Heapability, Interactive Particle Systems, Partial Orders: Results and Open Problems

Gabriel Istrate, Cosmin Bonchiş

HAL Id: hal-01633954
https://inria.hal.science/hal-01633954

Submitted on 13 Nov 2017

Heapability, interactive particle systems, partial orders: results and open problems.

Gabriel Istrate[1,2] and Cosmin Bonchiş[1,2]

[1] Dept. of Computer Science, West University of Timişoara, Timişoara, Romania.
[2] e-Austria Research Institute, Bd. V. Pârvan 4, cam. 045 B, Timişoara, RO-300223, Romania. Invited speaker's email: `gabrielistrate@acm.org`

Abstract. We outline results and open problems concerning partitioning of integer sequences and partial orders into heapable subsequences (previously defined and established by Byers et al.).

Keywords: Heapable sequences, posets.

1   Introduction

Suppose $a_1, a_2, \ldots, a_n$ is a sequence of integers. Can one insert the elements of the sequence, successively, as the leaves of a binary tree that satisfies the min heap property ? This is possible, for instance, for sequence 1 3 2 7 6 5 4 but not for sequence 5 4 3 2 1. Byers et al.[1] (who introduced the notion), called such a sequence *heapable*. They provided a polynomial time algorithm to recognize heapability (though, interestingly, *complete heapability*, i.e. heapability on a complete binary tree is NP-complete).

One can view the notion of heapability as a (parametric) relaxation of the notion of *monotonicity*. Indeed, heapability of a sequence requires the fact that the smallest element comes first. The next two elements may, however, arive in any order and the constraints on element ordering become progressively looser. The view of heapability as a generalization of monotonicity, connects the study of heapable sequences to the rich theory built in connection with longest increasing subsequence [2].

In [3] we studied the partition of random permutations into heapable sequences. Similar results were obtained independently in [4]. Perhaps the most exciting finding was the scaling of the number of classes in a partition of a random permutation into heapable subsequences, conjectured to scale as $\phi \cdot \ln(n)$, with ϕ the golden ratio: in Section 5 we explain and motivate this conjecture.

This extended abstract continues this line of inquiry. We present some results and outline several open questions related to the problem of extending notions related to heapability from numbers to partial orders. More topics will be mentioned in the conference presentation.

2 Preliminaries

A *(binary min-)heap* is a binary tree, not necessarily complete for the purposes of this paper, such that $A[parent[x]] \leq A[x]$ for every non-root node x. If instead of binary we require the tree to be k-ary we get the concept of k-ary min-heap.

A partially ordered set $P = (X, \prec)$ is called *k-heapable* if there exists some k-ary tree T whose nodes are in bijection with the elements of X, such that for every non-root node X_i and parent X_j, $X_j \prec X_i$ and $j < i$. In particular a 2-heapable partial order will simply be called *heapable*.

We easily recover the case of permutations, dealt with in [3], as follows: given permutation $\pi \in S_n$, we define partial order $\prec$ on $\{1, 2, \ldots, n\}$ by $i \prec j$ iff $i < j$ and $\pi[i] < \pi[j]$.

The *height of partial order P*, denoted by $h(P)$, is the length of the longest *chain* (totally ordered subset) of P. The *width of P* is defined as the size of the largest antichain of P. By Dilworth's Theorem [5], $w(P)$ is equal to the smallest number of elemenst in a partition of P into chains. Finally, the *dimension of P* is the smallest number r such that the partial order is the intersection of r permutations.

Example 1. Let $X = \{I_1, I_2, \ldots I_k\}$ be a finite set of closed intervals on the real line, with the partial order $I \preceq J$ given by $end(I) \leq start(J)$. By the Gallai theorems for intervals [6], $height(P)$ is equal to the minimal number of points that *pierce* (i.e. intesect) every interval in P. On the other hand $width(P)$ is equal to the maximum cardinality of a set of intervals with nonempty joint intersection.

