

Covering n -Permutations with $(n + 1)$ -Permutations

Taylor F. Allison*

Department of Mathematics
University of North Carolina
Chapel Hill, NC, U.S.A.

tfallis2@ncsu.edu

Anant P. Godbole[†]

Department of Mathematics & Statistics
East Tennessee State University
Johnson City, TN, U.S.A.

godbolea@etsu.edu

Kathryn M. Hawley[‡]

Department of Mathematics
University of Oregon
Eugene, OR, U.S.A.

kmhawley@gmail.com

Bill Kay[§]

Department of Mathematics and Computer Science
Emory University
Atlanta, GA, U.S.A.

bill.w.kay@gmail.com

Submitted: Mar 25, 2012; Accepted: Jan 3, 2013; Published: Jan 14, 2013

Mathematics Subject Classifications: 05D99, 05D40, 05A05

Abstract

Let S_n be the set of all permutations on $[n] := \{1, 2, \dots, n\}$. We denote by κ_n the smallest cardinality of a subset $\mathcal{A}$ of S_{n+1} that “covers” S_n , in the sense that each $\pi \in S_n$ may be found as an order-isomorphic subsequence of some π' in $\mathcal{A}$. What are general upper bounds on κ_n ? If we randomly select ν_n elements of S_{n+1} , when does the probability that they cover S_n transition from 0 to 1? Can we provide a fine-magnification analysis that provides the “probability of coverage” when ν_n is around the level given by the phase transition? In this paper we answer these questions and raise others.

1 Introduction

The problem discussed in this paper was posed by Prof. Robert Brignall during the Open Problem Session at the *International Permutation Patterns Conference* held at California State Polytechnic University in June 2011. The conference webpage may be found at

*Supported by NSF grant 1004624.

[†]Supported by NSF grant 1004624.

[‡]Supported by NSF grant 1004624.

[§]Supported by NSF grant 1004624.

Brignall asked, “If S_n denotes the set of all permutations on $[n] := \{1, 2, \dots, n\}$, what can we say about κ_n , the smallest cardinality of a subset $\mathcal{A}$ of S_{n+1} that “covers” S_n , in the sense that each $\pi \in S_n$ may be found as an order-isomorphic subsequence of some π' in $\mathcal{A}$.” Specifically, he asked about bounds, exact values, and asymptotics. Small values are easy to calculate; e.g., it is easy to see that $\kappa_1 = 1, \kappa_2 = 1$, and the permutation set $\{1342, 4213\}$ reveals that $\kappa_3 = 2$ – but the situation rapidly gets out of precise control.

In Section 2, we provide a trivial pigeonhole lower bound on κ_n and then use the “method of alterations” [1] to derive an upper general bound on κ_n that contains a logarithmic factor that often shows up when using such arguments in covering-type situations; see, e.g. the general upper bound on the size of covering designs that was proved by Erdős and Spencer in their early work [3]. We continue by showing that the *second and subsequent* coverings of the n -permutations are accomplished in linear $\log \log$ time, in a result that is reminiscent of the ones in [8] (covering designs) and [7] (t -covering arrays). Lastly, in Section 2, we make comparisons to the development in Spencer [12] to produce evidence that the asymptotic value of κ_n is (upto a $O(1)$ or perhaps $1 + o(1)$ factor) the same as that given by the lower bound; we conjecture that this is true.

In Section 3, we estimate the number of permutations π' that can be jointly covered, along with a fixed π , by some permutation $\rho \in S_{n+1}$, and then give a careful estimate of how many $\rho \in S_{n+1}$ can jointly cover two fixed $\pi, \pi' \in S_n$. Then, in Section 4, we use these results to answer the question “If we randomly select ν_n elements of S_{n+1} , when does the probability that they cover S_n transition from asymptotically 0 to asymptotically 1?” Our main result in this area, Theorem 7, is proved using the Janson exponential inequality ([1]). The threshold in Theorem 7 contains a small gap, but the result is the best possible. We prove this fact in Section 5, using the Stein-Chen method of Poisson approximation. Specifically it is proved that when ν_n is “in the gap,” the number of uncovered n -permutations X has a Poisson distribution with finite mean. In particular, this implies $\mathbb{P}(X = 0)$ is a finite constant that is bounded away from zero and one.

2 Bounds

Our first preliminary result provides a formula for the number $c(n, \pi)$ of permutations in S_{n+1} that cover a fixed $\pi \in S_n$.

