Asymptotic bounds for permutations containing many different patterns

Alison Miller*

Harvard University

Department of Mathematics

Cambridge, MA 02138

millerab@gmail.com

May 23, 2008

*Please send correspondence to the following address: 2255 Algonquin Rd., Niskayuna, NY 12309

Abstract

We say that a permutation $\sigma \in S_n$ contains a permutation $\pi \in S_k$ as a pattern if some subsequence of σ has the same order relations among its entries as π. We improve on results of Wilf, Coleman, and Eriksson et al. that bound the asymptotic behavior of $\mathrm{pat}(n)$, the maximum number of distinct patterns of any length contained in a single permutation of length n. We prove that $2^n - O(n^2 2^{n-\sqrt{2n}}) \leq \mathrm{pat}(n) \leq 2^n - \Theta(n 2^{n-\sqrt{2n}})$ by estimating the amount of redundancy due to patterns that are contained multiple times in a given permutation. We also consider the question of k-superpatterns, which are permutations that contain all patterns of a given length k. We give a simple construction that shows that L_k, the length of the shortest k-superpattern, is at most $\frac{k(k+1)}{2}$. This may lend evidence to a conjecture of Eriksson et al. that $L_k \sim \frac{k^2}{2}$.

1 Introduction

The field of permutation pattern research studies the set of permutations S_n of n elements as permutations of the totally ordered set $[n] = \{1, 2, \ldots, n\}$. We will represent a permutation $\sigma \in S_n$, by the sequence $\sigma(1), \sigma(2), \ldots, \sigma(n)$ (omitting commas when there is no ambiguity). We then say that $\sigma \in S_n$ contains a pattern $\pi \in S_k$, for some $1 \leq k \leq n$, if there exists a subsequence of σ with the same order relations as π (see Section 2 for a formal definition). For example, the permutation 25314 contains the pattern 312 because its subsequence 514 has the same order relations as 312. We wish to estimate the maximum number of distinct patterns that can be contained in a permutation $\sigma \in S_n$. This maximum, call it $\mathrm{pat}(n)$, is trivially bounded above by 2^n.

The main known results and conjectures about the asymptotic growth of $\mathrm{pat}(n)$ have compared $\mathrm{pat}(n)$ to 2^n. In [8], Wilf first attacked the question of determining its rate of growth. He found the exponential lower bound $\mathrm{pat}(n) \geq F_{n+1}$ (the $(n+1)$st Fibonacci number), but was unable to determine the value of $\limsup \sqrt[n]{\mathrm{pat}(n)}$, and asked whether it was less than 2. His question was answered by Coleman [4], who proved that $\mathrm{pat}(k^2) \geq 2^{(k-1)^2}$, implying that $\limsup \sqrt[n]{\mathrm{pat}(n)} = 2$. More recently, Albert et al. [1] showed that $\mathrm{pat}(n)$ approaches 2^n in a stronger manner. Their bound $\mathrm{pat}(n) \geq 2^n - O(\sqrt{n} 2^{n-\sqrt{n}/2})$ is strong enough to imply $\lim_{n\to\infty} \frac{\mathrm{pat}(n)}{2^n} = 1$, answering a question of Bóna. However, we will show that their correction term $O(\sqrt{n} 2^{n-\sqrt{n}/2})$ can be reduced substantially, to $O(n^2 2^{n-\sqrt{2n}})$.

In this paper, we determine just how close $\mathrm{pat}(n)$ is to 2^n by giving upper and lower bounds for the quantity $2^n - \mathrm{pat}(n)$ that differ only by an $O(n)$ factor. Our lower bound $\mathrm{pat}(n) \geq 2^n - O(n^2 2^{n-\sqrt{2n}})$ (Theorem 4.1) comes from "tilted checkerboard" permutations that have an easily analyzable checkerboard-like structure similar to that of the square grid permutations used in [4] and [1] and the constructions of Eriksson et al. [6] for the closely related superpattern problem. We show our upper bound, $\mathrm{pat}(n) \leq 2^n - O(n2^{n-\sqrt{2n}})$ (Theorem 6.5) by making rigorous Coleman's insight that to maximize the number of distinct patterns contained, we must space the entries of our permutation apart in the taxicab metric.

There is another natural problem in this area of permutation packings, loosely dual to that of finding $\mathrm{pat}(n)$. A *k-superpattern*, defined in [2], is a permutation containing all permutations of length k as patterns. We wish to bound the minimal length L_k of a k-superpattern. The best previously known asymptotic bounds for $\frac{L_k}{k^2}$ are

$$\frac{1}{e^2} \leq \lim_{k\to\infty} \frac{L_k}{k^2} \leq \frac{2}{3}.$$

The lower bound is trivial, much like the upper bound of 2^n for $\mathrm{pat}(n)$. However, Eriksson et al. ([5, Theorem 6.2]) require a nontrivial probabilistic argument to show that any pattern can be contained in their tilted checkerboard-like permutation of length $\frac{2k^2}{3} + o(k^2)$. In the same paper, they conjectured that asymptotically $L_k \sim \frac{k^2}{2}$. In Theorem 3.1 we will give a simple construction, related to the tilted checkerboard, that yields a large family of k-superpatterns of length $\frac{k(k+1)}{2}$.

The structure of this paper is as follows: Section 2 contains the definitions and constructions that will form the basis for our later arguments. Section 3 is a brief excursion into superpatterns; we use a "zigzag word" construction to produce numerous small superpatterns whose structure resembles that of a tilted checkerboard. We then focus our attention on finding lower and upper bounds for $\mathrm{pat}(n)$.

In Section 4 we find a constructive lower bound for $\mathrm{pat}(n)$ using tilted checkerboard permutations. Our goal is to show that a tilted checkerboard having length n has nearly $2^n - O(n^2 2^{n-\sqrt{2n}})$ patterns that it contains in only one way, which implies our desired result. Because the tilted checkerboard has only $O(\sqrt{n})$ descents, evenly spaced, we can use the descents of a contained pattern as benchmarks to locate it within the checkerboard and do the same with the inverse ascents. This tightly restricts the possibilities for subsets of the checkerboard whose corresponding permutation is not uniquely contained. The restriction we get is equivalent to saying that S is disjoint from each of $O(n^2)$ subsets of $[n]$ of size close to $\sqrt{2n}$, from which our bound follows (with some explicit counting postponed to Section 5).

Finally, we show our upper bound, which says that at least $\Theta(n 2^{n-\sqrt{2n}})$ of the 2^n subsets $S \subseteq [n]$ are redundant, in the sense that there is another subset of $[n]$ that corresponds to the same pattern. We do so by considering the geometry of the graph of a permutation σ, that is, the set of all pairs $(i, \sigma(i)) \in \mathbb{Z}^2$, equipped with the standard taxicab metric. In Section 6, we bound $\mathrm{pat}(n)$ from above in terms of a geometric quantity called the "swap-redundancy coefficient," which we then bound by an inclusion-exclusion argument, with the technical details of showing that our correction terms are small postponed to Section 7. (The same methods can also be used to give the bound $\mathrm{pat}(n) \leq 2^n - \Theta(2^{n-\sqrt{2n}})$ with much less work.)

2 Definitions and constructions

A pattern π of length k is a permutation in S_k, that is, a permutation of the elements of the set $[k] = \{1, 2, 3, \ldots, k\}$: in this sense, "pattern" and "permutation" are synonymous, but we will generally talk about smaller patterns being contained in larger permutations.

Definition. For $\sigma = \sigma(1), \sigma(2), \ldots, \sigma(n) \in S_n$ and $\pi = \pi(1), \pi(2), \ldots, \pi(k) \in S_k$ where $1 \leq k \leq n$, we say that σ *contains the pattern* π if there exist indices

$a_1, a_2, \ldots, a_k$ with $1 \leq a_1 < a_2 < \cdots < a_k \leq n$ such that for any $i, j \in [k]$, $\sigma(a_i) < \sigma(a_j)$ if and only if $\pi(i) < \pi(j)$. In such a case, we say that the subset $\{a_1, a_2, \ldots, a_k\}$ of $[n]$ *represents* π.

The quantity of interest is the number of distinct patterns contained in a given permutation.

Definition. For σ a permutation, let $f(\sigma)$ denote the number of distinct patterns of any positive length contained in σ. Furthermore, for a fixed positive integer n, let

$$\mathrm{pat}(n) = \max_{\sigma \in S_n} f(\sigma).$$

Any pattern π contained in a permutation $\sigma \in S_n$ is represented by one of the 2^n subsets of $[n]$. If $\mathrm{pat}(n)$ is close to 2^n, most of these patterns must be represented by only one of these 2^n subsets. We now consider the number of such "uniquely represented" patterns.

