Combinatorics, Probability and Computing (2018) 27, 69–83. © Cambridge University Press 2017
doi:10.1017/S0963548317000517

Monotone Subsequences in High-Dimensional Permutations

NATHAN LINIAL[1†] and MICHAEL SIMKIN[2]

[1]School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
(e-mail: nati@cs.huji.ac.il)
[2]Institute of Mathematics and Federmann Center for the Study of Rationality, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
(e-mail: menahem.simkin@mail.huji.ac.il)

Received 11 February 2016; revised 21 August 2017; first published online 16 October 2017

This paper is part of the ongoing effort to study high-dimensional permutations. We prove the analogue to the Erdős–Szekeres theorem: For every $k \geqslant 1$, every order-n k-dimensional permutation contains a monotone subsequence of length $\Omega_k(\sqrt{n})$, and this is tight. On the other hand, and unlike the classical case, the longest monotone subsequence in a random k-dimensional permutation of order n is asymptotically almost surely $\Theta_k(n^{k/(k+1)})$.

2010 *Mathematics subject classification*: Primary 05B15

1. Introduction

The study of monotone subsequences in permutations began with the famous Erdős–Szekeres theorem [5]. Since then numerous proofs and generalizations have emerged (see Steele's survey [14]). We recall the theorem.

Theorem 1.1. *Every permutation in S_n contains a monotone subsequence of length at least $\lceil\sqrt{n}\rceil$, and this is tight: for every n there exists some permutation in S_n in which all monotone subsequences are of length at most $\lceil\sqrt{n}\rceil$.*

In order to derive a high-dimensional analogue of Theorem 1.1 we need to define high-dimensional permutations and their monotone subsequences. If we view a permutation as a sequence of distinct real numbers, it is suggestive to consider sequences of points in $\mathbb{R}^k$, with coordinatewise monotonicity. The following argument is attributed by Kruskal [9] to de Bruijn: repeatedly apply Theorem 1.1 to conclude that every sequence $x_1, x_2, \ldots, x_n \in \mathbb{R}^k$ must have a

† Supported by ERC grant 339096 'High-Dimensional Combinatorics'.

coordinatewise monotone subsequence of length $n^{1/2^k}$, and this is tight up to an additive constant. In [9] Kruskal considers projections of the points to a line and defines the length of the longest monotone subsequence according to the line with the longest such subsequence. Szabó and Tardos [15] consider sequences in $\mathbb{R}^k$ that avoid at least one of the 2^k coordinatewise orderings.

Here we adopt the perspective of [11] of a high-dimensional analogue of permutation matrices, and monotone subsequences are defined by strict coordinatewise monotonicity. We show (Theorem 2.1) that every k-dimensional permutation of order n has a monotone subsequence of length $\Omega_k(\sqrt{n})$, and this is tight up to the implicit multiplicative constant.

A related question, posed by Ulam [16] in 1961, concerns the distribution of H_n^1, the length of the longest increasing subsequence in a random member of S_n. In 1972 Hammersley [6] showed that there exists some $C > 0$ such that $H_n^1 n^{-1/2}$ converges to C in probability. In 1977 Logan and Shepp [12] showed that $C \geqslant 2$ and Vershik and Kerov [17] demonstrated that $C \leqslant 2$. This yields the next theorem.

Theorem 1.2. *Let H_n^1 be the length of the longest increasing subsequence in a uniformly random member of S_n. Then $\lim_{n\to\infty} H_n^1 n^{-1/2} = 2$ in probability.*

This result was famously refined in 1999 by Baik, Deift and Johansson [1], who related the limiting distribution of H_n^1 to the Tracy–Widom distribution.

Using coordinatewise monotonicity, Bollobás and Winkler [3] extended Theorem 1.2 to show that the longest increasing subsequence among n independently random points in $[0,1]^k$ is typically of length $c_k n^{1/k}$ for some $c_k \in (0,e)$. We show (Theorem 4.1) that the longest monotone subsequence of a typical k-dimensional permutation of order n has length $\Theta_k(n^{k/(k+1)})$. A k-dimensional permutation can be viewed as a set of n^k points in $[0,1]^k$, and it is interesting to note this asymptotic match with Bollobás and Winkler's result.

2. Definitions and main results

Note. Throughout the paper all asymptotic expressions are in terms of $n \to \infty$ and k fixed.

As discussed in [11] and [10], we equate a permutation with the corresponding permutation matrix, that is, an $n \times n$ $(0,1)$-matrix in which each row or column (henceforth, *line*) contains a single 1. We correspondingly define an *order-n k-dimensional permutation* as an $[n]^{k+1}$ $(0,1)$-array in which each line contains precisely one 1. A *line* in an $[n]^{k+1}$ array comprises all the positions obtained by fixing k coordinates and varying the remaining coordinate. We denote the set of order-n k-dimensional permutations by L_n^k.

For a given $A \in L_n^k$ and $\alpha \in [n]^k$, there is a unique $t \in [n]$ such that $A(\alpha,t) = 1$. Since t is uniquely defined by α, we can write $t = f_A(\alpha)$. The function f_A has the property that if we fix $k-1$ coordinates and vary the remaining coordinate, the result is a permutation of $[n]$. In fact, the mapping $A \mapsto f_A$ is a bijection between L_n^k and the family of $[n]^k$ arrays in which every line is a permutation of $[n]$. In dimension one this is exactly the identification between permutation matrices and permutations. This shows in particular that two-dimensional permutations, that is, members of L_n^2, are order-n *Latin squares*.

We let G_A denote the *support* of $A \in L_n^k$, that is, the set of $\alpha \in [n]^{k+1}$ such that $A(\alpha) = 1$.

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

The next definition generalizes monotonicity to higher dimensions.

Definition 1. A length-m monotone subsequence in $A \in L_n^k$ is a sequence $\alpha^1, \alpha^2, \ldots, \alpha^m \in G_A$ such that for every $1 \leqslant j \leqslant k+1$ the sequence $\alpha_j^1, \alpha_j^2, \ldots, \alpha_j^m$ is *strictly* monotone.

In dimension one this clearly coincides with the definition of a monotone subsequence in a permutation $\pi \in S_n$.