We give a parametric generalization of $height(P)$ and $width(P)$ as follows:

Definition 1. *Given an integer $k \geq 1$, a subset $Q \subset P$ is a k-chain if nodes of Q are the vertices of a k-ary $\preceq$-ordered subtree of P (not necessarily induced).*

The k-height of P is defined to be the size of the largest k-ary chain of P. The k-width of P is defined as the minimal number of classes in a partition of P into k-chains.

We will employ random models of partial orders of fixed dimension. A complete discussion is beyond the scope of the paper [7]. Instead, we recall the following popular model $P_d(n)$ [8]: given constant $d \geq 1$ we choose random partial order $\prec$ as the intersection of d permutations $\pi_1, \pi_2, \ldots, \pi_d$ chosen uniformly at random (with repetitions from S_n. In other words, given $i, j \in \{1, 2, \ldots, n\}$ define

$$i \prec j \iff \pi_1(i) < \pi_1(j), \pi_2(i) < \pi_2(j), \ldots, \pi_d(i) < \pi_d(j).$$

An equivalent mode to generate a partial order P from $P_d(n)$ is the following: choose n points $P_1, P_2, \ldots P_n,$, $P_i = (x_1^i, \ldots, x_d^i)$, uniformly at random from the hypercube $[0, 1]^d$. Define

$$i \prec j \iff \pi_1(i) < \pi_1(j), \pi_2(i) < \pi_2(j), \ldots, \pi_d(i) < \pi_d(j).$$

We will refer to this alternate description as model (II).

3 The computational complexity of generalized height and width.

Open Problem 1 *What is the computational complexity of the following decision problem:*

- *[GIVEN:] Partial order $P = (X, \prec)$ and integer $r \geq 1$.*
- *[TO DECIDE:] Can X be partioned into at most r k-chains? That is, is inequality $k\text{-}w(P) \leq r$ true ?*

Even the case $k = 1$ (a.k.a. the *longest heapable subsequence* of a random permutation) is still open [1]. In contrast, the k-width of a finite partial order can be computed in polynomial time:

Theorem 1. *For every fixed $k \geq 1$ there is a polynomial time algorithm that, given finite partial order $P = (X, \preceq)$ as input, computes the value $k\text{-}w(P)$.*

Proof. Define the following boolean integer programming problem: define a variable $X_{p,q}$ for every pair $p \prec q \in P$. Intuitively $X_{p,q} = 1$ if p is the parent of q in the k-chain decomposition of P, 0 otherwise.

Every integral solution to this system correponds to a decomposition of P into k-ary trees: indeed, every node has at most one parent in the decomposition induced by variables $X_{p,q} = 1$, and at most k children.

Since in each tree the number of edges is one less than the number of vertices, in any decomposition of P into k-chains, the number of such chains is $n - \sum_{p \prec q} X_{p,q}$.

So to compute the k-width of P we have to solve the following integer program:

$$
\begin{cases}
max(\sum_{p \prec q} X_{p,q}) \\[2mm]
\sum_{q:p \prec q} X_{p,q} \leq k, \forall p \in X \\
\sum_{p:p \prec q} X_{p,q} \leq 1, \forall q \in X \\[2mm]
X_{p,q} \in \{0, 1\}
\end{cases}
$$

Consider the linear programming relaxation of the system above, obtained by replacing condition $X_{p,q} \in \{0, 1\}$ by $X_{p,q} \geq 0$. The matrix of the system is totally unimodular, since it coincides with the vertex-edge incidence matrix of the bipartite graph induced by partial order $\prec$. Such bipartite matrices are well-known to be totally unimodular [9]. So linear programming will find an integral solution to the system in polynomial time. $\square$

Remark 1. The argument above owes much to a discussion with János Balogh from Szeged: we told him a restricted version of the problem, that of scheduling

intervals on binary trees. This amounts to the setting of Example 1. At the time we had a direct (somewhat complicated) proof of this special case. He came up with a (different but related) argument, using network flows. Subsequently we came with this third proof for the general setting, obviously related to his.

