent random variables  $\{X_j\}$ . The event  $E'_L$  that  $[1, L-1] \cap \bigcap_{r=1}^4 S(\omega^r)$  is non-empty is also a increasing function of the coordinates  $X_j$  and has some fixed nonzero probability  $b_L$ ; for example,  $b_L$  is at least the probability that  $Z_L = 0$ , which is a simple Poisson event with probability asymptotic to  $e^{-4\gamma} L^{-4}$  when  $L$  is large. Harris's inequality says that any two increasing functions of independent random variables are nonnegatively correlated (see, e.g., [Gri99, Section 2.2]). Their complements are also nonnegatively correlated and this gives

$$\mathbb{P}^4 [E_L^c \cap (E'_L)^c] \geq \mathbb{P}(E_L^c) \mathbb{P}((E'_L)^c) \geq \frac{b_L}{2},$$

finishing the proof of Theorem 1.6.

#### 4. COMPUTATION OF THE MARGINAL PROBABILITIES

In this section we prove Theorem 1.7. This is not needed for Theorem 1.5 but serves to establish the so-called lottery effect, that is, the fact that  $p_k$  has a different exponent of decay from the quenched probabilities  $\tilde{p}_k$ . It is generally easier to prove estimates for the partial sums  $\sum_{k=1}^n p_k$  than for  $p_n$ . We take care of this first, which is the bulk of the work. Let  $A_n := \sum_{k=1}^n p_k = \mathbb{E}|S \cap [n]|$ .

**Lemma 4.1.**

$$A_n = n^1 + \eta + o(1),$$

where

$$\eta = \frac{1 - \log 2 - \log(1/\log 2)}{\log 2} \approx -0.08607 \dots$$

**PROOF:** For the upper bound, recall the notation  $Z_n = \sum_{j=1}^n X_j$  and observe that the cardinality of  $|S \cap [n]|$  is at most  $2^{Z_n} \wedge n$  (here  $\wedge$  is used to denote the binary operation of taking the minimum). This leads to

$$\begin{aligned} A_n &\leq \mathbb{E}(2^{Z_n} \wedge n) \\ &\leq 1 + \sum_{0 \leq k \leq \log_2 n} 2^k \mathbb{P}(Z_n > k) \\ &\leq C \log n \sup_{0 \leq x \leq 1/\log 2} \mathbb{P}(Z_n \geq x \log n) 2^{x \log n}. \end{aligned}$$

Recalling from (2.8) that  $\log \mathbb{P}(Z_n \geq x \log n) \leq \log n \cdot [x - 1 - x \log x]$  we obtain

$$\frac{\log A_n}{\log n} \leq o(1) + \sup_{0 \leq x \leq 1/\log 2} [x \log 2 + x - 1 - x \log x].$$

The supremum is achieved at the right endpoint, therefore

$$\frac{\log A_n}{\log n} \leq o(1) + \frac{1 - \log(1/\log 2)}{\log 2} = \eta + 1 + o(1).$$

For the reverse inequality, we begin by recalling the tilted laws  $\mathbb{P}_x$  from the proof of the lower bound in Lemma 2.3, fixing the value  $x = 1/\log 2$  for the remainder of this proof. The idea is that when  $Z_n \approx (1/\log 2)n$ , then  $S_n \cap [n]$  should have size roughly  $2^{Z_n} \approx n$ . Fix  $\varepsilon > 0$  and define

$$G_n^\varepsilon := G_n \cap \{|S_n \cap [n]| \geq n^{1-\varepsilon}\}.$$

The infimum of the Radon-Nikodym derivative  $d\mathbb{P}/d\mathbb{P}_x$  on  $G_n$ , computed in (2.10), is  $n^\eta + o(1)$ , so the proof is complete once we establish

$$(4.1) \quad \mathbb{P}_x(G_n^\varepsilon) = 1 - o(1) \text{ for each fixed } \varepsilon > 0.$$

