

LONGEST  $k$ -MONOTONE CHAINS

GERGELY AMBRUS

*Alfréd Rényi Institute of Mathematics*

ABSTRACT. We study higher order convexity properties of random point sets in the unit square. Given  $n$  uniform i.i.d random points, we derive asymptotic estimates for the maximal number of them which are in  $k$ -monotone position, subject to mild boundary conditions. Besides determining the order of magnitude of the expectation, we also prove strong concentration estimates. We provide a general framework that includes the previously studied cases of  $k = 1$  (longest increasing sequences) and  $k = 2$  (longest convex chains).

## 1. HIGHER ORDER CONVEXITY

Let  $X_n$  be a set of  $n$  uniform, independent random points in the unit square  $[0, 1]^2$ . It is a classical and well studied problem to determine the maximal number of points forming a *monotone increasing chain* in  $X_n$ , i.e. a set of points  $p_1, \dots, p_m$  in  $X_n$  so that both the  $x$ -coordinates and the  $y$ -coordinates of  $(p_i)_1^m$  form an increasing sequence. This is the geometric analogue of the famous question of *longest increasing subsequences* in random permutations, first mentioned in 1961 by Ulam [21], which has been studied extensively ever since (see e.g. [1, 6, 18]). Let  $L_n^1$  denote the maximum number of points of  $X_n$  forming a monotone increasing chain. The order of magnitude of the expectation of  $L_n^1$  was determined by Hammersley about half a century ago, with the exact value of the constant in the asymptotics determined five years later by Vershik and Kerov [23], and independently, by Logan and Shepp[16]:

**Theorem 1.1** ([13], [16], [23]). *As  $n \rightarrow \infty$ ,*

$$\mathbb{E}L_n^1 \sim 2n^{1/2}.$$

This result serves as the starting point for the current research. We are going to study point sets which satisfy a more general monotonicity criteria. We start off with a basic concept.

**Definition 1.2.** *A set of points  $p_1, \dots, p_m$  is a chain if their  $x$ -coordinates form a monotone increasing sequence. The length of the chain is the cardinality of the point set, that is,  $m$ .*

Next, one may study points of the random sample  $X_n$  forming a *convex* chain. The motivation is two-fold. On the one hand, convex polygons with vertices among a random sample have been studied extensively in the last 50 years (see e.g. the excellent survey of Bárány [7] or the monograph of Schneider and Weil [19]). On the other hand, given how

---

*E-mail address:* ambrus@renyi.hu.

*Date:* September 30, 2020.

Research of the author was supported by NKFIH grants PD125502 and K116451 and by the Bolyai Research Scholarship of the Hungarian Academy of Sciences.

fruitful and far-reaching the research of the monotone increasing subsequences has been, it is a natural attempt to transfer the results to the convex analogue.

The first steps in that direction were taken in our joint paper with I. Bárány [5], where we studied the order of magnitude of the maximal number of points of  $X_n$  forming a *convex chain* together with  $(0, 0)$  and  $(1, 1)$ , that is, a chain whose points are in convex position. Let  $L_n^2$  denote the maximal number of points of  $X_n$  in a convex chain lying under the diagonal  $y = x$ .

**Theorem 1.3** ([5]). *There exists a positive constant  $\alpha_2$ , so that as  $n \rightarrow \infty$ ,*

$$\mathbb{E}L_n^2 \sim \alpha_2 n^{1/3}.$$

We also proved a limit shape result for the longest convex chains and established upper and lower estimates for  $\alpha_2$ . Alternative proofs to some of the results are given in [3] and [4].

The goal of the present paper is to study the analogous questions for higher order convexity, and to describe a unified framework to the above results. Note that the properties studied above are equivalent to non-negativity of the first (monotone increasing property) and second (convexity property) “discrete derivatives” of the chains. Therefore, it is natural to define higher order convexity along this scheme. Eliás and Matoušek introduced the following concept in order to establish Erdős-Szekeres type results:

**Definition 1.4** (Eliás and Matoušek, [10]). *The  $(k+1)$ -tuple  $(p_1, \dots, p_{k+1})$  of distinct points in the plane is called positive, if it lies on the graph of a function whose  $k$ -th derivative exists everywhere, and is nowhere negative. The points  $(p_1, \dots, p_m)$  in the plane form a  $k$ -monotone chain if their  $x$ -coordinates are monotone increasing, and every  $(k+1)$ -tuple of them is positive.*

Note that in the present paper, “monotone” will always refer to *monotone increasing*. It would be an alternative to use the term “ $k$ -convex”. However, there are already various other concepts existing by that name, thus we stick to “ $k$ -monotone”.

A second, important remark points out the difference between cases of  $k = 1, 2$ , and larger values of  $k$ . In the above definition, positivity of different  $(k+1)$ -tuples may be demonstrated by different functions. For  $k = 1, 2$ , there exists a single monotone/convex function containing all the points on its graph. The same property was conjectured to hold also for larger values of  $k$  by Eliás and Matoušek [10]. However, Rote found a counterexample for  $k = 3$  [10].

An alternative but equivalent definition may be given, see Corollary 2.3 of [10]: a  $(k+1)$ -tuple is positive iff its  $k$ th divided difference is nonnegative, where divided differences are defined as follows. Assume  $p_1, \dots, p_n$  are points in the plane of the form  $p_i = (x_i, y_i)$  (note that here, the  $x$ -coordinates do not necessarily form an increasing sequence). The  $j$ th (forward) divided difference  $\Delta_j(p_i, \dots, p_{i+j+1})$  of the  $(j+1)$ -tuple  $p_i, \dots, p_{i+j}$  is defined recursively by

$$\begin{aligned} \Delta_0(p_i) &:= y_i \\ (1) \quad \Delta_j(p_i, \dots, p_{i+j}) &:= \frac{\Delta_{j-1}(p_{i+1}, \dots, p_{i+j}) - \Delta_{j-1}(p_i, \dots, p_{i+j-1})}{x_{i+j} - x_i} \end{aligned}$$

for every  $0 \leq i \leq n - j$ . Note that divided differences (and, hence, positivity of a  $(k+1)$ -tuple) are invariant under permutations.

Divided differences are used in polynomial approximation; in particular, they provide the coefficients for the summands of Newton’s interpolating polynomial:

**Lemma 1.5** (Newton interpolating polynomial). *Let  $p_1, \dots, p_{k+1}$  be points in the plane with distinct  $x$ -coordinates. Assume that  $p_i = (x_i, y_i)$ . The unique polynomial  $P(x)$  of degree  $k$  whose graph contains all the points  $p_i$  for  $i = 1, \dots, k+1$  may be expressed as*

$$(2) \quad P(x) = \sum_{j=0}^k \Delta_j(p_1, \dots, p_{j+1}) \prod_{i=1}^j (x - x_i).$$

Divided differences are also related to higher order derivatives by the following generalization of the mean value theorem (see [17], Eq. 1.33):

**Lemma 1.6** (Cauchy). *Assume that the points  $p_1, \dots, p_{k+1}$  have increasing  $x$ -coordinates  $a := x_1 < \dots < x_{k+1} =: b$ , and they lie on the graph of a function  $f$  which is  $k$  times differentiable everywhere on the interval  $[a, b]$ . Then there exists  $\xi \in (a, b)$  so that*

$$(3) \quad \Delta_k(p_1, \dots, p_{k+1}) = \frac{f^{(k)}(\xi)}{k!}.$$

Next, we extend the definition of divided differences to multisets of points. For a point  $p$ , introduce the notation

$$p^{\circ i} = \underbrace{\{p, \dots, p\}}_i,$$

that is, the multiset of  $p$  with multiplicity  $i$ . Assume  $p = (x, f(x))$  is a point on the graph of a function  $f$ , which is  $k$  times differentiable at  $x$ . In accordance with (3), we define the  $i$ th divided difference of  $p^{\circ(i+1)}$  with respect to  $f$  by

$$(4) \quad \Delta_i(p^{\circ(i+1)}; f) := \frac{f^{(i)}(x)}{i!}$$

for every  $i \leq k$ . Note that this agrees with the limit of  $\Delta_i(\tilde{p}_1, \dots, \tilde{p}_{i+1})$  as  $\tilde{p}_1, \dots, \tilde{p}_{i+1}$  converge to  $p$  along the graph of  $f$ .