Definition. If π is a pattern contained in σ such that exactly one subset $A = \{a_1, a_2, \ldots, a_k\}$ of $[n]$ represents π, we say that π *is uniquely contained in* σ and that the subset $A = \{a_1, a_2, \ldots, a_k\} \subseteq [n]$ is *uniquely determined* with respect to σ. Otherwise, we say that the set A is *redundant* for σ. Then let $u(\sigma)$ be the number of patterns uniquely contained in σ, which is also the number of uniquely determined subsets for σ. Additionally, let $r(\sigma)$, the *redundancy* of σ, be the number of nonempty subsets of σ that are redundant. Again we define the maximum number of unique patterns

$$\mathrm{uni}(n) = \max_{\sigma \in S_n} u(\sigma)$$

and the corresponding minimum number of redundant subsets,

$$\mathrm{red}(n) = \min_{\sigma \in S_n} r(\sigma).$$

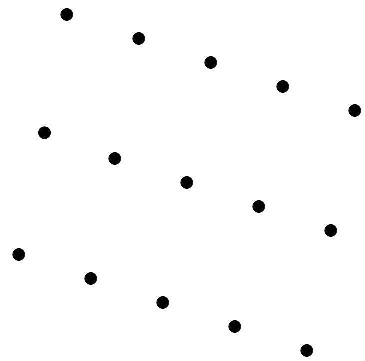

Figure 1: The graph of the 5×3 tilted rectangle permutation 5, 10, 15, 4, 9, 14, 3, 8, 13, 2, 7, 12, 1, 6, 11. In general, the $j \times k$ tilted rectangle consists of j staggered upward runs of length k, each an arithmetic progression of difference j, such that the first entry of the i-th run is $j + 1 - i$.

We note that both $u(\sigma)$ and $f(\sigma)$ are bounded above by 2^n. More precisely, $u(\sigma)+r(\sigma) = 2^n-1$ and $\mathrm{uni}(n)+\mathrm{red}(n) = 2^n-1$ because the redundant subsets of $[n]$ are exactly those that are not uniquely determined.

Proposition 2.1. *For any permutation σ of length n,*

$$2^n - 1 - r(\sigma) = u(\sigma) \leq f(\sigma) < 2^n - 1 - \frac{r(\sigma)}{2}$$

and so

$$2^n - 1 - \mathrm{red}(n) = \mathrm{uni}(n) \leq \mathrm{pat}(n) < 2^n - 1 - \frac{\mathrm{red}(n)}{2}.$$

Proof. The left hand inequality is trivial. For the right hand side, consider the sum $2u(\sigma) + r(\sigma) = 2^{n+1} - 2 - r(\sigma)$. It counts every pattern contained in σ at least twice, so $2^{n+1} - 2 - r(\sigma) \geq 2f(\sigma)$, and the desired inequality follows. $\square$

Our construction in Section 4 will make heavy use of a certain type of permutation, the *tilted square*, which is also closely related to the construction of Section 3.

Definition. The $j \times k$ *tilted rectangle* is the permutation of S_{jk} given by

$$j, 2j, 3j, \ldots, kj, j-1, 2j-1, \ldots, kj-1, \ldots 1, j+1, \ldots, kj-j+1.$$

(See Figure 1.) The *tilted square of size* k is the $k \times k$ tilted rectangle.

Remark. We follow Coleman's convention in [4] in defining the tilted square; in [5] tilted rectangles and tilted squares are defined to be the reflections of those given above.

The term "tilted rectangle" comes from the shape of the graph of this permutation. (For a formal definition of the graph of a permutation, see Section 6.) It is visually evident from the symmetry of the graph of the tilted rectangle (see Figure 1) that the inverse permutation of the $j \times k$ tilted rectangle is identical to the reverse, or reflection in a vertical axis, of the $k \times j$ tilted rectangle. In sequence form, the inverse of the $j \times k$ tilted rectangle is given by

$$kj-k+1, kj-2k+1, \ldots, 1, kj-k+2, kj-2k+2, \ldots, 2, \ldots, kj, kj-k, \ldots, k$$

which has k staggered downwards runs of length j. In the graph of the tilted rectangle, the j upward runs appear as tilted columns, and the k downward runs of the inverse appear as tilted rows. We can generalize this concept of rows and columns to arbitrary permutations.

Definition. A *column* of a permutation σ is a maximal upwards run of σ, that is, a maximal set of the form $\{i, i+1, i+2, \ldots, i+m-1, i+m\}$ satisfying $\sigma(i) < \sigma(i+1) < \cdots < \sigma(i+m)$. A *row* of a permutation σ is a maximal subset of $[n]$ corresponding to a downwards run in σ^{-1}, namely, a maximal set of the form $\{\sigma^{-1}(i), \sigma^{-1}(i+1), \ldots, \sigma^{-1}(i+m)\}$ with $\sigma^{-1}(i) > \sigma^{-1}(i+1) > \cdots > \sigma^{-1}(i+m)$.

Observe that the descents of σ form the dividing lines between adjacent columns, so that the number of columns of a permutation is one more than its number of descents. Similarly, σ has one more row than σ^{-1} has ascents. (Note: what we call a "column" is also called an "increasing run" or "ascending run" in the literature.)

We now show a simple but useful lemma on rows and columns in an arbitrary permutation.

Lemma 2.2. *Let R be a row and C be a column of a permutation σ. Then R and C intersect in at most one element.*

Proof. We note that if $a, b \in C$ and $a < b$, then $\sigma(a) < \sigma(b)$. Similarly, if $a, b \in R$ and $a < b$, then $\sigma(a) > \sigma(b)$. Hence it is impossible to have two distinct elements in both R and C. $\square$

Note that tilted rectangles have the special property that every row intersects every column, but this is not the case for general permutations. We number the columns of a permutation σ, starting with 1, from left to right as they appear in the graph of σ, and number the rows from bottom to top.

Although Albert et al. [1] used tilted squares in their lower bound for $\text{pat}(n)$, there is another closely related permutation with a higher diversity of patterns. We call it the "tilted checkerboard" because we will construct it by coloring the graph of the tilted square black and white in a checkerboard fashion so that the first entry is colored black, and then taking the subpermutation corresponding to all black squares.

Definition. Color each entry of the $j \times k$ tilted rectangle black or white according to whether the sum of its row number and column number is 0 or 1 mod 2. The $j \times k$ *tilted checkerboard* is the subpermutation of the tilted square induced by the set of all black elements (see Figure 2). In this paper we work mainly with the square $k \times k$ tilted checkerboard, which we call the *k-checkerboard* for brevity.

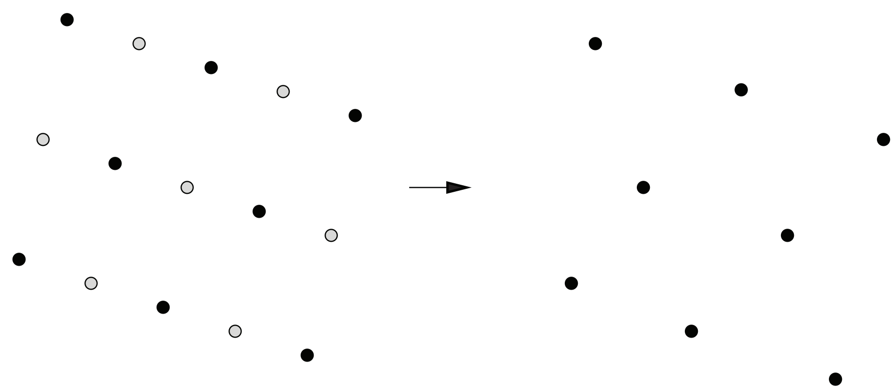

Figure 2: The set of black squares of the 5×3 tilted rectangle and the corresponding checkerboard permutation $(3,\ 8,\ 5,\ 2,\ 7,\ 4,\ 1,\ 6)$

To help us when working with patterns in the $k \times k$ tilted checkerboard, we make the following definitions.

Definition. For a permutation $\sigma \in S_n$, and a subset $S \subseteq [n]$, let π be the pattern of σ determined by S. We say that a subset S *avoids* the ith column of σ if it contains no elements from the ith column of σ. We also say that S *truncates* the ith column of σ if S does not avoid the ith column, and the rightmost element of S in the ith column of σ corresponds to the left end of an ascent of π (equivalently, if it lies below the next element of S). Similarly, S truncates the ith row of σ if the topmost entry of S in the ith row is the bottom end of an inverse descent of π. (See Figure 3.)