Both our original argument and his extend to the general case, and will be jointly presented somewhere else. In retrospect, the fact that there are several distinct proofs is not surprising: Theorem 1 is obviously related to Dilworth's Theorem, and the three existing proofs (direct, using network flows, using linear programming) can be seen as extensions of the corresponding arguments for proving this latter result.

4   The asymptotic behavior of the average k-height and k-width.

The problem of computing the 1-width of a random partial order of dimension 2 is a variant of the classical problem of computing the longest increasing subsequence of a random permutation. The correct asymptotic behavior is $2\sqrt{n}$, [10], [11], [12], [13] and substantially more is known.

The (1-)width and (1-)height of a partial order have also been studied in other dimensions: notable partial results are due to Winkler [8], who showed that the correct order of magnitude for the height of a partial order of dimension k is $\Theta(n^{1/k})$. Further results were obtained by Brightwell [14].

As for the height, the 1-height of a d-dimensional partial order was considered by Winkler [8], and then determined by Bollobás and Winkler [15] to be approximately $c_k \cdot n^{1/k}$ for some constant $c_k > 0$.

In [3] we gave a simple simple lower bound valid for all values of the k-width(P), where P is a random permutation of width 2. We extend this argument to all dimensions as follows:

Theorem 2. *For every fixed $k, n, d \geq 1$*

$$E_{P \in P_d(n)}[k\text{-}w(P)] \geq \frac{ln^{k-1}(n)}{(k-1)!} \cdot (1 + o(1)). \tag{1}$$

Proof. For $P \in P_d(n)$, generated according to model (II) as a sequence of random points $P = (P_1, P_2, \ldots, P_n) \in [0, 1]^d$ we define the set of its *minima* as

$$Min(P) = \{j \in [n] : P_i < P_j \text{ for no } 1 \leq i < j\}.$$

Clearly k-width(P)$\geq |Min(P)|$. Indeed, every minimum of P must determine the starting of a new heap, no matter what k is. Now we use an inequality proved by Winkler [8]:

$$E_{P \in P_d(n)}[|Min(P)|] \geq \frac{ln^{k-1}(n)}{(k-1)!} \cdot (1 + o(1)).$$

$\square$

Open Problem 2 *Is there a constant $c_{k,d} > 0$ such that*

$$\lim_{n \to \infty} \frac{E_{P \in P_d(n)}[\text{k-}w(P)]}{ln^{k-1}(n)} = c_{k,d} \ ? \tag{2}$$

As for the k-height, a result from Byers et al. can be recast as $h(P) = n - o(n)$ for almost all $\pi \in S_n$. We easily generalize this result to random d-dimensional partial orders as follows:

Theorem 3. *For all $d \geq 2, k \geq 1$ and almost all permutations $P \in P_d(n)$ we have k-$h(P) = n - o(n)$.*

Proof. A straightforward adaptation of the argument of Byers et al. [1]. Rather than with k-dimensional permutations, we will work with random points in $[0,1]^d$ (model II).

First one shows that w.h.p. k-$h(P) = \Omega(n)$, using a similar idea to the one in [1]: we consider division of P into subcubes $[0, 1/2]^d$ and $[1/2, 1]^d$, respectively. Let A_1 be the suborder of P determined by the restriction to the first $n/2$ elements and first subcube. W.h.p. $LHS(A_1) = \Theta(n^{1/d})$. This follows from the result of Bollobás and Winkler [15], together with the result of Bollobás and Brightwell [16], that provides concentration of measure for $LIS(A_1)$.

Now we organize the subsequence A_1 into a k-ary tree W with $\Omega(n^{1/d})$ leaves and continue to add elements of subsequence A_2, corresponding to points in the second half; we assume we add elements greedily, in the first possible subheap rooted at a node of A_1 on the frontier of W, stopping when we can no longer place a node in the tree. With high probability this happens after adding $\Omega(n)$ nodes from A_2: to see this we employ the observation that the stopping of the algorithm implies the existence of a decreasing sequence of A_2 of size $\Omega(n^{1/d})$. We then apply the concentration inequality [16] for $LDS(A_2)$.

