

COUNTING SMALL PERMUTATION PATTERNS

CHAIM EVEN-ZOHAR AND CALVIN LENG

ABSTRACT. A sample of n generic points in the xy -plane defines a permutation that relates their ranks along the two axes. Every subset of k points similarly defines a pattern, which occurs in that permutation. The number of occurrences of small patterns in a large permutation arises in many areas, including nonparametric statistics. It is therefore desirable to count them more efficiently than the straightforward $\tilde{O}(n^k)$ time algorithm.

This work proposes new algorithms for counting patterns. We show that all patterns of order 2 and 3, as well as eight patterns of order 4, can be counted in nearly linear time. To that end, we develop an algebraic framework that we call *corner tree formulas*. Our approach generalizes the existing methods and allows a systematic study of their scope.

Using the machinery of corner trees, we find twenty-three independent linear combinations of order-4 patterns, that can be computed in time $\tilde{O}(n)$. We also describe an algorithm that counts another 4-pattern, and hence all 4-patterns, in time $\tilde{O}(n^{3/2})$.

As a practical application, we provide a nearly linear time computation of a statistic by Yanagimoto (1970), Bergsma and Dassios (2010). This statistic yields a natural and strongly consistent variant of Hoeffding's test for independence of X and Y , given a random sample as above. This improves upon the so far most efficient $\tilde{O}(n^2)$ algorithm.

1. INTRODUCTION

In many situations it is interesting to investigate the *local* structure of a large combinatorial object. This may mean small subgraphs of a given large graph, short patterns in a long sequence, and so forth. Local views provide meaningful and efficient ways to analyze large combinatorial structures, both in theory and in practice. In the realm of big data, where data sets are too large to observe in full, local approaches are essentially inescapable.

Here we investigate the local structure of permutations. The restriction of a large permutation to some of its entries yields a *pattern*, a smaller permutation defined by the relative ordering of the observed entries. Namely, the order- k pattern σ occurs at positions $i_1 < i_2 < \dots < i_k$ of the order- n permutation π , if $\sigma(j) < \sigma(j')$ whenever $\pi(i_j) < \pi(i_{j'})$ for $j \neq j'$. For example, the pattern 132 occurs in $2\mathbf{3}6\mathbf{4}7\mathbf{5}1$ at the marked positions $\{2, 5, 6\}$. We throughout use the one line notation $\pi = \pi(1)\pi(2) \dots \pi(n)$ for permutations and their patterns.

We refer to the positions $i_1 < i_2 < \dots < i_k$ as an *occurrence* of $\sigma \in S_k$ in $\pi \in S_n$ and denote the number of occurrences of σ in π by $\#\sigma(\pi)$. For example, one can easily verify that $\#132(2364751) = 7$. The k -*profile* of $\pi \in S_n$ consists of the $k!$ numbers $\#\sigma(\pi)$ for all $\sigma \in S_k$. These numbers clearly sum up to $\binom{n}{k}$.

There is considerable combinatorial literature concerning the numbers $\#\sigma(\pi)$. For example, given σ , one seeks the maximum of $\#\sigma(\pi)$ over all $\pi \in S_n$ [Pri97,

AAH⁺02, Häs02, PS10, BH10, BHL⁺15, SS17]. It is also interesting to understand the distribution of $\#\sigma(\pi)$ for a randomly chosen $\pi \in S_n$ [Ful04, Bón07, Bón10, BH10, JNZ15, Hof17, Eve18]. Study of k -profiles leads to the notion of “permutons”, limits of sequences of permutations [HKM⁺13, GGKK15, KKRW19]. This theory includes the case of quasi-random permutations [Coo04, Coo06, KP13].

But the local viewpoint on permutations is of interest as well in a range of applications, most notably in nonparametric statistics of bivariate data. Consider a sequence of n independent paired samples (X_i, Y_i) drawn from a common continuous distribution (X, Y) on \mathbb{R}^2 . Various measures have been suggested to detect and quantify a relation between X and Y based on these samples. Often one prefers to rely only on the ranking of X_i and Y_i rather than on their numerical values. In this case it suffices to consider the permutation $\pi \in S_n$, defined by $\text{rank } Y_i = \pi(\text{rank } X_i)$, where the rank of X_1, \dots, X_n is their order preserving map to $\{1, \dots, n\}$.

Many nonparametric measures depend only on π 's local structure. Here are some classical examples. Kendall's τ correlation coefficient is based on the 2-profile: $\tau = [\#12(\pi) - \#21(\pi)] / \binom{n}{2}$ [Ken38]. Similarly Spearman's ρ is a function of the 3-profile of π [Spe04], and so is a rotation-invariant measure of correlation by Fisher and Lee [FL82, Jan84], and other correlation tests [Cro66].

Local features of π can even detect whether there is *any* relationship between X and Y , as demonstrated by Hoeffding's independence test [Hoe48]. This classical nonparametric test is based on the 5-profile of π , and so is a refined version with wider consistency [BKR61, Ros75, DW80]. A simpler variant that only relies on the 4-profile has been proposed by Bergsma and Dassios [Ber10, BD10, BD14], and goes back to a result of Yanagimoto [Yan70, DHS18]. A general family of independence tests expressible by pattern counts is suggested in [HHK⁺16].

Pattern counting also plays a role in *property testing* and *parameter testing* for permutations [HKMS11]. In these problems, one estimates features of a large permutation using a randomized algorithm that makes a small number of queries.

Applications to stack-sorting [Knu68] have inspired a significant amount of research around the case $\#\sigma(\pi) = 0$. In particular, there has been interest in the count of such σ -avoiding permutations of order n [SS85, MT04]. This line of research in Enumerative Combinatorics has introduced a wealth of generalizations to other permutation statistics, and connections to other branches of Mathematics [Ste10, Kit11, Bón12].

The extensive study of σ -avoidance has naturally led to the exploration of its algorithmic aspects. The most obvious question is *permutation pattern matching*: deciding whether a given pattern $\sigma \in S_k$ occurs anywhere in a given permutation $\pi \in S_n$. In this generality the problem is NP-complete [BBL98, BCKM15, CK16]. Straightforward exhaustive search solves the problem in time $\tilde{O}\left(\binom{n}{k} \binom{k}{2}\right)$, but faster algorithms have been found, improving the exponent by a constant factor [AR08, BL16, Koz19, BKM19]. Numerous works have achieved improved running times for patterns σ that are either short or have some structural properties [Sch61, CW92, Iba97, BBL98, AAAH01, YS05, GV09, BL12, BL13, BL16, GM14, HS14, JK17, HS18, Ber19]. Remarkably, this problem is “fixed-parameter tractable”. For $\sigma \in S_k$ and large n it can be solved in time $\tilde{O}(f(k) \cdot n)$ [GM14].

The question of *permutation pattern counting* seeks to determine $\#\sigma(\pi)$, the number of occurrences of a given pattern $\sigma \in S_k$ in a given permutation $\pi \in S_n$. This problem, which is $\#P$ -complete in general, has received less attention. It

has mostly been considered in papers on the decision problem, as it sometimes happens that a decision algorithm can be adapted to address as well the counting problem [BBL98, AAAH01, YS05, BKM19]. For example, a separable pattern σ can be counted in time $\tilde{O}(kn^6)$ [BBL98]. However, the linear time algorithm from [GM14] does not yield a counting method. A recent paper [BKM19] shows that if $\#\sigma(\pi)$ can be computed for all σ and π in time $f(k)n^{o(k/\log k)}$, then the exponential-time hypothesis fails to hold. For comparison, recall that the run time of the trivial algorithm is $n^{O(k)}$.

While the above hardness results virtually settle the counting problem for general $\sigma \in S_k$ and asymptotically in k , they leave open some questions on specific patterns and particular k . For instance, which patterns can be counted in linear time, and how? Since the decision algorithm of [GM14] does not help us here, new ideas are clearly needed.

From a practical standpoint, these questions are much relevant to the implementation of rank-based statistical tests. Spearman's ρ is computable in linear time by definition. Kendall's τ naively requires $\tilde{O}(n^2)$ time, but an algorithm by Knight computes it in $\tilde{O}(n)$ [Kni66]. By the same technique, Hoeffding's independence test is readily computable in $\tilde{O}(n)$, despite being stated and sometimes implemented with $\tilde{O}(n^2)$ formulas [Hoe48, HJD06, HWC13]. Several works have addressed the computation of the Bergsma–Dassios–Yanagimoto test, only bringing it down to $\tilde{O}(n^2)$ time [BD10, BD14, WDL16, HH16, WDM18]. It seems that these pattern-based statistics have only been treated ad hoc, by dedicated algorithms, without a unified or systematic approach.