By repeatedly applying (1), we may define divided differences up to order  $k$  with respect to a function  $f$  of any multiset of points lying on the graph of a  $k$ -times differentiable function  $f$ . Therefore, we may extend the  $k$ -monotonicity property to multisets of points with respect to  $f$ , provided that all points of multiplicity larger than 1 lie on the graph of  $f$ .

From now on, we assume that all  $(k+1)$ -tuples of  $X_n$  are in  $k$ -general position, that is, they do not lie on the graph of a polynomial of degree at most  $k-1$ . This property holds with probability 1.

Under this assumption, positivity of  $(k+1)$ -tuples is a *transitive* property:

**Lemma 1.7** ([10], Lemma 2.5). *Assume the distinct points  $p_1, \dots, p_{k+2}$  form a chain, they are in  $k$ -general position, and that both  $(k+1)$ -tuples  $(p_1, \dots, p_{k+1})$  and  $(p_2, \dots, p_{k+2})$  are positive. Then any  $(k+1)$ -element subset of  $(p_1, \dots, p_{k+2})$  is positive.*

In other words, the 2-coloring of  $(k+1)$ -tuples given by positivity/non-positivity is a transitive coloring (defined in [10] and [11]). This will prove to be crucial in the subsequent arguments. In particular, it implies that in order to check  $k$ -monotonicity of a chain, it suffices to check positivity of all of its intervals (i.e. sets of consecutive points) of length  $k+1$ .

## 2. RESULTS

Our goal is to determine the order of magnitude of the maximum number of points in a uniform random sample from the unit square which form a  $k$ -monotone chain. For technical reasons, we also impose boundary conditions on the chain – these conditions will ensure that a  $k$ -monotone chain is also  $l$ -monotone for every  $l \leq k$ . In the case  $k = 1$ , the boundary condition simply requires the monotone chain to start at  $(0, 0)$  and finish at  $(1, 1)$ . For convex chains, the points are required to lie in the triangle below the graph of  $y = x$ . For general  $k$ , we introduce the curve

$$\Gamma_k = (x, x^k), \quad x \geq 0$$

and for every  $x \geq 0$ , let

$$(5) \quad \gamma_k(x) = (x, x^k) \in \Gamma_k.$$

For the point  $\gamma_k(x)$ , we write

$$\Delta_i(\gamma_k(x)^{\circ(i+1)}) := \Delta_i(\gamma_k(x)^{\circ(i+1)}; x^k) = k(k-1)\dots(k-i+1)x^{k-i},$$

that is, we consider  $k$ -monotonicity with respect to  $f(x) = x^k$ . Note that for  $x = 0$  and  $x = 1$ ,  $\gamma_k(x)$  is the same for every  $k$ ; however, ambiguity is avoided by specifying the value of  $k$ .

The setup is the following. Fix  $k \geq 1$ . For any  $n \geq 1$ , as set before, let  $X_n$  be a set of  $n$  i.i.d. uniform random points in the unit square  $[0, 1]^2$ .

**Definition 2.1.** Let  $\mathcal{M}^k(X_n)$  be the set of all chains  $(p_1, \dots, p_m)$  of  $X_n$  so that

$$(6) \quad (\gamma_k(0)^{\circ k}, p_1, p_2, \dots, p_m, \gamma_k(1)^{\circ k})$$

is a  $k$ -monotone chain. Furthermore, let  $L^k(X_n) =: L_n^k$  denote the maximal cardinality of elements of  $\mathcal{M}^k(X_n)$ .

Thus,  $L_n^k$  is a random variable defined on the space of  $n$ -element i.i.d. uniform samples from the square. Note that the definition of divided differences and the boundary condition at  $\gamma_k(0)$  implies that up to order  $k$ , the  $j$ th divided differences of the consecutive  $(j+1)$ -tuples of the chain form a monotone increasing sequence, starting at 0. Therefore,

$$(7) \quad \mathcal{M}^k(X_n) \subset \mathcal{M}^j(X_n)$$

for every  $j \leq k$ .

We first extend Hammersley's result [13] by generalizing Theorem 1.1 and Theorem 1.3.

**Theorem 2.2.** For any  $k \geq 1$  there exists a positive constant  $\alpha_k$  so that

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{k+1}} \mathbb{E}L_n^k = \alpha_k.$$

Furthermore,  $n^{-\frac{1}{k+1}} L^k(X_n) \rightarrow \alpha_k$  almost surely, as  $n \rightarrow \infty$ .

The exact value of the constant is not known except for the case  $k = 1$ , where  $\alpha_1 = 2$  holds. However, we may estimate it from below:

**Proposition 2.3.** For every  $k \geq 1$ ,  $\alpha_k \geq \frac{1}{6}$ .

By utilizing the deviation estimates of Talagrand [20], we obtain the strong concentration property of  $L_n^k$ :

**Theorem 2.4.** *For every  $k \geq 1$ , and for every  $\varepsilon > 0$ ,*

$$\mathbb{P} \left( |L_n^k - \mathbb{E}L_n^k| > \varepsilon n^{\frac{1}{2(k+1)}} \right) \leq 5e^{-\varepsilon^2/5\alpha_k}$$

*holds for every sufficiently large  $n$ .*

Finally, we conjecture that similarly to the  $k = 2$  case [5], the above stochastic concentration property leads to *geometric concentration*: it implies that the *limit shape* of longest  $k$ -monotone chains satisfying the boundary conditions is  $\Gamma_k$ .

**Conjecture 2.5.** *For any  $k \geq 1$ , the longest  $k$ -monotone chains converge in probability to  $\Gamma_k$ . That is, for any  $\varepsilon > 0$ ,*

$$\mathbb{P}(\exists \text{ longest } k\text{-monotone chain with distance } > \varepsilon \text{ from } \Gamma_k) \rightarrow 0$$

*as  $n \rightarrow \infty$ .*

We stated the results for  $X_n$  being chosen from the unit square. As we will see in Section 3, the boundary conditions imply that only a small fraction of the square plays a role here: all members of  $\mathcal{M}^k(X_n)$  lie in  $C_k(0, 1)$ , see Definition 3.3. The area of the region is  $1/(k2^{k-1})$  (see (13)); therefore, switching the base domain from the square to  $C_k(0, 1)$  results in multiplying  $\alpha_k$  by a factor of 2, without changing the order of magnitude of  $\mathbb{E}L_n^k$ .