We are now ready to state a high-dimensional analogue of the Erdős–Szekeres theorem.

Theorem 2.1. *Every member of L_n^k contains a monotone subsequence of length $\Omega_k(\sqrt{n})$. The bound is tight up to the implicit multiplicative constant: for every n and k there exists some $A \in L_n^k$ such that every monotone subsequence in A has length $O_k(\sqrt{n})$.*

The next theorem is a high-dimensional analogue of Theorem 1.2.

Theorem 2.2. *Let H_n^k be the length of the longest monotone subsequence in a uniformly random element of L_n^k. Then $\mathbb{E}[H_n^k] = \Theta_k(n^{k/(k+1)})$ and $H_n^k = \Theta_k(n^{k/(k+1)})$ a.a.s.*

Remark 1. Aside from *strong monotonicity* as in Definition 1, it is interesting to consider *weak monotonicity*. A sequence of pairwise distinct $\alpha^1, \alpha^2, \ldots, \alpha^m$ in $[n]^{k+1}$ is called weakly monotone if it is weakly monotone in every coordinate. In the spirit of the Hales–Jewett theorem one may also consider the case where every coordinate is either strictly monotone or constant.

We strive throughout to deal with the harder of the two cases, namely to prove large lower bounds for strongly monotone subsequences and small upper bounds for the weakly monotone case. The one exception is that the proof of the upper bound in Theorem 2.1 applies only to the strongly monotone case. It remains an interesting open problem to determine the correct upper bound for weakly monotone subsequences.

Remark 2. Note the following symmetries of high-dimensional permutations.

(1) S_{k+1} acts on L_n^k by permuting the coordinates.
(2) For each $1 \leqslant i \leqslant k+1$, the group S_n acts on L_n^k by permuting the values of the ith coordinate of each $A \in L_n^k$. Actions on different coordinates commute, and so this defines an S_n^{k+1}-action on L_n^k.
(3) A special case of (2) is reversal, that is, applying the map $a \mapsto n+1-a$ on the ith coordinate.

Note that actions (1) and (3) preserve monotonicity.

3. A high-dimensional analogue of the Erdős–Szekeres theorem

We begin by proving Theorem 2.1. Due to the Erdős–Szekeres theorem it suffices to consider the case $k \geqslant 2$.

We define two partial orders on $[n]^{k+1}$. Let $\alpha, \beta \in [n]^{k+1}$. We write $\alpha <_1 \beta$ if, for all $1 \leqslant i \leqslant k+1$, $\alpha_i < \beta_i$, and we write $\alpha <_2 \beta$ if, for all $1 \leqslant i \leqslant k$, $\alpha_i < \beta_i$ and $\alpha_{k+1} > \beta_{k+1}$. For $\alpha, \beta \in [n]^k$ we write $\alpha < \beta$ if, for all $1 \leqslant i \leqslant k$, $\alpha_i < \beta_i$.

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

Recall that the *height* $h(P)$ of a poset P is the size of the largest chain in P and its *width* $w(P)$ is the size of its largest antichain. The next lemma is an easy consequence of Dilworth's theorem [4] or Mirsky's theorem [13].

Lemma 3.1. *For every finite poset P there holds $h(P) \cdot w(P) \geqslant |P|$.*

We use Lemma 3.1 to show that if A has no long monotone subsequences, then there is a large $S \subseteq G_A$ that is an anti-chain in both $<_1$ and $<_2$. On the other hand, the next two lemmas give an upper bound on the size of anti-chains common to $<_1$ and $<_2$. This yields the theorem.

Lemma 3.2. *Let X be an $M \times N$ matrix in which every two entries in the same column are distinct. Let S be a set of positions in X such that $X_a = X_b$ for every $a, b \in S$ with a to the left and above b. Then $|S| \leqslant M + 2N$.*

Proof. If either $M = 1$ or $N = 1$, this is obvious. We prove the claim inductively by showing that either S has at most two positions in the rightmost column of X or at most one element in the topmost row of X. Indeed, if S has at least three entries in the rightmost column, then at least two of them, say a and b, are not in the top row. But there are no repetitions in the same column, so $X_a \neq X_b$. It follows that the only element S may have in the top row is at the top-right corner, for any other such element must equal both X_a and X_b, which is impossible. □

We are now ready to prove Theorem 2.1.

Proof. For the lower bound, let $A \in L_n^k$ and consider the $n \times n$ matrix X defined by $X_{a,b} = f_A(a, b, b, \ldots, b)$. We define two partial orders on $[n]^2$. Let $\alpha, \beta \in [n]^2$. We write $\alpha <_1 \beta$ if $\alpha_i < \beta_i$, $i = 1, 2$ and $X_\alpha < X_\beta$. We write $\alpha <_2 \beta$ if $\alpha_i < \beta_i$, $i = 1, 2$ and $X_\alpha > X_\beta$. Clearly, a sequence $\alpha^1 <_1 \alpha^2 <_1 \cdots <_1 \alpha^m$ corresponds to a monotone subsequence in A, and similarly for $<_2$.

Assume $[n]^2$ contains no $<_1$-monotone subsequences of length $r = \lfloor \sqrt{n}/3 \rfloor$. By Lemma 3.1 there is an $<_1$-anti-chain $S_1 \subseteq [n]^2$ of size at least n^2/r. Order S_1 by $<_2$ and let $S \subseteq S_1$ be an anti-chain. S is an anti-chain with respect to both $<_1$ and $<_2$, hence if $\alpha \in S$ is above and to the left of $\beta \in S$ we have $X_\alpha = X_\beta$. Every column in X is a permutation of $[n]$, so X and S satisfy the conditions of Lemma 3.2 and therefore $|S| \leqslant 3n$. This is true for every anti-chain in S_1 and so $w(S_1) \leqslant 3n$. Applying Lemma 3.1 again we conclude that

$$h(S_1) \geqslant \frac{|S_1|}{w(S_1)} \geqslant \frac{n^2}{3nr} \geqslant r = \left\lfloor \frac{\sqrt{n}}{3} \right\rfloor.$$

The height of S_1 is realized by a monotone subsequence of length $h(S_1)$ in A, yielding the lower bound.