This work proposes new approaches for permutation pattern counting. In Section 2 we define the key notion of a *corner tree*, a structure that lets us apply dynamic programming to the problem. This leads to a class of permutation statistics that have a *corner tree formula*. We present an algorithm that evaluates such statistics efficiently, proving the following theorem.

Theorem 1. *Let $f : S_n \rightarrow \mathbb{Z}$ be a permutation statistic, given as a corner tree formula. There exists an algorithm that computes $f(\pi)$ for a given permutation $\pi \in S_n$ in running time $\tilde{O}(n)$.*

Although corner trees can come in any size, they prove particularly useful for counting small patterns. In Section 3, we give corner tree formulas for all 3-patterns, which implies the following.

Corollary 2. *The number of occurrences of any pattern of size 3 in a given permutation $\pi \in S_n$ can be computed in $\tilde{O}(n)$ time.*

Similarly, we show that corner tree formulas exist for the following eight patterns of size 4.

Corollary 3. *The occurrences of any of the patterns 1234, 1243, 2134, 2143, 3412, 3421, 4312, 4321 in a given permutation $\pi \in S_n$ can be counted in time $\tilde{O}(n)$.*

The total count of these eight patterns is sufficient for computing the Bergsma–Dassios–Yanagimoto statistic. This yields the following practical consequence.

Corollary 4. *The Bergsma–Dassios–Yanagimoto independence test for continuous bivariate data can be computed in $\tilde{O}(n)$ time.*

In section 3 we provide a review of this increasingly popular statistical test. Corollary 4 is a significant improvement over the best previous method, which requires $\tilde{O}(n^2)$ time and space [HH16]. Indeed, a near linear time performance is crucial when dealing with large amounts of data. We thus expect our method to be beneficial in real world scenarios.

We derive explicit procedures for fast evaluation of this statistic. See Proposition 10. The paper is accompanied by a Python package that includes, among other things, a routine for the Bergsma–Dassios–Yanagimoto statistic’s computation [EL19].

Another important ingredient of our approach is treating the k -profile vector as a whole, rather than looking on each pattern separately. The permutation statistics that we study are linear combinations of pattern counts:

$$F(\pi) = \sum_{\sigma} f_{\sigma} \# \sigma(\pi)$$

Here the sum is over patterns, with finitely many nonzero $f_{\sigma} \in \mathbb{Q}$. The corner tree formulas that we define span a vector space of such combinations. This viewpoint provides much more information on the k -profile, as demonstrated in the following proposition with $k = 4$.

Proposition 5. *There exist twenty-three linearly independent combinations,*

$$F_i(\pi) = \sum_{\sigma \in S_4} f_{i\sigma} \# \sigma(\pi) \quad i \in \{1, 2, \dots, 23\}$$

that can be evaluated for a given $\pi \in S_n$ in time $\tilde{O}(n)$.

Since the 4-profile is a 24-dimensional vector, only one additional combination is needed to fully reveal it. In other words, one can efficiently compute this vector up to a single linear degree of freedom. All the remaining problems of counting 4-patterns are thus equivalent to each other.

Clearly, Proposition 5 was established with computer assistance. At the end of Section 3 we provide more details about the implementation.

Our code provides general-purpose routines for finding and using corner tree formulas, and it is scalable to larger patterns and combinations. For example, it yields a 100-dimensional space of linear functions of the 120-dimensional 5-profile, that are computable in $\tilde{O}(n)$ time.

In Section 4 we explore new counting methods beyond corner trees. We naturally focus on the missing piece of the 4-profile. An $\tilde{O}(n^2)$ time computation is straightforward from the existing methods. We present two algorithms that do somewhat better, as follows.

Theorem 6. *The occurrences of any 4-pattern in a given permutation $\pi \in S_n$ can be counted in $\tilde{O}(n^{5/3})$ time and $\tilde{O}(n)$ space.*

Theorem 7. *The occurrences of any 4-pattern in a given permutation $\pi \in S_n$ can be counted in $\tilde{O}(n^{3/2})$ time and space.*

Finally, in Section 5, we discuss the scope of our methods, and questions for future work.

Remark. Since our results reduce the degree of some polynomial run times, we are less concerned with logarithmic factors. Therefore, we throughout use the soft O notation, $\tilde{O}(f(n))$, which means $O(f(n) \log^c n)$ where c is some constant. In words, we sometimes write “nearly linear”, “nearly quadratic”, and so on.

Moreover, some algorithms in the literature, such as [GM14], already assume a computational model that operates on $(\log n)$ -bit numbers in $O(1)$ time. Thus, by ignoring logarithmic factors, our statements become more robust to different choices of the abstract machine.

In contrast, some results in this area are machine dependent. In the case of counting the pattern 12 , an improvement of order $\sqrt{\log n}$ was given specifically assuming a Word RAM model [CP10].

2. TREES

Fix a pattern $\sigma \in S_k$. Given a permutation $\pi \in S_n$ as input, our goal is to compute the pattern count $\#\sigma(\pi)$, also denoted $\#\sigma$ whenever no confusion arises. Simple $\tilde{O}(n)$ time algorithms are known for $\#12$, and extend to larger increasing patterns $\#12\dots k$. We first review these techniques, as they are relevant to our general approach.

The task of computing Kendall’s τ rank correlation coefficient is essentially, after sorting the samples, counting either increasing or decreasing pairs in a permutation.

$$\tau = \frac{\#12 - \#21}{\binom{n}{2}} = 2 \cdot \frac{\#12}{\binom{n}{2}} - 1$$

A naive implementation iterates through every two entries and compares their values, which yields an answer in $\tilde{O}(n^2)$ time.

Knight improved this n^2 to $n \log n$ with a divide-and-conquer approach, motivated by merge sort [Kni66]. While performing the usual merge sort on a permutation, treating it as a list to be ordered, we record some numbers that add up to $\#12$. Namely, when we merge two sub-lists and select an element from the second one, we accumulate the number of elements selected so far from the first one. Every occurrence of 12 contributes one to the result.

A different approach uses dynamic programming [CLRS09, ex 14.1-7]. We iterate once through π . At step x , we *insert* one at position $y = \pi(x)$ of an array, and we add its current *prefix sum* of size $y - 1$ to the result.

We briefly describe the *sum tree* data structure [SV82], that performs these two operations in logarithmic time, as it is simple and will serve us well later. Place the array entries at the leaves of a complete binary tree of depth $\lceil \log n \rceil$, from left to right. Insertion at y is done by an increment to the y -th leaf and to all its ancestors. Every node hence always contains the sum of its children. The y -*prefix*, i.e. the first $y - 1$ leaves, may be efficiently summed using left-siblings of y ’s ancestors.

In fact, the two presented methods add exactly the same numbers in a different order. We also remark that similar tree-based structures are used in [AAAH01] for permutation pattern matching.

Counting larger increasing patterns, such as $\#123$, is similarly done with more sum trees [Cha18]. Let the sum tree A record the already-observed values as before. In the next sum tree B , we insert at $B[y]$ the number of 12 such that y corresponds to the 2 . For example, $B[5] = 3$ for $\pi = \underline{2}8\underline{1}76\underline{4}\underline{5}3$. We remark that the array B eventually sums to $\#12$, and that π can be reconstructed from B , which is closely

related to the *inversion table* of π [Sta11, page 36], but these facts are not needed here.

Similar to the computation of $\#12$ by summing queries to A , the computation of $\#123$ sums queries to B . To sum up, at every step x the algorithm will

- (a) insert one to $A[\pi(x)]$,
- (b) insert the $\pi(x)$ -prefix sum of A to $B[\pi(x)]$,
- (c) add the $\pi(x)$ -prefix sum of B to the result.

An occurrence of 123 at $i_1 < i_2 < i_3$ contributes one to $A[\pi(i_1)]$ at (a) of step i_1 , which we pass along to $B[\pi(i_2)]$ at (b) of step i_2 , and to the result at (c) of step i_3 .

One can similarly compute $\#12\dots k$ in $\tilde{O}(k^2n)$ time and space with $k - 1$ sum trees. Every sum tree A_m will hold counts of $12\dots m$ based on prefix sums of A_{m-1} . Note that this algorithm can also yield $\#k\dots 21$, by reversing the input π .

Corner Trees

The graph of an order- n permutation π in the xy -plane consists of n points. In the above dynamic programming, one scans these points from left to right, and queries a $\pi(x)$ -prefix sum in order to integrate some function over the points to the left and below $(x, \pi(x))$. From this perspective, the extension to other statistics will be done by looking at all four quadrants around a point, combining several of them together, and repeating recursively. These ideas motivate the following definitions.

