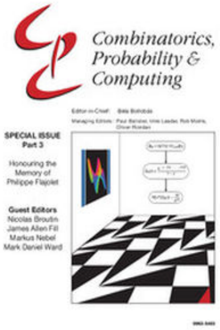
Combinatorics, Probability and Computing

http://journals.cambridge.org/CPC



Longest Increasing Subsequences of Randomly Chosen Multi-Row Arrays

MARCOS KIWI and JOSÉ A. SOTO

Combinatorics, Probability and Computing / Volume 24 / Special Issue 01 / January 2015, pp 254 - 293
DOI: 10.1017/S0963548314000637, Published online: 02 October 2014

Link to this article: http://journals.cambridge.org/abstract_S0963548314000637

How to cite this article:
MARCOS KIWI and JOSÉ A. SOTO (2015). Longest Increasing Subsequences of Randomly Chosen Multi-Row Arrays. Combinatorics, Probability and Computing, 24, pp 254-293 doi:10.1017/S0963548314000637

Combinatorics, Probability and Computing (2015) 24, 254–293. © Cambridge University Press 2014
doi:10.1017/S0963548314000637

Longest Increasing Subsequences of Randomly Chosen Multi-Row Arrays

MARCOS KIWI[1†] and JOSÉ A. SOTO[2‡]

[1]Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático UMI 2807, Universidad de Chile
(e-mail: mk@dim.uchile.cl

[2]Departamento de Ingeniería Matemática and Centro de Modelamiento Matemático UMI 2807, Universidad de Chile
(e-mail: jsoto@dim.uchile.cl)

Received 15 March 2013; revised 11 March 2014; first published online 2 October 2014

To Philippe Flajolet, a mathematical discontinuity, a tamer of singularities.

A two-row array of integers

$$\alpha_n = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ b_1 & b_2 & \cdots & b_n \end{pmatrix}$$

is said to be in lexicographic order if its columns are in lexicographic order (where character significance decreases from top to bottom, *i.e.*, either $a_k < a_{k+1}$, or $b_k \leqslant b_{k+1}$ when $a_k = a_{k+1}$). A length ℓ (strictly) increasing subsequence of α_n is a set of indices $i_1 < i_2 < \cdots < i_\ell$ such that $a_{i_1} < a_{i_2} < \cdots < a_{i_\ell}$ and $b_{i_1} < b_{i_2} < \cdots < b_{i_\ell}$. We are interested in the statistics of the length of a longest increasing subsequence of α_n chosen according to $\mathcal{D}_n$, for different families of distributions $\mathcal{D} = (\mathcal{D}_n)_{n\in\mathbb{N}}$, and when n goes to infinity. This general framework encompasses well-studied problems such as the so-called longest increasing subsequence problem, the longest common subsequence problem, and problems concerning directed bond percolation models, among others. We define several natural families of different distributions and characterize the asymptotic behaviour of the length of a longest increasing subsequence chosen according to them. In particular, we consider generalizations to d-row arrays as well as symmetry-restricted two-row arrays.

2010 *Mathematics subject classification*: 60C05

[†] Gratefully acknowledges the support of CONICYT via Basal in Applied Mathematics and FONDECYT 1090227, and Núcleo Milenio Información y Coordinación en Redes ICM/FI P10-024F.

[‡] Gratefully acknowledges the support of CONICYT via Basal in Applied Mathematics and Núcleo Milenio Información y Coordinación en Redes ICM/FI P10-024F.

1. Introduction

Suppose that we select uniformly at random a permutation π of $[n] \overset{\text{def}}{=} \{1,\ldots,n\}$. We can associate to this permutation the two-row lexicographically (column) sorted array

$$\alpha_\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(n) \end{pmatrix}.$$

We denote by $\mathsf{lis}(\pi)$ the length of a longest increasing subsequence of α_π. The determination, as $n \to \infty$, of the first moments of $\mathsf{lis}(\pi)$ has been a problem of much interest for a long time (for surveys see [1, 26, 30] and references therein). This line of research led to what is considered a major breakthrough: the determination by Baik, Deift and Johansson [7] of, after proper scaling, the distribution of $\mathsf{lis}(\cdot)$. Baik and Rains [8] studied variations where, instead of permutations of $[n]$, involutions, signed permutations, and signed involutions are selected at random. Generalizations where $d-1$ random permutations are selected can be restated as problems concerning longest increasing subsequences of d-row arrays. Furthermore, there are other quite relevant instances, that go far beyond those relating to permutations, where the general problem formulated in the abstract also arises. In order to illustrate this claim, as well as to stress the adequacy of the level of generality at which we have chosen to frame our work, our next section describes two other scenarios encompassed by the general framework concerning multi-row arrays that we consider.