For the second part of the theorem, for every n and k we construct $A \in L_n^k$ with all monotone subsequences having length $O(\sqrt{n})$. We first assume n is prime, and use a simple construction similar to one that shows the tightness of the Erdős–Szekeres theorem. We later modify the construction to deal with composite n. Assuming n is prime, let $M = \lfloor \sqrt{n/(k+1)} \rfloor$, and define

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

A as follows:

$$A(\alpha_1,\alpha_2,\ldots,\alpha_{k+1})=1 \iff M\sum_{i=1}^{k}\alpha_i+\alpha_{k+1}=0 \pmod n.$$

Since n is prime it follows easily that A is a k-dimensional permutation.

We will show that if $\alpha,\beta\in G_A$ differ in every coordinate then $\|\alpha-\beta\|_1\geqslant M$. This is sufficient, since if $\alpha^1,\alpha^2,\ldots,\alpha^m\in G_A$ is a monotone subsequence, then for every $1\leqslant j<m$, α^j,α^{j+1} differ on every coordinate and so

$$M(m-1)\leqslant\sum_{j=1}^{m-1}\|\alpha^{j+1}-\alpha^j\|_1.$$

On the other hand, by monotonicity we have

$$\sum_{j=1}^{m-1}\|\alpha^{j+1}-\alpha^j\|_1=\|\alpha^m-\alpha^1\|_1\leqslant(k+1)n.$$

It follows that

$$m\leqslant\sqrt{(k+1)n}+1=O(\sqrt{n}).$$

Assume $\alpha,\beta\in G_A$ differ in every coordinate. We have

$$M\sum_{i=1}^{k}(\alpha_i-\beta_i)+(\alpha_{k+1}-\beta_{k+1})=0 \pmod n.$$

Now $Mx+y=0 \pmod n$ implies either $|y|\geqslant M$, $|x|\geqslant n/M-1\geqslant M$ or $x=y=0$. Setting $x=\sum_{i=1}^{k}(\alpha_i-\beta_i)$ and $y=(\alpha_{k+1}-\beta_{k+1})$, we have by assumption $y\neq 0$ and so $\|\alpha-\beta\|_1\geqslant|x|+|y|\geqslant M$.

In this construction we need M and n to be relatively prime. For composite n this is not necessarily the case, and we offer two remedies. The first is an appeal to number theory to produce $M\approx\sqrt{n/(k+1)}$ coprime to n. It is known [2] that for large x, there is always a prime in the interval $[x-x^{0.525},x]$. Therefore, we can find three distinct primes in an interval

$$\left[\sqrt{\frac{n}{k+1}},(1+o(1))\sqrt{\frac{n}{k+1}}\right].$$

At least one of these must be coprime to n, since their product exceeds n for large n. This implies that all monotone subsequences have length $\leqslant(2+o(1))\sqrt{(k+1)n}$.

The second approach is easy to generalize, as is done in the proof of Theorem 3.4. Take $M=\lfloor\sqrt{n/(k+1)}\rfloor$ as before. Let $g=\gcd(M,n)$ and define the permutation $\pi\in S_n$ as follows (all values are taken modulo n):

$$\pi=\left(M,2M,\ldots,\frac{n}{g}M,1+M,\ldots,1+\frac{n}{g}M,\ldots,g-1+M,\ldots,g-1+\frac{n}{g}M\right).$$

Set

$$f_A(\alpha_1,\alpha_2,\ldots,\alpha_k)=-\pi\left(\sum_{i=1}^{k}\alpha_i\right).$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

Note that if $\gcd(M,n)=1$, this coincides with the construction above. As before, we show that if $\alpha,\beta\in G_A$ differ on all coordinates then $\|\alpha-\beta\|_1\geqslant M$, which is enough.

Assume $\alpha,\beta\in G_A$ differ on all coordinates. We then have

$$M\sum_{i=1}^{k}(\alpha_i-\beta_i)+(\alpha_{k+1}-\beta_{k+1})=r\ (\mathrm{mod}\ n)$$

for some $|r|<g\leqslant M$. If $r=0$ we have the same situation as before, and we may conclude $\|\alpha-\beta\|_1\geqslant M$. Otherwise, by definition of π, we must have either

$$\|\alpha-\beta\|_1\geqslant\left|\sum_{i=1}^{k}(\alpha_i-\beta_i)\right|\geqslant\frac{n}{g}-1\geqslant\frac{n}{M}-1\geqslant M$$

or else $|\alpha_{k+1}-\beta_{k+1}|\geqslant M$. □

Most proofs of Theorem 1.1 actually yield the following, more general statement.

Theorem 3.3. *Let r, s and n be positive integers with $rs<n$. Then every permutation in S_n contains either an increasing subsequence of length $r+1$ or a decreasing subsequence of length $s+1$. The bound is tight: if $rs\geqslant n$ then there is a permutation in S_n with neither an increasing subsequence of length $r+1$ nor a decreasing subsequence of length $s+1$.*

It is possible to extend Theorem 2.1 in a similar fashion. To this end we refine our notion of monotonicity. In dimension one we distinguish between ascending and descending subsequences, and we need something similar in higher dimensions.

Definition 2. A vector $\vec{c}\in\{0,1\}^{k+1}$ induces a partial order $x<_{\vec{c}}y$ on $\mathbb{R}^{k+1}$ as follows: $x<_{\vec{c}}y$ if for every $1\leqslant i\leqslant k+1$ such that $c_i=1$, $x_i<y_i$, and $y_i<x_i$ otherwise.

Theorem 3.4. *Let $\vec{c},\vec{d}\in\{0,1\}^{k+1}$ differ in exactly one coordinate. Let $rs<n/(3(k-1))$. Then every $A\in L_n^k$, contains either a $<_{\vec{c}}$-monotone subsequence of length r or a $<_{\vec{d}}$-monotone subsequence of length s.*

The bound is tight up to the multiplicative constants: If $r,s\geqslant 9(k+10)$ and $rs>5kn$, then there exists $A\in L_n^k$ with neither a $<_{\vec{c}}$-monotone subsequence of length r nor a $<_{\vec{d}}$-monotone subsequence of length s.