Definition. A *corner tree* is a rooted tree, whose vertices are labeled by

$$\{\star, \text{SW}, \text{SE}, \text{NW}, \text{NE}\}$$

where \star is designated as the label for the root.

Example. See Figure 1.

Definition. An *occurrence* of a corner tree T in a permutation $\pi \in S_n$ is a map

$$i : \text{vertices}(T) \rightarrow \{1, \dots, n\}$$

such that the image of every vertex v relative to its parent is compatible with its label in the following sense:

- *W $i(v) < i(v.\text{parent})$
- *E $i(v) > i(v.\text{parent})$
- S* $\pi(i(v)) < \pi(i(v.\text{parent}))$
- N* $\pi(i(v)) > \pi(i(v.\text{parent}))$

Remark. Here *W means that the label of v is either SW or NW, and similarly in the other cases.

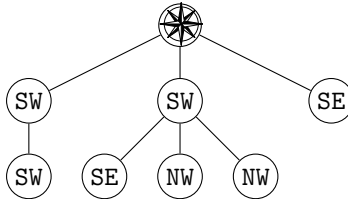


FIGURE 1. A corner tree

Definition. Let $\#T(\pi)$ be the *count* of T , which is the number of different occurrences of T in π . When π is not important we write $\#T$.

We give some examples of corner trees and their counts, showing that the latter are also pattern counts. This relation justifies our reuse of terminology.

Example. Let $T = \textcircled{\ast}$. Then $\#T(\pi) = |\pi|$, hence $\#T = \#1$.

Example. Let $T' = \textcircled{\ast} - \textcircled{\text{NE}}$. Then $\#T' = \#12$.

Indeed, an occurrence of T' is a map of its two vertices to two positions $i_1 < i_2$ such that $\pi(i_1) < \pi(i_2)$, which is an occurrence of the pattern 12 in π . Similarly,

Example. $\# \textcircled{\ast} - \textcircled{\text{NE}} - \textcircled{\text{NE}} = \#123 = \# \textcircled{\text{SW}} - \textcircled{\ast} - \textcircled{\text{NE}}$

Note that these two corner trees counting 123 are essentially the same one with different choices of the root. The role of the root will become clear later, in the algorithm. Some corner trees yield combinations of pattern counts:

Example. $\# \textcircled{\ast} - \textcircled{\text{SE}} - \textcircled{\text{NE}} = \#213 + \#312$

Indeed, either 213 or 312 occur at $i_1 < i_2 < i_3$ whenever $\pi(i_1) > \pi(i_2) < \pi(i_3)$.

Example. $\# \textcircled{\ast} - \textcircled{\text{NE}} - \textcircled{\text{SW}} = 2 \cdot \#123 + 2 \cdot \#213 + \#12$

Here we have considered all the ways the vertex images may be related, including the case that the root and the leaf SW map to the same entry.

In general, as we now show, the count of a corner tree is a linear combination of counts of patterns.

Lemma 8. *Let T be a corner tree. Then there exists a finite list of integer coefficients $\{\widehat{\#T}(\sigma)\}_{\sigma \in \Sigma(T)}$ such that for every permutation $\pi \in \bigcup_{n=1}^{\infty} S_n$*

$$\#T(\pi) = \sum_{\sigma \in \Sigma(T)} \widehat{\#T}(\sigma) \#\sigma(\pi)$$

Moreover, the coefficients $\widehat{\#T}(\sigma)$ are uniquely determined by the recursive formula

$$\widehat{\#T}(\sigma) = \#T(\sigma) - \sum_{|\rho| < |\sigma|} \widehat{\#T}(\rho) \#\rho(\sigma)$$

where the sum is over all patterns ρ that are shorter than σ .

Proof. Let $\widehat{\#T}(\sigma)$ be the number of occurrences of T in σ which are onto the set of positions $\{1, \dots, |\sigma|\}$.

Every occurrence i of a tree T in a permutation π , defines a unique pattern σ that occurs in π at the image of i . In other words, we can uniquely represent the occurrence as $i = i'' \circ i'$ where i' is an occurrence of T onto σ and i'' is an occurrence of σ in π . Conversely, any such composition, of occurrences of T onto σ and σ in π , gives rise to an occurrence of T in π . The lemma now follows by dividing into cases according to σ .

Clearly $|\sigma| \leq |T|$, where $|\sigma|$ is the length of the pattern, and $|T|$ is the number of vertices in the tree. Thus the set of potential patterns $\Sigma(T)$ is finite and does not depend on π .

The formula for the coefficients follows by renaming the variables π, σ as σ, ρ and noting that for $|\rho| \geq |\sigma|$, we have $\#\rho(\sigma) = 1$ if $\rho = \sigma$ and 0 otherwise. \square

In the other direction, counting a pattern *may* reduce to counting corner trees. By the above examples, the pattern count $\#12$ is also a corner tree count, and same for $\#123$ and all increasing patterns. Since $2\#123 + 2\#213 + \#12$ is also a corner-tree count, one can solve for $\#213$ and express it as a combination of three corner-tree counts.

Remark. Lemma 8 lets us work with finite combinations of pattern counts and tree counts, and derive relations that hold for any input of any length. To emphasize this point, we may identify formal sums of trees with formal sums of patterns.

Let V be the linear space of all formal linear combinations of patterns, that is, expressions of the form $A = \sum_{\sigma} a_{\sigma} \sigma$ where $a_{\sigma} \in \mathbb{Q}$. Denote the resulting permutation statistics $\#A = \sum_{\sigma} a_{\sigma} \#\sigma$. By Lemma 8 every corner tree T is an element of V . A *corner tree formula* is a combination of the form $\sum_T c_T T$, and we are interested in representing patterns by such formulas. Clearly, corner tree formulas form a subspace U of V .


One can also define a multiplication that satisfies $\#(A \cdot B) = \#A \#B$, thereby giving both U and V a commutative ring structure. The unit would be the empty tree/pattern ϕ , with $\#\phi \equiv 1$, a constant permutation statistics. Such a definition shows that corner tree formulas are not more restrictive than expressions that involve products of tree counts. Therefore, we do not pursue this idea here.

Proof of Theorem 1

We now show that counting a corner tree in a permutation can be done in nearly linear time in its length. The pseudo-code to compute $\#T(\pi)$ is presented in Algorithm 1. The input to the algorithm is the permutation π , the corner tree T , and its root vertex v .

Algorithm 1 Count a corner tree

```

1: function COUNT-CT(permutation  $\pi$ , tree  $T$ , vertex  $v$ )
2:    $A \leftarrow \text{ARRAY}(1, \dots, 1)$  ▷ of size  $n = |\pi|$ 
3:   for  $w$  in  $v$ .children do
4:      $A \leftarrow A \odot \text{COUNT-CT}(\pi, T, w)$  ▷ term-wise product of arrays
5:   if  $v$ .label =  then
6:     return SUM( $A$ ) ▷ of the  $n$  values
7:    $C \leftarrow \text{ARRAY}(0, \dots, 0)$  ▷ size  $n$ 
8:    $B \leftarrow \text{SUM-TREE}(0, \dots, 0)$  ▷ size  $n$ 
9:   for  $x$  in  $\begin{cases} (1, \dots, n) & \text{if } v.\text{label} = *W \\ (n, \dots, 1) & \text{if } v.\text{label} = *E \end{cases}$  do ▷ two cases
10:     $C[x] \leftarrow \begin{cases} B.\text{prefix-sum}(\pi[x]) & \text{if } v.\text{label} = S* \\ B.\text{suffix-sum}(\pi[x]) & \text{if } v.\text{label} = N* \end{cases}$  ▷ two cases
11:     $B[\pi[x]] \leftarrow A[x]$ 
12:   return  $C$  ▷ the whole array

```

Lines 2-4 produce an array A , such that $A[x]$ is the number of occurrences of the subtree rooted at v , that map v to x . This number is computed for every x as a product over v 's children, as we have to choose valid maps for their subtrees. We thus proceed by recursion on the tree, and call the function once for every child. Note that if v is a leaf then there is a single such map and $A[x] = 1$ for every x . If v is the root, then we return the total count, on lines 5-6, as it doesn't matter whither it maps.

Otherwise, we create and return an array of counts C , so that $C[x]$ is the number of valid maps that will send v 's parent to the position x . This is a sum of A over all positions x' such that the relation between the two entries is compatible with v 's label. For example, $C[x]$ will be the sum of $A[x']$ with $x' < x$ and $\pi(x') < \pi(x)$ in the case that v 's label is SW. The other three cases are similar. To avoid a double loop and save time, we compute all these sums in one iteration over A , by inserting it into a sum-tree array B , from which we properly retrieve $C[x]$.