The rationale for this nomenclature is that in such a case the ith column runs into the next column to make a single ascending run in π.

3 Superpatterns

The gridlike structure of the tilted rectangle and checkerboard permutations makes it easy to locate specific permutations as patterns in them. However,

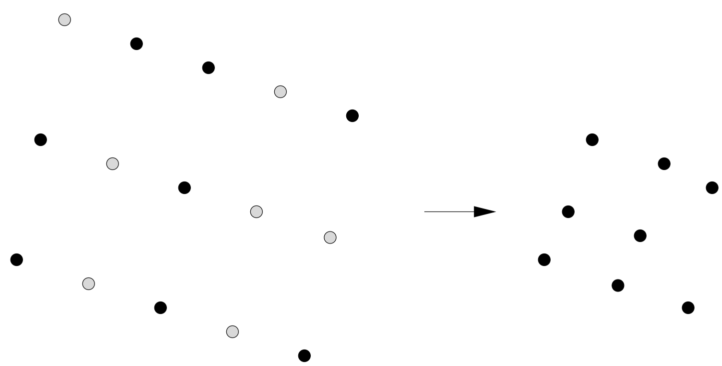

Figure 3: If S is the set of black entries of the 5×3 tilted rectangle shown above, S truncates the first column of the rectangle and avoids the fourth column. By Lemma 4.3 the resulting pattern has $5-1-1=3$ columns, as seen in the picture.

they are not an optimal construction for superpatterns because they contain so many more ascents than descents. In this section we construct patterns with a structure similar to the tilted checkerboards, but with a more equal number of ascents and descents.

We will produce superpatterns of length $\frac{k(k+1)}{2}$ by first constructing a word Z of that length that contains all permutation patterns of length k. We can then easily convert Z into a permutation with the same property. A *word* $w = w(1)w(2)\dots w(n)$ is a sequence of positive integers of arbitrary length, allowing repetitions. We generalize the idea of pattern containment to words: a word w is said to contain a permutation $\pi \in S_k$ as a pattern if there is a subsequence $w(a_1)w(a_2)\dots w(a_k)$ such that $w(a_i) < w(a_j)$ if and only if $\pi(i) < \pi(j)$. We say that a permutation σ *represents* a word w if they have the same length and if for any $i, j \in [n]$, $w(i) < w(j)$ implies $\sigma(i) < \sigma(j)$ (but the converse need not be true). If a permutation σ represents a word w, any pattern contained in w is also contained in σ. Also, if w contains a_i appearances of each positive integer i, there are $\prod_{i \in \mathbb{N}} a_i!$ permutations representing w, because for each i

the a_i elements corresponding to the appearances of i can be ordered in any manner. We also say that a word v is a *subword* of a word w if v is contained in w as a subsequence (not just up to order-preserving relabeling), and we can also consider a permutation $\pi \in S_n$ as the word $\pi(1)\pi(2)\ldots\pi(n)$, so that it makes sense to say that a permutation is a subword of a word.

Now, for a positive integer k, let Z be the *zigzag superword* made by alternating between $\lceil \frac{k}{2} \rceil$ copies of the odd uprun $1\,3\,5\ldots(2\lfloor\frac{k}{2}\rfloor-1)\,(2\lfloor\frac{k}{2}\rfloor+1)$ and $\lfloor\frac{k}{2}\rfloor$ copies of the even downrun $(2\lfloor\frac{k+1}{2}\rfloor)\,(2\lfloor\frac{k+1}{2}\rfloor-2)\,\ldots\,4\,2$. It has a total of $\lceil\frac{k}{2}\rceil\left(\lfloor\frac{k}{2}\rfloor+1\right)+\lfloor\frac{k}{2}\rfloor\lfloor\frac{k+1}{2}\rfloor=\frac{k(k+1)}{2}$ entries. When we refer to a *run* of Z, we will mean one of those $k+1$ odd upruns or even downruns. This zigzag superword has a pattern of runs similar to that of the tilted checkerboard, but it has the advantage over the tilted checkerboard of having roughly equal numbers of ascents and descents instead of being "biased" towards ascents. (See Figure 4).

Theorem 3.1. *For any $\pi \in S_k$, either π or $\pi+1$ (the word produced by adding 1 to each entry of π) is a subword of Z. As a corollary, Z, and so any permutation representing Z, contains any $\pi \in S_k$ as a pattern.*

Proof. We first make a few definitions. An *even ascent* of a word w is an ascent with both entries even: that is, an i for which $w(i) < w(i+1)$ and both $w(i)$ and $w(i+1)$ are even. Let $A_0(w)$ be the set of even ascents of w. Similarly we can define $A_1(w)$, $D_0(w)$, and $D_1(w)$ as sets of odd ascents, even descents, and odd descents of w respectively. Note that odd entries in π correspond to even entries in $\pi+1$, so that

$$A_0(\pi) = A_1(\pi+1), \qquad A_1(\pi) = A_0(\pi+1), \tag{1}$$

$$D_0(\pi) = D_1(\pi+1), \qquad D_1(\pi) = D_0(\pi+1). \tag{2}$$

We can consider Z as an inital segment of an infinite word Z_∞ that alternates

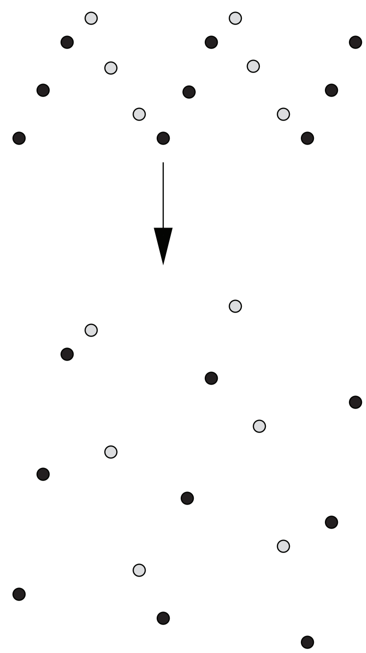

Figure 4: The graphs of the zigzag word
$Z = (1,\ 3,\ 5,\ 6,\ 4,\ 2,\ 1,\ 3,\ 5,\ 6,\ 4,\ 2,\ 1,\ 3,\ 5)$ for $k = 5$ and of one of the $(3!)^3(2!)^3 = 1728$ permutations that represent it.

between infinitely many even upruns and odd downruns. For any π, we can obtain π as a subsequence of Z_∞ by a greedy algorithm: take the first appearance in Z of $\pi(1)$, say it is in the $m_1(\pi)$th run, the first succeeding appearance of $\pi(2)$ in the $m_2(\pi)$th run, and so on, so that $\pi(i)$ lands in the $m_i(\pi)$th run for each i. This algorithm will construct π as a subsequence of Z if the subsequence of Z_∞ corresponding to π is contained in the first k runs of Z_∞, that is, if $m_k(\pi) \le k$. Similarly, $\pi + 1$ will be a subsequence of Z if $m_k(\pi + 1) \le k$.

However, one can easily verify that $m_i(\pi)$ satisfies the conditions

$$m_1(\pi) = \begin{cases} 1 & \text{if } \pi(1) \text{ is odd} \\ 2 & \text{if } \pi(1) \text{ is even} \end{cases} \tag{3}$$

and

$$m_{i+1}(\pi) = \begin{cases} m_i(\pi) + 1 & \text{if } \pi(i+1) \text{ differs in parity from } \pi(i) \\ m_i(\pi) & \text{if } i \in A_0(\pi) \text{ or } i \in D_1(\pi) \\ m_i(\pi) + 2 & \text{if } i \in A_1(\pi) \text{ or } i \in D_0(\pi) \end{cases} \tag{4}$$

and hence

$$m_k(\pi) = m_1(\pi) + k - 1 + (|A_1(\pi)| + |D_0(\pi)| - |A_0(\pi)| - |D_1(\pi)|). \tag{5}$$

Similarly,

$$m_k(\pi{+}1) = m_1(\pi{+}1){+}k{-}1{+}(|A_1(\pi{+}1)|{+}|D_0(\pi{+}1)|{-}|A_0(\pi{+}1)|{-}|D_1(\pi{+}1)|). \tag{6}$$

Now we add (5) and (6) together, using (1) and (2) to cancel terms:

$$m_k(\pi) + m_k(\pi + 1) = m_1(\pi) + m_1(\pi + 1) + 2(k - 1) = 2k + 1 \tag{7}$$

using (3) in the final step. Because $m_k(\pi)$ and $m_k(\pi+1)$ are integers summing to $2k+1$, one of them must be at most k, and we conclude that either π or $\pi+1$ is a subsequence of Z. $\square$

4 Lower bound by identifying uniquely determined subsets of the tilted checkerboard

Our goal in this section will be to give the following lower bound on $\mathrm{pat}(n)$ by analyzing square checkerboards.