Lemma 1. *Let $c(n, \pi)$ denote the number of permutations in S_{n+1} that cover a fixed $\pi \in S_n$. Then $c(n, \pi) = c(n, \pi') = n^2 + 1$ for each $\pi, \pi' \in S_n$.*

Proof. It is clear that any permutation pattern $\pi \in S_n$ may be represented in $\binom{n+1}{n} = n + 1$ ways, one for any choice of n numbers from $\{1, 2, \dots, n + 1\}$. Arrange these ways lexicographically (for example if $n = 3$, we can represent the pattern 132 as 132, 142, 143, 243). Note that the r th and $(r + 1)$ st lex-orderings of π differ in a single bit. Now, given any representation of π , the $(n + 1)$ st position may clearly be inserted in $(n + 1)$ ways to create an $(n + 1)$ -covering permutation; however, for any $1 \leq r \leq n$, the list of

covering $(n + 1)$ -permutations for the r th and $(r + 1)$ st lex-orderings have an overlap of magnitude 2, corresponding to whether the $(n + 1)$ st position is inserted before or after the non-matching bit. Thus $c(n, \pi) = c(n, \pi') = (n + 1)^2 - 2n = n^2 + 1$, as asserted. $\square$

Lemma 1 can now be used to prove

Theorem 2.

$$\frac{(n + 1)!}{n^2} \left(1 - \frac{2n + 1}{(n + 1)^2} \right) \leq \kappa_n \leq \frac{(n + 1)!}{n^2 + 1} \left(1 + \log \left(\frac{n^2 + 1}{n + 1} \right) \right).$$

Proof. The lower bound is elementary. Each of the κ_n permutations in a covering, “takes care,” with repetition, of $\binom{n+1}{n} = n + 1$ n -permutations. Since we have a covering, clearly, $\kappa_n(n + 1) \geq n!$, or

$$\kappa_n \geq \frac{n!}{(n + 1)} = \frac{(n + 1)!}{n^2} \left(1 - \frac{2n + 1}{(n + 1)^2} \right) \asymp \frac{(n + 1)!}{n^2}.$$

For the upper bound, we use the method of alterations [1] as follows: Choose a random number Y of $(n + 1)$ -permutations by “without replacement” sampling. The expected number of uncovered n -permutations, by Lemma 1 and linearity of expectation, is

$$\mathbb{E}(X) = n! \frac{\binom{(n+1)! - n^2 - 1}{Y}}{\binom{(n+1)!}{Y}}.$$

We choose a realization with $X = X_Y \leq \mathbb{E}(X)$ and cover these with at most

$$n! \left(1 - \frac{n^2 + 1}{(n + 1)!} \right)^Y \leq n! \exp\{-Y(n^2 + 1)/(n + 1)!\}$$

additional $(n + 1)$ -permutations, yielding, for any initial size Y , a covering with at most

$$Y + n! \exp\{-Y(n^2 + 1)/(n + 1)!\}$$

members. Minimizing over Y yields an initial choice of size

$$\frac{(n + 1)!}{(n^2 + 1)} \log \left(\frac{n^2 + 1}{n + 1} \right),$$

and an upper bound of

$$\kappa_n \leq \frac{(n + 1)!}{(n^2 + 1)} \left(1 + \log \left(\frac{n^2 + 1}{n + 1} \right) \right) \asymp \log n \frac{(n + 1)!}{n^2},$$

as claimed. $\square$

We next ask how many more $(n + 1)$ -permutations are required to cover the n permutations multiple times. Here the situation is often nuanced, and leads to the question as to whether the logarithmic factor, present in the upper bound for the first covering, is extraneous. Typically, in other covering contexts, we find that the second and subsequent coverings need an appropriately normalized $\log \log n$ additional elements in the cover; see, e.g. [8] for a *covering design* analogy, [7] for an occurrence of this phenomenon in *t-covering arrays*, and [4], [9] for the $\log \log$ behavior in the *coupon collection problem*. We briefly describe the parallel in the context of covering designs: A collection $\mathcal{A}$ of sets of size k of $[n]$ is said to form a t -covering design if each t -set is contained in at least one k -set in $\mathcal{A}$. If $m(n, k, t)$ denotes the smallest size of a t -covering design $\mathcal{A}$ then it is clear that $m(n, k, t) \geq \binom{n}{t} / \binom{k}{t}$; Erdős and Spencer proved in [3] that for each n, k, t ,

$$m(n, k, t) \leq \frac{\binom{n}{t}}{\binom{k}{t}} \left(1 + \log \binom{k}{t} \right);$$

it was shown furthermore in [8] that the minimum number $m(n, k, t, \lambda)$ of k -sets needed to cover each t -set λ times satisfied

$$m(n, k, t, \lambda) \leq \frac{\binom{n}{t}}{\binom{k}{t}} \left(1 + \log \binom{k}{t} + (\lambda - 1) \log \log \binom{k}{t} + O(1) \right),$$

$n, k, t \rightarrow \infty$. This was the $\log \log$ result. Also, the Erdős-Hanani conjecture, namely that for fixed k, t ,

$$\lim_{n \rightarrow \infty} \frac{m(n, k, t)}{\binom{n}{t}} = \frac{1}{\binom{k}{t}}$$

was proved by Rödl [11] and, later, by Spencer [12]. This showed that the logarithmic factor in the Erdős-Spencer bound could be asymptotically dispensed with. Finally, see [5] for a corresponding threshold result. It is these questions we seek to address, in our context, in the rest of this section and the next.