Note that this algorithm reduces to the above method for counting increasing patterns, in the case that the tree T is a path of SW.

In conclusion, Algorithm 1 correctly counts corner trees. Overall, the function is called $k = |T|$ times. The arrays A, B, C hold $O(n)$ numbers. On lines 4, 6, 9 there are loops of n iterations. Thanks to sum trees, lines 10 and 11 are performed with at most $\log n$ arithmetic operations each. All the numbers involved require at most $k \log(n + 1)$ bits. Thus for a k -vertex corner tree the function runs in $\tilde{O}(k^2 n)$ time. It follows that any fixed permutation statistic expressible by corner trees can be computed in $\tilde{O}(n)$ time, and this proves Theorem 1.

3. SMALL PATTERNS

Before using corner trees to count 3-patterns and 4-patterns, we mention some well-known symmetries that reduce the number of cases. Let $\text{rev} : S_n \rightarrow S_n$ be the reverse map, $\text{rev}(\pi(1) \dots \pi(n)) = \pi(n) \dots \pi(1)$. Also, let $\text{inv}(\pi) = \pi^{-1}$. These two transformations act as reflections on the graph of the permutation in the plane, and generate D_4 , the order-eight dihedral group of symmetries of the square.

Clearly these eight symmetries respect pattern occurrence, in the sense that $\#[\text{rev } \sigma](\text{rev } \pi) = \#\sigma(\pi)$ for example. Thus we obtain linear time reductions between counting σ and any other pattern in its orbit under the action of D_4 . We may also apply these symmetries to corner trees, where rev swaps W and E, and inv swaps NW and SE. This yields direct reductions between corner tree formulas.

For example, the problem of computing $\#321$ is equivalent to $\#123$. Similarly, $\#132, \#231, \#213$ and $\#312$ are pairwise equivalent as rotations of each other. In Section 2 we have seen that $\#123$ and $\#213$ are computable with corner trees. It follows that every 3-pattern can be counted in $\tilde{O}(n)$ time, as claimed by Corollary 2.

These cases extend to the orbits of $123\dots k$ and $213\dots k$ for every k . For example, we have the following formulas for counting 4-patterns.

$$\begin{aligned} \#1234 &= \# \left(\text{diagram: a square with a star in the center, and three circles labeled SW connected in a horizontal line} \right) \\ \#2134 &= - \#1234 - \frac{1}{2} \#123 + \frac{1}{2} \# \left(\text{diagram: a square with a star in the center, and two circles labeled SW connected in a horizontal line, with a third circle labeled SW connected to the right one by a diagonal line} \right) \end{aligned}$$

We can give another corner tree formula for a 4-pattern, which is remarkably more complicated. Let

$$\begin{aligned}
A = & \begin{array}{c} \text{SE} \\ \diagup \\ \text{SW} \text{---} \text{SE} \\ \text{---} \end{array} - \begin{array}{c} \text{SE} \\ \diagup \\ \text{SE} \text{---} \text{SW} \\ \text{---} \end{array} + \begin{array}{c} \text{SE} \\ \diagup \\ \text{SE} \text{---} \text{NE} \\ \text{---} \end{array} \\
& - 2 \begin{array}{c} \text{---} \text{SW} \text{---} \text{SE} \text{---} \text{SE} \\ \text{---} \end{array} - 2 \begin{array}{c} \text{---} \text{SW} \text{---} \text{SE} \text{---} \text{SW} \\ \text{---} \end{array} - \begin{array}{c} \text{---} \text{SW} \text{---} \text{SE} \\ \text{---} \end{array} \\
& + \frac{1}{3} \begin{array}{c} \text{NE} \text{---} \text{NE} \\ \diagup \quad \diagup \\ \text{NE} \\ \text{---} \end{array} - \begin{array}{c} \text{NE} \\ \diagup \\ \text{NE} \text{---} \text{NE} \\ \text{---} \end{array} - \frac{1}{2} \begin{array}{c} \text{NE} \\ \diagup \\ \text{NE} \\ \text{---} \end{array} + \frac{1}{6} \begin{array}{c} \text{---} \text{NE} \\ \text{---} \end{array}
\end{aligned}$$

Lemma 9. $\#A = \#2143$

Proof. This follows by a routine inspection of pattern occurrences counted by these trees. Denote the trees in A by T_1, \dots, T_{10} in order of appearance. The desired pattern 2143 is counted once by T_1 and nowhere else. Every other pattern is cancelled out when summing all the terms. We include a brief account for the interested reader.

- (a) $2413, 2431$ are also counted by T_1 , but subtracted back by T_2 .
- (b) $4123, 4132, 4213, 4231, 4312, 312$ only appear in T_2 and T_3 , the same number of times with opposite signs.
- (c) $3124, 3142, 3214, 3241, 3412, 3421$ are counted twice among T_1 and T_3 , and $231, 213$ counted once. They are cancelled out by T_4, T_5 and T_6 .
- (d) $1324, 1342, 1423, 1432, 132$ are added twice by T_7 to compensate for their unnecessary subtraction by T_2 or T_5 .
- (e) $1234, 1243, 123, 132, 12$ appear in T_7 and are balanced by T_8, T_9, T_{10} . \square

In conclusion, we have presented nearly linear time algorithms that count all patterns from the following three orbits: $\{1234, 4321\}$, $\{1243, 2134, 3421, 4312\}$, $\{2143, 3412\}$. This proves Corollary 3.

The Bergsma–Dassios–Yanagimoto Statistic

Let $(X_1, Y_1), \dots, (X_k, Y_k)$ be independent samples in the real plane, drawn from a common bivariate distribution (X, Y) with non-atomic marginals. Recall that such samples induce, with probability one, a random permutation $\rho \in S_k$ such that $\text{rank } Y_1 = \rho(\text{rank } X_1)$, $\text{rank } Y_2 = \rho(\text{rank } X_2)$, \dots , $\text{rank } Y_k = \rho(\text{rank } X_k)$. Equivalently, the samples can be written as $(X_{(1)}, Y_{(\rho(1))}), \dots, (X_{(k)}, Y_{(\rho(k))})$ with the usual order statistics notation, $X_{(1)} < X_{(2)} < \dots < X_{(k)}$. As noted above in the introduction, many nonparametric measures of bivariate data can be described in terms of this random permutation.

In a typical statistical test, the null hypothesis states that the two variables X and Y are independent. Under this assumption, it is easy to see that for all $k \in \mathbb{N}$,

$$(\star) \quad \forall \sigma \in S_k \quad P(\rho = \sigma) = \frac{1}{k!}$$

A classical result by Hoeffding gives a strong converse [Hoe48]. If (\star) holds for some $k \geq 5$ with X and Y jointly absolutely continuous, then they are independent.

Indeed, the well-known nonparametric independence test by Hoeffding, which can be expressed with counts of 5-pattern, is consistent against all alternative joint densities.

Hoeffding’s test has been extended in later works to variables in \mathbb{R}^d , to three or more random variables, and in other directions. Some variants that relax the continuity assumption have been studied. Yanagimoto has shown that independence actually follows from (\star) with $k = 4$, and this is optimal since (\star) with $k = 3$ is not sufficient [Yan70]. Bergsma and Dassios have suggested a corresponding nonparametric independence test, based on 4-patterns, which has gained popularity in recent years [Ber10, BD14]. Up to normalization, the Bergsma–Dassios–Yanagimoto statistic has the following simple form:

$$T^* = \#1234 + \#1243 + \#2134 + \#2143 + \#3412 + \#3421 + \#4312 + \#4321$$

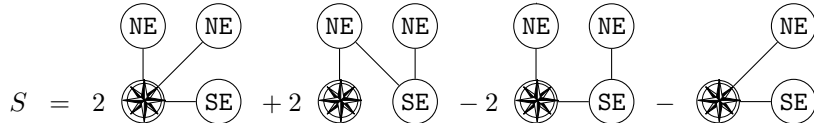
Under independence, $T^*/\binom{n}{4}$ is concentrated at $1/3$, while for any other bivariate distribution with continuous marginals these eight patterns appear at a strictly higher overall rate [Yan70, BD10, DHS18]. Natural interpretations of this statistical measure have been discussed in recent works [BD14, WDM18, Eve18].

We remark that the relation between independence and pattern frequencies has been independently discovered in the theory of so-called *permutons*, limit objects for permutations. In that context, it means that quasi-randomness of permutations is “finitely forcible” [KP13, GGKK15].