Theorem 4.1. *The maximum number of distinct patterns contained in some permutation in S_n is bounded below by*

$$\mathrm{pat}(n) \geq 2^n - O(n^2 2^{n-\sqrt{2n}}).$$

We note that this bound is significantly stronger than the lower bound $2^n - O(\sqrt{n}2^{n-\sqrt{n}/2})$ given in [1].

Our main method will be to estimate $\mathrm{pat}(n)$ by way of $\mathrm{uni}(n)$, or equivalently, by way of $\mathrm{red}(n)$. For this section, we will set $N = \lceil \frac{k^2}{2} \rceil$, the length of the k-checkerboard. Our first step is to show for certain large subpermutations $\sigma \in S_n$ of the k-checkerboard, the non-uniquely determined subsets of $[n]$ can only be of a certain restricted type: namely, those that truncate or avoid at least two of the rows and columns of the k-checkerboard. First we prove two useful lemmas about rows and columns that we will need to characterize those non-uniquely determined subsets.

Lemma 4.2. *Suppose that $\pi \in S_m$ is a pattern represented in a permutation $\sigma \in S_n$ by a subset $S \subseteq [n]$. Say that σ has k columns (rows) and that π has $k-d$ columns (rows). Then if an element a of $[m]$ lies in the jth column (row)* *of π, the corresponding element a' of $S \subseteq [n]$ appears somewhere between the jth column (row) and the $(j+d)$th column (row) of σ, inclusive.*

Proof. We do the argument for columns; the argument for rows is identical. If a is in the jth column of π, there must be $j-1$ descents of π before a, and $k-d-j$ descents of π after a. We can find $j-1$ corresponding descents of σ before a', so a' must lie in at least the jth column of π. Similarly, we can find $k-d-j$ corresponding descents of σ after a', so at least $k-d-j$ of the k columns of σ must come after a', implying that a' can be in at the latest the $k-(k-d-j) = (j+d)$th column of σ. □

We can determine the number of descents or ascents of a pattern from its subset S using the concept of truncation.

Lemma 4.3. *Suppose that a set $S \subseteq [n]$ represents the pattern π in some permutation $\sigma \in S_n$. Say that σ has k total columns, and that S avoids e of the columns (rows) of σ and truncates t columns (rows). Then π has $k-e-t-1$ descents (inverse ascents).*

Proof. We first reduce to the case $e = 0$ by letting σ' be the subpermutation of σ derived by removing the columns avoided by S. The new permutation σ' has $k-e$ columns, corresponding to the remaining columns of S, and the columns of σ' truncated by the new set S' correspond to the columns of σ truncated by S. We now count the number of descents of π. None of them can occur within the columns of σ', so they must all occur between columns of σ. The values of j for which there is a descent of π between the jth and $(j+1)$st columns of σ are exactly those j for which the jth column is not truncated, so π has $(k-e-1)-t = k-e-t-1$ descents. The argument for rows and inverse ascents is exactly the same. □

We now characterize a large family of uniquely contained patterns in a tilted checkerboard.

Proposition 4.4. *Let π be a pattern contained in a tilted checkerboard of size k such that the total number of descents and inverse ascents in π is at least $2k-3$. Then π is uniquely contained. If S is a subset of $[N]$ representing the pattern π, this condition is equivalent to saying that S truncates or avoids at most one of the rows and columns of the tilted checkerboard.*

Proof. We first need to show that any permutation $\pi \in S_m$ contained in the k-checkerboard that has at least $2k-3$ combined descents and inverse ascents is contained uniquely. Note that π can have at most as many ascents and at most as many inverse descents as the k-checkerboard containing it does, that is, at most $k-1$ of each. Hence either it has $k-1$ of both ascents or descents, or it has $k-1$ of one and $k-2$ of the other. In the first case it has k rows and k columns, while in the second it has k of one and $k-1$ of the other. We first consider the case where π has k rows and k columns. By Lemma 4.2, if an element of $[m]$ lies in the ith row and jth column of π, the corresponding element of $[n]$ must also lie in the ith row and jth column of the k-checkerboard. This uniquely determines the position of each element in $[n]$ because any row and any column intersect in at most one element.

Now we consider the case when π has k rows and $k-1$ columns. Applying Lemma 4.3 again, if an element $a \in [m]$ lies in the ith row and jth column of π, the corresponding element a' of $[n]$ lies in the ith row and either the jth or $(j+1)$st column of the k-checkerboard. But the uth row and vth column of the checkerboard only intersect when u and v have the same parity. Because exactly one of $i+j$ and $i+j+1$ is even, this is enough to determine the location of any given element. Hence π is uniquely contained.

The formulation in terms of S then follows directly from Lemma 4.3. $\square$

We now deduce an upper bound on the size of the set of subsets that do not yield unique patterns. Instead of just working with k-checkerboard permutations, which can only have lengths that are numbers of the form $N = \lceil \frac{k^2}{2} \rceil$, we will work with any subpermutation σ of the k-checkerboard, represented by a subset T of $[N]$ that omits at most one element from each row and each column of the k-checkerboard. The advantage of this is that we can find such permutations σ of any possible length n, allowing us to prove our lower bound for all n, not just for n of the special form $\lceil \frac{k^2}{2} \rceil$.

Any subset of $[n]$ that is redundant for σ corresponds to a subset S of T that is redundant for the k-checkerboard. By Proposition 4.4, S must truncate or avoid at least two among the rows and columns of the k-checkerboard. The idea behind counting these is very simple: if S avoids or truncates two rows or columns of σ, S must not contain any members from at least one of a family of $O(n^2)$ subsets of $[n]$, each of which has $k = \sqrt{2n} + O(1)$ elements. For each of these $O(n^2)$ subsets, there are $O(2^{n-\sqrt{2n}})$ choices of S that do not intersect it, combining for a total of $O(n^2 2^{n-\sqrt{2n}})$ redundant subsets S. We leave the details of this counting to the following section: for now we merely state the lemma we need.

Lemma 4.5. *Let T be a subset of $[N]$ that omits at most one element from each row and each column of the board, and let $n = |T|$. The number of subsets $S \subseteq T$ of $[N]$ that have at least two of the rows and columns of the k-checkerboard truncated or vanishing is $O(n^2 2^{n-\sqrt{2n}})$.*

The subsets just counted are exactly those that fail the criterion of Lemma 4.4, so we can now bound the number of redundant subsets.

Corollary 4.6. *Let σ be a subpermutation of the k-checkerboard represented by a subset T of $[N]$ that omits at most one entry from each row and each column* *of the k-checkerboard, and let $n = |T|$ be the length of σ. Then*

$$r(\sigma) \leq O(n^2 2^{n-\sqrt{2n}}).$$

Proof. By Proposition 4.4, any subset of $[N]$ that is redundant for the k-checkerboard must truncate or avoid at least two of the rows and columns of the checkerboard. As seen above, the redundant subsets of $[n]$ for σ must all correspond to subsets of such type that are also contained in T, so we may apply Lemma 4.5 to deduce $r(\sigma) \leq O(n^2 2^{n-\sqrt{2n}})$. □

Theorem 4.7. *Let $\sigma \in S_n$ be a subpermutation of the k-checkerboard represented by a subset T satisfying the conditions of Lemma 4.5. Then $f(\sigma) \geq 2^n - O(n^2 2^{n-\sqrt{2n}})$.*

Proof. By Corollary 4.6 $r(\sigma) \leq O(n^2 2^{n-\sqrt{2n}})$. Hence $u(\sigma) \geq 2^n - O(n^2 2^{n-\sqrt{2n}})$, and also $f(\sigma) \geq 2^n - O(n^2 2^{n-\sqrt{2n}})$. □

From this, Theorem 4.1 follows directly.