Theorem 3. *Let $\kappa_{n,\lambda}$ denote the minimum number of $(n + 1)$ -permutations needed to cover each n -permutation $\lambda \geq 2$ times. Then,*

$$\kappa_{n,\lambda} \leq \frac{(n + 1)!}{n^2} (\log n + (\lambda - 1) \log \log n + O(1)).$$

Proof. We first choose an unspecified number Y of $(n + 1)$ -permutations randomly and *with* replacement. This might lead to replication with very small probability, but the proof is far more streamlined - and can easily be adapted to the case where we choose Y distinct permutations. The probability that any permutation is covered just j times; $0 \leq j \leq \lambda - 1$ is

$$\binom{Y}{j} \left(\frac{n^2 + 1}{(n + 1)!} \right)^j \left(1 - \frac{n^2 + 1}{(n + 1)!} \right)^{Y-j},$$

so that the expected number of such permutations is

$$\mathbb{E}(X_j) = n! \binom{Y}{j} \left(\frac{n^2 + 1}{(n+1)!} \right)^j \left(1 - \frac{n^2 + 1}{(n+1)!} \right)^{Y-j}; 0 \leq j \leq \lambda - 1.$$

We can cover each such permutation in any *ad hoc* way by choosing $\lambda - j$ additional $(n+1)$ -permutations, and so the number Z_Y of $(n+1)$ -permutations in this successful λ -covering is

$$Z_Y = Y + \sum_{j=0}^{\lambda-1} (\lambda - j) X_j,$$

and thus

$$\mathbb{E}(Z_Y) = Y + n! \sum_{j=0}^{\lambda-1} (\lambda - j) \binom{Y}{j} \left(\frac{n^2 + 1}{(n+1)!} \right)^j \left(1 - \frac{n^2 + 1}{(n+1)!} \right)^{Y-j}$$

for any initial choice of Y permutations. Set

$$p = \frac{n^2 + 1}{(n+1)!}; q = 1 - p,$$

and, given Y Bernoulli trials with success probability p , denote the cumulative and point binomial probabilities by $B(Y, p, k)$ and $b(Y, p, k)$ respectively, $0 \leq k \leq Y$. It is easy to verify that for $Yp \geq \lambda$

$$\begin{aligned} \mathbb{E}(Z_Y) &= Y + n! \lambda B(Y, p, \lambda - 1) - n! Y p B(Y - 1, p, \lambda - 2) \\ &= Y + n! \lambda b(Y, p, \lambda - 1) + n! \{ \lambda B(Y, p, \lambda - 2) - Y p B(Y - 1, p, \lambda - 2) \} \\ &\leq Y + n! \lambda b(Y, p, \lambda - 1) + n! \lambda \{ B(Y, p, \lambda - 2) - B(Y - 1, p, \lambda - 2) \} \\ &= Y + n! \lambda b(Y, p, \lambda - 1) + n! \lambda \sum_{j=0}^{\lambda-2} \binom{Y-1}{j} p^j q^{Y-1-j} \left(\frac{Yq}{Y-j} - 1 \right) \\ &\leq Y + n! \lambda b(Y, p, \lambda - 1) \\ &\leq Y + n! \lambda \left(\frac{Yp}{q} \right)^{\lambda-1} \frac{e^{-pY}}{(\lambda-1)!}. \end{aligned} \tag{1}$$

We do not attempt to optimize carefully in (1); rather we set

$$Y = \frac{(n+1)!}{n^2 + 1} (\log n + (\lambda - 1) \log \log n)$$

and see that the second term T_2 in (1) reduces as

$$\begin{aligned} T_2 &= n! \lambda \left(\frac{\log n + (\lambda - 1) \log \log n}{q} \right)^{\lambda-1} \frac{\exp\{-\log n - (\lambda - 1) \log \log n\}}{(\lambda - 1)!} \\ &= \frac{(n+1)!}{n^2 + 1} \frac{\lambda}{(\lambda - 1)!} \frac{1}{(\log n)^{\lambda-1}} \left(\frac{\log n + (\lambda - 1) \log \log n}{q} \right)^{\lambda-1} (1 + o(1)) \\ &= \frac{(n+1)!}{n^2 + 1} \frac{\lambda}{(\lambda - 1)!} (1 + o(1)), \end{aligned}$$

so that (1) yields

$$\mathbb{E}(Z_Y) \leq \frac{(n+1)!}{n^2+1} \left(\log n + (\lambda-1) \log \log n + (1+o(1)) \frac{\lambda}{(\lambda-1)!} \right).$$