A straightforward evaluation of T^* requires $\tilde{O}(n^4)$ time, though it was noted right away that this can be reduced to $\tilde{O}(n^3)$ [BD10]. Therefore, it was suggested at first to approximate this statistic by a sufficiently large random sample of subsets of four entries [BD14, Sect. 4]. This method is applicable, but due to the degenerate nature of the null distribution [NWD16, DDB16, Eve18] it incurs a substantial loss of information even if as much as $O(n^2)$ random samples are used [Jan84]. Two recent papers have been dedicated to an exact computation of T^* in $\tilde{O}(n^2)$ time [WDL16, HH16]. These methods might still be infeasible in big data applications, where only near linear-time work is considered practical.

Corollary 4, stating that indeed the Bergsma–Dassios–Yanagimoto statistic can be evaluated in $\tilde{O}(n)$ time, is immediate from Corollary 3. Nevertheless, we derive a more direct formula, based on the invariance of T^* under the symmetries mentioned above. Let the action of the dihedral group linearly extend from patterns and trees to their formal sums, so that $\#S(\pi) = \#[g \cdot S](g \cdot \pi)$ for a corner tree formula S and a reflection or a rotation $g \in D_4$.

Proposition 10. *The Bergsma–Dassios–Yanagimoto permutation statistic is obtained from averaging the corner tree formula,*



over all eight reflections and rotations, as follows:

$$T^*(\pi) = \binom{|\pi|}{4} - \frac{1}{8} \sum_{g \in D_4} \#[g \cdot S](\pi)$$

Proof. This is a straightforward calculation. In spirit of the discussion following Lemma 8, expand S in patterns:

$$\begin{aligned} S &= 2(2 \cdot 2134 + 2 \cdot 2143 + 2 \cdot 2314 + 2 \cdot 2341 + 2 \cdot 2413 + 2 \cdot 2431 + 213 + 231) \\ &\quad + 2(1324 + 1423 + 2314 + 2413 + 3412) \\ &\quad - 2(2 \cdot 2134 + 2 \cdot 2143 + 2314 + 2413 + 3124 + 3142 + 3412 + 213) \\ &\quad - (213 + 231) \end{aligned}$$

Sort the patterns into their D_4 orbits:

$$\begin{aligned} S &= (231 - 213) \\ &\quad + 2(1324) + 4(2341) + 2(2 \cdot 2413 - 3142) + 8(\frac{1}{4}1423 + \frac{1}{2}2314 + \frac{1}{2}2431 - \frac{1}{4}3124) \end{aligned}$$

Let $S^* = \frac{1}{8} \sum_{g \in D_4} [g \cdot S]$. Then $(231 - 213)$ vanishes, and the other orbits equalize:

$$\begin{aligned} S^* &= (1324 + 4231) + (1432 + 2341 + 3214 + 4123) + (2413 + 3142) \\ &\quad + (1342 + 1423 + 2314 + 2431 + 3124 + 3241 + 4132 + 4213) \end{aligned}$$

The proposition now follows by the definition of T^* , noting that the total count of all twenty-four order-4 patterns in an order- n permutation is $\binom{n}{4}$. \square

Remark. The symmetrized corner tree formula S^* has 20 nonequivalent terms. To simplify the computation, one can equivalently write $\#S(g^{-1} \cdot \pi)$ in the proposition instead of $\#[g \cdot S](\pi)$, and so always deal with the four corner trees in S . Note that these trees have some overlapping branches, which may be exploited by a dedicated algorithm in practical computations. We provide implementations of these different approaches in the python package, see details below.

Corner Tree Spaces

As noted after Lemma 8, formal sums of corner trees constitute a subspace of formal sums of patterns. This means that certain linear combinations of pattern counts may be computable by corner trees, even if individual patterns in those combinations are not so.

For example, we have seen eight order-4 patterns, out of 24, that are contained in the subspace spanned by order-4 corner trees. However, a computer-assisted calculation shows that this subspace is actually 23-dimensional, proving Proposition 5. This means that the 4-profile can be computed in near linear time up to only one freedom degree, an unknown multiple of one particular vector.

What is the missing element? Endow the space of formal sums of 4-patterns with the unique inner product that makes the single 4-patterns an orthonormal basis. One can solve for a combination perpendicular to the pattern expansions of corner trees as above, and obtain:

$$\begin{aligned} N &= (1324 + 4231) + (1432 + 2341 + 3214 + 4123) + (2413 + 3142) \\ &\quad - (1342 + 1423 + 2314 + 2431 + 3124 + 3241 + 4132 + 4213) \end{aligned}$$

Therefore, the subspace of linear combinations of 4-patterns for which we have corner tree formulas is characterized as N^\perp , with respect to the above inner product.

Remark. Here is an application of these computations. In the setting of Hoeffding's test, one may look for other independence-detecting sums $F = \sum_{\sigma \in S_k} f_\sigma \# \sigma$, where $E[F/\binom{n}{k}] = \alpha$ assuming independence, and $E[F/\binom{n}{k}] > \alpha$ under all dependent

alternatives. Such combinations form a cone, studied for 4-patterns in a recent paper [CKN⁺19]. Ten so-called Σ -forcing combinations with such a property are provided there, including the Bergsma–Dassios–Yanagimoto statistic. We note that six out of the ten are contained in the subspace of corner trees, and hence can be computed efficiently.

We conclude our discussion of corner tree formulas with some brief details on their implementation. The Python package `cornertree` accompanies this article [EL19]. It includes the general-purpose classes `SumTree`, `CornerTree`, and `CornerTreeFormula`, as well as an implementation of Algorithm 1, with a plenty of examples. The function `Tstar` performs an optimized computation of the Bergsma–Dassios–Yanagimoto statistic. The iterator `canonical_corner_trees` generates all corner trees of a given order, and the routine `expand_corner_trees` expands them as formal sums in pattern. Proposition 5 and the other observations in the previous paragraphs follow from the output by standard linear algebra.

The same procedure shows that corner trees span a 100-dimensional subspace of 5-patterns, a 463-dimensional subspace of 6-patterns, and a 2323-dimensional subspace of 7-patterns. The latter computation required several hours of CPU time.

4. FASTER 4-PATTERNS

We turn to explore other methods for counting patterns, beyond those provided by corner trees. This section focuses on the intriguing problem of computing the full 4-profile. Proposition 5 asserts that corner trees yield 23 independent linear functionals of this 24-dimensional vector. Therefore, counting any 4-pattern that is not already spanned by corner trees will accomplish the computation.

Here is a simple $\tilde{O}(n^2)$ solution, achieved by counting 3214 in a given $\pi \in S_n$. Let $i_4 \in \{1, \dots, n\}$ be a position that will correspond to 4 , the last entry in an occurrence of 3214 . We count i_1, i_2, i_3 that complete an occurrence. This is done by restricting π to smaller positions $x < i_4$ with smaller values $\pi(x) < \pi(i_4)$, and counting 321 exactly as described in the beginning of Section 2. To obtain $\#3214$ we sum over i_4 , adding up n numbers, each computed in $\tilde{O}(n)$ time.

Faster $\#3214$

The above method gives rise to much overlapping work, as the same occurrences of 321 are counted over and over, for different values of i_4 . Our approach is to combine these tasks together for adjacent i_4 and adjacent $\pi(i_4)$.

We thus divide $\{1, \dots, n\}$ into chunks of size m , where m is a parameter to be suitably chosen later. We define two properties for an occurrence of 3214 at positions i_1, i_2, i_3, i_4 :

- (A) $\pi(i_1) \leq am < \pi(i_4)$ for some integer a .
- (B) $i_3 \leq bm < i_4$ for some integer b .

To describe the properties geometrically, divide the graph of π in the plane into horizontal and vertical strips of width m . Property A asserts that the 3 and the 4 of the 3214 are contained in different horizontal strips, and property B similarly asserts that the 1 and the 4 are in different vertical strips.

We first count type A occurrences in $\tilde{O}(n^2/m)$ time, as follows. For every integer $a < n/m$, we scan π once, for the purpose of counting occurrences with 4 in the

a -th horizontal strip. Our action at position x depends on the value of $\pi(x)$. If $\pi(x) > (a+1)m$ then we do nothing. If $\pi(x) \leq am$ then we use it to count 321 with two sum-trees as above. If $\pi(x) \in [am+1, (a+1)m]$, then we add the number of 321 counted so far to the total count of 3214 .

By applying the same method to the transposed permutation π^{-1} , we also count occurrences of 3214 that satisfy B.