Proof of Theorem 4.1. For each n, we must find a $\sigma_n \in S_n$ such that $f(\sigma_n) \geq 2^n - O(n^2 2^{n-\sqrt{2n}})$. Let k be the unique natural number with $(k-1)^2 < 2n \leq k^2$, and let $N = \lceil \frac{k^2}{2} \rceil$ be the size of the k-checkerboard, so $N - n \leq k$. Now, the k-checkerboard has a diagonal with k elements, no two in the same row or same column. Deleting $N - n$ of those k elements yields a subset T_n of $[n]$ satisfying the condition of Lemma 4.5. Theorem 4.7 then applies to the corresponding permutation σ_n, as desired. □

This lower bound is remarkably close to the upper bound $2^n - \Theta(n 2^{n-\sqrt{2n}})$ found in Section 6, but it is possible that one might be able to find an improved lower bound for $\mathrm{pat}(n)$ either by closer scrutiny of the patterns analyzed above or by other means.

5 Proof of Lemma 4.5

We complete the proof of our lower bound by proving Lemma 4.5, which bounds the number of subsets of the tilted checkerboard that satisfy our necessary condition for being redundant.

Proof of Lemma 4.5. We consider the two essentially different ways that a subset S of $[N]$ ($N = \lceil \frac{k^2}{2} \rceil \geq n = |T|$) contained in T can truncate or avoid two among the rows and the columns of the tilted checkerboard: either it can avoid or truncate two rows or two columns, or it can have one of each. Before doing so, however, we observe that if S either avoids or truncates the ith column of the checkerboard, we can find a set A_i of at least $\alpha = \lceil \frac{k}{2} \rceil - 1$ consecutive entries from the ith and $(i+1)$st columns, none of which is in S. Indeed, if S avoids the ith column, the ith column will serve as our sequence. If S truncates the ith column, each of the α integers following the last element of S in the ith column must either be in the ith column or below the last element of the ith column: in either case, they can't be in S, because the last element of the ith column corresponds to an ascent in π. Analogously, if S truncates or avoids the ith row, we can find a set B_i of integers of size at least α that take on consecutive values, none of which is in S.

Two rows or two columns: Without loss of generality, we do the "two columns" case: suppose that S avoids or truncates the ith and jth columns of the checkerboard, say $i < j$. We claim that in this case S misses two disjoint sequences of consecutive elements of length at least α. We know that A_i and A_j are two such sequences, so they will suffice unless they intersect. This can only happen in the configuration where S truncates column i and avoids column $j = i + 1$.

In this case, we claim the 2α entries following the last element of S in the ith column cannot be in S. For those that are in column i come after the last element of S in that column; none can be in column $i+1=j$, which vanishes for S; and those in column $i+2$ must have values below that of the last element of S in the ith column.

We now count the possibilities. Suppose that the first run of missing elements starts at position r, and the second at position s (where we choose r and s to be minimal). Given r and s, we have specified that the α elements starting with r and the α starting with s are not in S, and all other elements of T may or may not be in S. However, some of these specifications may be redundant due to the fact that $S \subseteq T$. Each run, though, contains elements from at most two columns, so at most four of these can have been already specified. Since there are at most N^2 ways to choose the pair (r,s), and any such choice of a pair specifies that S must be disjoint from a given subset of at least $2\alpha - 4$ elements of T, there are at most

$$N^2 2^{n-2\alpha+4}$$

non-uniquely determined subsets of this form in T. Because $T \subseteq [N]$ misses at most one element of each of the k rows of the k-checkerboard, $N-k \le n = |T| \le N$. Because $N = \lceil \frac{k^2}{2} \rceil$, the number N is $O(n)$, and also $\alpha = \sqrt{\frac{n}{2}} - O(1)$. Hence the above expression is asymptotically $O(n^2 2^{n-\sqrt{2n}})$.

One of each: We use essentially the same argument as above, this time supposing that S truncates the ith row and jth column. This gives us two sequences of elements in consecutive positions or with consecutive values having length at least α. We shorten them if necessary so that both have length α. These sequences will not in general be disjoint, but we claim that their intersection will be small. In fact, the run of α consecutive missing positions intersects at most two rows of the k-checkerboard because those rows are each of length at most α. Similarly, the run of consecutive missing values intersects at most two columns of the k-checkerboard. By the checkered structure of the k-checkerboard, two consecutive rows and two consecutive columns intersect in only two elements. Hence we find in this case that there are at most $N^2 2^{n-2\alpha+6}$ possibilities for S, which is again $O(n^2 2^{n-\sqrt{2n}})$. $\square$

6 Proof of the upper bound

In a sense, the underlying idea behind our upper bound $\mathrm{pat}(n) \leq 2^n - O(n2^{n-\sqrt{2n}})$ is simpler than that for the lower bound. We will continue to focus on the redundancy. One common way in which a pattern is repeated in a permutation is by two subsets that differ in only one element. In [4], Coleman observed that two entries of a permutation are most often "interchangeable" in this way when the sum of the difference of their positions and that of their values is small. Our argument will quantify Coleman's insight, which will let us view the problem in terms of sphere packing with respect to the standard taxicab metric on $\mathbb{R}^2$.

We start by considering the permutation σ from the following geometric viewpoint. In previous sections, we have used the graph of the permutation as a visual aid to illustrate the argument, but we will now use its geometry directly.

Definition. The *graph* of a permutation σ is the subset of $[n]^2$ given by $\mathcal{G}(\sigma) = \{(i, \sigma(i)) \mid i \in [n]\}$.

Another geometric concept that will be useful is the rectangle with opposite corners at two points of $\mathcal{G}(\sigma)$.

Definition. For $i, j \in [n]$, the *rectangle* $R_\sigma(i, j)$ is the rectangle in $\mathbb{R}^2$ with sides parallel to the axes and opposite corners at the points $(i, \sigma(i))$ and $(j, \sigma(j))$.

This geometric viewpoint and Coleman's observation motivate the following definition.

Definition. The *taxicab distance* $d_\sigma(i, j)$ between the ith and jth entries of a permutation σ is given by

$$d_\sigma(i,j) = |i-j| + |\sigma(i) - \sigma(j)|.$$

(We will drop the subscript σ when the permutation in question is clear.)

The following definitions and lemma will be our main tools in formalizing and quantifying the amount of redundancy observed by Coleman, which we will call *swap-redundancy*.

Definition. The ith and jth entries ($i \neq j$) of a permutation σ are said to be *interchangeable* for a subset S of $[n] - \{i,j\}$ if the subsets $S \cup \{i\}$ and $S \cup \{j\}$ represent the same pattern π in σ in such a way that the ith and jth entries of σ correspond to the same element of π.

Definition. For $i, j \in [n]$ and a permutation $\sigma \in S_n$, the set $S_\sigma(i,j)$ is defined by

$$S_\sigma(i,j) = \{a \in [n] \mid a \notin [i,j] \text{ and } \sigma(a) \notin [\sigma(i), \sigma(j)]\}.$$

where $[i,j]$ denotes the closed interval with endpoints i and j.

Again, we will drop the σ when the permutation referred to is clear.

Lemma 6.1. *The following are equivalent for a permutation $\sigma \in S_n$, integers $i, j \in [n]$ and a set $S \subseteq [n] - \{i,j\}$:*

(i) the ith and jth entries of σ are interchangeable for S.

(ii) there is no $a \in S$ such that either a lies between i and j or $\sigma(a)$ lies between $\sigma(i)$ and $\sigma(j)$.

(iii) $S \subseteq S_\sigma(i,j)$.

Furthermore, there are $2^{|S_\sigma(i,j)|} \geq 2^{n-d_\sigma(i,j)}$ such sets S. Equality holds when $R_\sigma(i,j)$ has no elements of the graph of σ in its interior.

Proof. The equivalence of (ii) and (iii) is true by definition, so we need only show that (i) $\Leftrightarrow$ (ii).

(i) $\Rightarrow$ (ii): If the ith and jth entries of σ are interchangeable, they have the same position in π; this means that there must be the same number of elements of S before and after them. As a result, there can be no element of S positioned between i and j, and analogously there can be no element of S with values between $\sigma(i)$ and $\sigma(j)$.

(ii) $\Rightarrow$ (i): The condition (ii) implies that the map from $S \cup \{i\}$ to $S \cup \{j\}$ fixing each element of S and sending i to j is order-preserving with respect to both the standard ordering on $[n]$ and with respect to the ordering induced by σ. Hence the permutations represented are identical.

Condition (iii) implies that the number of such sets is $2^{|S_\sigma(i,j)|}$. It remains to show that $|S_\sigma(i,j)| \geq n - d_\sigma(i,j)$. Note that there are $|i-j|-1$ elements of $[n]-\{i,j\}$ that lie in $[i,j]$, and $|\sigma(i)-\sigma(j)|-1$ possible values for $a \in [n]-\{i,j\}$ with $\sigma(a) \in [\sigma(i),\sigma(j)]$. All other elements of $[n]-\{i,j\}$ are in $S_\sigma(i,j)$, so $S_\sigma(i,j) \geq (n-2)-(|i-j|-1)-(|\sigma(i)-\sigma(j)|-1)$ with equality if and only if no elements of $[n]$ are double-counted, and the result follows. $\square$

This lemma yields a family of redundant subsets of $[n]$ for any pair (i,j) of indices, as seen in the following corollary.