To show (4.1), we begin with some definitions. Let  $\tau_j := \inf\{n : Z_n = j\}$  be the  $j^{\text{th}}$  smallest value in the multiset  $M$ . Let  $\mathcal{F}_j := \sigma(X_i \wedge (j - Z_{i-1}))$  be the  $\sigma$ -field containing the

values of the  $j$  elements of the multiset  $M$ . This is a natural filtration on which the random variables  $x_j := X_{\tau_j}$  form an adapted sequence. Given  $x_j$ , we may easily compute

$$\begin{aligned} \mathbb{P}_x \left( \log \frac{x_{j+1}}{x_j} > u \right) &= \left( 1 + O\left(\frac{1}{x_j}\right) \right) \prod_{x_j < k < x_j e^u} e^{-1/(k \log 2)} \\ &= \left( 1 + O\left(\frac{1}{x_j}\right) \right) e^{-u/\log 2}; \end{aligned}$$

It is not hard to see from this that the conditional distribution of  $\log(x_{j+1}/x_j)$  given  $x_j$  is stochastically bounded between exponentials of means  $\log 2 + O(1/x_j)$ , where the fudge term accounts for the possibility that  $x + 1 = x_j$  and for the discretization. Define

$$\begin{aligned} s_j &:= \text{sumset}(x_1, \dots, x_j); \\ Y_j &:= \log |s_j| - \log x_j; \\ \Delta_j &:= Y_{j+1} - Y_j = U_j - V_j, \end{aligned}$$

where

$$\begin{aligned} U_j &:= \log |s_{j+1}| - \log |s_j|; \\ V_j &:= \log x_{j+1} - \log x_j. \end{aligned}$$

**Lemma 4.2.**

$$(4.2) \quad \mathbb{P}_x(Y_j \leq -j/4) \rightarrow 0 \text{ as } j \rightarrow \infty.$$

Assuming the lemma and plugging  $j = \log_2 n - \log \log n$  into (4.2), it follows that

$$\mathbb{P}_x(Y_{\log_2 n - \log \log n} \leq -\varepsilon \log n) = o(1)$$

as  $n \rightarrow \infty$  for any  $\varepsilon > 0$ . Another event whose probabilities goes to zero is the event that  $W_j \geq n$  (recall that  $W_{\tau_j}$  is the sum of the elements of  $M$  up to  $X_{\tau_j} = x_j$ ). On the complement of this event,  $s_j \subseteq S \cap [n]$ . Finally, the event  $\log_2 x_j < j - j^{2/3}$  also goes to zero. On the complement of the union of these three small events,  $|S \cap [n]| \geq |s_j| = x_j e^{Y_j} \geq 2^{j-j^{2/3}} n^{-\varepsilon} \geq n^{1-(\log n)^{-1/3-\varepsilon}}$ . Because  $\varepsilon > 0$  is arbitrary, this proves (4.1) and finishes the proof of Lemma 4.1 modulo Lemma 4.2.  $\square$

The proof of Lemma 4.2 requires the following standard deviation estimate for supermartingales with bounded exponential moment.

**Lemma 4.3.** *Let  $\{S_i\}_{i \geq 0}$  be a supermartingale with respect to the filtration  $\{\mathcal{F}_i\}$ , with  $S_0 = 0$ . Suppose that the increments  $\xi_{i+1} := S_{i+1} - S_i$  satisfy  $\mathbb{E}(e^{\xi_{i+1}} | \mathcal{F}_i) \leq B$  for all  $i \geq 0$ . Then for all integer  $\ell > 0$  and real  $R \in [0, 2\ell B]$ , we have*

$$(4.3) \quad \mathbb{P}(S_\ell > R) \leq e^{-R^2/(4\ell B)}.$$

PROOF: By Lemma 3.1 from [Fre75], the positive function  $g(t) = (e^t - 1 - t)/t^2$  (where  $g(0) = 1/2$ ) is increasing in  $\mathbb{R}$ . Thus for all  $\lambda \in [0, 1]$  and  $\xi \in \mathbb{R}$ , we have

$$(4.4) \quad (\lambda \xi)^2 g(\lambda \xi) \leq (\lambda \xi)^2 \max\{g(0), g(\xi)\} \leq \lambda^2 e^{|\xi|}.$$