Finally, we choose a sample outcome for which $|\mathcal{A}| \leq \mathbb{E}(Z_Y)$ to complete the proof. $\square$

We now describe the hypergraph formulation of [10] that was used in [12] to prove the Erdős-Hanani conjecture using a method that involved branching processes, dynamical algorithms, hypergraph theory, and differential equations. In this formulation the vertices of the hypergraph consisted of the ensemble of t -sets; for us they would be the class of permutations in S_n . The edges in [12] were the collections of t -subsets of the k -sets, so that the hypergraph was $\binom{k}{t}$ uniform. If analogously, we let edges be the set of n -permutations covered by an $(n+1)$ -permutation, then the hypergraph is no longer uniform. It is not too hard to prove, however, that each $(n+1)$ -permutation π covers $n+1-s_\pi$ n -permutations, where s_π is the number of *successions* in π , where a succession is defined as an episode $\pi(i+1) = \pi(i) \pm 1$. Moreover, we know [6] that the number of successions in a random permutation is approximately Poisson with parameter ~ 2 , so that it is reasonable to assert that most hypergraph edges consist of $n - O(1)$ vertices. This is the first deviation from the Pippenger-Spencer model of [10], which we consider to be *not too serious* insofar as the lack of uniformity of the hypergraph is concerned but *rather serious* due to the fact that the uniformity level $n - O(1)$ is not finite. Lemma 1 above shows that the degree of each vertex is $O(n^2)$, and we will prove in Lemma 6 below that the codegree of two vertices π and π' is at most $O(1)$, so that the codegree is an order of magnitude smaller than the degree. This is good. The above problems with the hypergraph formulation notwithstanding, we make the following conjecture:

Conjecture 4. *For some constant A ,*

$$\limsup_{n \rightarrow \infty} \frac{\kappa_n}{(n+1)!/n^2} = A,$$

and possibly $A \leq 2$.

3 Covering Two Permutations

Lemma 5. *For any $\pi \in S_n$, the set*

$$\mathcal{J}_\pi := \{\pi' \in S_n : \pi \text{ and } \pi' \text{ can be jointly covered by } \rho \in S_{n+1}\}$$

has cardinality at most n^3 .

Proof. Fix π . For an $(n+1)$ -permutation to be able to successfully cover another $\pi' \in S_n$ (in addition to π), π must contain an $(n-1)$ -subpattern of π' . This subpattern may be present in $\binom{n}{n-1} = n$ possible positions of π , and can be represented, using the numbers $\{1, 2, \dots, n\}$, in n ways. Finally, the n th letter of π' can be inserted into this subpattern in n ways. This proves the first part of the lemma. $\square$

Before proving the next result on how many $\rho \in S_{n+1}$ can jointly cover two fixed $\pi, \pi' \in S_n$, we give some motivation and examples:

(i) First, note that in the covering $(n+1)$ -permutation ρ , the numbers 1 and $n+1$ may only play the role of 1 and n respectively in either of the permutations π, π' , while any other number t in ρ may play the role of either t or $t-1$. Thus whenever one can jointly embed π and π' into some permutation ρ , there are some points of π and π' which match up by position *and* value, and all remaining points can match up by position but be one off by value, or vice-versa. For example, if $\pi = 126983745$; $\pi' = 127938645$, then an alignment (σ, σ') of π and π' such as

$$\begin{array}{rcccccccccc} \sigma : & 1 & 2 & 6 & 9 & 8 & 3 & 7 & \emptyset & 4 & 5 \\ \sigma' : & 1 & 2 & 7 & 9 & \emptyset & 3 & 8 & 6 & 4 & 5 \\ \rho : & 1 & 2 & 7 & 10 & 9 & 3 & 8 & 6 & 4 & 5, \end{array}$$

allows for joint coverage by ρ as indicated (we call such an alignment “legal”), whereas a “bad” alignment of the same two permutations would be, e.g.,

$$\begin{array}{rcccccccccc} \sigma : & 1 & 2 & 6 & 9 & \emptyset & 8 & 3 & 7 & 4 & 5 \\ \sigma' : & 1 & 2 & 7 & 9 & 3 & 8 & \emptyset & 6 & 4 & 5, \end{array}$$

since, starting with 10 representing the two 9’s and 9 representing the two 8’s, we find that there is no position, in ρ , where the 8 can be placed.

