Distribution Of Maximal Layers Of Random Orders

Indranil Banerjee, Dana Richards

George Mason University
Department Of Computer Science
Fairfax Virginia 22030, USA
ibanerje@gmu.edu,richards@cs.gmu.edu

Abstract. In this paper we address the problem of computing the expected size of the different maximal layers of a random partial order. That is, given a point set $P = \{p_1, ..., p_n\}$ with $p_i \in [0, 1]^k$ picked uniformly at random, we try to determine the expected size of successive maximal layers of P. We present an enumerative expression for this quantity when $k = 2$ and study its behavior for higher dimensions using Monte-Carlo based simulations.

Keywords: Maximal Layers, Random Order, Complexity

1 Introduction

The problem of computing various parameters of random k-dimensional partial orders had seen a lot of attention in the 1980s. Notably authors have studied height, width, number of isolated points, number of linear extensions, size of the first maximal layer and other such properties of a random partial order [1–4]. Some of the more general aspects of these type of results were encapsulated as first order statements on partial orders by Winkler[1]. For example we already have a tight upper and lower bound for the height and width of P. Winkler gave the first non-trivial bound for the expected height h as $\Theta(n^{1/k})$ and showed that the expected width w to satisfy: $\frac{1}{e}n^{1-1/k} \leq w(P) \leq (\ln n)n^{1-1/k}$[1]. Later Brightwell tightened the upper bound to $4kn^{1-1/k}$[2].

However, to the best of our knowledge, there has not been any attempt at computing the distribution of points into different layers of a random order. We only have a bound on the size of the first layer, which is $O(\log^{k-1} n)$ [3]. Let $w_m(n, k)$ be the expected size of the m^{th} layer ($1 \leq m \leq n$). We want to know how $w_m(n, k)$ behaves as a function of n, k and m. We already have a bound on w_1 as stated earlier. In the ideal scenario we want to get similar bounds for $w_m(n, k)$ as we have for w and h. Trivially, $w_m \leq w$ for all m. We also note that $w_n = 0$ unless P is a total order. Since, $w = \Theta(n^{1-1/k})$

and $w_1 = O(\log^{k-1} n)$ it can be conjectured that w_m initially grows and then falls as m goes from 1 to n. This is also evident from our empirical results. Before proceeding on to empirical results we shall first attempt to come up with an expression for $w_m(n, 2)$. In this regard we take an enumerative approach. This is in contrast to a recursive one taken by the authors in [3] to compute $w_1(n, k)$.

The paper is organized as follows: in Section 2 we introduce some preliminary definitions regarding posets and random orders. Section 3 describes our enumerative strategy. Section 4 concludes with empirical results describing the behavior of $w_m(n, k)$ based on Monte-Carlo type simulations.

2 Preliminaries

We denote $P = \{p_1, ..., p_n\}$ as the input set of n points in E^k. The j^{th} coordinate of a point p is denoted as $p[j]$. For any points $p, q \in P$, we define an ordering relation $\succ$, such that $p \succ q$ if $p[j] \geq q[j]\ \forall j \in [1...k]$. Clearly, $(P, \succ)$ defines a partial order. If $p \succ q$ then we say that p precedes (or dominates) q in the partial order and that they are comparable. We say that p and q are incomparable (denoted by $p \parallel q$) if $p \not\succ q$ and $q \not\succ p$. The height h of P is defined as the number of non-empty maximal layers of P and the width w of P as the size of the largest subset of P of mutually incomparable elements. Now we can define maximal layers (or simply layers) of P: Given P its first maximal layer is defined to be the set $\mathcal{M}_1$ of points $q \in P$ such that for any other $p \in P \setminus q$, $p \not\succ q$. The l^{th} maximal layer $\mathcal{M}_l$ is recursively defined as the first maximal layer of remainder of P upon removing from P all the elements of layers from 1 to $l-1$. Note that $\mathcal{M}_l$ could be empty and that the maximum size of any layer is $\leq w$.

There are several different but related models of random partial orders, interested readers are referred to [5]. In this paper we shall work with the model as defined in [1]. That is, we build P by picking points uniformly at random from $[0, 1]^k$. This is equivalent[1] to saying that $(P, \succ)$ is the intersection of k linear orders $T_1 \times ... \times T_k$ where the k-tuple $(T_1, ..., T_k)$ is chosen uniformly at random from $(n!)^k$ such tuples. Here, each T_j is a linear ordering (permutation) of $\{1, 2, ..., n\}$.