### 3. GEOMETRIC PROPERTIES

We start with a geometric characterization of positivity. Let  $(p_1, \dots, p_{k+1})$  be a  $(k+1)$ -tuple. We define its *sign* to be the sign of  $\Delta_k(p_1, \dots, p_{k+1})$ .

**Lemma 3.1** ([10], Lemma 2.4.). *Assume that  $\mathcal{P} = (p_1, \dots, p_{k+1})$  is a chain in  $k$ -general position. For any  $i \in [k+1]$ , let  $P_i$  be the unique polynomial of degree  $k-1$  containing all the points  $p_j$ ,  $j \neq i$ . The  $(k+1)$ -tuple  $\mathcal{P}$  has sign  $(-1)^{k-i}$  if  $p_i$  lies below the graph of  $P_i$ , and has sign  $(-1)^{k-i+1}$  if  $p_i$  lies above the graph.*

We may naturally extend the above statement for chains containing multiple points. In the next lemma, we illustrate this for the special case when the chain consists of only 3 points, with the two endpoints having multiplicity larger than 1. The same method can be applied for the general case as long as the points of multiplicity lie on the graph of a function  $f$  with the necessary differentiability properties: if  $p = (x, f(x))$  has multiplicity  $\beta$  in the chain, then the approximating polynomial  $P$  is required to have derivatives agreeing with those of  $f$  up to order  $\beta - 1$  at  $x$ .

Below, we define  $f^{(0)}(x) := f(x)$ .

**Lemma 3.2.** *Let  $q = (a, f(a))$  and  $\bar{q} = (b, f(b))$ ,  $a < b$  be points on the graph of a function  $f$ , which is  $(k-1)$ -times differentiable in the interval  $[a, b]$ . Assume furthermore that  $f^{(k-1)}$  does not vanish on  $[a, b]$ . Let  $1 \leq i \leq k+1$ , and denote by  $\Phi_{i,k}(a, b, f)$  the unique polynomial of degree  $k-1$  which satisfies*

$$(8) \quad \begin{aligned} \Phi_{i,k}^{(j)}(a, b, f)(a) &= f^{(j)}(a) \text{ for every } 0 \leq j \leq i-2, \text{ and} \\ \Phi_{i,k}^{(j)}(a, b, f)(b) &= f^{(j)}(b) \text{ for every } 0 \leq j \leq k-i. \end{aligned}$$

*Assume that for the point  $p \neq q, \bar{q}$ , the  $(k+1)$ -tuple  $\mathcal{P}' = (q^{\circ(i-1)}, p, \bar{q}^{\circ(k-i+1)})$  is a chain. Then  $\mathcal{P}'$  has sign  $(-1)^{k-i}$  if  $p$  lies below the graph of  $\Phi_{i,k}^{(j)}(a, b, f)$ , and has sign  $(-1)^{k-i+1}$  if  $p$  lies above the graph.*

*Proof.* The statement follows directly from Lemma 3.1 by letting the points  $p_1, \dots, p_{i-1}$  converge to  $q$ , and  $p_{i+1}, \dots, p_{k+1}$  converge to  $\tilde{q}$ , along the graph of  $f$ . By Lemma 1.6 and (4), all the divided differences of  $\mathcal{P}$  converge to the corresponding divided differences of  $(q^{\circ(i-1)}, p, \tilde{q}^{\circ(k-i+1)})$ . Moreover, the polynomial  $P_i$  converges to  $\Phi_{i,k}(a, b, f)$ , which is shown by convergence of the derivatives. Therefore, the statement follows.  $\square$

Note that the  $(k-1)$ -times differentiability of  $f$  on the interval  $(a, b)$  is not fully used; we only prescribe the derivatives of  $\Phi_{i,k}(a, b, f)$  at  $a$  and  $b$  up to order  $i-1$  and  $k-i-1$ , respectively.

As a consequence, consider the special  $(k+1)$ -tuple of the form  $(q^{\circ k}, p)$ , and let  $\Phi_k(a, f) := \Phi_{k,k-1}(a, b, f)$  denote the above defined polynomial. If  $k$  is even, then for any other point  $p$ , the  $(k+1)$ -tuple  $(q^{\circ k}, p)$  is positive, if  $p$  lies above the graph of  $\Phi_k(a, f)$ . If  $k$  is odd, then for any other point  $q = (x, y)$ , the  $(k+1)$ -tuple  $(q^{\circ k}, p)$  is positive, if  $x < a$  and  $p$  lies below the graph of  $\Phi_k(a, f)$ , or if  $x > a$ , and  $p$  lies above the graph of  $\Phi_k(a, f)$ .

Next, we apply Lemma 3.2 for the multisets of the form  $(\gamma_k(a)^{\circ k}, p, \gamma_k(b)^{\circ k})$ , where  $0 \leq a < b$ , and  $k \geq 1$ , see (5), (6). Some notations are in order. Let  $a, b \geq 0$ . Define the polynomials

$$(9) \quad \Phi_k(a)(x) = x^k - (x-a)^k$$

$$(10) \quad \Psi_k(a, b)(x) = x^k - (x-a)^{k-1}(x-b).$$

With a slight abuse of notation, we are going to denote the graphs of these polynomials by the same symbols. It will always be clear from the context which meaning do we refer to.

**Definition 3.3.** For  $0 \leq a < b$ , the cell  $C_k(a, b)$  is defined as follows:

- For  $k = 1$ :  $C_1(a, b) = [a, b]^2$ , the square with diagonal vertices  $\gamma_1(a)$  and  $\gamma_1(b)$ ;
- For  $k = 2$ :  $C_2(a, b)$  is the triangle bounded by  $\Phi_k(a)$ ,  $\Phi_k(b)$ , and  $\Psi_k(a, b) = \Psi_k(b, a)$ ;
- For  $k \geq 3$ :  $C_k(a, b)$  is the 4-vertex cell bounded by  $\Phi_k(a)$ ,  $\Phi_k(b)$ ,  $\Psi_k(a, b)$  and  $\Psi_k(b, a)$ .

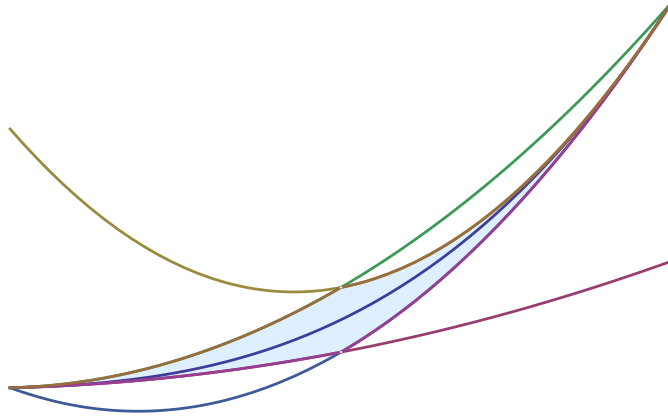


FIGURE 1. The cell  $C_k(a, b)$  between  $\gamma_k(a)$  (left endpoint) and  $\gamma_k(b)$  (right endpoint),  $k$  odd, shaded with blue. The curve  $\Gamma_k$  runs in the middle of the cell. The lower boundary consists of  $\max\{\Phi_k(a), \Psi_k(b, a)\}$ , the upper boundary is defined by  $\min\{\Psi_k(a, b), \Phi_k(b)\}$ .