Proof. Using the symmetries from Remark 2 we may assume without loss of generality that $\vec{c}=(1,1,\ldots,1)$ and $\vec{d}=(1,1,\ldots,1,0)$.

The proof of the lower bound is similar to the proof of the lower bound in Theorem 2.1, and we provide only a sketch. As in the proof of Theorem 2.1, consider the matrix X and the partial orders $<_1,<_2$. Lemma 3.2 gives an upper bound of $3n$ on the size of any anti-chain under both $<_1$ and $<_2$. Two applications of Lemma 3.1 yield the lower bound.

For the upper bound, assume without loss of generality that $r\geqslant s$. We construct $\pi\in S_n$ and $A\in L_n^k$ as before, with $M=\lfloor s/2k\rfloor$. Let $\alpha^1,\alpha^2,\ldots,\alpha^m\in G_A$ be a $<_{\vec{c}}$-monotone subsequence.

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

Then the sequence is increasing in every coordinate. For all j, if $\alpha_{k+1}^{j+1} - \alpha_k^j < M$ then

$$\sum_{i=1}^{k} (\alpha_i^{j+1} - \alpha_i^j) \geqslant \frac{n}{g} \geqslant \frac{n}{M}.$$

Thus

$$m \leqslant \frac{n}{M} + \frac{kn}{n/M} + 1 = \frac{n}{M} + kM + 1 \leqslant \frac{2kn}{s}\left(1 + \frac{2k}{s}\right) + \frac{s}{2} + 1.$$

Using the assumptions that $r/5k > n/s$ and $r \geqslant s \geqslant 9(k+10)$, we have

$$m \leqslant r\left(\frac{2}{5}\left(1 + \frac{2}{9}\right) + \frac{1}{2} + \frac{1}{r}\right) \leqslant r.$$

Now, let $\alpha^1, \alpha^2, \ldots, \alpha^m \in G_A$ be a $<_{\vec{d}}$-monotone subsequence. For $1 \leqslant j \leqslant m$ define

$$s_j = M \sum_{i=1}^{k} \alpha_i^j.$$

This is an increasing sequence and $s_{j+1} - s_j \geqslant M$ for all j. By definition of A, $\alpha_{k+1}^j = s_j \pmod n + r_j$ for some $0 \leqslant r_j < M$. Because $\alpha_{k+1}^1, \alpha_{k+1}^2, \ldots, \alpha_{k+1}^m$ is decreasing, if for some j, s_j and s_{j+1} fall in the same interval of the form $[dn+1, (d+1)n]$ (for $d \in \mathbb{Z}$), then

$$s_j + r_j > s_{j+1} \implies s_{j+1} - s_j < r_j < M,$$

a contradiction. Therefore the s_j fall into distinct intervals of the form $[dn+1, (d+1)n]$. But for every j, $0 < s_j \leqslant Mkn$. Since $[0, Mkn]$ contains only $\lceil (Mkn)/n \rceil \leqslant Mk+1$ intervals of length n, we have $m \leqslant Mk+1 \leqslant s/2+1 < s$. □

4. Monotone subsequences in random high-dimensional permutations

As mentioned in the Introduction, the longest monotone subsequence of a random permutation is typically of length $2\sqrt{n}$. In view of the Erdős–Szekeres theorem this means that the random case and the worst case are of the same order of magnitude and differ by only a constant factor. In higher dimensions this is no longer the case. The longest monotone subsequence of a typical element in L_n^k has length $\Theta_k(n^{k/(k+1)})$.

We define the random variable H_n^k – the length of the longest monotone subsequence in a uniformly random element of L_n^k – and prove the next theorem.

Theorem 4.1. *For every $k \in \mathbb{N}$:*

(1) *for every $\varepsilon > 0$,*

$$H_n^k n^{-k/(k+1)} \in \left[\frac{1}{k+1}, e + \varepsilon\right]$$

asymptotically almost surely,

(2)

$$1 - \frac{\ln k + 1}{k+1} - o_k(1) \leqslant \mathbb{E}[H_n^k n^{-k/(k+1)}] \leqslant e + o_k(1).$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

There are 2^{k+1} distinct order types of monotone subsequences, indexed by binary vectors $\vec{c} \in \{0,1\}^{k+1}$. By reversing some of the coordinates (operation (3) in Remark 2) we see that the distribution of the longest $<_{\vec{c}}$-monotone subsequence in a random element of L_n^k is independent of $\vec{c}$. Thus it suffices to prove Theorem 4.1 for $<_{(1,1,\ldots,1)}$-monotone subsequences. For brevity of notation we write $<$ in place of $<_{(1,1,\ldots,1)}$.

The following lemmas are useful in dealing with uniformly random elements of L_n^k.

Lemma 4.2. *Given $A \in L_n^k$ and $\pi = (\pi_1, \pi_2, \ldots, \pi_{k+1}) \in S_n^{k+1}$, let $\pi(A) \in L_n^k$ be the k-dimensional permutation given by*

$$\pi(A)(x_1, x_2, \ldots, x_{k+1}) = A(\pi_1(x_1), \pi_2(x_2), \ldots, \pi_{k+1}(x_{k+1}))$$

(equivalently, $\pi(A)$ is obtained by permuting the ith coordinate of G_A according to π_i^{-1}). If A is chosen uniformly at random from L_n^k and π is independently chosen from any *distribution on S_n^{k+1}, then $\pi(A)$ is uniformly distributed in L_n^k.*

Proof. This follows immediately from the fact that S_n^{k+1} acts on L_n^k in the way described. □

Lemma 4.3. *Let $\alpha^1, \alpha^2, \ldots, \alpha^m \in [n]^{k+1}$ be a weakly monotone sequence of positions. For a uniformly drawn $A \in L_n^k$,*

$$\mathbb{P}[A(\alpha^1) = A(\alpha^2) = \cdots = A(\alpha^m) = 1] \leqslant \frac{(n-m)!}{n!}.$$

Proof. Assume without loss of generality that the sequence is weakly monotone according to $<$.