(ii) Second, note that if the alignment (σ, σ') is legal, then its components may be laid atop each other, by simultaneous rearrangement, so that each is left-right non-decreasing; and the covering $(n+1)$ -permutation is id_{n+1} . If the alignment is not legal, then such a rearrangement leads to sequences that are not both non-decreasing. For instance, the legal example above can be rearranged as follows:

$$\begin{array}{rcccccccccc} \sigma : & 1 & 2 & 3 & 4 & 5 & \emptyset & 6 & 7 & 8 & 9 \\ \sigma' : & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & \emptyset & 9 \end{array}$$

It is always possible as above to align any two co-coverable π and π' in $n-1$ spots (using two $\emptyset$ symbols), because there exists a ρ which covers both π and π' , and deletion of two symbols from ρ yields a $(n-1)$ -subpattern shared by π, π' . The question is when this process is non-unique. To facilitate understanding of this issue we indicate an *alternative* way of identifying the $(n-1)$ -pattern that is shared by such π, π' , *without* the use of the $\emptyset$ symbols. Consider the following situations. In the first example,

$$\begin{array}{rcccccccccc} \pi : & 3 & 8 & 7 & 4 & 2 & 9 & 6 & 10 & 5 & 1 \\ \pi' : & 3 & 4 & 8 & 5 & 2 & 9 & 7 & 10 & 6 & 1, \end{array}$$

Example (i)

the “mismatch” (i.e., the pair of symbols that do not contribute to the common sub-pattern) could be considered to be between the 8 (in π) and the 4 (in π'); in Example (ii)

$$\begin{array}{rcccccccc} \pi : & 2 & 1 & 9 & 3 & 4 & 7 & 8 & 6 & 5 \\ \pi' : & 2 & 1 & 9 & 3 & 7 & 4 & 8 & 6 & 5, \end{array}$$

Example (ii)

the mismatch could be viewed as being between the two 7s or between the two 4s; and finally in the last example

$$\begin{array}{rcccccccc} \pi : & 1 & 2 & 6 & 9 & 8 & 3 & 7 & 4 & 5 \\ \pi' : & 1 & 2 & 7 & 9 & 3 & 8 & 6 & 4 & 5, \end{array}$$

Example (iii)

the 8 in π and the 6 in π' are mismatched. The point is that the mismatched pair might be atop each other, or arbitrarily far apart. Example (iii) contains the same permutations as the ones introduced earlier; utilizing the notion of common subpatterns and mismatches, we have the matching pairs consisting of

$$(1, 1), (2, 2), (6, 7), (9, 9), (3, 3), (7, 8), (4, 4), (5, 5).$$

The skipped symbols, 8 in π and 6 in π' , provide the mismatch – and the corresponding covering $(n+1)$ -permutation can be constructed from the above matching by introducing the $\emptyset$ symbols to appear below or above the mismatched numbers exactly as before. Thus, the two ways of representing co-coverable permutations, using either the $\emptyset$ symbols or mismatches, are equivalent, and we will use one or the other as appropriate.

Lemma 6. *For any $\pi, \pi' \in S_n$, the cardinality of*

$$\mathcal{C}_{\pi, \pi'} := \{\rho \in S_{n+1} : \rho \text{ covers both } \pi \text{ and } \pi'\}$$

is at most 4.

Proof. We shall be proving that for any two co-coverable permutations π, π' , there are at most two legal insertions of the $\emptyset$ symbols, for each of which there can be at most two covering permutations, thus giving a total of four possibilities.

Given a mismatch in which the two numbers are the same, say a , they may be represented by either a or $a+1$ in the covering $(n+1)$ -permutation. This gives two coverings depending on where a is placed. But the common mismatched number may be chosen in up to two ways, as in Example (ii), for a total of 4 covering permutations; this maximal possibility exists if and only if π' is obtained from π by a single swap of adjacent elements. If the two mismatched numbers a, b are different, the case where these are “atop” each other (as in Example (i)) yields two covering permutations, since the $\emptyset$ symbols may be inserted in the two alternate forms

$$\begin{array}{cc} a & \emptyset \\ \emptyset & b \end{array} \quad \text{or} \quad \begin{array}{cc} \emptyset & a \\ b & \emptyset \end{array},$$

which lead to different covering ρ s. If the mismatched numbers are farther apart and different, then it is clear that there is only one covering permutation.

In what follows, we will denote the positions of potential mismatches by $(\alpha, \beta), (\gamma, \delta)$, and set

$$a = \pi_\alpha, b = \pi'_\beta; c = \pi_\gamma, d = \pi'_\delta.$$

We now need to show that all possibilities have been covered, in other words, that there do not exist two sets of matching pairs with mismatched pairs at positions (α, β) and at (γ, δ) , where c, d are not obtained from a, b by the processes described in the previous paragraph, and where c, d yield genuinely different covering permutations. First observe that alternative choices of c and d *do* exist. For instance, in Example (i), c and d may be taken to be 7 and 4; or 7 and 3; or 8 and 3, but these *do not* yield ρ s that are different from the $\begin{smallmatrix} \emptyset & 8 \\ 4 & \emptyset \end{smallmatrix}$ alignment.