Corollary 6.2. *For an ordered pair of indices (i,j), the family*

$$\mathcal{F}_{i,j} = \{S \cup \{i\} \mid S \subseteq S_\sigma(i,j)\}$$

consists of $2^{|S_\sigma(i,j)|} \geq 2^{n-d_\sigma(i,j)}$ distinct subsets of $[n]$, all of which are redundant.

This immediately tells us that there are at least $2^{n-\min_{i,j} d_\sigma(i,j)}$ redundant subsets. By taxicab metric packing arguments such as the ones we will soon use, it is easy to show that in an n-by-n square grid of points, given any n points, some two of them must be distance at most $\sqrt{2n}+3$ apart in the taxicab metric. The previous two sentences immediately give us a bound of the form $\mathrm{red}(n) \geq \Theta(2^{n-\sqrt{2n}})$. However, this bound can be improved, because we can show that there are $\Theta(n)$ pairs of points of $\mathcal{G}$ that are distance at most $\sqrt{2n}+3$ apart, each of which yields a family of redundant subsets of size $\Theta(2^{n-\sqrt{2n}})$. In order to finish up this argument, we must show that these families of subsets have small overlap.

Any pair (i,j) of indices creates an amount of redundancy roughly proportional to $2^{n-d_\sigma(i,j)}$, so we will focus on the pairs (i,j) where $d_\sigma(i,j)$ is small. For our purposes, "small" will mean at most $\ell+3$, where $\ell=\sqrt{2n}$. This motivates our next definition. Let $\mathcal{P}_\sigma$ be the set of all ordered pairs (i,j) of elements of $[n]$ with $d(i,j) < \ell+3$.

Definition. The *swap-redundancy coefficient* $K(\sigma)$ of a permutation $\sigma \in S_n$ is the sum $K(\sigma) = \sum_{(i,j)\in\mathcal{P}_\sigma} 2^{-d(i,j)}$.

Because we intend to use $K(\sigma)$ to estimate $r(\sigma)$, we estimate $K(\sigma)$ first. We bound $K(\sigma)$ by considering it as a problem of spacing points apart in the taxicab metric on the square. Namely, we construct taxicab balls around the points of the graph $\mathcal{G}(\sigma)$ so that their total area exceeds that of the square they cover by a certain amount, forcing us to have sufficiently many pairs sufficently close together. This gives the following bound.

Proposition 6.3. *For a permutation $\sigma \in S_n$,*

$$K(\sigma) \geq \Theta(n2^{-\ell}).$$

Proof. We have viewed $\mathcal{G}_\sigma$, the graph of the permutation, as a subset of $[n]^2$, but it is also contained in $\mathbb{R}^2$. For each $i \in [n]$, construct a ball B_i in the taxicab metric on $\mathbb{R}^2$ centered at the point $(i, \sigma(i))$ and of radius $\frac{\ell+3}{2}$. If two of these balls intersect, their centers must be a distance at most $\ell + 3$ apart. Provided that the center of a ball is neither in one of the first or last $\lfloor \frac{\ell}{2} \rfloor + 1$ rows nor in one of the first or last $\lfloor \frac{\ell}{2} \rfloor + 2$ columns, the ball will be contained in the square $Q = [1/2, (n+1)/2]^2$. There are at least $n - 2\ell - 6$ such balls; we now consider how they cover Q.

Each ball is a square diamond with diagonal of length $\ell+3$, and area $\frac{(\ell+3)^2}{2} = n + 3\ell + \frac{9}{2}$, giving a total area of at least $(n - 2\ell - 6)(n + 3\ell + \frac{9}{2}) = n^2 + \Theta(n\ell)$. Because the area of Q is only n^2, the sum of the areas of overlap between pairs of balls must be at least $\Theta(n\ell)$.

The next step of our analysis takes us from the size of the overlap to an estimate of $K(\sigma)$. Take any pair of points of $\mathcal{G}(\sigma)$, say $\{(i, \sigma(i)), (j, \sigma(j))\}$; without loss of generality, assume that $i < j$ and $\sigma(i) < \sigma(j)$. Then if the point (x, y) is in $B_i \cap B_j$, by the triangle inequality for absolute value $|x+y-i-\sigma(i)| \leq \frac{\ell+3}{2}$. Similarly $|x + y - j - \sigma(j)| \leq \frac{\ell+3}{2}$, and combining the two results in the chain of inequalities yields $j + \sigma(j) - \frac{\ell+3}{2} \leq x + y \leq i + \sigma(i) + \frac{\ell+3}{2}$. By the same method we can also get $i - \sigma(i) - \frac{\ell+3}{2} \leq x - y \leq i - \sigma(i) + \frac{\ell+3}{2}$. This system of inequalities shows that (x, y) must lie within a rectangle with side lengths $j + \sigma(j) - i - \sigma(i) + \ell + 3 = \ell + 3 - d(i, j)$ and $\ell + 3$. That is, it has area at most $(\ell + 3 - d(i, j))(\ell + 3)$ when this is positive, and otherwise when $d(i, j) \geq \ell + 3$ the overlap is a line or is empty.

We have just bounded the contribution of a particular pair of points that are close together. We now add everything up. We know that the total overlap is at least $\Theta(n\ell)$, so summing up the above yields

$$(\ell + 3) \sum_{(i,j) \in \mathcal{P}_\sigma} (\ell + 3 - d(i, j)) \geq \Theta(n\ell).$$

Because $\ell + 3 - d$ is positive, we can use the inequality $\ell + 3 - d \leq 2^{\ell+2-d}$

to obtain

$$(\ell+3)2^{\ell+2}K(\sigma) = (\ell+3)\sum_{(i,j)\in\mathcal{P}_\sigma} 2^{\ell+2-d(i,j)}$$
$$\geq (\ell+3)\sum_{(i,j)\in\mathcal{P}_\sigma} (\ell+3-d(i,j)) \geq \Theta(n\ell).$$

Dividing through, we find that $K(\sigma) \geq \Theta(n2^{-\ell})$. □

We now wish to get an $O(K(\sigma))$ lower bound on $r(\sigma)$. This need not be true in general, but we will prove it in a "worst case" scenario, when no two points in $\mathcal{G}_\sigma$ are closer than $\ell - \log_2 n$. The non-"worst case" scenario is easily disposed of, because a pair of points that are closer than $\ell - \log_2 n$ automatically gives a family of $O(2^{n-\ell+\log_2 n}) = O(n2^{n-\ell})$ redundant subsets by Corollary 6.2. The advantage of considering such a scenario is that we can get more precise bounds on the size of the overlap between any two families, as we will see in Section 7.

We will need to use the fact that the amount of double-counting caused by overlapping families of redundant subsets generated by different pairs in $\mathcal{P}^*_\sigma$ is small. This follows from the following technical lemma, whose proof we postpone to the following section. We recall that we have defined

$$\mathcal{F}_{i,j} = \{S \cup \{i\} \mid S \subseteq S_\sigma(i,j).$$

Lemma 6.4. *There is a constant c such that the following holds: For all permutations $\sigma \in S_n$ such that $d_\sigma(i,j) \geq \ell - \log_2 n$ for all $(i,j) \in \mathcal{P}_\sigma$, there exists a subset $\mathcal{P}^*_\sigma \subseteq \mathcal{P}_\sigma$ satisfying: (i) the total contribution $K^*(\sigma)$ to $K(\sigma)$ from pairs in $\mathcal{P}^*_\sigma$ is at least $\frac{1}{20}K(\sigma)$, and (ii) for all pairs of ordered pairs $(i_1,j_1),(i_2,j_2) \in \mathcal{P}^*_\sigma$,*

$$|\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}| \leq c2^{n-2\ell}.$$

From this we can deduce the result we want:

Theorem 6.5. *We have the following bounds:*

$$\mathrm{red}(n) \geq \Theta(n2^{n-\sqrt{2n}}) \tag{8}$$

$$\mathrm{pat}(n) \leq 2^n - \Theta(n2^{n-\sqrt{2n}}) \tag{9}$$

for any positive integer n.