Let us elaborate on the  $k \geq 2$  case (see Figure 1). When  $k$  is even, the lower boundary of  $C_k(a, b)$  consists of two arcs:  $\Phi_k(a)$  for  $a \leq x \leq (a+b)/2$ , and  $\Phi_k(b)$  for  $(a+b)/2 \leq x \leq b$ . The upper boundary again consists of two arcs:  $\Psi_k(a, b)$  for  $a \leq x \leq (a+b)/2$ , and  $\Psi_k(b, a)$

for  $(a+b)/2 \leq x \leq b$  (for  $k = 2$ , these coincide with each other). When  $k$  is odd, the lower and upper boundaries of  $C_k(a, b)$  are the same as in the even case; however, for  $(a+b)/2 \leq x \leq b$ ,  $\Phi_k(b)$  is the upper boundary, while  $\Psi_k(b, a)$  is the lower boundary. Thus, in any case,  $\gamma_k(a)$  and  $\gamma_k(b)$  are two opposite vertices of  $C_k(a, b)$ , and other two vertices both have  $x$ -coordinates  $(a+b)/2$ .

The importance of the cell  $C_k(a, b)$  is given by the following statement.

**Lemma 3.4.** *Let  $k \geq 1$ , and  $0 \leq a < b$ . The set of points  $p$  in the plane for which*

$$(\gamma_k(a)^{\circ k}, p, \gamma_k(b)^{\circ k})$$

*is a  $k$ -monotone chain is exactly  $C_k(a, b)$ .*

*Proof.* We prove the statement first assuming  $a > 0$ . The case  $a = 0$  may be obtained by a standard limit argument.

Let  $\Phi_{i,k}(a, b, x^k)$  be the polynomial defined by (8) for  $f(x) = x^k$ . By comparing derivatives, we obtain that for every  $1 \leq i \leq k+1$ ,

$$(11) \quad \Phi_{i,k}(a, b, x^k) = x^k - (x-a)^{i-1}(x-b)^{k+1-i}.$$

By Lemma 3.2, our goal is to determine the intersection of the regions above  $\Phi_{i,k}(a, b, x^k)$  for  $i = k+1, k-1, k-3, \dots, 1 + \text{mod}(k, 2)$  and the regions below  $\Phi_{i,k}(a, b, x^k)$  for  $i = k, k-2, k-4, \dots, 2 - \text{mod}(k, 2)$  (see Figure 2). Thus, we have to determine

$$\max_{\substack{i \equiv k+1 \pmod{2} \\ 1 \leq i \leq k+1}} \Phi_{i,k}(a, b, x^k)$$

and

$$\min_{\substack{i \equiv k \pmod{2} \\ 1 \leq i \leq k+1}} \Phi_{i,k}(a, b, x^k)$$

for every  $x \in [a, b]$ . By (11), we obtain that for  $x \in [a, (a+b)/2]$ , the above extrema occur when  $i = k+1$ , and  $i = k$ , respectively, while for  $x \in [(a+b)/2, b]$ , the extrema are taken when  $i = 1, 2$ . Therefore, the boundary of  $C_k(a, b)$  is constituted by the polynomials of the form (9) and (10).  $\square$

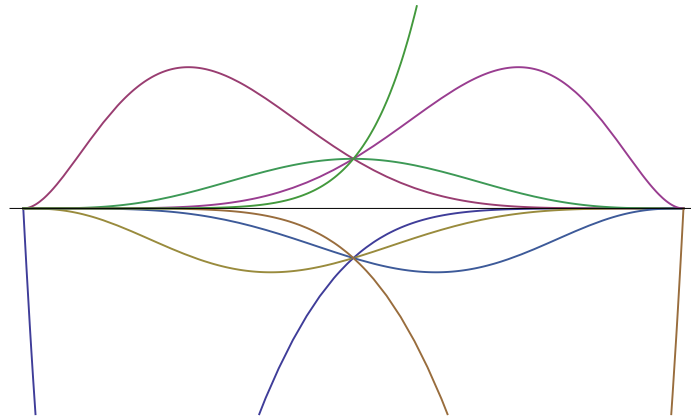


FIGURE 2. Graphs of the polynomials  $x^k - \Phi_{i,k}(a, b, x^k)$ ,  $1 \leq i \leq k+1$ , plotted between  $a$  and  $b$ , in the  $k = 7$  case.

The next lemma states that not only  $C_k(a, b)$  is the location of the  $k$ -monotone chains, but given two  $k$ -monotone chains in neighboring cells, we may concatenate them.

**Lemma 3.5.** *Assume  $0 \leq a < b < c$ , and that the points  $p_1, \dots, p_l$  and  $p_{l+1}, \dots, p_m$  are so that  $(\gamma_k(a)^{\circ k}, p_1, \dots, p_l, \gamma_k(b)^{\circ k})$  and  $(\gamma_k(b)^{\circ k}, p_{l+1}, \dots, p_m, \gamma_k(c)^{\circ k})$  are  $k$ -monotone chains in  $C_k(a, b)$  and  $C_{b,c}$ , respectively. Then*

$$(\gamma_k(a)^{\circ k}, p_1, \dots, p_m, \gamma_k(c)^{\circ k})$$

is a  $k$ -monotone chain in  $C_k(a, c)$ .

*Proof.* By the remark following Lemma 1.7,

$$(\gamma_k(a)^{\circ k}, p_1, \dots, p_l, \gamma_k(b)^{\circ k}, p_{l+1}, \dots, p_m, \gamma_k(c)^{\circ k})$$

is a  $k$ -monotone chain. By the transitivity property provided by Lemma 1.7, we may delete from this chain any point, still maintaining  $k$ -monotonicity. Therefore,

$$(\gamma_k(a)^{\circ k}, p_1, \dots, p_l, \gamma_k(b)^{\circ(k-1)}, p_{l+1}, \dots, p_m, \gamma_k(c)^{\circ k})$$

is also  $k$ -monotone. By iterating the erasure process, we finally erase all copies of  $\gamma_k(b)$ , yielding the statement.  $\square$

Next, we introduce a transformation mapping  $C_k(a, b)$  to  $C_k(c, d)$  which preserves  $k$ -monotonicity, showing the equivalence of the cells  $C_k(a, b)$  with respect to problems regarding  $k$ -monotone chains.