We proceed by contradiction, and by considering all possible ways in which the mismatched pairs (a, b) and (c, d) can interact. Assume without loss of generality that $\pi'_{p_i} \geq \pi_{m_i}$ for all i , where p_i and m_i are the positions of the i th match. This forces $a \geq b$. Moreover, if $a = b$ then the $\pi - \pi'$ match is between identical elements.

CASE 1: Mismatches a and b are atop each other; mismatches c and d are also atop each other; c and d are to the right of a and b respectively.

Since a and b are atop each other, we must have $\pi^{-1}(i) = \pi'^{-1}(i)$ for each $i \leq b-1; i \geq a+1$; $\pi^{-1}(i) \neq \pi'^{-1}(i)$ for each $b \leq i \leq a$ (assuming without loss of generality that $a \neq b$). Similarly, i and i are matched for each $i \leq d-1; i \geq c+1$ and not for other i s. This forces $[b, a] \cap [d, c] = \emptyset$, and thus $\pi = \pi'$, a contradiction.

CASE 2: a and b are atop each other; c and d are to the right of a and b respectively (and not atop each other).

Assume that d is to the right of c . c and d are initially matched to elements below and above it respectively. When the mismatched pair becomes (c, d) , all matches between c and d are “tilted to the right”, e.g., the number below the c is now matched to the number to the right of c . Since $c \geq d$, and since $i \in \pi'$ may only be matched with either i or $i-1$ in π , we see that the sequences of integers between γ and δ are identical in π and π' , and consist of a monotone decreasing block of consecutive integers. It follows that all integers lower than d in π are matched to themselves in π' , as are all integers higher than c . This forces both π and π' to be the same permutation of the identity, a contradiction.

CASE 3: $\alpha, \beta, \gamma, \delta$ are in increasing order from left to right.

We proceed as in Case 2, except that there are now two monotone blocks – between a and b ; and between c and d .

CASE 4: The positions appear in the order $\alpha, \gamma, \beta, \delta$ or one of its variants.

Assume that the match corresponding to the (a, b) mismatch has oblique lines that go from bottom left to top right. We see that the oblique matching lines corresponding to the (c, d) -mismatch start between α and β but continue beyond β . If, however, $\gamma \geq \alpha+1$, or $\delta \geq \beta+1$, we see that there are again one or two monotone blocks, forcing a contradiction. Thus $\alpha = \gamma$ and $\beta = \delta$, which reduces us to one mismatch. We thus need to consider

when $\alpha = \beta = \gamma$ or $\alpha = \beta = \delta$, as in Example (i), with the (7,4) or (8,3) mismatches. In this case, either α has to be compatible with both β, δ ; or β has to be compatible with both α, γ . The same covering permutations thus result with both sets of mismatches.

CASE 5: The positions appear in the order $\gamma, \alpha, \beta, \delta$ or one of its variants.

Here a contradiction arises if $\gamma \leq \alpha - 2$ or $\delta \geq \beta + 2$. Assume, therefore, that $\gamma = \alpha - 1; \delta = \beta + 1$; α and β could be arbitrarily far apart, but the point is that the $a - b$ mismatch could be rewired to become a $c - b$ or $a - d$ or even a $c - d$ mismatch, as illustrated in Example (i), with the 7-3 mismatch. The fact that α and β could be arbitrarily far apart is illustrated by the example below:

$$\begin{aligned}\pi : & \quad 5 \quad 6 \quad 9 \quad 10 \quad 11 \quad 12 \quad 4 \quad 7 \quad 8 \quad 1 \quad 2 \quad 3 \\ \pi' : & \quad 6 \quad 9 \quad 10 \quad 11 \quad 12 \quad 3 \quad 5 \quad 7 \quad 8 \quad 1 \quad 2 \quad 4,\end{aligned}$$

Example (iv)

where the (6,3) mismatch can be rewired to become a (5,3) mismatch without changing the covering permutation. If $\gamma = \alpha - 1; \delta = \beta + 1$; a $c - d$ rewiring may lead to up to two entirely new ρ s, as in Example (ii).

This completes the proof of Lemma 5. □

4 Thresholds

Theorem 7. *Consider the probability model in which each $\pi \in S_{n+1}$ is independently picked with probability p . Let the resulting random collection of permutations be denoted by $\mathcal{A}$. Then,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{A} \text{ is a cover of } S_n) = \begin{cases} 0 & \text{if } p \leq \frac{\log n}{n}(1 + o^*(1)) \\ 1 & \text{if } p \geq \frac{\log n}{n}(1 + o(1)). \end{cases}$$