Because  $\{S_i\}$  is a supermartingale,  $\mathbb{E}_\ell(\xi) \leq 0$  where (just for this proof) we abbreviate  $\mathbb{E}_\ell(\cdot) = \mathbb{E}(\cdot | \mathcal{F}_\ell)$ . Taking expectations in (4.4), we infer that

$$\mathbb{E}_\ell(e^{\lambda \xi}) \leq 1 + \lambda \mathbb{E}_\ell(\xi_{\ell+1}) + \lambda^2 \mathbb{E}_\ell(e^{|\xi_{\ell+1}|}).$$

Thus  $\mathbb{E}_\ell(e^{\lambda \xi_{\ell+1}}) \leq 1 + B\lambda^2 < e^{B\lambda^2}$ , whence  $\mathbb{E}_\ell(e^{\lambda S_{\ell+1}}) \leq e^{\lambda S_\ell + B\lambda^2}$ . A simple induction then leads to  $\mathbb{E}(e^{\lambda S_\ell}) \leq e^{\ell B\lambda^2}$ . We conclude that  $\mathbb{P}(S_\ell \geq R) \leq e^{\ell B\lambda^2 - \lambda R}$ . To minimize the right-hand side, we take  $\lambda = R/(2\ell B)$ , which yields the assertion of the lemma.  $\square$

PROOF OF LEMMA 4.2: The quantity  $\Delta_j$  is the difference of positive variables  $U_j$  and  $V_j$ . Conditional on  $\mathcal{F}_j$ , the variable  $U_j$  is stochastically greater than  $-E_j$  where  $E_j$  is an exponential of mean  $\log 2 + 0.01$ . The variable  $V_j$  is of necessity in the interval  $[0, \log 2]$ . We begin by showing that

$$(4.5) \quad \mathbb{E}V_j \geq \log 2 (1 - e^{Y_j \wedge 0}) .$$

Let  $R$  denote the size of the overlap  $R := |s_j \cap (x_{j+1} \oplus s_j)|$  where the  $\oplus$  symbol in this case denotes translation of the set  $s_j$  by  $x_{j+1}$ . We may then express

$$\log \frac{s_{j+1}}{s_j} = \log \frac{2|s_j| - R}{|s_j|} = \log 2 + \log \left( 1 - \frac{R}{2|s_j|} \right) .$$

Using the fact that  $R/(2|s_j|) \in [0, 1/2]$  and the bound  $\log(1 - u) \geq -u \log 4$  for  $u \in [0, 1/2]$  then gives

$$\mathbb{E}_x(V_j | \mathcal{F}_j) \geq \log 2 - \log 4 \frac{\mathbb{E}_x(R | \mathcal{F}_j)}{2|s_j|} + O\left(\frac{1}{x_j}\right) .$$

But

$$\mathbb{E}_x(R | \mathcal{F}_j) = \sum_{a, b \in s_j} \mathbb{P}_x(x_{j+1} = b - a | \mathcal{F}_j) \leq |s_j|^2 / x_j$$

because  $\mathbb{P}_x(x_{j+1} = k | \mathcal{F}_j) \leq 1/x_j$  for any  $k$ . Also trivially  $R \leq |s_j|$ , whence

$$\frac{\mathbb{E}_x(R | \mathcal{F}_j)}{2|s_j|} \leq \frac{s_j}{2x_j} \wedge \frac{1}{2} .$$

Replacing  $|s_j|/x_j$  by  $\exp(Y_j)$  then proves (4.5).

The event  $G := \{Y_j \leq -j/4\}$  can be covered by the union over  $0 \leq i \leq j$  of the events  $G_i$  defined as follows. Let  $G_0$  be the event that for some  $i \leq j$  we have  $U_i \leq -\varepsilon j/4$ . For  $1 \leq i \leq j$  define  $G_i$  to be the event that  $Y_i \in [-\varepsilon j/2, -\varepsilon j/4]$ ,  $Y_j \leq -\varepsilon j$ , and  $Y_t \leq -\varepsilon/4$  for every  $t \in [i, j]$ . To see that  $G \subseteq \bigcup_{i=0}^j G_i$ , observe that if no jump is less than  $-\varepsilon/4$  then the last time  $i \leq j$  that  $Y_i \geq -\varepsilon j/2$ , we must have  $Y_i \leq -\varepsilon j/4$ .