For the second, rescaled part of the proof, we search for constants $\alpha, \beta > 0$ such that w.h.p. the subsequence B_1, consisting of points among the first n^α ones that belong to the rectangle $[0, n^{-\beta}]^d$ has w.h.p. k-width $\Omega(n^{1/d+\epsilon})$. For this to happen, we take α, β so that $\alpha - d \cdot \beta > 1/d$. It is always possible to find some positive α, β with this property, e.g. $\alpha = 1 - \frac{1}{2d^2}, \beta = \frac{1}{2d^3}$. Now subsequence B_2 consisting of numbers in the rectangle $[n^{-\beta}, 1]^d$ among the last $n - n^\alpha$ ones has w.h.p. its LDS of size $\Theta(n^{1/d})$. Thus sequence B_2 can w.h.p. be placed in its entirety on the tree W. Ther remaining parallelipipeds have $o(1)$ volume, hence a sublinear number of points. The rest of the details are as in [1]. □

Let us note that a random d-dimensional partial order P can be regarded, by definition, as a subset (thinning) of a $(d-1)$-dimensional partial order Q: if $P_1, P_2, \ldots, P_d$ are the permutations defining P, simply define Q to be the intersection of $P_1, P_2, \ldots, P_{d-1}$. So the previous result can be interpreted as the statement that no constant amount of thinning is enough to reduce the width of a random permutation to sublinear.

5 The special case $d = 2$.

In the special case of heapable sequences and random permutations ($d = 2$) we have better insights on the constants $c_{k,d}$ from the above open problem:

Conjecture 1. We have $c_{2,2} = \phi$, with $\phi = \frac{1+\sqrt{5}}{2}$ the golden ratio. More generally

$$c_{k,2} = \frac{1}{\phi_k}, \tag{3}$$

where ϕ_k is the unique root in $(0, 1)$ of equation $X^k + X^{k-1} + \ldots + X = 1$.

Open Problem 3 *Prove this conjecture.*

In the next session we sketch some of the experimental and nonrigorous theoretical evidence for this result. The calculations are nonrigorous, "physics-like", and have yet to be converted to a rigorous argument.

5.1 The connection with the multiset Hammersley process

One of the most rewarding ways to analyze the asymptotic behavior of the LIS of a random permutation is the connection with a model from Nonequilibrium Statistical Physics called *the Hammersley process*.

The easiest way to describe the Hammersley process is via a sequence of random numbers $X_1, X_2, \ldots, X_n \ldots \in (0, 1)$ (note that this combinatorial description is good for our purposes; the general Hammersley process assumes a unit intensity Poisson process on the real line).

We interpret X_i's as *particles*. At each moment the insertion of a new particle removes (*kills*) the smallest (if any) particle X_j, $X_j > X_i$. Intuitively, particles correspond to pile heads in *patience sorting*, a well-known algorithm for computing LIS. The piles are nondecreasing, hence putting a new particle on a pile with head X_j "kills" X_j. Particles that are the largest at the moment when inserted do not kill any particle but simply start a new pile.

A sequence Y of n random particles corresponds naturally to a random n-dimensional permutation. The live particles in the Hammersley process correspond to piles in patience sorting. Therefore $LIS(Y)$ is equal to the number of live particles.

The correspondance between live particles and trees in an optimal decomposition of a random permutation carries on to the framework of heapability as well, with a twist: the multiset generalization of the Hammersley process (defined in [3] and denoted by HAD_k) sees every particle come with a fixed number of k *lives*. A particle does X_i does not kill outright the smallest particle $X_j > X_i$: it simply removes one of its lives.

The infinite-time limit of the multiset Hammersley process with two lives (so-called *hydrodynamic behavior* [17]) seems experimentally to be the so-called *compound Poisson process*. This can be understood combinatorially as follows:

- At stage n the "typical" configuration of the HAM_2 process is characterized by n particles holding 0,1 or two lives.
- The number of particles holding λ lives, for $\lambda \in \{0, 1, 2\}$ is approximately equal to $d_\lambda \cdot n$, for some constants $0 < d_\lambda < 1$. That is, the global density of particles with λ lives converges asymptotically to d_λ.
- Moreover, particles with λ lives are distributed approximately uniformly at random throughout interval $(0, 1)$, so that the relative densities are valid not only globally, but throughout each bin.