Let  $0 < a < b$  and  $0 < c < d$ , and define the transformation  $T_{a,b,c,d} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  by

$$(12) \quad T_{a,b,c,d}(x, y) = \left( c + (x - a) \frac{d - c}{b - a}, \left( c + (x - a) \frac{d - c}{b - a} \right)^k + (y - x^k) \left( \frac{d - c}{b - a} \right)^k \right).$$

**Lemma 3.6.** *For any  $0 < a < b$  and  $0 < c < d$ , the map  $T_{a,b,c,d}$  preserves  $k$ -monotonicity, keeps  $\Gamma_k$  fixed, and maps the uniform distribution on  $C_k(a, b)$  onto the uniform distribution on  $C_k(c, d)$ .*

*Proof.* Notice that

$$T_{a,b,c,d}(a + t(b - a), (a + t(b - a))^k + \tilde{y}) = \left( c + t(d - c), (c + t(d - c))^k + \tilde{y} \left( \frac{d - c}{b - a} \right)^k \right).$$

Thus, it is immediate that  $T_{a,b,c,d}(\Gamma_k) = \Gamma_k$  (as this corresponds to the case  $\tilde{y} = 0$ ). On the boundary of  $C_k(a, b)$ ,  $\tilde{y}$  is either  $\pm(t(b - a))^k$ ,  $\pm((1 - t)(b - a))^k$ ,  $\pm t^{k-1}(1 - t)(b - a)^k$ , or  $\pm t(1 - t)^{k-1}(b - a)^k$ , which are mapped to the same expressions with  $(d - c)$  in place of  $(b - a)$ . Therefore,  $T_{a,b,c,d}$  maps  $C_k(a, b)$  onto  $C_k(c, d)$  by mapping the boundary curves to the corresponding ones. Measure invariance is seen by writing

$$T_{a,b,c,d} = G_k^{-1} \circ A \circ G_k,$$

where  $G_k : (x, y) \mapsto (x, y - x^k)$ , and  $A$  is an affine map. Thus, it only remains to check the invariance of  $k$ -monotonicity under  $T_{a,b,c,d}$ . To this end, we may write  $T_{a,b,c,d}$  as the composition of three maps:

$$T_{a,b,c,d} = T_{0,d-c,c,d} \circ T_{0,b-a,0,d-c} \circ T_{a,b,0,b-a}.$$

Assume  $f(x)$  has the form  $f(x) = x^k + g(x)$ , then  $f^{(k)}(x) = k! + g^{(k)}(x)$ . The first and third of the above maps do keep this derivative fixed, as  $g(x)$  is preserved by them. Finally, the map  $T_{0,b-a,0,d-c}$  is a linear map scaling the  $x$  and  $y$  coordinates independently, therefore, it preserves the sign of the derivatives. Therefore, using Definition 1.4,  $T_{a,b,c,d}$  preserves positivity of  $(k + 1)$ -tuples, hence it also preserves positivity.  $\square$



We conclude this section by calculating the area of the base cell  $C_k(a, b)$ . By (9), (10) and the discussion afterwards, the distance between the upper and lower boundary of  $C_k(a, b)$  is  $(x - a)^{k-1}(b - x)$ . Therefore,

$$(13) \quad \begin{aligned} A(C_k(a, b)) &= \int_a^{(a+b)/2} (x - a)^{k-1}(b - a) dt + \int_{(a+b)/2}^b (b - x)^{k-1}(b - a) dt \\ &= \frac{(b - a)^{k+1}}{k 2^{k-1}}. \end{aligned}$$

#### 4. THE POISSON MODEL

In order to make the problem more approachable, in this section we switch to the *Poisson model*, that has by now become an industry standard (see e.g. [18]).

Let  $\Pi$  be a planar homogeneous Poisson process with intensity 1. Given any domain  $D$  of area  $A(D)$  in the plane, the number of points of  $\Pi$  in  $D$  has Poisson distribution with parameter  $A(D)$ . That is, its probability mass function is given by

$$(14) \quad \mathbb{P}(|D \cap \Pi| = k) = \frac{A(D)^k e^{-A(D)}}{k!},$$

and its expectation is  $A(D)$ .

We are going to use the following standard tail estimate for Poisson random variables, see Proposition 1 of [12]. Assume that  $X$  has Poisson distribution with parameter  $\lambda$ . Then

$$(15) \quad \mathbb{P}(X \geq m) \leq \frac{m + 1}{m + 1 - \lambda} \mathbb{P}(X = k) = \frac{m + 1}{m + 1 - \lambda} e^{-\lambda} \frac{\lambda^m}{m!}.$$

For arbitrary  $0 \leq a < b$ , let  $N_{\Pi}^k(a, b)$  be the cardinality of  $\Pi \cap C_k(a, b)$ . By (13) and (14),  $N_{\Pi}^k(a, b)$  has a Poisson distribution with parameter (and mean)  $(b - a)^{k+1}/(k 2^{k-1})$ . Moreover, conditioning on the event  $N_{\Pi}^k(a, b) = N$ , the joint distribution of the points of  $\Pi$  falling in  $C_k(a, b)$  is the same as the joint distribution of  $N$  i.i.d. uniform points in  $C_k(a, b)$ .

As the analogue of Definition 2.1, we introduce

**Definition 4.1.** Let  $\mathcal{M}_{\Pi}^k(a, b)$  be the set of all chains  $(p_1, \dots, p_m) \subset \Pi \cap C_k(a, b)$  so that

$$(16) \quad (\gamma_k(a)^{\circ k}, p_1, p_2, \dots, p_m, \gamma_k(b)^{\circ k})$$

is a  $k$ -monotone chain. Furthermore, let  $L^k(a, b)$  denote the maximal cardinality of elements of  $\mathcal{M}_{\Pi}^k(a, b)$ .

By Lemma 3.6 and the invariance property of the Poisson process, the distribution of  $L^k(a, b)$  depends solely on  $(b - a)$ ; therefore, the results below involving  $L^k(0, n)$  remain also valid for the general variables  $L^k(a, b)$ .

Next, we establish the link between the Poisson and the uniform models. By (13), the area of  $C_k(0, n)$  is  $n^{k+1}/(k 2^{k-1})$ , therefore,  $N_{\Pi}^k(0, n)$  has a Poisson distribution with parameter  $n^{k+1}/(k 2^{k-1})$ . On the other hand, let us denote by  $N_n^k$  the number of points of  $X_n$  in  $C_k(0, 1)$ . Then,  $N_n^k$  has binomial distribution with parameters  $n$  and  $1/(k 2^{k-1})$ , and its mean is  $n/(k 2^{k-1})$ . Standard Chernoff type concentration estimates for binomial and Poisson random variables (see e.g. Chapter 2 of [9]) yield the following quantitative bound.

**Proposition 4.2.** For any  $k \geq 1$ , and for any  $c > 0$ ,

$$\mathbb{P} \left( |N_{\Pi}^k(0, n) - N_{n^{k+1}}^k| > c \sqrt{\frac{n^{k+1}}{k 2^{k-1}}} \right) < 4e^{-c^2/3}$$

holds for every sufficiently large  $n$ .

This also implies that the random variable  $L^k(0, n)$  is a good approximation of  $L_{n^{k+1}}^k$ . Applying Proposition 4.2 with  $c = \varepsilon n^{(k+1)/2}$  allows us to transfer the statement of Theorem 2.2 to the Poisson model. We are going to prove the following theorem, which readily implies Theorem 2.2.

**Theorem 4.3.** *For any  $k \geq 1$ , there exists a positive constant  $\alpha_k$  so that*

$$(17) \quad \lim_{n \rightarrow \infty} n^{-1} \mathbb{E} L^k(0, n) = \alpha_k .$$

Furthermore,  $n^{-1} L^k(0, n) \rightarrow \alpha_k$  almost surely, as  $n \rightarrow \infty$ .