From the fact that  $U_j \geq -E_j$  and  $V_j \geq 0$  we see easily that

$$\mathbb{P}_x(G_0) \leq j \exp\left(-\frac{\varepsilon j}{4(\log 2 + 0.01)}\right) .$$

A sufficient condition to imply the lemma is that there is some  $c > 0$  such that  $\mathbb{P}_x(G_i) < e^{-cj}$  for all  $i, j$  with  $1 \leq i < j$ . This follows from an application of Lemma 4.3

Fix  $i$  and  $j$  and for  $i \leq k \leq j$  let  $M_k = -(Y_{k \wedge \tau} - Y_i) - k \wedge \tau \varepsilon/4$ , where  $\tau$  is the least  $t$  for which  $Y_y \geq -\varepsilon j/4$ . On the event  $G_i$ , the value of  $M_j - M_i$  is at least  $R := \varepsilon j/4$ . The expected increment  $\Delta M_k := \mathbb{E}(M_{k+1} - M_k | \mathcal{F}_k)$  is zero when  $k \geq \tau$  and otherwise is at most  $\log 2 + 0.01 - \mathbb{E}V_k - \varepsilon/4$ . By (4.5), this is at most  $0.01 + (\log 2)e^{-\varepsilon j/4} - \varepsilon/4$  which is less than zero, hence  $\{M_k\}$  is supermartingale. For any  $\lambda < (\log 2 + 0.01)^{-1}$ , and in particular for  $\lambda = 1$ , the quantity  $\mathbb{E}e^{\lambda \Delta M_k}$  is bounded above by some constant,  $B$ , independent of  $i$  and  $j$ . Applying Lemma 4.3 to the supermartingale  $\{S_t := M_{i+t} - M_i\}$  with  $R = \varepsilon j/4$  and

$\ell = j - i \leq j$ , we see that

$$\mathbb{P}_x \left( (M_j - M_i \geq \frac{\varepsilon j}{4}) \right) \leq \exp \left( -\frac{(\varepsilon j/4)^2}{4jB} \right) = \exp \left( \frac{\varepsilon^2}{64B} j \right).$$

This completes the proof of Lemma 4.2.  $\square$

*Proof of Theorem 1.7:* Typically one requires some kind of regularity to get from an estimate on the partial sums to an estimate on the individual summands. Here instead we copy the proof of Lemma 3.1, using the large-index summands, rather than some kind of monotonicity, to do the smoothing.

Recall that  $m(n) := \lfloor n/\log n \rfloor$  and bound  $p_n$  from below by one minus the probability that all attempts to make  $n$  using a part of size between  $n - m$  and  $n$  fail.

$$\begin{aligned} p_n &\geq 1 - \mathbb{E} \prod_{k=n-m}^n \left( 1 - \frac{1}{k} \mathbf{1}_{n-k \in S} \right) \\ &\geq 1 - \mathbb{E} \exp \left( - \sum_{k=n-m}^n \frac{\mathbf{1}_{n-k \in S}}{k} \right) \\ &\geq 1 - \mathbb{E} \exp \left( -\frac{1}{n} |S \cap [n]| \right). \end{aligned}$$

By convexity of the exponential, the maximum value of  $\mathbb{E} e^{-Y/n}$  over all variables with mean  $A_m$  taking values in  $[0, m]$  is achieved when  $Y$  is equal to  $m$  times a Bernoulli with mean  $(A_m/m)$ . This yields

$$(4.6) \quad p_n \geq \frac{A_m}{m} (e^{m/n} - 1) \geq \frac{A_m}{n}.$$

On the other hand, by Lemma 2.2,

$$(4.7) \quad p_n \leq n^{1-\log \log n} + \sum_{k=m}^n \frac{1}{k} p_k \leq \frac{\log n}{n} A_n.$$