Occurrences that satisfy both A and B have now been counted twice, and need to be subtracted. We hence make the following adjustments to the counting of type A. Whenever i reaches a multiple of m , we save the count of 321 at that time. Upon observing $\pi(i)$ in the a -th strip, we add the current count of 321 as before, but subtract the saved count from the previous m -multiple. Thereby, we subtract all occurrences of type $A \cap B$, compensating for their double counting.

It is left to deal with occurrences of 3214 satisfying neither A nor B. Given any position i_4 , there are at most m relevant $i_3 < i_4$ in the same vertical strip, and at most m relevant i_1 with $\pi(i_1) < \pi(i_4)$ in the same horizontal strip. We iterate through all $O(nm^2)$ such (i_1, i_3, i_4) , and count i_2 that complete an occurrence of 3214 . This count is the number of points $(x, \pi(x))$ in the box $(i_1, i_3) \times (\pi(i_3), \pi(i_1))$. We turn to describe a data structure that efficiently answers such queries.

An n -by- n two-dimensional sum-tree uses a Cartesian product of two complete binary trees of depth $\lceil \log n \rceil$. The pairs of leafs (l, l') hold an $n \times n$ array of numbers. A general pair of vertices (v, v') always holds the sum over pairs of leafs from the corresponding subtrees. Insertion makes access to $O(\log^2 n)$ pairs, as well as querying a box sum. If the array is sparse then we may use a hash table, and only store the $O(s \log^2 n)$ nonzero entries, where s is the number of nonzero array cells. See [Ben75, Mat94, DCVO08, CLP11] for various general data structures of this kind.

Using a two-dimensional sum-tree, with ones at all $(x, \pi(x))$ and zeros elsewhere, the remaining occurrences are counted in $\tilde{O}(nm^2)$ time and $\tilde{O}(n)$ space, as explained above. In order to balance n^2/m and nm^2 , we set $m = \lfloor n^{1/3} \rfloor$, so that the total running time is $\tilde{O}(n^{5/3})$. This algorithm together with Proposition 5 prove Theorem 6.

Faster #3241

Theorem 7 is proven by another algorithm based on similar ideas, which computes #3241 in $\tilde{O}(n^{3/2})$ time and space. As before we divide the occurrences $i_1 < i_2 < i_3 < i_4$ of 3241 into cases. This time we ask whether $2,1$ and whether $3,4$ are contained in the same horizontal strips of height m .

- (A) $\pi(i_4) \leq am < \pi(i_2)$ for some integer a .
- (B) $\pi(i_1) \leq bm < \pi(i_3)$ for some integer b .

Since this algorithm is somewhat more involved, we present it in full detail, see Algorithm 2. In the following paragraphs we outline the key steps.

The first part of the algorithm counts occurrences of types A and B. We thus iterate over horizontal strips as before and count 3241 with: (a) the 1 in the strip and the 324 above it; (b) the 4 in the strip and the 321 below it.

When counting type B, we have to make sure that the 4 falls between the 2 and the 1 . Therefore, as we go over the entries of the permutation, we increment the

Algorithm 2 Compute $\#3241$

```

1: function COUNT-3241(permutation  $\pi$  of size  $n$ , integer  $m$ )
2:    $t \leftarrow 0$ 
3:   for  $chunk$  in  $\{1 \dots m\}, \{m+1 \dots 2m\} \dots \{\lfloor \frac{n}{m} \rfloor m + 1 \dots n\}$  do
4:      $c \leftarrow 0$ 
5:      $A_1, A_{12}, B_1, B_{21} \leftarrow 4 \times \text{SUM-TREE}(0, \dots, 0)$  ▷ size  $n$  each
6:     for  $y$  in  $\{\pi(1), \pi(2), \dots, \pi(n)\}$  do
7:       if  $y < \min(chunk)$  then
8:          $B_1[y] \leftarrow 1$ 
9:          $B_{21}[y] \leftarrow B_1[> y]$  ▷ means suffix-sum
10:         $c \leftarrow c + B_1[> y] \cdot (y - 1 - B_1[< y]) - B_{21}[> y]$ 
11:        if  $y > \max(chunk)$  then
12:           $A_1[y] \leftarrow 1$ 
13:           $A_{12}[y] \leftarrow A_1[< y]$ 
14:           $y' \leftarrow \lfloor \frac{y-1}{m} \rfloor m$  ▷  $m$ -multiple below  $y$ 
15:           $c \leftarrow c + \binom{A_1[< y]}{2} - A_{12}[< y] - \binom{A_1[\leq y']}{2} + A_{12}[\leq y']$ 
16:          if  $y \in chunk$  then  $t \leftarrow t + c$ 
17:         $P \leftarrow \text{PRODUCT-TREE}(n \times n)$  ▷ 2-dim array, all 0
18:        for  $y$  in  $\{1, 2, \dots, n\}$  do
19:           $y' \leftarrow \lfloor \frac{y-1}{m} \rfloor m$ 
20:          for  $z$  in  $\{y' + 1, y' + 2, \dots, y - 1\}$  do
21:             $P[\pi^{-1}(z), \pi^{-1}(y)] \leftarrow 1$ 
22:            for  $z$  in  $\{y + 1, y + 2, \dots, y' + m\}$  do
23:              if  $\pi^{-1}(y) < \pi^{-1}(z)$  then
24:                 $t \leftarrow t + P.\text{sum-box}(\pi^{-1}(z), n) \times (\pi^{-1}(y), \pi^{-1}(z))$ 
25:   return  $t$ 

```

count of 321 upon observing the 2 and decrease it back upon observing the 1 . See the two corresponding terms on line 10 of the algorithm.

As for type A, we count 324 by adding $234+324$ and subtracting 234 . This is similar to the treatment of $\#213$ in Section 2. See the first two terms on line 15. The other two terms with y' further subtract those type A occurrences that are already counted as type B for another strip.

The second part counts those 3241 that are neither A nor B, so that both the 21 and the 34 are contained in horizontal strips. We use a two-dimensional sum tree again, this time keeping account of all same-strip pairs of entries. Their coordinates in the tree are the position of the 1 and the position of the 2 .

We scan the graph of π bottom to top, inserting around $\frac{n}{m} \binom{m}{2}$ such pairs to the two-dimensional sum tree on line 21. In parallel, for every same-strip increasing

pair 34 , we lookup those 21 below it that properly merge to 3241 . This number is efficiently retrieved by a suitable box sum query, on line 24.

The two parts of the algorithm run in time $\tilde{O}(n^2/m)$ and $\tilde{O}(nm)$. We hence set $m = \lfloor \sqrt{n} \rfloor$ to obtain $\tilde{O}(n^{3/2})$ time complexity. Since we make at most nm insertions into the two-dimensional sum tree, it is sparse and only needs $\tilde{O}(n^{3/2})$ memory space. More generally, we obtain a tradeoff between $\tilde{O}(n^{2-a})$ computation time and $\tilde{O}(n^{1+a})$ space, where a ranges from 0 to $1/2$.

We note that this algorithm still seems to contain much overlapping work. For example, the same 321 and 324 are counted many times as we iterate over horizontal strips. Further improvements are yet to be sought, perhaps by dividing into more cases.

See also the python implementation, `count_3214` and `count_3241`, in our package `cornertree.py` [EL19]. It includes a general-purpose class `ProductTree` implementing two-dimensional sum-trees.

5. DISCUSSION

The following open problems and directions for future research naturally arise from the results presented in this paper.

- (1) Are there pattern counts, or linear combinations of them, that are not spanned by corner trees, but still can be computed in $\tilde{O}(n)$ time?
- (2) What is the computational complexity of finding the full 4-profile of a given permutation of size n ?
- (3) Is there a simple characterization of the linear space of pattern combinations that are spanned by corner trees?
- (4) In particular, what is the dimension of the restriction of that space to patterns of size k ?
- (5) What other patterns can be counted, say in $\tilde{O}(n^{3/2})$ time, using the ideas presented in Section 4?

In the spirit of the last question, it would be interesting to initiate a systematic study, similar to our treatment of corner trees, of other algorithmic ideas and the linear spaces of patterns that they yield.

This may also be relevant to the remaining gap in the asymptotic complexity of k -pattern counting. Given $k \in \mathbb{N}$, what is the smallest $d(k)$ such that the k -profile can be computed in $\tilde{O}(n^{d(k)})$ time as n grows? For example $d(2) = d(3) = 1$ and $d(4) \leq \frac{3}{2}$ by our results. Recall that Berendsohn et al. showed that for k large $d(k) \leq \frac{k}{4} + o(k)$, while $o(k/\log k)$ is not likely [BKM19].