First, we need an upper bound on  $\mathbb{E} L^k(0, n)$ . In the uniform model for  $k = 2$ , this is fairly easy to establish. The probability that a random chain of length  $n$  is convex may be calculated exactly [8], based on a beautiful argument of Valtr [22]. The probability that  $n$  uniform independent random points in the unit square form a convex chain is exactly

$$\frac{1}{n!(n+1)!} .$$

The calculation is based on rearranging convex chains while keeping the underlying probability space invariant. Unfortunately, this approach brakes down for larger values of  $k$ , and thus, such a sharp result does not hold in the more general setting. However, we may still prove that the probability of the existence of very long  $k$ -monotone chains is minuscule.

**Lemma 4.4.** *For every  $k \geq 1$  there exists a constant  $c_k$  so that*

$$(18) \quad \mathbb{E} L^k(0, n) < c_k n$$

holds for every  $n \geq 1$ .

*Proof.* As the statement is known for the cases  $k = 1$  and  $k = 2$ , we may assume that  $k \geq 3$ . Also, it is sufficient to prove (18) for sufficiently large values of  $n$ . Let  $C$  be a constant whose value we are going to specify later. Set  $N = n^{k+1}$ . As we noted before,  $N_{\Pi}^k(0, n)$  has Poisson distribution with parameter  $n^{k+1}/(k2^{k-1})$ . Therefore, using (15),

$$(19) \quad \begin{aligned} \mathbb{E} L^k(0, n) &= \int_0^{\infty} \mathbb{P}(L^k(0, n) \geq x) dx \\ &\leq Cn + \int_{Cn}^N \mathbb{P}(L^k(0, n) \geq x) dx + \int_N^{\infty} \mathbb{P}(L^k(0, n) \geq x) dx \\ &\leq Cn + N \mathbb{P}(L^k(0, n) \geq Cn) + \int_N^{\infty} \mathbb{P}(N_{\Pi}^k(0, n) \geq x) dx \\ &\leq Cn + N \mathbb{P}(L^k(0, n) \geq Cn) + 2 \sum_{i=N}^{\infty} \mathbb{P}(N_{\Pi}^k(0, n) = i) \\ &\leq Cn + N \mathbb{P}(L^k(0, n) \geq Cn) + \frac{1}{4^N} . \end{aligned}$$

Thus, it suffices to show that for suitably large  $C$  (depending on  $k$  only),

$$\mathbb{P}(L^k(0, n) \geq Cn) = o(n^{-k}).$$

Call a  $k$ -monotone chain in  $\Pi \cap C_k(0, n)$  *long* if its cardinality is at least  $Cn$ . To every such long  $k$ -monotone chain  $\mathcal{C}$  we assign its *skeleton* as follows. Assume that  $\mathcal{C} = \{p_1, \dots, p_m\}$

(with the points ordered according to their  $x$ -coordinates), where  $m \geq Cn$ . Then

$$(20) \quad \begin{aligned} \text{skel}(\mathcal{C}) &= \{\gamma_k(0)^{\circ k}, p_{\lfloor \frac{m}{n} \rfloor}, p_{2\lfloor \frac{m}{n} \rfloor}, \dots, p_{(n-1)\lfloor \frac{m}{n} \rfloor}, \gamma_k(n)^{\circ k}\} \\ &=: \{\gamma_k(0)^{\circ k}, s_1, \dots, s_{n-1}, \gamma_k(n)^{\circ k}\}, \end{aligned}$$

that is,  $s_i = p_{i\lfloor m/n \rfloor}$  for every  $i = 1, \dots, n-1$ . Also, set  $s_i = \gamma_k(0)$  for  $i \leq 0$  and  $s_j = \gamma_k(n)$  for  $j \geq n$ . Any long chain is cut into  $n$  intervals of length at least  $C$  by its skeleton.

The free part of the skeleton is a chain of length  $n-1$  contained in  $C_k(0, n)$ . The distribution of the long chains in  $\Pi \cap C_k(0, n)$  induces a probability distribution  $\mu$  on the space of skeletons. By the law of total probability,

$$\mathbb{P}(L^k(0, n) \geq Cn) = \int \mathbb{P}(\exists \text{ a long } k\text{-monotone chain } \mathcal{C} \mid \text{skel}(\mathcal{C}) = S) d\mu(S),$$

where the integral is taken over the space of possible skeleta. Thus, (19), implies (18) as long as

$$\mathbb{P}(\exists \text{ a long } k\text{-monotone chain } \mathcal{C} \mid \text{skel}(\mathcal{C}) = S) < o(n^{-k})$$

holds true for every possible skeleton  $S$ , with the constants of the asymptotic estimate being independent of  $S$ . This is what we are going to prove.

Let us now fix  $S$  of the form (20) and assume that  $\mathcal{C}$  is a long  $k$ -monotone chain with  $\text{skel}(\mathcal{C}) = S$ . Let  $p \in \mathcal{C} \setminus \text{skel}(\mathcal{C})$ . For any point  $u \in \mathbb{R}^2$ , let  $x(u)$  denote its  $x$ -coordinate. There exists a unique index  $i$  so that  $x(p) \in [x(s_i), x(s_{i+1})]$ . Then, by Definition 1.4 of  $k$ -monotone chains, the  $(k+1)$ -tuples

$$(s_{i-k+1}, s_{i-k+2}, \dots, s_i, p)$$

and

$$(s_{i-k+2}, s_{i-k+3}, \dots, s_i, p, s_{i+1})$$

are positive. Let  $P_1$  be the unique polynomial of degree  $k-1$  whose graph contains the points  $s_{i-k+1}, s_{i-k+2}, \dots, s_i$ , and similarly, let  $P_2$  be the unique polynomial of degree  $k-1$  whose graph contains the points  $s_{i-k+2}, s_{i-k+3}, \dots, s_i, s_{i+1}$ , possibly using the extended definition for multisets discussed in Section 1. That is, if  $\gamma_k(0)$  appears with multiplicity  $\beta$  among the nodes for  $P_i$  for  $i = 1$  or  $2$ , then the derivatives up to order  $\beta-1$  of  $P_i$  at 0 are required to agree with those of  $x^k$  at 0.

Lemma 3.1 and its generalization to multisets implies that the point  $p$  lies in the region  $R_i$  bounded by the graphs of the polynomials  $P_1$  and  $P_2$  over the interval  $[x(s_i), x(s_{i+1})]$ . Lemma 1.5 and formula (2) shows that

$$\begin{aligned} P_1(x) - P_2(x) &= \left( \Delta_{k-1}(s_{i-k+1}, s_{i-k+2}, \dots, s_i) \right. \\ &\quad \left. - \Delta_{k-1}(s_{i-k+2}, s_{i-k+3}, \dots, s_i, s_{i+1}) \right) \prod_{j=1}^{k-1} (x - x(s_{i+1-j})). \end{aligned}$$

Therefore,

$$(21) \quad \begin{aligned} A(R_i) &= \int_{x(s_i)}^{x(s_{i+1})} |P_1(x) - P_2(x)| \\ &\leq (x(s_{i+1}) - x(s_i))(x(s_{i+1}) - x(s_{i-k+2}))^{k-1} D_i \\ &\leq (x(s_{i+1}) - x(s_{i-k+2}))^k D_i \end{aligned}$$

with

$$D_i = \Delta_{k-1}(s_{i-k+2}, s_{i-k+3}, \dots, s_i, s_{i+1}) - \Delta_{k-1}(s_{i-k+1}, s_{i-k+2}, \dots, s_i).$$