Together, (4.6) and (4.7) show that  $A_n = n^{1+\eta+o(1)}$  implies  $p_n = n^{\eta+o(1)}$ ; this proves Theorem 1.7 modulo Lemma 4.1.  $\square$

## 5. RANDOM PERMUTATIONS

In this section we prove Theorem 1.5. The starting point is a coupling between the Poisson variables in the Poisson model and permutations in the group theoretic model. The underlying space for the coupling will be the space  $(\Omega, \mathcal{F}, \mathbf{Q})$  and its fourfold product where

$$\Omega := \left( \prod_{N=1}^{\infty} \mathcal{S}_N \right) \times \prod_{j=1}^{\infty} \mathbb{Z}^+$$

and  $\mathcal{F}$  is the product of the complete  $\sigma$ -fields (the power set of  $\mathcal{S}_N$  or  $\mathbb{Z}^+$ ) in each coordinate. For  $\omega = (s_1, s_2, \dots, x_1, x_2, \dots) \in \Omega$ , define the coordinate functions  $X_j(\omega) := x'_j \in \mathbb{Z}^+$  and  $\sigma_N(\omega) := s_N \in \mathcal{S}_N$ .

Let  $\Delta_N := \|Q_{N,m(N)} - \nu_{m(n)}\|_{TV}$  be as in the statement of Lemma 1.8. To re-iterate,  $\Delta_N$  denotes the total variation distance between the joint distribution of number of cycles of sizes  $1, \dots, m(N)$  in a uniform random permutation from  $\mathcal{S}_N$  and the product Poisson

measure on  $(\mathbb{Z}^+)^{m(N)}$  whose  $j^{\text{th}}$  coordinate has mean  $1/j$ . The lemma of Arratia and Tavaré states that

$$(5.1) \quad \Delta_N \leq \exp(-C \log N \log \log N) = N^{-c \log \log N}.$$

**Lemma 5.1.** *There is a probability measure  $\mathbf{Q}$  on  $\Omega$  such that the laws of the random variables  $\sigma_N$  and  $X_j$  have the following properties for all  $N$  and  $j$ :*

- (i)  $\sigma_N \sim \mathbb{P}_N$  (the uniform measure on  $\mathcal{S}_N$ );
- (ii)  $X_j \sim \mathcal{P}(1/j)$  (a Poisson with mean  $1/j$ );
- (iii) with probability  $1 - \Delta_N$ , for all  $N$  and all  $j \leq m(N)$ , the permutation  $\sigma_N$  has exactly  $X_j$  cycles of length  $j$ .

PROOF: There is a coupling  $\mathbf{Q}_N$  of  $\mathbb{P}_N$  and  $\nu_{m(N)}$  giving measure  $1 - \Delta_N$  to the set of  $(\sigma_N, \{X_n : n \geq 1\})$  such that there are  $X_j$  cycles of  $\sigma_N$  of length  $j$  for all  $j$ . The grand coupling  $\mathbf{Q}$  may be constructed by first making  $(x_1, x_2, \dots)$  independent Poissons with means  $1/j$  and then giving  $\sigma_N$  the conditional distribution of  $\mathbf{Q}_N$  given  $(x_1, x_2, \dots)$ .  $\square$

PROOF OF THEOREM 1.5: Fix  $L$  and  $b_L$  as in the end of the proof of Theorem 1.6. Choose  $N_0$  such that  $N_0/(\log N_0)^2 > L$ . Let  $(\Omega, \mathcal{F}, \mathbf{Q})^4$  be the fourfold product of the measure constructed in Lemma 5.1. The notation is a bit unwieldy but we will denote the generic element  $\omega \in \Omega^4$  by  $\langle s_j^r, x_j^r : j \in \mathbb{Z}^+, 1 \leq r \leq 4 \rangle$ . Let  $X_j^r, 1 \leq r \leq 4$  denote the  $(j, r)$  coordinate  $x_j^r$  and  $\sigma_j^r$  the  $(j, r)$  permutation coordinate  $s_j^r$  of  $\omega$ . We let  $\mathbf{X}^r$  denote the sequence  $(X_j^r : j \geq 1)$ .