The heuristic explanation given above is confirmed experimentally by Figure 1. Here we have divided interval (0,1) into 200 bins, and we plot the relative densities (for each bin, represented on the x axis as the corresponding point in [0,1]) of average number of particles in that bin holding 0,1,2 lives, respectively. We simulated each realization of the HAM_2 process for 100.000 steps, and average each value over 100 realizations. The densities seem to be approximately constant among bins. Moreover $d_0 = d_2 \sim 0.38...$, whereas $d_1 \sim 0.23...$. End bin differences appear to be simulation artifacts: larger simulations reduce this difference.
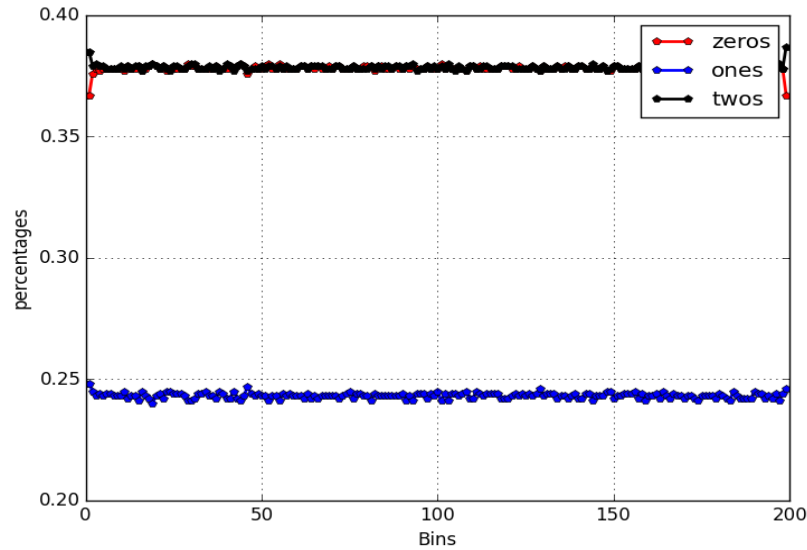

Fig. 1. Relative densities of particles in the HAM_2 process.

But what are constants d_0, d_1, d_2 ? Clearly $d_0 + d_1 + d_2 = 1$. The number of particles with two lives grows by one at each step. On the other hand, except in

the (probabilistically rare) cases the new particle is the largest live one, it takes a life from a particle counted by d_1 or d_2. Assuming well-mixing the probability that it takes a life of particle with two lives is $\frac{d_2}{d_1+d_2}$. We get, therefore, a "mean-field" equation for d_2:

$$d_2 = 1 - \frac{d_2}{d_1 + d_2}. \tag{4}$$

As for d_1, the flow into d_1 has rate $\frac{d_2}{d_1+d_2}$. However, with probability $\frac{d_1}{d_1+d_2}$ there is a flow from d_1 to d_0, decreasing d_1. The "mean-field" equation for d_1 is:

$$d_1 = \frac{d_2 - d_1}{d_1 + d_2} \tag{5}$$

Solving the system of equations for d_0, d_1, d_2 yields

$$d_0 = d_2 = \frac{3 - \sqrt{5}}{2} \sim 0.381\ldots, d_1 = \sqrt{5} - 2 \sim 0.236\ldots \tag{6}$$

a prediction matching the experimental evidence in Figure 1.

So how does this hydrodynamical limit predict the claimed scaling behavior, $E[2 - w(P)] \sim \frac{1+\sqrt{5}}{2}$?