Since

$$\sum_{i=0}^{n-1} x(s_{i+1}) - x(s_{i-k+2}) = \sum_{j=0}^{k-2} x(s_{n-j}) - x(s_{-j}) < kn,$$

there are at least  $(2n)/3$  indices  $i$  in the interval  $[0, n-1]$  so that

$$(22) \quad x(s_{i+1}) - x(s_{i-k+2}) \leq 3k.$$

On the other hand, the  $k$ -monotonicity of  $\mathcal{C}$  implies that  $(\Delta_{k-1}(s_{i-k+1}, s_{i-k+2}, \dots, s_i))_{i=0}^{n+k-1}$  is a monotone increasing sequence, which, by (4), satisfies

$$\Delta_{k-1}(s_{-k+1}, s_{-k+2}, \dots, s_0) = 0$$

and

$$\Delta_{k-1}(s_n, s_{n+1}, \dots, s_{n+k-1}) = kn.$$

Thus, there exist at least  $(2n)/3$  indices  $j$  in the interval  $[0, n-1]$  so that

$$(23) \quad \Delta_{k-1}(s_{j-k+1}, s_{j-k+2}, \dots, s_j) \leq 3k.$$

Combining (22) and (23) with (21), we obtain that there at least  $n/3$  indices  $i \in [0, n-1]$  so that

$$A(R_i) \leq (3k)^{k+1}.$$

In order for  $\mathcal{C}$  to be long, each of these regions must contain at least  $C$  points of  $\Pi$ . Pick such a region  $R$ . By (15) and Stirling's approximation,

$$\mathbb{P}(|R \cap \Pi| \geq C) \leq 2 \left( \frac{e(3k)^{k+1}}{C} \right)^C$$

holds for any sufficiently large  $C$ . Therefore, for any given  $\varepsilon > 0$ , there exists a corresponding  $C$  so that the above probability is bounded above by  $\varepsilon$ . For that choice of  $C$ ,

$$\begin{aligned} \mathbb{P}(\exists \text{ a long } k\text{-monotone chain } \mathcal{C} \mid \text{skel}(\mathcal{C}) = S) &\leq \prod_{i=0}^{n-1} \mathbb{P}(|R_i \cap \Pi| \geq C) \\ &\leq \varepsilon^{n/3}, \end{aligned}$$

where the independence property of the Poisson process is used in the first inequality. The proof is finished by noting that the above expression is of order  $o(n^{-k})$  for sufficiently small values of  $\varepsilon$ , and all the above estimates depend on  $k$  only.  $\square$

## 5. EXPECTATION AND CONCENTRATION ESTIMATES

In this section, we show that the order of magnitude of the length of the longest  $k$ -monotone chains among  $n$  random points is  $n^{1/(k+1)}$ . We are going to prove this in the Poisson model, Theorem 4.3, which implies Theorem 2.2 of the uniform model. The proof builds on Kingman's subadditive ergodic theorem. Below, we present a version of it along with an important extension by Liggett.

**Theorem 5.1** (Kingman's subadditive ergodic theorem with Liggett's extension [14, 15]). *Assume  $X_{n,m}$ ,  $n, m \in \mathbb{N}$ , is a family of random variables satisfying the following conditions:*

- S1)  $X_{l,n} \leq X_{l,m} + X_{m,n}$  whenever  $0 \leq l < m < n$ ;
- S2) For every  $s \geq 0$  integer, the joint distributions of the process  $\{X_{m+s, n+s}\}$  are the same as those of  $\{X_{m,n}\}$ ;
- S3) For each  $n$ ,  $\mathbb{E}|X_{0,n}| < \infty$  and  $\mathbb{E}X_{0,n} > -cn$  for some constant  $c$ .

Then

$$\gamma = \lim_{n \rightarrow \infty} \frac{\mathbb{E} X_{0,n}}{n}$$

exists,

$$X = \lim_{n \rightarrow \infty} \frac{X_{0,n}}{n}$$

exists almost surely, and  $\mathbb{E} X = \gamma$ .

Furthermore, if the stationary sequences  $(X_{in,(i+1)m})_{i=1}^{\infty}$  are ergodic for any  $m \geq 1$ , then  $X = \gamma$  almost surely.

*Proof of Theorem 4.3.* We show that Conditions S1), S2) and S3) of Theorem 5.1 hold for the family of random variables  $X_{m,n} := -L^k(n, m)$ ,  $n, m \in \mathbb{N}$ , see Definition 4.1. Lemma 3.5 shows that

$$L^k(a, c) \geq L^k(a, b) + L^k(b, c)$$

for every  $0 \leq a < b < c$ , showing the validity of S1). The invariance property S2) follows from Lemma 3.6. Finally, S3) is implied by Lemma 4.4.

Therefore, we may apply Theorem 5.1 to obtain that  $\mathbb{E} L^k(0, n) \approx \alpha_k n$  with some positive constant  $\beta_k$ . Moreover,  $(L^k(in, (i+1)n))_{i=1}^{\infty}$  is a sequence of independent, identically distributed random variables, hence it is ergodic. Therefore,  $n^{-1}L^k(0, n)$  converges to  $\alpha_k$  almost surely.  $\square$

Theorem 4.3 and Proposition 4.2 implies that

$$\frac{L_n^k}{n^{1/(k+1)}} \rightarrow \alpha_k$$

almost surely, proving Theorem 2.2.

Next, we derive a lower bound on the constant  $\alpha_k$  of (17).

*Proof of Proposition 2.3.* We are going to prove the statement in the Poisson model by showing that for sufficiently large  $n$ ,

$$\mathbb{E} L^k(0, n) \geq \frac{n}{6}$$

holds.

Set  $a_i = 3i$  for every  $i \in [0, \lfloor n/3 \rfloor]$ . By (13), the area of  $C_k(a_i, a_{i+1})$  (see Definition 3.3) is

$$A(C_k(a_i, a_{i+1})) = \frac{3^k}{k2^{k-1}} > 1.$$

Since the number of points of  $\Pi$  in  $C_k(a_i, a_{i+1})$  has Poisson distribution with parameter  $A(C_k(a_i, a_{i+1}))$ ,

$$\mathbb{P}(|\Pi \cap C_k(a_i, a_{i+1})| = 0) = e^{-A(C_k(a_i, a_{i+1}))} < \frac{1}{e}.$$

Let  $Y$  be the number of cells of the form  $C_k(a_i, a_{i+1})$  in which  $\Pi$  has at least one point. Then  $Y \sim B(\lfloor \frac{n}{3} \rfloor, p)$  with  $p > 1 - 1/e > 1/2$ . Let  $\lambda := \mathbb{E} Y$ , then  $\lambda > n(1 - 1/e)/3$ . By a standard Chernoff-type bound for binomial random variables (see Theorem A.1.12 of [2]),

$$\mathbb{P}\left(Y \leq \lambda - c\sqrt{\lambda \log \lambda}\right) < \lambda^{-c^2/2}.$$