Proof. For the first inequality, we need to show that for any permutation $\sigma \in S_n$, $r(\sigma) \geq \Theta(n2^{n-\sqrt{2n}})$. We first take care of the case when there exist $i, j \in [n]$ with $d_\sigma(i,j) < \ell - \log_2 n$. Then

$$r(\sigma) \geq |\mathcal{F}_{i,j}| \geq 2^{n-d_\sigma(i,j)} > 2^{n-\ell+\log_2 n} = \Theta(n2^{n-\sqrt{2n}}).$$

Suppose instead that this is not the case: then the condition of Lemma 6.4 holds, and we can proceed as follows.

We will estimate the size of the set $\mathcal{F}^* = \left|\bigcup_{(i,j)\in\mathcal{P}_\sigma^*} \mathcal{F}_{i,j}\right|$, where $\mathcal{P}_\sigma^*$ is as in Lemma 6.4. Because this set contains only redundant subsets of $[n]$, this will give a lower bound for $r(\sigma)$. We can use Bonferroni's inequality to give the following lower bound on $\mathcal{F}$:

$$|\mathcal{F}^*| = \left|\bigcup_{(i,j)\in\mathcal{P}_\sigma^*} \mathcal{F}_{i,j}\right| \geq \sum_{(i,j)\in\mathcal{P}_\sigma^*} |\mathcal{F}_{i,j}| - \sum_{\{(i_1,j_1),(i_2,j_2)\}\subseteq\mathcal{P}_\sigma^*} |\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}|. \tag{10}$$

We apply Corollary 6.2 and sum over all pairs to bound the first term of (10) by

$$\sum_{(i,j)\in\mathcal{P}_\sigma^*} |\mathcal{F}_{i,j}| \geq \sum_{(i,j)\in\mathcal{P}_\sigma^*} 2^{n-d(i,j)} = 2^n K^*(\sigma) \geq \Theta(n2^{n-\ell}). \tag{11}$$

We can use Proposition 6.4 to bound each of the terms in the other sum:

$$\sum_{\{(i_1,j_1),(i_2,j_2)\}\subseteq\mathcal{P}_\sigma^*} |\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}| \leq \sum_{\{(i_1,j_1),(i_2,j_2)\}\subseteq\mathcal{P}_\sigma^*} c2^{n-2\ell} \tag{12}$$

which is at most $O(n^4 2^{n-2\ell})$ because the number of terms in the sum is bounded above by the total number of possible ways of choosing two pairs (i_1, j_1) and (i_2, j_2) of indices in $[n]$, which is $O(n^4)$.

Subtracting (12) from (11) yields

$$\begin{aligned} |\mathcal{F}^*| &= \sum_{(i,j)\in\mathcal{P}_\sigma^*} |\mathcal{F}_{i,j}| - \sum_{\{(i_1,j_1),(i_2,j_2)\}\subseteq\mathcal{P}_\sigma^*} |\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}| \\ &\geq \Theta(n2^{n-\ell}) - O(n^4 2^{n-2\ell}) = \Theta(n2^{n-\ell}). \end{aligned} \tag{13}$$

We conclude from the above that we can always find at least $\Theta(n2^{n-\ell})$ redundant subsets for any given pattern. Hence $\mathrm{red}(n) \geq \Theta(n2^{n-\ell})$, as desired.

Equation 9 then follows by Proposition 2.1. $\square$

7 Proof of Lemma 6.4

We first need the following technical results, which allow us to use the power of our "worst-case" assumption that no pair of points is closer than $\ell - \log_2 n$. The following lemma tells us that under this assumption, the equality case of Corollary 6.2 holds for all pairs in $\mathcal{P}_\sigma$.

Lemma 7.1. *Suppose $\sigma \in S_n$ (n sufficiently large) is a permutation such that for any $i, j \in [n]$, $d_\sigma(i,j) \geq \ell - \log_2 n$. If $i, j \in [n]$ with $d_\sigma(i,j) \leq \ell + 3$, then the rectangle $R_\sigma(i,j)$ contains no points of $\mathcal{G}(\sigma)$ in its interior.*

Proof. Suppose that $R_\sigma(i,j)$ contained some point $(a, \sigma(a))$. Then we would have $d(i,j) = d(i,a) + d(a,j)$. But the left hand side is at most $\ell + 3$, while the right hand side is at least $2(\ell - \log_2 n)$, which cannot hold for sufficiently large n. $\square$

For the remainder of this section, we will assume that n is sufficiently large that Lemma 7.1 holds, which we can do because we are proving an asymptotic result.

Instead of using all the pairs in $\mathcal{P}_\sigma$, we restrict to a smaller subset. Without loss of generality, at least $\frac{1}{4}$ of the total contribution to $K(\sigma)$ comes from pairs (i,j) with $i < j$ and $\sigma(i) < \sigma(j)$: let $\mathcal{P}'_\sigma$ denote the subset of such pairs and $K'(\sigma)$ be the contribution to $K(\sigma)$ from pairs in $\mathcal{P}'_\sigma$. We claim that we can shrink $\mathcal{P}'_\sigma$ further to a set $\mathcal{P}^*_\sigma$ such that $\mathcal{P}'_\sigma$ contains no two distinct pairs (i_1, j_1) and (i_2, j_2) with $i_1 = i_2$ or $j_1 = j_2$, and such that the contribution $K^*(\sigma)$ to $K(\sigma)$ from pairs in $\mathcal{P}^*_\sigma$ is at least $\frac{1}{5}K'(\sigma) \geq \frac{1}{20}K(\sigma)$. To do this, we need the following lemma.

Lemma 7.2. *Suppose $\sigma \in S_n$ (n sufficiently large) is a permutation such that for any $i, j \in [n]$, $d_\sigma(i,j) \geq \ell - \log_2 n$. For any $i \in [n]$, there are at most 3 values of j for which $(i,j) \in \mathcal{P}'_\sigma$. Likewise, for any $j \in [n]$, there are at most 3 values of i for which (i,j) in $\mathcal{P}'_\sigma$.*

Proof. We argue by contradiction. If not, suppose that there exists $j_1 < j_2 < j_3 < j_4$ such that $(i, j_r) \in \mathcal{P}'_\sigma$, for each $r = 1,2,3,4$. The definition of $\mathcal{P}'_\sigma$ implies that $i < j_r$ and $\sigma(i) < \sigma(j_r)$ for each r. If for some pair r, s with $r < s$, it were the case that $\sigma(j_r) < \sigma(j_s)$, it would follow that $i < j_r < j_s$ and $\sigma(i) < \sigma(j_r) < \sigma(j_s)$, contradicting Lemma 7.1. Hence $\sigma(j_1) > \sigma(j_2) > \sigma(j_3) > \sigma(j_4)$. We also know that $0 < j_r - i < d(j_r, i) < \ell + 3$ and $0 < \sigma(j_r) - \sigma(i) < \ell + 3$. It follows from the previous two inequalities with $r = 1$ and $r = 4$ that

$$(j_4 - j_1) + (\sigma(j_1) - \sigma(j_4)) < 2\ell + 6$$

but also

$$\begin{aligned}(j_4 - j_1) + (\sigma(j_1) - \sigma(j_4)) =& (j_4 - j_3) + (\sigma(j_3) - \sigma(j_4)) \\ &+ (j_3 - j_2) + (\sigma(j_2) - \sigma(j_3)) \\ &+ (j_2 - j_1) + (\sigma(j_1) - \sigma(j_2)) \\ =& d(j_4, j_3) + d(j_3, j_2) + d(j_2, j_1) \geq 3\ell - 3\log_2 n\end{aligned}$$

by assumption. Combining the above yields $2\ell + 6 > 3\ell - 3\log_2 n$, which is impossible for n sufficiently large.

The second half of the result follows analogously. □

We return to proving our claim that we can pick $\mathcal{P}'_\sigma$ as specified above. We do this by the greedy algorithm. First pick $(i_1, j_1) \in \mathcal{P}'_\sigma$ such that $d_\sigma(i_1, j_1)$ is minimal. By Lemma 7.2, $\mathcal{P}'_\sigma$ can contain at most 2 other pairs of the form (i_1, j) with $j \neq j_1$, and only 2 other pairs of the form (i, j_1) with $i \neq i_1$. Discard all such pairs, and repeat the process: letting (i_2, j_2) be one of the remaining pairs with $d_\sigma(i_2, j_2)$ minimal, discard the at most 4 total pairs that share a left or right endpoint with (i_2, j_2). We repeat until we run out of pairs to pick, and we let $\mathcal{P}^*_\sigma = \{(i_1, j_1), (i_2, j_2), \dots\}$ be the set of all chosen pairs. At each stage the pair we choose contributes the maximum amount possible to $K'(\sigma)$, and we then discard at most 4 pairs that each contribute the same or a smaller amount. Hence the total contribution $K^*(\sigma)$ to $K(\sigma)$ from pairs in $\mathcal{P}^*_\sigma$ is at least $\frac{1}{5}K'(\sigma)$, which is in turn at least $\frac{1}{20}K(\sigma)$.