3 An Enumerative Strategy To Compute $w_m(n, k)$

The case when $m = 1$ is a special one. We know that an element $x \in P$ belongs to the first layer iff it is not dominated by any element in $P \setminus x$. It does not matter how these elements in $P \setminus x$ are related to each other: that

[1] Since the event that two points in $[0, 1]^k$ share the same coordinate has a null support.

is the sub-poset formed by these elements does not decide the membership of x in $\mathcal{M}_1$. Only their individual relations with x does. However, this notion cannot be extended to compute expected sizes of other layers. From definitions, $x \in \mathcal{M}_m$ iff there exists a chain of size $m-1$ above x. Hence, the relationships between elements that dominate x determines which layer x will belong to. For example the recursive formulation given in [4] fails for $m > 1$ precisely because of this reason.

In this section, using enumeration, we will derive an expression for $w_m(n,2)$. However, the computation depends on whether we can compute the number of permutations having some increasing subsequence of a given length or greater. Let us define $T(n,l)$ to be the number of permutations (of $[1, ..., n]$) whose largest increasing subsequence has length at least l. The following theorem gives an expression for $w_m(n,2)$ in terms of $T(n,l)$.

Theorem 1. *The expected size of* $w_m(n,2)$ *is given by:*

$$w_m(n,2) = W_m(n) - W_{m+1}(n) \tag{1}$$

where, $W_m(n) =$

$$\left(\frac{1}{n!}\right) \sum_{i=1}^{n} (n-i)! \sum_{l=m-1}^{i-1} \binom{i-1}{l} T(l, m-1) \sum_{j=m}^{n} \binom{j-1}{l} (n-j)_{i-1-l}$$

and $(n)_k = \prod_{i=0}^{k-1} (n-i)$ *is the falling factorial.*

Proof. We first sort the vectors in P dimension-wise in *descending* order. Relabel the vectors according to their ranks in the sorted order in the first dimension. Now with respect to this labelling, the (sorted) ordering of elements in the other dimension is just some permutation of $[1, ..., n]$. Here, the first dimension is just the identity permutation (see figure below).

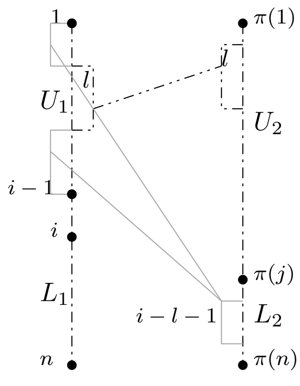

Fig. 1. Visualizing the four sets U_1, L_1, U_2, L_2

Let us consider the i^{th} vector in this labelling, that is, the vector which has a rank i in dimension 1. Let it have rank j in the other. Our strategy is to compute the probability that the vector with rank i in the first dimension will belong to the m^{th} layer or below. Let this probability be $P(i, n)$. Then,

$$W_m(n) = \sum_{i=1}^{n} P(i, n)$$

will be the expected number of points which are on or below the m^{th} layer. Hence, $w_m(n, 2) = W_m(n) - W_{m+1}(n)$.

We compute $P(i, n)$ by counting the number of instances (of the random order P) in which this event occurs. In the first dimension the element has a rank i and in the second it has a rank j. Clearly, for the element to belong some layer $\geq m$, there must exists an increasing subsequence of length at-least $m-1$ in U_2 consisting of elements only from U_1. Here, $U_1 = (1, ..., i-1)$ and $U_2 = (\pi(1), ..., \pi(j-1))$. This in turn can be computed by considering each possible subset of U_1 of size $m-1$ to $i-1$ and asking how many ways we can map it to U_2. For each l-subset of U_1, we can place it in U_2 in $\binom{j-1}{l}$ ways and whose elements can be permuted among themselves in exactly $T(l, m-1)$ ways. Now the rest of U_1 (those elements which were not mapped to U_2) is mapped to L_2 in $(n-j)_{i-1-l}$ ways. Lastly, the elements of L_1 can be placed into $U_2 \cup L_2$ in $(n-i)!$ way. We do this for each subset of U_1 of size $\geq m-1$ and each rank j from m to n. Thus yielding,

$$P(i, n) = \left(\frac{1}{n!}\right) \sum_{l=m-1}^{i-1} \binom{i-1}{l} T(l, m-1) \sum_{j=m}^{n} \binom{j-1}{l} (n-j)_{i-1-l} (n-i)! \tag{2}$$

Which immediately gives the expression for $w_m(n, 2)$ as stated in the theorem. □

In order to compute $w_m(n, 2)$ from Theorem 4 we need to compute $T(n, k)$. Authors have found that expressing $T(n, k)$ in terms of some generating function is quite difficult. In [6] authors prove that l_n converges in distribution to the famous Tracy-Widom distribution [7], where l_n is the size of the largest increasing sequence of a random permutation. Thus it is possible to numerically approximate $T(n, k)$ (and consequently $w_m(n, 2)$) by first numerically approximating the cumulative probability $P(l_n \geq k)$ of the Tracy-Widom distribution.