Proof. We use the Janson inequalities, see [1]. By the lower Janson inequality, we have, with X denoting as before the number of uncovered $\pi \in S_n$,

$$\begin{aligned}\mathbb{P}(X = 0) & \geq \prod_{i=1}^{n!} \mathbb{P}(\text{permutation } i \text{ is covered}) \\ & \geq \prod_{i=1}^{n!} 1 - (1 - p)^{n^2} \\ & \geq \prod_{i=1}^{n!} \exp\{-(1 - p)^{n^2} / (1 - (1 - p)^{n^2})\} \\ & = \exp\{-n!(1 - p)^{n^2} / (1 - (1 - p)^{n^2})\} \rightarrow 1\end{aligned}$$

if $\mathbb{E}(X) \rightarrow 0$, or, using Stirling's formula, if $C\sqrt{n} \left(\frac{n}{e}\right)^n e^{-n^2 p} \rightarrow 0$, i.e., if

$$p = \frac{\log n}{n} - \frac{1}{n} + \frac{\log n}{2n^2} + \frac{\omega(n)}{n^2} = \frac{\log n}{n}(1 + o(1)),$$

where $\omega(n) \rightarrow \infty$ is arbitrary. This proves the second part of the result. To prove the first part, we invoke the upper Janson inequality to give

$$\mathbb{P}(X = 0) \leq \exp\{-\lambda + \Delta\} \quad (2)$$

where $\lambda = \mathbb{E}(X)$ and

$$\Delta = \sum_i \sum_{j \sim i} \mathbb{P}(i \text{ and } j \text{ are both uncovered}),$$

and $i \sim j$ if permutations i and j can be covered by the same $(n+1)$ -permutation. Since, by Lemmas 5 and 6, for any i there are at most n^3 permutations j that can be jointly covered with i , and for any i, j there are at most 4 co-covering $(n+1)$ -permutations, it follows that

$$\Delta = n!n^3(1-p)^{2n^2+2-4},$$

so that (2) yields

$$\mathbb{P}(X = 0) \leq \exp\{-n!(1-p)^{n^2+1} + n!n^3(1-p)^{2n^2-2}\} \rightarrow 0$$

if $n!(1-p)^{n^2} \rightarrow \infty$, i.e., if

$$p = \frac{\log n - 1 + \frac{1}{2} \frac{\log n}{n} - \frac{\omega(n)}{n}}{n} = \frac{\log n}{n}(1 + o^*(1)), \quad (3)$$

where $\omega(n) \rightarrow \infty$ is arbitrary. This completes the proof of Theorem 7. $\square$

5 Poisson Approximation in the “Gap”

The proof of Theorem 7 reveals that $\mathbb{E}(X)$ undergoes a rapid transition when p is around the level given by (3). In fact, if

$$p = \frac{\log n - 1 + \frac{1}{2} \frac{\log n}{n} - \frac{K}{n}}{n}, K \in \mathbb{R},$$

then for large n , $\mathbb{E}(X) \sim \sqrt{2\pi}e^{-K}$ and $\mathbb{P}(X = 0) \sim \exp\{-\sqrt{2\pi}e^{-K}\}$. Much more is true, however, as we shall show next: The entire probability distribution $\mathcal{L}(X)$ of X can be approximated, in the total variation sense, by that of a Poisson random variable with mean $\lambda = \mathbb{E}(X)$ in a range of p s that allows for large means. This result can be thought of as being a probabilistic counterpart to Theorem 7, and is proved using the Stein-Chen method of Poisson approximation [2]:

Theorem 8. Consider the model in which each $\pi \in S_{n+1}$ is independently chosen with probability p , thus creating a random ensemble $\mathcal{A}$ of $(n+1)$ -permutations. Then $d_{\text{TV}}(\mathcal{L}(X), \text{Po}(\lambda)) \rightarrow 0$ if $p \geq \frac{\log n}{n^2}(1 + \epsilon)$, where $\text{Po}(\lambda)$ denotes the Poisson distribution with parameter λ , $\epsilon > 0$ is arbitrary, and the total variation distance d_{TV} is defined by

$$d_{\text{TV}}(\mathcal{L}(Y), \mathcal{L}(Z)) = \sup_{A \subseteq \mathbb{Z}^+} |\mathbb{P}(Y \in A) - \mathbb{P}(Z \in A)|.$$

Proof. As before, we set $X = \sum_{j=1}^{n!} I_j$, with $I_j = 1$ if the j th n -permutation is uncovered by permutations in $\mathcal{A}$ ($I_j = 0$ otherwise), and let $\lambda = \mathbb{E}(X) = n!(1-p)^{n^2+1}$. Consider the following coupling, for each j : If $I_j = 1$, we “do nothing,” setting $J_i = J_{ji} = I_i$, $1 \leq i \leq n!$. If, on the other hand, the j th permutation is covered by one or more $(n+1)$ -permutations in $\mathcal{A}$, we “deselect” these permutations, setting $J_i = 1$ if the i th permutation is uncovered after this change is made; $J_i = 0$ otherwise. Now it is clear that $J_i \geq I_i$ for $i \neq j$, since there is no way that an uncovered permutation can magically get covered after a few $(n+1)$ -permutations are deselected. Also, setting $N = n!$, we have for each j ,

$$\mathcal{L}(J_1, J_2, \dots, J_N) = \mathcal{L}(I_1, \dots, I_N | I_j = 1).$$

Corollary 2.C.4 in [2] thus applies, telling us that

$$d_{\text{TV}}(\mathcal{L}(X), \text{Po}(\lambda)) \leq \frac{1 - e^{-\lambda}}{\lambda} \left(\mathbb{V}(X) - \lambda + 2 \sum_j \mathbb{P}^2(I_j = 1) \right), \quad (4)$$