We define a distribution $\mathcal{D}$ on S_n^{k+1} such that if $\pi \sim \mathcal{D}$ and A is drawn independently and uniformly from L_n^k, then

$$\mathbb{P}[\pi(A)(\alpha^1) = \pi(A)(\alpha^2) = \cdots = \pi(A)(\alpha^m) = 1] \leqslant \frac{(n-m)!}{n!}.$$

The conclusion follows from Lemma 4.2.

In order to define $\mathcal{D}$ we construct distributions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_m$ on S_n^{k+1}, and we let $\pi = \pi_m \pi_{m-1} \cdot \ldots \cdot \pi_1$ where for each i, π_i is drawn independently from $\mathcal{D}_i$. We then define $\pi(A)$ via

$$A \to A_1 = \pi_1(A) \to A_2 = \pi_2(A_1) \to \cdots \to A_m = \pi_m(A_{m-1}) = \pi(A).$$

We will define the distributions $\mathcal{D}_i$ such that the following properties hold.

- For all $1 \leqslant i < j \leqslant m$, $A_j(\alpha^i) = A_i(\alpha^i)$, so the value at position α^i remains fixed from stage i onward.
- For $1 \leqslant i \leqslant m$,

$$\mathbb{P}[A_i(\alpha^1) = A_i(\alpha^2) = \cdots = A_i(\alpha^i) = 1] \leqslant \frac{(n-i)!}{n!}.$$

Let $\mathcal{D}_1$ be uniformly distributed on $S_n \times \{I\}^k$, where $I \in S_n$ is the identity element. There is a unique x such that $A(x, \alpha_2^1, \ldots, \alpha_{k+1}^1) = 1$, and therefore

$$\mathbb{P}[A_1(\alpha^1) = 1] = \mathbb{P}[A(\pi_1(\alpha_1^1), \alpha_2^1, \ldots, \alpha_{k+1}^1) = 1] = \mathbb{P}[\pi_1(\alpha_1^1) = x] = \frac{1}{n}.$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

Now suppose that $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_i$ are already defined and have the properties above. The sequence $\alpha^1, \alpha^2, \ldots, \alpha^m$ is weakly increasing so there exists some coordinate $1 \leqslant j \leqslant k+1$ such that $\alpha^i_j < \alpha^{i+1}_j$. Let $T \subseteq S_n$ be the set of permutations that fix $\{\alpha^1_j, \alpha^2_j, \ldots, \alpha^i_j\}$, and let $\mathcal{D}_{i+1}$ be the uniform distribution on $\{I\}^{j-1} \times T \times \{I\}^{k+1-j}$. We write $\pi_{i+1} = (I, \ldots, I, \tau, I, \ldots, I)$ and verify the properties above.

- For $1 \leqslant \ell \leqslant i$, by definition
$$A_{i+1}(\alpha^\ell) = A_i(\alpha^\ell_1, \ldots, \alpha^\ell_{j-1}, \tau(\alpha^\ell_j), \alpha^\ell_{j+1}, \ldots, \alpha^\ell_{k+1}).$$
But τ fixes α^ℓ_j, so $A_{i+1}(\alpha^\ell) = A_i(\alpha^\ell) = A_\ell(\alpha^\ell)$, where the last equality follows by induction.
- We have
$$\begin{aligned}\mathbb{P}[A_i(\alpha^1) = A_i(\alpha^2) = \cdots = A_{i+1}(\alpha^{i+1}) = 1] \\ = \mathbb{P}[A_{i+1}(\alpha^{i+1}) = 1 | A_{i+1}(\alpha^1) = A_{i+1}(\alpha^2) = \cdots = A_{i+1}(\alpha^i) = 1] \\ \times \mathbb{P}[A_{i+1}(\alpha^1) = A_{i+1}(\alpha^2) = \cdots = A_{i+1}(\alpha^i) = 1].\end{aligned}$$
By the inductive assumption,
$$\begin{aligned}\mathbb{P}[A_{i+1}(\alpha^1) = A_{i+1}(\alpha^2) = \cdots = A_{i+1}(\alpha^i) = 1] \\ = \mathbb{P}[A_i(\alpha^1) = A_i(\alpha^2) = \cdots = A_i(\alpha^i) = 1] \leqslant \frac{(n-i)!}{n!}.\end{aligned}$$
Now, $\alpha^j_{i+1} \notin \{\alpha^j_1, \alpha^j_2, \ldots, \alpha^j_i\}$, so that $\tau(\alpha^j_{i+1})$ is distributed uniformly on a set of cardinality $\geqslant n-i$, and is independent of $A_{i+1}(\alpha^1), A_{i+1}(\alpha^2), \ldots, A_{i+1}(\alpha^i)$. Thus
$$\mathbb{P}[A_{i+1}(\alpha^{i+1}) = 1 | A_{i+1}(\alpha^1) = A_{i+1}(\alpha^2) = \cdots = A_{i+1}(\alpha^i) = 1] \leqslant \frac{1}{n-i}.$$
We conclude that
$$\mathbb{P}[A_i(\alpha^1) = A_i(\alpha^2) = \cdots = A_{i+1}(\alpha^{i+1}) = 1] \leqslant \frac{1}{n-i}\frac{(n-i)!}{n!} = \frac{(n-(i+1))!}{n!},$$
as desired. □

We first prove the upper bounds in Theorem 4.1.

Proposition 4.4.
(1) *For every* $\varepsilon > 0$ *there holds* $\mathbb{P}[H^k_n n^{-k/(k+1)} > e + \varepsilon] = o(1)$.
(2) $\mathbb{E}[H^k_n] n^{-k/(k+1)} \leqslant e + o(1)$.