4 Empirical Approximation of $w_m(n, k)$

Using Monte-Carlo based simulation we study the behavior of $w_m(n, k)$ and $w^*(n, k) = \max_m w_m(n, k)$. Results of these simulations are shown in the figures below.

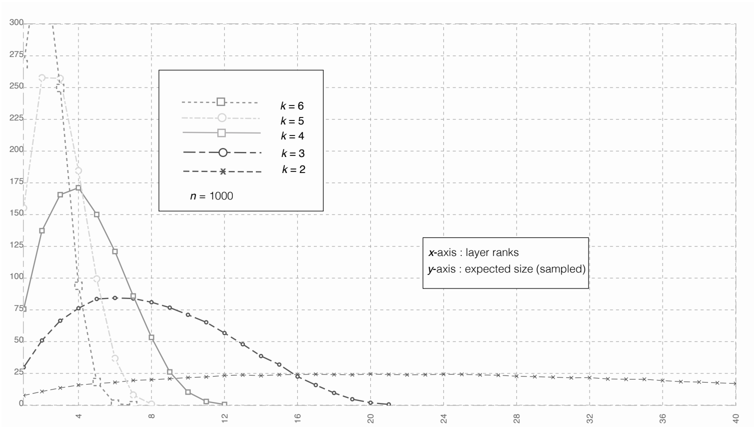

Fig. 2. Plot showing the behavior of $w_m(n, k)$ for different values of k. Here n is fixed to 1000.

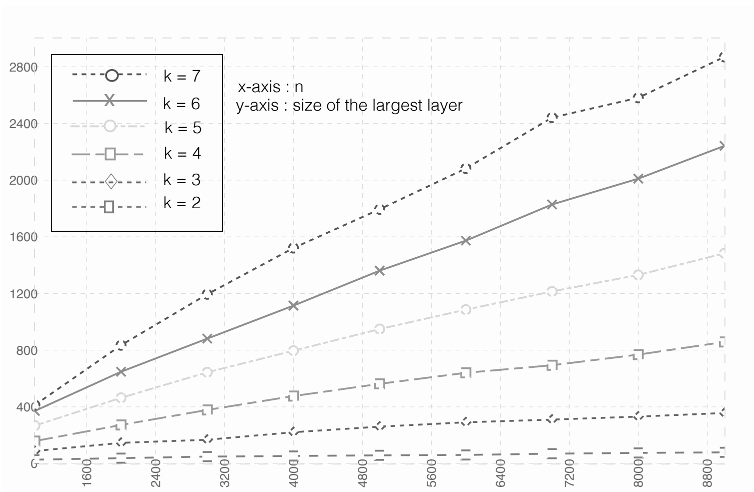

Fig. 3. Here we see that $w^*(n, k)$ tends to a linear function of n as k increases.

Not surprisingly, we see that the expected number of non-empty layers decreases quite sharply, supporting the fact that the expected height is bounded by $O(n^{\frac{1}{k}})$. It is also interesting to note the behavior of $w^*(n,k)$ as k increases. It is evident from Fig. 3 that $w^*(n,k)$ tends to a linear function of n (for a fixed k). This is consistent with the fact that $\lim_{k\to\infty} w^*(n,k) = w$, as the number of layers in P decreases and already know that the expected width w is bounded by $O(n^{1-\frac{1}{k}})$ for fixed k. Thus the shape of a random order looks to be that of an onion.

References

1. Winkler P. Random Orders, Order, 1-4, 317-331 (1985)
2. Brightwell G. Random k-Dimensional Orders: Width and Number of Linear Extensions, Order , 9, 333-342 , (1992)
3. Bentley J. L., Kung H. T., Schkolnick M., Thompson C. D. On the Average Number of Maxima in a Set of Vectors and Applications, Journal of ACM, 25-4, 536-543 (1978)
4. Golin M. J. How many maxima can there be?. Computational Geometry, 2(6), 335-353. (1993)
5. Brightwell G. Models of random partial orders. Surveys in combinatorics, 53-83. (1993)
6. Baik J., Deift P., Johansson K. On the distribution of the length of the longest increasing subsequence of random permutations. Journal of the American Mathematical Society, 12(4), 1119-1178. (1999)
7. Tracy C.A., Widom H. Level-Spacing distributions and the Airy kernel, Comm. Math. Phys., 159, 151-174, (1994).