The dimension mentioned in question 4 is currently known for $k \leq 7$. It is clearly bounded by the number of k -vertex corner trees, which is $O(11.07^k)$ [OEI19, A052763]. Since the k -profile is $k!$ -dimensional, a super-exponential family of such objects would be necessary to span all k -patterns.

As for more practical research directions, one may improve the computation time of other rank-based statistical tests, applying or extending our methods. Some potential examples are the general families of independence tests recently proposed in [HHG12, HHK⁺16, WDM18].

By working with permutations, we always assume distinct inputs and distinct outputs, but the applications to data analysis also make sense for distributions with atoms, where ties occur with positive probability. It may hence be useful to extend our methods to such data sets, treating ties appropriately.

Another potential extension concerns other families of permutation statistics studied in Combinatorics. In particular, can one adapt corner trees to count so-called generalized permutation patterns, such as vincular patterns [BS00], and mesh patterns [BC11]?

Finally, we remark that while this work focuses on the problem of exact counting of patterns, also questions of approximation seem interesting and useful. Several works have been recently devoted to property testing of permutations [HKMS11, KK14, FW18, BC18, NRRS19], which may be relevant to the study of approximated pattern counting.

REFERENCES

- [AAAH01] Michael H Albert, Robert EL Aldred, Mike D Atkinson, and Derek A Holton. Algorithms for pattern involvement in permutations. In *International Symposium on Algorithms and Computation*, pages 355–367. Springer, 2001.
- [AAH⁺02] Michael H Albert, Mike D Atkinson, Chris C Handley, Derek A Holton, and Walter Stromquist. On packing densities of permutations. *Electron. J. Combin.*, 9(1), 2002.
- [AR08] Shlomo Ahal and Yuri Rabinovich. On complexity of the subpattern problem. *SIAM Journal on Discrete Mathematics*, 22(2):629–649, 2008.
- [BBL98] Prosenjit Bose, Jonathan F Buss, and Anna Lubiw. Pattern matching for permutations. *Information Processing Letters*, 65(5):277–283, 1998.
- [BC11] Petter Brändén and Anders Claesson. Mesh patterns and the expansion of permutation statistics as sums of permutation patterns. *The Electronic Journal of Combinatorics*, 18(2):5, 2011.
- [BC18] Omri Ben-Eliezer and Clément L Canonne. Improved bounds for testing forbidden order patterns. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2093–2112. SIAM, 2018.
- [BCKM15] Ivan Bliznets, Marek Cygan, Paweł Komosa, and Lukáš Mach. Kernelization lower bound for permutation pattern matching. *Information Processing Letters*, 115(5):527–531, 2015.
- [BD10] Wicher Bergsma and Angelos Dassios. A consistent test of independence based on a sign covariance related to Kendall’s tau. *arXiv preprint arXiv:1007.4259*, 2010.
- [BD14] Wicher Bergsma and Angelos Dassios. A consistent test of independence based on a sign covariance related to Kendall’s tau. *Bernoulli*, 20(2):1006–1028, 2014.
- [Ben75] Jon Louis Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.
- [Ber10] Wicher Bergsma. Nonparametric testing of conditional independence by means of the partial copula. *arXiv preprint, arXiv:1101.4607*, 2010.
- [Ber19] Benjamin Aram Berendsohn. Complexity of permutation pattern matching. *Master Thesis, FU, Berlin*, 2019.
- [BH10] Alexander Burstein and Peter Hästö. Packing sets of patterns. *European Journal of Combinatorics*, 31(1):241–253, 2010.
- [BHL⁺15] József Balogh, Ping Hu, Bernard Lidický, Oleg Pikhurko, Balázs Udvari, and Jan Volec. Minimum number of monotone subsequences of length 4 in permutations. *Combinatorics, Probability and Computing*, 24(04):658–679, 2015.
- [BKM19] Benjamin Aram Berendsohn, László Kozma, and Dániel Marx. Finding and counting permutations via CSPs. *arXiv preprint arXiv:1908.04673*, 2019.
- [BKR61] Julius Robin Blum, Jack Carl Kiefer, and Murray Rosenblatt. Distribution free tests of independence based on the sample distribution function. *The Annals of Mathematical Statistics*, pages 485–498, 1961.

- [BL12] Marie-Louise Bruner and Martin Lackner. A fast algorithm for permutation pattern matching based on alternating runs. In *Scandinavian Workshop on Algorithm Theory*, pages 261–270. Springer, 2012.
- [BL13] Marie-Louise Bruner and Martin Lackner. The computational landscape of permutation patterns. *arXiv preprint arXiv:1301.0340*, 2013.
- [BL16] Marie-Louise Bruner and Martin Lackner. A fast algorithm for permutation pattern matching based on alternating runs. *Algorithmica*, 75(1):84–117, 2016.
- [Bón07] Miklós Bóna. The copies of any permutation pattern are asymptotically normal. *arXiv:0712.2792*, 2007.
- [Bón10] Miklós Bóna. On three different notions of monotone subsequences. *Permutation Patterns*, 376:89–114, 2010.
- [Bón12] Miklós Bóna. *Combinatorics of Permutations*. Chapman and Hall / CRC Press, New York, US, 2012.
- [BS00] Eric Babson and Einar Steingrímsson. Generalized permutation patterns and a classification of the Mahonian statistics. *Séminaire Lotharingien de Combinatoire [electronic only]*, 44:1–18, 2000.
- [Cha18] Aayush Chaturvedi. Count inversions of size k in a given array. *Geeks for Geeks*, www.geeksforgeeks.org, 2018.
- [CK16] Joshua Cooper and Anna Kirkpatrick. The complexity of counting poset and permutation patterns. *Australasian Journal of Combinatorics*, 64(1):154–165, 2016.
- [CKN⁺19] Timothy Chan, Daniel Kral, Jonathan A Noel, Yanitsa Pehova, Maryam Sharifzadeh, and Jan Volec. Characterization of quasirandom permutations by a pattern sum. *arXiv preprint arXiv:1909.11027*, 2019.
- [CLP11] Timothy M Chan, Kasper Green Larsen, and Mihai Pătraşcu. Orthogonal range searching on the RAM, revisited. In *Proceedings of the twenty-seventh annual symposium on Computational geometry*, pages 1–10. ACM, 2011.
- [CLRS09] Thomas H Cormen, Charles E Leiserson, Ronald L Rivest, and Clifford Stein. *Introduction to algorithms*. MIT press, 2009.
- [Coo04] Joshua N Cooper. Quasirandom permutations. *Journal of Combinatorial Theory, Series A*, 106(1):123–143, 2004.
- [Coo06] Joshua N Cooper. A permutation regularity lemma. *the Electronic Journal of Combinatorics*, 13(1):22, 2006.
- [CP10] Timothy M Chan and Mihai Pătraşcu. Counting inversions, offline orthogonal range counting, and related problems. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 161–173. Society for Industrial and Applied Mathematics, 2010.
- [Cro66] Casparus F Crouse. Distribution free tests based on the sample distribution function. *Biometrika*, 53(1/2):99–108, 1966.
- [CW92] Maw-Shang Chang and Fu-Hsing Wang. Efficient algorithms for the maximum weight clique and maximum weight independent set problems on permutation graphs. *Information Processing Letters*, 43(6):293–295, 1992.
- [DCVO08] Mark De Berg, Otfried Cheong, Marc Van Kreveld, and Mark Overmars. Orthogonal range searching. *Computational Geometry: Algorithms and Applications*, pages 95–120, 2008.
- [DDB16] Subhra Sankar Dhar, Angelos Dassios, and Wicher Bergsma. A study of the power and robustness of a new test for independence against contiguous alternatives. *Electronic Journal of Statistics*, 10(1):330–351, 2016.
- [DHS18] Mathias Drton, Fang Han, and Hongjian Shi. High dimensional independence testing with maxima of rank correlations. *arXiv preprint arXiv:1812.06189*, 2018.
- [DW80] Tertius De Wet. Cramér-von Mises tests for independence. *Journal of Multivariate Analysis*, 10(1):38–50, 1980.
- [EL19] Chaim Even-Zohar and Calvin Leng. *GitHub repository: corners*, 2019. <http://github.com/chaim-e/corners>.
- [Eve18] Chaim Even-Zohar. Patterns in random permutations. *arXiv preprint arXiv:1811.07883*, 2018.
- [FL82] Nicholas I Fisher and Alan J Lee. Nonparametric measures of angular-angular association. *Biometrika*, pages 315–321, 1982.