Proof. We bound the expected number of length-m (weakly) monotone subsequences in a random k-dimensional permutation. For any increasing sequence of positions $\alpha = \alpha^1, \alpha^2, \ldots, \alpha^m \in [n]^{k+1}$ and $A \in L^k_n$, we define
$$X_\alpha(A) = \begin{cases} 1 & A(\alpha^1) = A(\alpha^2) = \cdots = A(\alpha^m), \\ 0 & \text{otherwise.} \end{cases}$$
By Lemma 4.3,
$$\mathbb{E}[X_\alpha(A)] = \mathbb{P}[X_\alpha(A) = 1] \leqslant \frac{(n-m)!}{n!}$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

for a uniform $A \in L_n^k$. Let S be the set of all length-m increasing sequences of positions in $[n]^k$. Clearly,

$$|S| \leqslant \binom{n+m-1}{m}^{k+1},$$

so by linearity of expectation,

$$\begin{aligned}\mathbb{P}[H_n^k \geqslant m] &= \mathbb{P}\left[\sum_{\alpha\in S} X_\alpha(A) > 0\right] \leqslant \mathbb{E}\left[\sum_{\alpha\in S} X_\alpha(A)\right] \\ &\leqslant \binom{n+m-1}{m}^{k+1} \frac{(n-m)!}{n!} \leqslant \left(\frac{e(n+m)}{m}\right)^{(k+1)m} \frac{1}{(n-m)^m}.\end{aligned}$$

Let $c = e+\varepsilon$ for some $\varepsilon > 0$, and let $m = \lceil cn^{k/(k+1)}\rceil$. Then

$$\begin{aligned}\mathbb{P}[H_n^k n^{-k/(k+1)} > c] = \mathbb{P}[H_n^k \geqslant m] &\leqslant \left((1+o(1))e^{k+1}\frac{n^k}{m^{k+1}}\right)^m \\ &\leqslant \left((1+o(1))\frac{e}{c}\right)^{(k+1)cn^{k/(k+1)}} = o(1),\end{aligned}$$

proving the first claim in the proposition. Further,

$$\begin{aligned}\mathbb{E}[H_n^k]n^{-k/(k+1)} &\leqslant (m\mathbb{P}[H_n^k < m] + n\mathbb{P}[H_n^k \geqslant m])n^{-k/(k+1)} \\ &\leqslant c + n^{1/(k+1)}\left(\frac{e}{c}\right)^{(k+1)cn^{k/(k+1)}} + o(1) = c + o(1),\end{aligned}$$

which proves the second claim. □

The proof of the lower bounds is more intricate. Fix some $C > 0$ and let $m = \lceil Cn^{1/(k+1)}\rceil$. For $1 \leqslant i \leqslant \lfloor n/m\rfloor$, let $D_i = [(i-1)m+1, im]^{k+1}$ be the diagonal subcubes of $[n]^{k+1}$. For a uniformly random $A \in L_n^k$, let Z_i be the indicator random variable of the event that A is not all zero on D_i. Clearly, $H_n^k \geqslant \sum_{1\leqslant i\leqslant n/m} Z_i$, since $\alpha < \beta$ if $\alpha \in D_i, \beta \in D_j$, and $i < j$. Indeed we prove lower bounds on H_n^k by bounding $\sum_{1\leqslant i\leqslant n/m} Z_i$. It is convenient to express everything in terms of the random variable

$$Y_n = n^{-k/(k+1)} \sum_{1\leqslant i\leqslant n/m} Z_i.$$

We show that for an appropriate choice of C (see below), Y_n converges in probability to a constant in $(0,1)$. These are our main steps.

(1) Note that $Y_n \leqslant 1/C + o(1)$ (trivially).
(2) Prove that $\mathbb{E}[Y_n] \geqslant C^k/(C^{k+1}+1) - o(1)$ (Proposition 4.6).
(3) Show that if $C < 1$, then $\mathbb{P}[Y_n > C^{k+1}+\varepsilon] = o(1)$ for every $\varepsilon > 0$ (Corollary 4.9).
(4) By letting $1 > C > 0$ be the unique solution to $C^k/(1+C^{k+1}) = C^{k+1}$, conclude that $\mathbb{P}[Y_n < C^{k+1}-\varepsilon] = o(1)$ for every $\varepsilon > 0$ (Proposition 4.10). Hence $\lim_{n\to\infty} Y_n = C^{k+1}$ in probability.

In step (1) we assume only that $C > 0$. The claim in step (2) applies to all $C > 0$, and we optimize the bound on $\mathbb{E}[Y_n]$ by a particular choice of C. Step (3) applies to all $1 > C > 0$. Finally in step (4) we assign a value to C to derive the conclusion that Y_n converges in probability to C^{k+1}.

We start with step (2), a lower bound on $\mathbb{E}[Y_n]$.

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

Lemma 4.5. *For* $1 \leqslant i \leqslant n/m$,

$$\mathbb{P}[Z_i = 1] \geqslant \frac{C^{k+1}}{C^{k+1}+1} - o(1).$$

Proof. Let $X_i = \sum_{\alpha \in D_i} A(\alpha)$ be the number of non-zero entries in D_i. Note that $X_i > 0 \iff Z_i = 1$. We prove a lower bound on the probability of this event by a second moment argument.

Clearly,

$$\mathbb{E}[X_i] = \frac{|D_i|}{n} = C^{k+1} + o(1),$$

since $\mathbb{P}[A(\alpha) = 1] = 1/n$ for every $\alpha \in [n]^{k+1}$.

We next seek an upper bound on $\mathbb{E}[X_i^2]$:

$$\mathbb{E}[X_i^2] = \sum_{\alpha,\beta \in D_i} \mathbb{E}[A(\alpha)A(\beta)] = \sum_{\alpha,\beta \in D_i} \mathbb{P}[A(\alpha)A(\beta) = 1].$$

There are m^{k+1} terms with $\alpha = \beta$, each being $1/n$. For $\alpha \neq \beta$, Lemma 4.3 gives

$$\mathbb{P}[A(\alpha)A(\beta) = 1] \leqslant \frac{1}{n(n-1)}.$$

There are fewer than $m^{2(k+1)}$ such pairs $\alpha, \beta \in D_i$, so

$$\mathbb{E}[X_i^2] = \sum_{\alpha,\beta \in D_i} \mathbb{P}[A(\alpha)A(\beta) = 1] \leqslant m^{k+1}\left(\frac{1}{n} + \frac{m^{k+1}}{n(n-1)}\right) = \frac{m^{k+1}}{n}\left(1 + \frac{m^{k+1}}{n-1}\right).$$