We are now ready to prove Lemma 6.4.

Proof of Lemma 6.4. Let $\mathcal{P}^*_\sigma$ be as constructed above: we already know that $K^*(\sigma) \geq \frac{1}{20}K(\sigma)$. It remains to prove (ii).

The general member of the set $\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}$ is of the form $S \cup \{i_1, i_2\}$, and represents the same pattern as $S \cup \{j_1, i_2\}$ and $S \cup \{i_1, j_2\}$. By Lemma 6.1, the interchangeability condition implies that $S \cup \{i_2\} \subseteq S_\sigma(i_1, j_1)$ and $S \cup \{i_1\} \subseteq S_\sigma(i_2, j_2)$.

For one thing, this means that $\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}$ will be empty unless it is the case both that $i_2 \in S_\sigma(i_1, j_1)$ and $i_1 \in S_\sigma(i_2, j_2)$. These conditions imply that i_2 lies outside the interval $[i_1, j_1]$ and i_1 lies outside the interval $[i_2, j_2]$. Hence the intervals $[i_1, j_1]$ and $[i_2, j_2]$ are disjoint. Likewise, the intervals $[\sigma(i_1), \sigma(j_1)]$ and $[\sigma(i_2), \sigma(j_2)]$ must also be disjoint for the overlap to be nonempty: hence we can assume that we are in this case. The rest of our independence condition tells us that the sets S that work are exactly those contained in $S_\sigma(i_1, j_1) \cap S_\sigma(i_2, j_2)$, so

$$|\mathcal{F}_{i_1,j_1} \cap \mathcal{F}_{i_2,j_2}| = 2^{|S_\sigma(i_1,j_1) \cap S_\sigma(i_2,j_2)|}. \tag{14}$$

We now show that under these conditions, $|S_\sigma(i_1, j_1) \cap S_\sigma(i_2, j_2)| \leq n-2\ell+O(1)$ and so there are only $O(2^{n-2\ell})$ possibilities for S.

We define two subsets of the plane, which we call *crosses*, as follows:

$$\begin{aligned} C_1 &= \{x, y \in \mathbb{R}^2 \mid i_1 < x < j_1 \text{ or } \sigma(i_1) < y < \sigma(j_1)\}, \\ C_2 &= \{x, y \in \mathbb{R}^2 \mid i_2 < x < j_2 \text{ or } \sigma(i_2) < y < \sigma(j_2)\}. \end{aligned}$$

Then the members of $S_\sigma(i_1, j_1)$ are exactly the members of $\mathcal{G}(\sigma)$ outside C_1, and the members of $S_\sigma(i_2, j_2)$ are exactly those outside C_2. We now show that $C_1 \cap C_2$ contains at most 18 points of $\mathcal{G}(\sigma)$. This intersection is composed of two rectangles, $X = [i_1, j_1] \times [\sigma(i_2), \sigma(j_2)]$ and $Y = [i_2, j_2] \times [\sigma(i_1), \sigma(j_1)]$. The sum of the length and width of X (that is, its taxicab diameter) is $j_1 - i_1 + \sigma(j_2) - \sigma(i_2) < d(i_1, j_1) + d(i_2, j_2) < 2(\ell + 3)$, so if we partition X into 9 subrectangles of one-third the length and width, each subrectangle will have diameter $< \frac{2}{3}(\ell + 3)$. For n large enough, $\frac{2}{3}(\ell + 3) < \ell - \log_2 n$, so each subrectangle can contain at most one point of $\mathcal{G}(\sigma)$. Hence X, and likewise Y, can each contain at most 9 points of $\mathcal{G}(\sigma)$, and the two can contain at most 18 between them.

By the equality case of Lemma 6.1 and Lemma 7.1, $S_\sigma(i_1, j_1)$ contains exactly $n - (d(i_1, j_1) - 2)$ points, which are exactly the points of $\mathcal{G}(\sigma)$ outside C_1. Likewise, $S_\sigma(i_2, j_2)$ contains exactly the $d(i_2, j_2) - 2$ points of $\mathcal{G}(\sigma)$ that are outside C_2. Because $C_1 \cap C_2$ contains at most 18 points of $\mathcal{G}(\sigma)$, we can apply inclusion-exclusion to conclude that

$$\begin{aligned} |S_\sigma(i_1, j_1) \cap S_\sigma(i_2, j_2)| &= n - |\mathcal{G}(\sigma) \cap C_1| - |\mathcal{G}(\sigma) \cap C_2| + |\mathcal{G}(\sigma) \cap C_1 \cap C_2| \\ &\leq n - (d(i_1, j_1) - 2) - (d(i_2, j_2) - 2) + 18 \\ &\leq n - 2(\ell + 3) + 22 = n - 2\ell + O(1). \end{aligned} \tag{15}$$

Combining this with (14), we conclude that

$$|\mathcal{F}_{i_1, j_1} \cap \mathcal{F}_{i_2, j_2}| = O(2^{n-2\ell}),$$

as desired. □

8 Conclusion and further directions

Our arguments have sandwiched $\operatorname{pat}(n)$ between the two bounds

$$2^n - O(n^2 2^{n-\sqrt{2n}}) \leq \operatorname{pat}(n) \leq 2^n - \Theta(n 2^{n-\sqrt{2n}}).$$

These bounds are quite close, but it is still possible that one or the other could be improved to give the exact rate of growth of $2^n - \operatorname{pat}(n)$, or equivalently, for $\operatorname{red}(n)$. Besides this, there are other questions that can be asked about $\operatorname{pat}(n)$. It has been observed ([3]) based on experimental data of Micah Coleman that for $n \geq 10$, the ratio $\frac{\operatorname{pat}(n)}{2^n}$ is monotone increasing: the bounds given here make such behavior plausible, but it seems impossible to prove it by the methods given here. An inductive approach may be more fruitful in this direction.

Another direction of generalization would be to patterns in words or permutations of multisets. Constructions like those used for the lower bound might again be close to optimal. The approach used for our upper bound breaks down when multiple entries can have the same value, but there may be some way to fix it.

In the field of superpatterns, on the other hand, the bounds are still wide open. Our upper bound of $L_k \leq \frac{k(k+1)}{2}$, while a significant improvement on the previous result of $L_k \leq (\frac{2}{3} + o(1))(k^2)$, is still some way off from the lower bound $L_k \geq \left(\frac{k}{e}\right)^2$, and it is not apparent how one might improve either bound. Although this paper makes substantial steps towards understanding the asymptotics of pattern packing, there is still much new territory to be explored.

9 Acknowledgments

This research was done at the University of Minnesota Duluth with the support of the National Science Foundation (grant number DMS-00447070) and the National Security Agency (grant number H98230-06-1-0013). I wish to thank David Arthur, Reid Barton, Philip Matchett Wood, Melanie Wood, and the anonymous referees for their helpful comments. I also wish to thank Joe Gallian for suggesting this problem and for his encouragement.

References

[1] M. H. Albert, M. Coleman, R. Flynn, I. Leader, Permutations containing many patterns, (2006), arxiv.org/pdf/math.CO/0609771

[2] R. Arratia, On the Stanley-Wilf conjecture for the number of permutations avoiding a given pattern, *Electron. J. Combin*, **6**(1) (1999), #N1.

[3] M. Bóna and M. Coleman, personal correspondence.

[4] M. Coleman, An answer to a question by Wilf on packing distinct patterns in a permutation, *Electron. J. Combin.*, **11**(1) (2004), #N8.

[5] H. Eriksson, K. Eriksson, S. Linusson, J. Wästlund, Dense packing of patterns in a permutation, Proceedings of the 15th Conference on Formal Power Series and Algebraic Combinatorics, Melbourne, Australia, 8-12 July 2002, paper #26.

[6] H. Eriksson, K. Eriksson, S. Linusson, J. Wästlund, Dense packing of patterns in a permutation, *Ann. Comb.*, accepted for publication.

[7] A. Flaxman, A. W. Harrow, G. B. Sorkin, Strings with Maximally Many Distinct Subsequences and Substrings, *Electron. J. Combin.* **11**(1) (2004), #R8.

[8] H. S. Wilf, Problem 414, *Discrete Math.* 272 (2003), 303.