Let  $G \in \mathcal{F}^4$  be the event that  $\bigcap_{r=1}^4 S(\mathbf{X}^r)$  is empty. By Theorem 1.6 and the identification of the constant  $b_L$ , we know that  $\mathbb{P}^4(G) \geq b_L/2$ . Choose  $N_1 \geq N_0$  so that  $\Delta_{N_1} \leq b_L/40$  and also  $\mathbb{P}_N(T \geq N/(\log N)^2) \leq b_L/40$  for  $N \geq N_1$ . Let  $H_N^r$  denote the uncoupling event, namely the event that for some  $j \leq m(N)$ , the permutation  $\sigma_N^r$  has a number of  $j$ -cycles different from  $X_j^r$ . This has probability  $\Delta_N$ , hence for  $N \geq N_1$ , at most  $b_L/40$ . Therefore, the event  $G_N := G \setminus (H_N^1 \cup H_N^2 \cup H_N^3 \cup H_N^4)$  has probability at least  $(2/5)b_L$ . On  $G_N$ , the common intersection of  $S(\sigma_N^r)$  for  $a \leq r \leq 4$  cannot contain any elements less than  $m(N)$  because the cycle counts of  $\sigma_N^r$  agree with  $\{X_j^r\}$  for all cycles of length at most  $m(N)$  and on  $G$ , the sumsets of these counts have no common intersection.

**Lemma 5.2.** *There is an  $N_2$  such that for all  $N \geq N_2$ ,*

$$(5.2) \quad \mathbb{P}^4(E_N) \leq \frac{3b_L}{10}$$

where the event  $E_N$  is defined by

$$E_N := G_N \cap \left\{ \bigcap_{r=1}^4 S(\sigma_N^r) \cap [m(N), N] \neq \emptyset \right\}.$$

Theorem 1.5 follows from this: for any  $N \geq N_2$ ,  $\mathbb{P}(G_N \setminus E_N) \geq (3/10)b_L - (1/5)b_L$ . On this event, the four sets  $S(\sigma_N^r)$  have no common intersection. Letting  $b$  be the minimum of  $b_L/10$  and the least probability of no common intersection over all  $N < N_2$  then proves the theorem. It remains to prove the lemma.

PROOF OF LEMMA 5.2: The outline is very similar to the outline of the proof of Theorem 1.6. Fix  $N \geq N_1$ . The analogue to Lemma 3.1 is to define, for  $m(N) \leq n \leq N$ , a quantity  $\tilde{q}_n$  analogous to  $\tilde{p}_n$ . This is the probability that  $n \in S(\sigma_n^r)$  while also  $T^r < m(n)$ ; this probability



clearly does not depend on  $r$ . We will show that  $\tilde{q}_n^A$  is summable. To see that this is enough, assume it is true and pick  $N_2$  to make the tail sum sufficiently small:

$$\sum_{n=m(N_2)}^{\infty} \tilde{q}_n \leq \frac{b_L}{10}.$$

If  $E_n$  occurs then either some  $T^r \geq m(n)$  or  $E_n$  occurs without this. The first of these two probabilities is limited to  $b_L/10$  by choice of  $N_1$ :  $m(n)$  can be no less than  $N_1/(\log N_1)^2$ , guaranteeing that  $T^r \geq m(n)$  with probability at most  $b_L/40$  and hence  $T > m(n)$  with probability at most  $b_L/10$ . The second of the two probabilities is limited to  $b_L/10$  as long as  $N \geq N_2$  because the sum of  $\tilde{q}_n^A$  as  $n$  ranges over  $[m(N), N]$  will be at most the tail sum of  $\tilde{q}_n$  from  $m(N_2)$ . This makes (5.2) the sum of two quantities each at most  $b_L/10$  and finishes the proof of the lemma with one tenth to spare.