Noting that

$$\mathbb{E}[X_i] = \frac{m^{k+1}}{n} = C^{k+1} + o(1),$$

we have

$$\mathbb{E}[X_i^2] \leqslant \mathbb{E}[X_i]\left(1 + \frac{n}{n-1}\mathbb{E}[X_i]\right) = C^{k+1}\left(1 + \frac{n}{n-1}C^{k+1}\right) + o(1).$$

The second moment method yields

$$\mathbb{P}[Z_i = 1] = \mathbb{P}[X_i > 0] \geqslant \frac{\mathbb{E}[X_i]^2}{\mathbb{E}[X_i^2]} = \frac{C^{k+1}}{C^{k+1}+1+o(1)} \geqslant \frac{C^{k+1}}{C^{k+1}+1} - o(1). \qquad \square$$

Proposition 4.6. $\mathbb{E}[Y_n] \geqslant C^k/(C^{k+1}+1) - o(1)$, *and consequently*

$$\mathbb{E}[n^{-k/(k+1)}H_n^k] \geqslant 1 - \frac{\ln k + 1}{k+1} - o(1).$$

Proof. As observed earlier,

$$\mathbb{E}[Y_n] = \mathbb{E}\left[n^{-k/(k+1)} \sum_{1 \leqslant i \leqslant n/m} Z_i\right] = n^{-k/(k+1)} \left\lfloor \frac{n}{m} \right\rfloor \mathbb{P}[Z_i = 1].$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

So, by Lemma 4.5,

$$\mathbb{E}[Y_n] \geqslant \frac{C^k}{C^{k+1}+1} - o(1).$$

For all C, $\mathbb{E}[n^{-k/(k+1)}H_n^k] \geqslant \mathbb{E}[Y_n]$. The optimal bound is attained when $C = k^{1/(k+1)}$, yielding

$$\mathbb{E}[n^{-k/(k+1)}H_n^k] \geqslant \frac{k^{k/(k+1)}}{k+1} - o(1) \geqslant 1 - \frac{\ln k + 1}{k+1} - o(1). \qquad \square$$

To prove the lower bound in Theorem 4.1 part (1), we apply a Chernoff bound to the events $\{Z_i = 1\}_{1 \leqslant i \leqslant n/m}$. To overcome the dependencies among these events we utilize the following version of the Chernoff inequality from [7] (Theorem 1.1).

Theorem 4.7. *Let $0 \leqslant \alpha \leqslant \beta \leqslant 1$ and let $\{X_i\}_{i\in[N]}$ be Boolean random variables such that for all $S \subseteq [N]$,*

$$\mathbb{P}\left[\prod_{i\in X} X_i = 1\right] \leqslant \alpha^{|S|}.$$

Then

$$\mathbb{P}\left[\sum_{i\in[N]} X_i \geqslant \beta N\right] \leqslant e^{-ND(\beta\|\alpha)},$$

where

$$D(\beta \parallel \alpha) = \beta \ln\left(\frac{\beta}{\alpha}\right) + (1-\beta)\ln\left(\frac{1-\beta}{1-\alpha}\right)$$

is the relative entropy function.

Lemma 4.8. *Assume $C < 1$. Let $S \subseteq \{1, 2, \ldots, \lfloor n/m \rfloor\}$. Then*

$$\mathbb{P}\left[\prod_{i\in S} Z_i = 1\right] \leqslant \alpha^{|S|}$$

for all $C^{k+1} < \alpha < 1$ and large enough n.

Proof. Note that $Z_i = 1$ for all $i \in S$ if and only if there exist positions $\{\beta^i\}_{i\in S}$ such that $\beta^i \in D_i$ for all $i \in S$ and $A_{\beta^i} = 1$ for all i. We bound the probability of this occurrence using a union bound.

Let $\{\beta^i\}_{i\in S}$ be positions such that $\beta^i \in D_i$ for all $i \in S$. If the indices in S are taken in order this is a monotone subsequence, and so by Lemma 4.3

$$\mathbb{P}[\wedge_{i\in S} A(\beta^i) = 1] \leqslant \frac{(n-|S|)!}{n!}.$$

There are $m^{(k+1)|S|}$ such coordinate sequences, and so, by a union bound,

$$\mathbb{P}\left[\prod_{i\in S} Z_i = 1\right] \leqslant m^{(k+1)|S|}\frac{(n-|S|)!}{n!} \leqslant \left(\frac{m^{k+1}}{n-|S|}\right)^{|S|}.$$

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

We have

$$|S| \leqslant \frac{n}{m} = \frac{1}{C} n^{k/(k+1)} + o(1).$$

Thus

$$\mathbb{P}\left[\prod_{i \in S} Z_i = 1\right] \leqslant \left((1+o(1)) \frac{C^{k+1} n}{n - (1/C) n^{k/(k+1)}}\right)^{|S|} = ((1+o(1)) C^{k+1})^{|S|},$$

and the result follows. □

Lemma 4.8 allows us to apply Theorem 4.7 to the variables $\{Z_i\}_{1 \leqslant i \leqslant n/m}$ to obtain the next corollary.

Corollary 4.9. *For all $\beta > C^{k+1}$, for large enough n we have*

$$\mathbb{P}[Y_n > \beta] \leqslant \exp(-n^{k/(k+1)} \gamma)$$

for some $\gamma > 0$.

We are now ready to complete the proof of Theorem 4.1.

Proposition 4.10. *Let $1 > C > 0$ be the unique solution to the equation $C(1 + C^{k+1}) = 1$. Then $\mathbb{P}[Y_n < 1/(k+2)] = o(1)$.*

Proof. By Proposition 4.6,

$$\mathbb{E}[Y_n] \geqslant \frac{C^k}{C^{k+1} + 1} - o(1) = C^{k+1} - o(1). \tag{4.1}$$

For an integer n and $0 < x < C^{k+1}$, let $p_n = \mathbb{P}[Y_n \leqslant x]$. Since $Y_n \leqslant 1/C + o(1)$ for every $\varepsilon > 0$,

$$\mathbb{E}[Y_n] \leqslant p_n x + (1 - p_n)(C^{k+1} + \varepsilon) + \left(\frac{1}{C} + o(1)\right) \mathbb{P}[Y_n \geqslant C^{k+1} + \varepsilon]. \tag{4.2}$$

Corollary 4.9 yields

$$\mathbb{P}[Y_n \geqslant C^{k+1} + \varepsilon] = o(1).$$

Combining inequalities (4.1) and (4.2) and rearranging,

$$p_n (C^{k+1} - x) \leqslant \varepsilon (1 - p_n) + o(1).$$

But this holds for all $\varepsilon > 0$, so $\lim_{n \to \infty} p_n = 0$.