Therefore, for sufficiently large  $n$ ,

$$(24) \quad \mathbb{P}(Y > n/6) \approx 1.$$

Let us now take a point  $p$  of  $\Pi$  in each of the non-empty cells, and let  $\mathcal{S} = \{s_1, \dots, s_Y\}$  be the collection of these points ordered with respect to their  $x$ -coordinates. By the construction,

$$\left( \gamma_k(0)^{\circ k}, s_1, \gamma_k(a_1)^{\circ k}, s_2, \dots, \gamma_k(a_{\lfloor n/3 \rfloor})^{\circ k}, \gamma_k(n)^{\circ k} \right)$$

is a  $k$ -monotone chain, where each  $s_i$  is placed in its corresponding interval so that we obtain a chain. By repeatedly applying Lemma 3.5, we deduce that

$$\left( \gamma_k(0)^{\circ k}, s_1, s_2, \dots, s_Y, \gamma_k(n)^{\circ k} \right)$$

is also a  $k$ -monotone chain. Therefore,  $L^k(0, n) \geq Y$ . The proof is finished by referring to (24), which shows that  $\mathbb{E} L^k(0, n) \geq n/6$ .  $\square$

We finish this section by establishing the exponential concentration estimate for  $L_n^k$ . Theorem 2.4 is straightforward consequence of Talagrand's strong concentration inequality.

**Theorem 5.2** (Talagrand [20]). *Suppose  $Y$  is a real-valued random variable on a product probability space  $\Omega^{\otimes n}$ , and that  $Y$  is 1-Lipschitz with respect to the Hamming distance, meaning that*

$$|Y(x) - Y(y)| \leq 1$$

*whenever  $x$  and  $y$  differ in one coordinates. Moreover assume that  $Y$  is  $f$ -certifiable. This means that there exists a function  $f : \mathbb{N} \rightarrow \mathbb{N}$  with the following property: for every  $x$  and  $b$  with  $Y(x) \geq b$  there exists an index set  $I$  of at most  $f(b)$  elements, such that  $Y(y) \geq b$  holds for every  $y$  agreeing with  $x$  on  $I$ . Let  $m$  denote the median of  $Y$ . Then for every  $s > 0$  we have*

$$\mathbb{P}(Y \leq m - s) \leq 2 \exp\left(\frac{-s^2}{4f(m)}\right)$$

and

$$\mathbb{P}(Y \geq m + s) \leq 2 \exp\left(\frac{-s^2}{4f(m + s)}\right).$$

The conditions of Theorem 5.2 are clearly satisfied by the random variable  $L_n^k$  with the certificate function  $f(b) = b$ , by fixing the points of the longest  $k$ -monotone chain in  $X_n$ . Since  $L_n^k \leq n$ , exponential concentration ensures that the mean and the median are within a distance of  $O(n^{1/2(k+1)})$  of each other. Thus, in the above estimates,  $m \approx \alpha_k n^{1/(k+1)}$ , and setting  $s = \varepsilon n^{1/2(k+1)}$ , we obtain Theorem 2.4.

The same proof yields the analogous concentration estimate for  $L^k(0, n)$ :

**Theorem 5.3.** *For every  $k \geq 1$ , and for every  $\varepsilon > 0$ ,*

$$\mathbb{P}\left(|L^k(0, n) - \mathbb{E} L^k(0, n)| > \varepsilon \sqrt{n}\right) \leq 5e^{-\varepsilon^2/5\alpha_k}.$$

*holds for every sufficiently large  $n$ .*

Summarizing the results proved in this section, we showed that  $L_n^k$  is a random variable exponentially concentrated in a neighbourhood of radius  $O(n^{1/2(k+1)})$  around its mean, which converges to  $\alpha_k n^{1/(k+1)}$ .

## REFERENCES

- [1] D. Aldous and P. Diaconis, *Longest increasing subsequences: from patience sorting to the Baik-Deift-Johansson theorem*. Bull. Amer. Math. Soc. **36** (1999), 413–432.
- [2] N. Alon, J. Spencer, *The probabilistic method*. 2nd ed. John Wiley & Sons, New York (2000).
- [3] G. Ambrus, *Analytic and Probabilistic Problems in Discrete Geometry*. Ph.D. Thesis, University College London, 2009.
- [4] G. Ambrus, *Longest convex chains and subadditive ergodicity*. In: G. Ambrus, K.J. Böröczky, and Z. Füredi (eds.), *Discrete geometry and convexity, in honour of Imre Bárány*. Typotex, Budapest (2017).
- [5] G. Ambrus and I. Bárány, *Longest convex chains*, Random Structures Algorithms **35** (2009), no. 2., 137–162.
- [6] J. Baik, P. Deift, and K. Johansson, *On the distribution of the length of the longest increasing subsequence of random permutations*. J. Amer. Math. Soc. **12** (1999), 1119–1178.
- [7] I. Bárány, *Random points and lattice points in convex bodies*. Bull. Amer. Math. Soc. **45** (2008), no. 3., 339–365.
- [8] I. Bárány, G. Rote, W. Steiger, and C.-H. Zhang, *A central limit theorem for convex chains in the square*. Discrete Comput. Geom. **23** (2000), 35–50.
- [9] S. Boucheron, G. Lugosi, and P. Massart, *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, Oxford, 2013.
- [10] M. Eliáš and J. Matoušek, *Higher order Erdős-Szekeres theorems*. Adv. Math. **244** (2013), 1–15.
- [11] J. Fox, J. Pach, B. Sudakov, and A. Suk, *Erdős-Szekeres-type theorems for monotone paths and convex bodies*. Proc. London Math. Soc. **105** (2012), no. 5., 953–982.
- [12] P. W. Glynn, *Upper bounds on Poisson tail probabilities*. Oper. Res. Lett. **6** (1987), no. 1, 9–14.
- [13] J.M. Hammersley, *A few seedlings of research*. In: L. M. LeCam, J. Neyman, and E.L. Scott (eds.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Theory of Statistics*. University of California Press (1972), 345–394.
- [14] J.F.C. Kingman, *Subadditive ergodic theory*, Ann. Probab. **1** (1973), no. 6., 883–909.
- [15] T.M. Liggett, *An improved subadditive ergodic theorem*. Ann. Probab. **13** (1985), 1279–1285.
- [16] B.F. Logan and L.A. Shepp, *A variational problem for random Young tableaux*. Adv. Math. **26** (1977), 206–222.
- [17] G.M. Phillips, *Interpolation and approximation by polynomials*. Springer Verlag, Berlin (2003).
- [18] D. Romik, *The surprising mathematics of longest increasing subsequences*. Cambridge University Press (2014).
- [19] R. Schneider and W. Weil, *Stochastic and Integral Geometry*. Springer Verlag (2008).
- [20] M. Talagrand, *A new look at independence*. Ann. Probab. **24** (1996), 1–34.
- [21] S. Ulam, *Monte Carlo calculations in problems of mathematical physics*. In: E.F. Beckenbach (ed.), *Modern Mathematics For the Engineer, Second Series*. McGraw-Hill (1961), 261–281.
- [22] P. Valtr, *The probability that  $n$  points are in convex position*. Discrete Comput. Geom. **13** (1995), 637–643.
- [23] A.M. Vershik and S.V. Kerov, *Asymptotics of the Plancherel measure of the symmetric group and the limiting form of Young tables*. Soviet Math. Dokl., **18** (1977), 527–531. Translation of Dokl. Acad. Nauk. SSSR **233** (1977), 1024–1027.