Fourth power summability of  $\tilde{q}_n$  is proved via an estimate very similar to the estimate in Lemma 3.1. Because  $E_N \subseteq G_N$ , the coupling is unbroken and it is therefore not possible to have  $n \in S(\sigma_n^r)$  equal to  $\sum j y_j$  with  $\mathbf{y}$  supported on  $[1, m(n)]$ . Hence, as before,  $\mathbf{y}$  decomposes into  $\mathbf{y}' + \mathbf{y}''$  with  $\mathbf{y}'$  supported on  $[1, m(n)]$  and  $\mathbf{y}''$  supported on  $[m(n) + 1, n]$  and not identically zero. Also as before we have the upper bound

$$(5.3) \quad \tilde{q}_n \leq \left( \sum_{k=1}^{m(n)} q'_k \right) \cdot \max_{m(n)+1 \leq k \leq n} q''_k$$

where  $q'_k$  is the probability  $\tilde{q}_k$  but using only cycles of size at most  $m(n)$  and  $q''_k$  is the analogue of  $\tilde{q}_k$  when only cycles of size at least  $m(n) + 1$  are used.

Analogously to (3.4), the first factor is at most  $m(n)^{1.01 \log^2}$  because the coupling is unbroken and we already proved this bound for the Poisson variables. It suffices therefore to prove the bound

$$(5.4) \quad q''_n \leq C \frac{\log^3 n}{n} \text{ for } n \geq m(N)$$

analogous to (3.5). Here the proof diverges from the proof of Lemma 3.1 because the constraint on the vector  $\mathbf{y}''$  in  $\sum j y_j''$  is that  $y_j''$  be at most the number  $Y_j$  of  $j$ -cycles in the actual permutation  $\sigma_N^r$ , rather than being at most the Poisson variable  $X_j^r$ . The variables  $Y_j$  as were the variables  $X_j^r$  are not independent so instead we argue as follows.

Recall that  $N_2/(\log N_2)^2 \leq N/(\log N)^2 \leq m(n) \leq N$  and observe that the quantity  $q''_n$  is at most the sum over  $j \geq m(n)$  of

$$\mathbb{P}_N \left[ Y_j \geq 1 \text{ and } n - j = \sum_i i y_i'' \text{ for some } \mathbf{y}'' \leq \mathbf{Y} - \delta_j \text{ supported on } [m(n) + 1, N] \right].$$

Here we have denoted by  $\mathbf{Y}$  the vector whose components are  $Y_j$ . The actual elements in the cycles of the permutation  $\sigma_N^r$  are exchangeable given the cycle lengths, so the probability that the element 1 is in the cycle of length  $j$  is at least  $N/m(n)$ , which is a least  $(\log n)^{-2}$ . Therefore, we may write a new upper bound

$$q''_n \leq (\log n)^2 \sum_{j \geq m(n)} \mathbb{P}_N \left[ 1 \text{ is in a cycle of length } j \text{ and } n - j \in S^*(\sigma_N^r) \right]$$

where the superscript  $S^*$  denotes that we count only sums of cycle sizes at least  $m(n)$ .

The reason for going through this trouble is that conditioned on 1 being in a cycle of size  $j$ , the remainder of the permutation is uncontaminated: its cycle sizes are distributed as those of a uniform pick from  $\mathcal{S}_{n-j}$ . Also, the probability of 1 being in a cycle of size  $j$  is precisely  $1/N$ . Therefore,

$$(5.5) \quad q_n'' \leq \frac{(\log n)^2}{N} \sum_{j \geq m(n)} \mathbb{P}_N \left[ n-j \in S^* P(\sigma_{N-j}^r) \right].$$