In the compound Poisson process the density of live particles is $d_1 + d_2 = \frac{\sqrt{5}-1}{2}$. If the first n particles were sampled *exactly* from this distribution, the expected value of the largest live particle would be $1 - \frac{\sqrt{5}+1}{2} \cdot \frac{1}{n}$. A new particle would start a new heap precisely when it is larger than all live particles (hence it does not kill anyone). The probability of this happening is $\frac{\sqrt{5}+1}{2} \cdot \frac{1}{n}$. Thus, "on the average", in the first $n + 1$ stages the number of created heaps is

$$1 + \frac{\sqrt{5}+1}{2} \cdot H_n = \phi \ln(n) + O(1),$$

with H_n the Harmonic number. Since process HAM_2 is asymptotically a compound Poisson process, we expect the high-order terms to be correct. Similar but more complicated calculations can be performed in the case $d = 2$ with k arbitrary.

6 High-dimensional permutations.

Linial has initiated [18], under the slogan of "high dimensional combinatorics", a multidimensional analog of permutations. A *p-dimensional permutation of order n* is a $n \times n \times \ldots \times n = [n]^{p+1}$ array of 0/1 values in which each *line* (obtained by setting p indices to values in $[n]$ and leaving free the remaining coordinate) contains *exactly* a one. Ordinary permutations correspond to the one-dimensional case, whereas two-dimensional permutations are essentially latin squares.

Recently, Linal and Simkin [19] have considered notions of monotonicity in high-dimensional permutations, proving a high-dimensonal analog of the

Erdős-Székeres theorem. They studied afterwards the scaling of LIS of a random multidimensional permutation, obtaining the scaling $E[LIS(\pi)] = \Theta(n^{p/p+1})$ for a random p-dimensional permutation.

Open Problem 4 *Study the heapability (2-width and 2-height) of random high-dimensional permutations.*

7 Partition into (un)equal parts: entropy and compression.

So far we have been interested into the partition of a sequence of numbers into a minimal number of k-chains.

One may want, instead, a partition that insists on parts as equal/unequal as possible. Porfilio [4] showed that the problem of dividing a sequence of integers into a number of equal parts is NP-complete.

One may look for the opposite kind of division, that into mostly unbalanced parts. One way to measure the imbalance is via *entropy* of the distribution induced on the poset by a partition into k-chains. Of course, of all distributions with finite support the uniform distribution has the largest entropy. Minimizing entropy is an objective of recent interest in combinatorial optimization [20–26].

Open Problem 5 *Study the complexity of partitioning a poset P into k-chains leading to a distribution of minimal entropy.*

The open problem is easily seen to be related to the minimum entropy coloring problem for interval graphs. Chromatic entropy is a natural measure with important applications to coding [27, 28, 20].

On the other hand we can state the following natural greedy algorithms:

- for $k = 1, d = 2$: compute a longest increasing subsequence L_1 of P using patience sorting (or dynamic programming).
- for other values of pair (k, d): use instead the Byers et al. algorithm for finding a longest heapable subsequence with $n - o(n)$ elements.
- remove L_1 from P and proceed recursively.

Open Problem 6 *Can one give guarantees on the approximation performance of these algorithms ?*

Finally, the decomposition of permutations into components (e.g. runs) forms the basis of the recent theory of data structures and methods for compressing permutations [29, 30] and partial orders. A question that arose during a conversation with Travis Gagie at CPM'2015, and that we would like to state here as an open question is

Open Problem 7 *Is the decomposition of sequences into trees, of the sort employed in computing the 2-width of a partial order, relevant to compression as well ?*