- [Ful04] Jason Fulman. Stein’s method and non-reversible Markov chains. In *Stein’s Method*, pages 66–74. Institute of Mathematical Statistics, 2004.
- [FW18] Jacob Fox and Fan Wei. Fast property testing and metrics for permutations. *Combinatorics, Probability and Computing*, 27(4):539–579, 2018.
- [GGKK15] Roman Glebov, Andrzej Grzesik, Tereza Klimošová, and Daniel Král’. Finitely forcible graphons and permutons. *Journal of Combinatorial Theory, Series B*, 110:112–135, 2015.
- [GM14] Sylvain Guillemot and Dániel Marx. Finding small patterns in permutations in linear time. In *Proceedings of the twenty-fifth annual ACM-SIAM symposium on Discrete algorithms*, pages 82–101. SIAM, 2014.
- [GV09] Sylvain Guillemot and Stéphane Vialette. Pattern matching for 321-avoiding permutations. In *International Symposium on Algorithms and Computation*, pages 1064–1073. Springer, 2009.
- [Häs02] Peter A Hästö. The packing density of other layered permutations. *Electronic Journal of Combinatorics*, 9(2):1, 2002.
- [HH16] Yair Heller and Ruth Heller. Computing the Bergsma Dassios sign-covariance. *arXiv preprint arXiv:1605.08732*, 2016.
- [HHG12] Ruth Heller, Yair Heller, and Malka Gorfine. A consistent multivariate test of association based on ranks of distances. *Biometrika*, 100(2):503–510, 2012.
- [HHK⁺16] Ruth Heller, Yair Heller, Shachar Kaufman, Barak Brill, and Malka Gorfine. Consistent distribution-free k-sample and independence tests for univariate random variables. *The Journal of Machine Learning Research*, 17(1):978–1031, 2016.
- [HJD06] Frank E Harrell Jr and Maintainer Charles Dupont. The Hmisc package. *R package version*, 3(0-12):3, 2006.
- [HKM⁺13] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, Balázs Ráth, and Rudini Menezes Sampaio. Limits of permutation sequences. *Journal of Combinatorial Theory, Series B*, 103(1):93–113, 2013.
- [HKMS11] Carlos Hoppen, Yoshiharu Kohayakawa, Carlos Gustavo Moreira, and Rudini Menezes Sampaio. Testing permutation properties through subpermutations. *Theoretical Computer Science*, 412(29):3555–3567, 2011.
- [Hoe48] Wassily Hoeffding. A non-parametric test of independence. *The annals of mathematical statistics*, pages 546–557, 1948.
- [Hof17] Lisa Hofer. A central limit theorem for vincular permutation patterns. *arXiv preprint 1704.00650*, 2017.
- [HS14] Yijie Han and Sanjeev Saxena. Parallel algorithms for testing length four permutations. In *2014 Sixth International Symposium on Parallel Architectures, Algorithms and Programming*, pages 81–86. IEEE, 2014.
- [HS18] Yijie Han and Sanjeev Saxena. Algorithms for testing occurrences of length 4 patterns in permutations. *Journal of Combinatorial Optimization*, 35(1):189–208, 2018.
- [HWC13] Myles Hollander, Douglas A Wolfe, and Eric Chicken. *Nonparametric statistical methods*, volume 751. John Wiley & Sons, 2013.
- [Iba97] Louis Ibarra. Finding pattern matchings for permutations. *Information Processing Letters*, 61(6):293–295, 1997.
- [Jan84] Svante Janson. The asymptotic distributions of incomplete U-statistics. *Probability Theory and Related Fields*, 66(4):495–505, 1984.
- [JK17] Vít Jelínek and Jan Kynčl. Hardness of permutation pattern matching. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 387–396. Society for Industrial and Applied Mathematics, 2017.
- [JNZ15] Svante Janson, Brian Nakamura, and Doron Zeilberger. On the asymptotic statistics of the number of occurrences of multiple permutation patterns. *Journal of Combinatorics*, 6:117–143, 2015.
- [Ken38] Maurice G Kendall. A new measure of rank correlation. *Biometrika*, 30(1):81–93, 1938.
- [Kit11] Sergey Kitaev. *Patterns in Permutations and Words*. Springer Science & Business Media, 2011.

- [KK14] Tereza Klimošová and Daniel Král'. Hereditary properties of permutations are strongly testable. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1164–1173. Society for Industrial and Applied Mathematics, 2014.
- [KKRW19] Richard Kenyon, Daniel Král', Charles Radin, and Peter Winkler. Permutations with fixed pattern densities. *Random Structures & Algorithms*, doi:10.1002/rsa.20882:1–31, 2019.
- [Kni66] William R Knight. A computer method for calculating Kendall's tau with ungrouped data. *Journal of the American Statistical Association*, 61(314):436–439, 1966.
- [Knu68] Donald E Knuth. Fundamental algorithms, volume 1 of the art of computer programming. *Addision Wesley, Reading, MA*, 1968.
- [Koz19] László Kozma. Faster and simpler algorithms for finding large patterns in permutations. *arXiv preprint arXiv:1902.08809*, 2019.
- [KP13] Daniel Král' and Oleg Pikhurko. Quasirandom permutations are characterized by 4-point densities. *Geometric and Functional Analysis*, 23(2):570–579, 2013.
- [Mat94] Jiří Matoušek. Geometric range searching. *ACM Computing Surveys (CSUR)*, 26(4):422–461, 1994.
- [MT04] Adam Marcus and Gábor Tardos. Excluded permutation matrices and the Stanley–Wilf conjecture. *Journal of Combinatorial Theory, Series A*, 107(1):153–160, 2004.
- [NRRS19] Ilan Newman, Yuri Rabinovich, Deepak Rajendraprasad, and Christian Sohler. Testing for forbidden order patterns in an array. *Random Structures & Algorithms*, 55(2):402–426, 2019.
- [NWD16] Preetam Nandy, Luca Weihs, and Mathias Drton. Large-sample theory for the Bergsma-Dassios sign covariance. *Electronic Journal of Statistics*, 10(2):2287–2311, 2016.
- [OEI19] OEIS Foundation Inc. The on-line encyclopedia of integer sequences, <http://oeis.org>, 2019.
- [Pri97] Alkes Long Price. Packing densities of layered *patterns. *Dissertation available from ProQuest. AA19727276*, 1997.
- [PS10] Cathleen Battiste Presutti and Walter Stromquist. Packing rates of measures and a conjecture for the packing density of 2413. *Permutation Patterns*, 376:287–316, 2010.
- [Ros75] Murray Rosenblatt. A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *The Annals of Statistics*, pages 1–14, 1975.
- [Sch61] Craige Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of Mathematics*, 13:179–191, 1961.
- [Spe04] Charles Spearman. The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1):72–101, 1904.
- [SS85] Rodica Simion and Frank W Schmidt. Restricted permutations. *European Journal of Combinatorics*, 6(4):383–406, 1985.
- [SS17] Jakub Sliacan and Walter Stromquist. Improving bounds on packing densities of 4-point permutations. *arXiv:1704.02959*, 2017.
- [Sta11] Richard P Stanley. *Enumerative Combinatorics*, volume 1. Cambridge University Press, 2011.
- [Ste10] Einar Steingrímsson. Generalized permutation patterns – a short survey. *Permutation patterns*, 376:137–152, 2010.
- [SV82] Yossi Shiloach and Uzi Vishkin. An $O(n^2 \log n)$ parallel max-flow algorithm. *Journal of Algorithms*, 3(2):128–146, 1982.
- [WDL16] Luca Weihs, Mathias Drton, and Dennis Leung. Efficient computation of the Bergsma–Dassios sign covariance. *Computational Statistics*, 31(1):315–328, 2016.
- [WDM18] Luca Weihs, Mathias Drton, and Nicolai Meinshausen. Symmetric rank covariances: a generalized framework for nonparametric measures of dependence. *Biometrika*, 105(3):547–562, 2018.
- [Yan70] Takemi Yanagimoto. On measures of association and a related problem. *Annals of the Institute of Statistical Mathematics*, 22(1):57–63, 1970.
- [YS05] V Yugandhar and Sanjeev Saxena. Parallel algorithms for separable permutations. *Discrete Applied Mathematics*, 146(3):343–364, 2005.

CHAIM EVEN-ZOHAR, THE ALAN TURING INSTITUTE, LONDON, NW1 2DB, UK
E-mail address: `chaim@ucdavis.edu`

CALVIN LENG, UNIVERSITY OF SOUTHERN CALIFORNIA, LOS ANGELES, CA 90089, US
E-mail address: `cleng@usc.edu`