Bounding $1 - e^{-\lambda}$ by one, (4) yields

$$d_{\text{TV}}(\mathcal{L}(X), \text{Po}(\lambda)) \leq \frac{\mathbb{V}(X)}{\lambda} - 1 + 2(1-p)^{n^2+1}. \quad (5)$$

The last term in (5), namely $(1-p)^{n^2+1}$ can easily be verified to tend to zero as long as $p \gg 1/n^2$, so we turn to a computation of $\mathbb{V}(X)$:

$$\mathbb{V}(X) = \sum_j \{\mathbb{E}(I_j) - \mathbb{E}^2(I_j)\} + \sum_{i \sim j} \{\mathbb{E}(I_i I_j) - \mathbb{E}(I_i)\mathbb{E}(I_j)\},$$

so that

$$\begin{aligned} \frac{\mathbb{V}(X)}{\lambda} - 1 &\leq \frac{\sum_{i \sim j} \{\mathbb{E}(I_i I_j) - \mathbb{E}(I_i)\mathbb{E}(I_j)\}}{\lambda} \\ &\leq n^3 \left((1-p)^{n^2-3} - (1-p)^{n^2+1} \right) \\ &= n^3 (1-p)^{n^2+1} \{ (1-p)^{-4} - 1 \} \\ &\leq 5pn^3 \exp\{-n^2 p\} \rightarrow 0 \end{aligned}$$

provided that $p \geq A \frac{\log n}{n^2}$, with $A > 1$. □

6 Open Problems

Proving the best possible result along the lines of Conjecture 4 would clearly be a most desirable outcome of this work. Can, e.g. we write “lim” instead of “lim sup”? Is $A = 1$? Secondly, can the bound in Lemma 5 be improved? (Note, however, that such an improvement would provide only marginal improvements in Theorems 7 and 8.) Finally, extending the results of this paper to encompass coverings of n permutations by $(n + k)$ -permutations would lead to several interesting questions and new techniques; for example, Theorem 2 can readily be generalized for $k \geq 2$. Also, the maximum degree of the dependency graph induced by the indicator random variables $\{I_i\}$ is low even for $k \geq 2$, so that these variables are almost independent, and one may even envision tight results as $k \rightarrow \infty$.

Acknowledgements

Thanks to the anonymous referee for many suggestions that helped improve the paper, both in form and in content.

References

- [1] N. Alon and J. Spencer, 1992. *The Probabilistic Method*, John Wiley, New York.
- [2] A. Barbour, L. Holst, and S. Janson, 1992. *Poisson Approximation*, Oxford Uni. Press.
- [3] P. Erdős and J. Spencer, 1974. *Probabilistic Methods in Combinatorics*, Academic Press, New York.
- [4] P. Flajolet and R. Sedgewick, 2009. *Analytic Combinatorics*, Cambridge Uni. Press.
- [5] A. Godbole and S. Janson (1996). “Random covering designs,” *J. Combinatorial Theory, Series A* **75**, 85–98.
- [6] A. Godbole and P. Sissokho (1999). “Near-matches and successions in random permutations,” *Congressus Numerantium* **135**, 159–170.
- [7] A. Godbole, D. Skipper, and R. Sunley (1996). “ t -covering arrays: upper bounds and Poisson approximations,” *Combinatorics, Probability and Computing* **5**, 105–118.
- [8] A. Godbole, S. Thompson, and E. Vigoda (1996). “General upper bounds for covering numbers,” *Ars Combinatoria* **42**, 211–221.
- [9] A. Myers and H. Wilf (2003). “Some new aspects of the coupon collector’s problem,” *SIAM J. Discrete Mathematics* **17**, 1–17.
- [10] N. Pippenger and J. Spencer (1989). “Asymptotic behaviour of the chromatic index for hypergraphs,” *J. Combinatorial Theory, Series A* **51**, 24–42.
- [11] V. Rödl (1985). “On a packing and covering problem,” *European J. Combinatorics* **5**, 69–78.
- [12] J. Spencer (1995). “Packing and covering via a branching process,” *Rand. Structures Algorithms* **7**, 167–172.