8   Acknowledgments

References

1. J. Byers, B. Heeringa, M. Mitzenmacher, and G. Zervas. Heapable sequences and subseqeuences. In *Proceedings of ANALCO*, pages 33–44, 2011.
2. D. Romik. *The surprising mathematics of longest increasing subsequences*, volume 4. Cambridge University Press, 2015.
3. G. Istrate and C. Bonchiş. Partition into heapable sequences, heap tableaux and a multiset extension of hammersleys process. In *Combinatorial Pattern Matching*, pages 261–271. Springer, 2015.
4. J. Porfilio. A combinatorial characterization of heapability. Master's thesis, Williams College, May 2015. available from http://library.williams.edu/theses/pdf.php?id=790.
5. R. P Dilworth. A decomposition theorem for partially ordered sets. *Annals of Mathematics*, pages 161–166, 1950.
6. A. Gyárfás and J. Lehel. Covering and coloring problems for relatives of intervals. *Discrete Mathematics*, 55(2):167–180, 1985.
7. G. Brightwell. Models of random partial orders. *Surveys in combinatorics*, pages 53–83, 1993.
8. P. Winkler. Random orders. *Order*, 1(4):317–331, 1985.
9. A. Schrijver. *Theory of linear and integer programming*. John Wiley & Sons, 1998.
10. J. M Hammersley et al. A few seedlings of research. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics*, 1972.
11. B. F Logan and L. A Shepp. A variational problem for random Young tableaux. *Advances in mathematics*, 26(2):206–222, 1977.
12. A. Vershik and S. V Kerov. Asymptotics of Plancherel measure of symmetrical group and limit form of Young tables. *Doklady Akademii Nauk SSSR*, 233(6):1024–1027, 1977.
13. D. Aldous and P. Diaconis. Hammersley's interacting particle process and longest increasing subsequences. *Probability theory and related fields*, 103(2):199–213, 1995.
14. G. Brightwell. Random k-dimensional orders: Width and number of linear extensions. *Order*, 9(4):333–342, 1992.
15. B. Bollobás and P. Winkler. The longest chain among random points in euclidean space. *Proceedings of the American Mathematical Society*, 103(2):347–353, 1988.
16. B. Bollobás and G. Brightwell. The height of a random partial order: concentration of measure. *The Annals of Applied Probability*, pages 1009–1018, 1992.
17. P. Groeneboom. Hydrodynamical methods for analyzing longest increasing subsequences. *Journal of Computational and Applied Mathematics*, 142(1):83–105, 2002.
18. N. Linial and Z. Luria. An upper bound on the number of high-dimensional permutations. *Combinatorica*, 34(4):471–486, 2014.
19. N. Linial and M. Simkin. Monotone subsequences in high-dimensional permutations. *arXiv preprint arXiv:1602.02719*, 2016.
20. J. Cardinal, S. Fiorini, and G. Van Assche. On minimum entropy graph colorings. In *Proceedings of ISIT'04, the International Symposium on Information Theory (2004)*, page 43, 2004.

21. E. Halperin and R. Karp. The minimum entropy set cover problem. *Theoretical Computer Science*, 348(2–3):340–350, 2005.
22. J. Cardinal, S. Fiorini, and G. Joraet. Tight results on minimum entropy set cover. *Algorithmica*, 51(1):49–60, 2008.
23. J. Cardinal, S. Fiorini, and G. Joret. Minimum entropy orientations. *Operations Research Letters*, 36(6):680–683, 2008.
24. J. Cardinal, S. Fiorini, and G. Joret. Minimum entropy combinatorial optimization problems. *Mathematical Theory and Computational Practice*, pages 79–88, 2009.
25. M. Kovačević, I. Stanojević, and V. Šenk. On the entropy of couplings. *Information and Computation*, 242:369–382, 2015.
26. G. Istrate, C. Bonchiş, and L. P. Dinu. The minimum entropy submodular set cover problem. In *Language and Automata Theory*, Springer, 2016.
27. N. Alon and A. Orlitsky. Source coding and graph entropies. In *IEEE Trans. Inform. Theory*. Citeseer, 1995.
28. V. Doshi, D. Shah, M. Médard, and M. Effros. Functional compression through graph coloring. *IEEE Transactions on Information Theory*, 56(8):3901–3917, 2010.
29. J. Barbay and J Ian Munro. Succinct encoding of permutations: Applications to text indexing. In *Encyclopedia of Algorithms*, pages 915–919. Springer, 2008.
30. J. Barbay and G. Navarro. On compressing permutations and adaptive sorting. *Theoretical Computer Science*, 513:109–123, 2013.