The result follows by taking $x = 1/(k+1) < C^{k+1}$. □

5. Concluding remarks and open problems

(1) As mentioned in Section 2, we do not know what the analogous statement of Theorem 2.1 is for weakly monotone subsequences.

https://doi.org/10.1017/S0963548317000517 Published online by Cambridge University Press

(2) What are the best constant factors in Theorems 2.1 and 3.4? For the sake of clarity we have neglected to optimize the constants, and our bounds can certainly be somewhat improved with some additional effort. However, we suspect that getting the correct bounds would require some new ideas. While we find the correct exponent of n in the problems addressed here, we are still unable to determine the dependency of the relevant coefficients on the dimension k. Perhaps the most pressing question of this sort is to derive a sharp result on the existence of long monotone subsequences in Latin squares.

(3) For $A \in L_n^k$ and $\vec{c} \in \{0,1\}^{k+1}$, let $\ell_{\vec{c}}(A)$ be the length of the longest $<_{\vec{c}}$-monotone subsequence in A. Let

$$\ell(A) = (\ell_{\vec{c}}(A))_{\vec{c}\in\{0,1\}^{k+1}}.$$

We seek a better description of the set $\ell_n^k = \{\ell(A) : A \in L_n^k\}$. By Theorem 2.1 we know that $\min_{x\in\ell_n^k} \|x\|_\infty = \Theta(\sqrt{n})$. Theorem 3.4 gives fairly tight sufficient conditions under which we can conclude that $x_{\vec{c}} \geqslant r \vee x_{\vec{d}} \geqslant s$ for $\vec{c},\vec{d} \in \{0,1\}^{k+1}$ that differ in precisely one coordinate.

(4) The proof of Theorem 4.1 uses only a very limited amount of randomness. Recall that L_n^k splits into *isotopy classes* where permutations are reachable from each other by applications of symmetries (2) in Remark 2. That theorem applies even when the high-dimensional permutation is drawn uniformly from a particular isotopy class, rather than from all of L_n^k. Beyond the randomness inherent in these symmetries, we have little insight concerning the structure of random high-dimensional permutations. In our view, it is a major challenge in this field to understand (fully) random high-dimensional permutations. In particular, we do not know how to uniformly sample elements of L_n^k. Even for Latin squares, the best known method is Jacobson and Matthews' Markov chain [8], which is not known to be rapidly mixing.

(5) We believe that Theorem 4.1 can be strengthened, and there exist constants c_k such that $H_n^k n^{-k/(k+1)} \to c_k$ in probability. This is borne out by numerical experiments, which indicate that $H_n^2 n^{-2/3}$ is concentrated in a small interval. We do not know how to prove this, but perhaps an approach based on super-additive ergodic theorems *à la* Hammersley [6] may apply. If these constants c_k do in fact exist, their dependence on k is of interest. We note that analogous results for random points in $[0,1]^k$ are known [3].

Acknowledgements

The authors wish to thank the anonymous referee for her thorough review and insightful comments.

References

[1] Baik, J., Deift, P. and Johansson, K. (1999) On the distribution of the length of the longest increasing subsequence of random permutations. *J. Amer. Math. Soc.* **12** 1119–1178.

[2] Baker, R. C., Harman, G. and Pintz, J. (2001) The difference between consecutive primes II. *Proc. London Math. Soc.* **83** 532–562.

[3] Bollobás, B. and Winkler, P. (1988) The longest chain among random points in Euclidean space. In *Proc. Amer. Math. Soc.* **103** 347–353.
[4] Dilworth, R. P. (1950) A decomposition theorem for partially ordered sets. *Ann. of Math.* **51** 161–166.
[5] Erdős, P. and Szekeres, G. (1935) A combinatorial problem in geometry. *Compositio Math.* **2** 463–470.
[6] Hammersley, J. (1972) A few seedlings of research. In *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 345–394.
[7] Impagliazzo, R. and Kabanets, V. (2010) Constructive proofs of concentration bounds. In *Approximation, Randomization, and Combinatorial Optimization: Algorithms and Techniques* (M. Serna *et al.*, eds), Springer, pp. 617–631.
[8] Jacobson, M. T. and Matthews, P. (1996) Generating uniformly distributed random Latin squares. *J. Combin. Designs* **4** 405–437.
[9] Kruskal, J. B. (1953) Monotonic subsequences. *Proc. Amer. Math. Soc.* **4** 264–274.
[10] Linial, N. and Luria, Z. (2014) On the vertices of the d-dimensional Birkhoff polytope. *Discrete Comput. Geom.* **51** 161–170.
[11] Linial, N. and Luria, Z. (2014) An upper bound on the number of high-dimensional permutations. *Combinatorica* **34** 471–486.
[12] Logan, B. F. and Shepp, L. A. (1977) A variational problem for random Young tableaux. *Adv. Math.* **26** 206–222.
[13] Mirsky, L. (1971) A dual of Dilworth's decomposition theorem. *Amer. Math. Monthly* **78** 876–877.
[14] Steele, J. M. (1995) Variations on the monotone subsequence theme of Erdős and Szekeres. In *Discrete Probability and Algorithms* (D. Aldous *et al.*, eds), Springer, pp. 111–131.
[15] Szabó, T. and Tardos, G. (2001) A multidimensional generalization of the Erdős–Szekeres lemma on monotone subsequences. *Combin. Probab. Comput.* **10** 557–565.
[16] Ulam, S. M. (1961) Monte Carlo calculations in problems of mathematical physics. In *Modern Mathematics for the Engineer: Second Series* (E. F. Beckenbach, ed.), McGraw-Hill, pp. 261–281.
[17] Vershik, A. M. and Kerov, S. V. (1977) Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables. *Doklady Akademii Nauk SSSR* **233** 1024–1027.