To evaluate the summand, first observe that replacing  $S^*$  with  $S$ , the expected number of invariant sets of  $\sigma_{N-j}^r$  of size  $n-j$  is precisely 1 for any  $N, n, j$ . Next, consider any invariant set of  $\sigma_{N-j}^r$  of size  $n-j$  and bound from above the probability that it consists entirely of cycles larger than  $m(n)$ . This is the same as the probability that a random element of  $\mathcal{S}_{n-j}$  has only cycles of length at least  $m(n)$ . We may evaluate this via the Arratia-Tavaré lemma: it is at most equal to  $\mathbb{P}(Z_{m(n)} = 0)$  (the probability that a Poisson ensemble with  $\mathbb{E}X_j = 1/j$  takes only value zero up to  $j = m(n)$ ) plus the total variation distance between the Poisson product measure and the actual counts of cycle sizes up to  $m(n)$ . By the bound in Lemma 1.8, this total variation distance is at most  $\exp[Cn/m(n) \log(n/m(n))]$  which is bounded above by  $n^{-C \log \log n}$  and hence decays faster than any polynomial in  $n$ . The probability of  $Z_{m(n)} = 0$  is  $e^{-H_{m(n)}} \leq 1/m(n)$ , whence the summand in (5.5) is therefore at most  $1/m(n) \sim \log n/n$  and the whole sum is at most  $\log n(N/n)$ . This makes the right-hand side of (5.5) at most  $\log^3 n/n$ , establishing (5.4) and completing the proof of Lemma 5.2 and hence of Theorem 1.5.  $\square$

## REFERENCES

- [AT92] R. Arratia and S. Tavaré. The cycle structure of random permutations. *Ann. Probab.*, 3:1567–1591, 1992.
- [Dix92] John D. Dixon. Random sets which invariably generate the symmetric group. *Discrete Math.*, 105(1-3):25–39, 1992.
- [DS00] J. H. Davenport and G. C. Smith. Fast recognition of alternating and symmetric Galois groups. *J. Pure Appl. Algebra*, 153(1):17–25, 2000.
- [Fre75] D. Freedman. On tail probabilities for martingales. *Ann. Probab.*, 3:100–118, 1975.
- [Gra06] Andrew Granville. Cycle lengths in a permutation are typically Poisson. *Electron. J. Combin.*, 13(1):Research Paper 107, 23, 2006.
- [Gri99] G. Grimmet. *Percolation*, volume 321 of *Grundlehren der mathematischen Wissenschaften*. Springer, New York, second edition, 1999.
- [HT88] R. Hall and G. Tenenbaum. *Divisors*, volume 90 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1988.
- [LO77] J. C. Lagarias and A. M. Odlyzko. Effective versions of the Chebotarev density theorem. In *Algebraic number fields: L-functions and Galois properties (Proc. Sympos., Univ. Durham, Durham, 1975)*, pages 409–464. Academic Press, London, 1977.
- [LP93] Tomasz Luczak and Łászló Pyber. On random generation of the symmetric group. *Combin. Probab. Comput.*, 2(4):505–512, 1993.
- [Mus78] David R. Musser. On the efficiency of a polynomial irreducibility test. *J. Assoc. Comput. Mach.*, 25(2):271–282, 1978.
- [Oes79] Joseph Oesterlé. Versions effectives du théorème de chebotarev sous lhypothese de riemann généralisée. *Astérisque*, 61:165–167, 1979.
- [Riv13] Igor Rivin. Large Galois groups with applications to Zariski density. *arXiv preprint arXiv:1312.3009*, 2013.
- [Ser81] Jean-Pierre Serre. Quelques applications du théorème de densité de Chebotarev. *Inst. Hautes Études Sci. Publ. Math.*, (54):323–401, 1981.

- [vdW34] Bartel Leendert van der Waerden. Die seltenheit der gleichungen mit affekt. *Mathematische Annalen*, 109(1):13–16, 1934.
- [Win13] Bruno Winckler. Théorie de chebotarev effectif. *arXiv preprint arXiv:1311.5715*, 2013.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF PENNSYLVANIA, 209 SOUTH 33RD STREET, PHILADELPHIA, PA 19104, USA

*E-mail address:* pemantle@math.upenn.edu

MICROSOFT RESEARCH, 1 MICROSOFT WAY, REDMOND, WA, 98052, USA

*E-mail address:* peres@microsoft.com

TEMPLE UNIVERSITY, 1805 N BROAD ST, PHILADELPHIA, PA

*Current address:* Mathematics Department, Brown University

*E-mail address:* igor.rivin@temple.edu