1.1. Two more examples

Suppose that we select uniformly at random two words μ and ν from Σ^n, where Σ is some finite alphabet of size k. We can associate to (μ,ν) the two-row lexicographically sorted array $\alpha_{\mu,\nu}$ where $\binom{i}{j}$ is a column of $\alpha_{\mu,\nu}$ if and only if the ith character of μ is the same as the jth character of ν (for an example, see Figure 1). The length of a longest common subsequence of μ and ν, denoted by $\mathsf{lcs}(\mu,\nu)$, is defined as the length of a longest increasing subsequence of $\alpha_{\mu,\nu}$. Since the mid-1970s, it has been known [16] that the expectation of $\mathsf{lcs}(\mu,\nu)$, when normalized by n, converges to a constant γ_k (the so-called Chvátal–Sankoff constant). The determination of the exact value of γ_k, for k fixed, remains a challenging open problem. To the best of our knowledge, the asymptotic distribution theory of the longest common subsequence problem is essentially uncharted territory. Generalizations where d random words of length n are chosen from a finite alphabet Σ can also be restated as problems concerning longest increasing subsequences of d-row arrays.

We now discuss yet another instance, previously considered by Seppäläinen [29], and encompassed by the framework described above. Fix a parameter $0 < p < 1$ and let n be a positive integer. For each site of the lattice $[n]^2$, let a point be present (the site is occupied) with probability p and absent (the site is empty) with probability $q = 1 - p$, independently of all the other sites. Let $\omega : [n]^2 \to \{0,1\}$ be an encoding of the occupied/empty sites (1 representing an occupied site and 0 a vacant one). We can associate to ω a two-row lexicographically sorted array α_ω where $\binom{i}{j}$ is a column of α_ω if and only if site $(i,j) \in [n]^2$ is occupied. Let $L(\omega)$ equal the number of sites on a longest strictly increasing path of occupied sites according to ω, where a path $(x_1,y_1),(x_2,y_2),\ldots,(x_m,y_m)$ of points on $[n]^2$ is

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 3 & 3 & 4 & 5 & 5 \\ 3 & 5 & 1 & 2 & 3 & 5 & 4 & 3 & 5 \end{pmatrix}$$

Figure 1. Lexicographically ordered two-row array $\alpha_{\mu,\nu}$ associated with words $\mu = abaca$ and $\nu = bbaca$ (note in particular that $\mathsf{lcs}(\mu,\nu) = \mathsf{lis}(\alpha_{\mu,\nu}) = 4$).

strictly increasing if $x_1 < x_2 < \cdots < x_m$ and $y_1 < y_2 < \cdots < y_m$. Observe that $L(\omega)$ equals the length of a longest increasing subsequence of α_ω. Subadditivity arguments easily imply that the expected value of $L(\omega)$, when normalized by n, converges to a constant $\delta_{p,2}$. Via a reformulation of the problem as one of interacting particle systems, Seppäläinen [29] shows that $\delta_{p,2} = 2\sqrt{p}/(1+\sqrt{p})$. Also worth noting is that the same object α_ω arises in the study of the asymptotic shape of a directed bond percolation model (see [29, § 1] for details). Symmetric variants, where for example site (i,j) is occupied if and only if (j,i) is occupied, can be easily formulated. Generalizations where d-dimensional lattices are considered can also be restated as problems concerning longest increasing subsequences of d-row arrays. However, to the best of our knowledge, neither of the latter two variants has been considered in the literature.

1.2. Reformulation

Thus far, we have described well-studied scenarios where the general problem formulated in the abstract naturally arises. This motivates our work. However, for the sake of clarity of exposition and in order to use more convenient notation, it will be preferable to reformulate the issues we are interested in as one concerning hyper-graphs. To carry out this reformulation, below we introduce some useful terminology and then address in this language the problem of determining the statistics of the length of a longest increasing subsequence of a randomly chosen lexicographically sorted d-row array.

Let $A_1,\ldots,A_d$ be d disjoint (finite) sets, also called *colour classes*. We assume that over each A_i there is a total order relation, which by some abuse of notation, we denote $\leqslant$ in all cases. When we consider subsets of a totally ordered colour class we always assume the subset inherits, and thus respects, the original order. A *d-partite hyper-graph* over totally ordered colour classes $A_1,\ldots,A_d$ with edge set $E \subseteq A_1 \times \cdots \times A_d$ is a tuple $H = (A_1,\ldots,A_d;E)$, and its edge set is denoted by $E(H)$. The set $A_1 \cup \cdots \cup A_d$ is called the vertex (or node) set of H and is denoted by $V(H)$. For $A'_i \subseteq A_i$ with $1 \leqslant i \leqslant d$ and hyper-graph $H = (A_1,\ldots,A_d;E)$, we denote by $H|_{A'_1\times\cdots\times A'_d}$ the *hyper-subgraph of H restricted to* $A'_1 \times \cdots \times A'_d$, *i.e.*, the hyper-graph with node set $V' = A'_1 \cup \cdots \cup A'_d$ and edge set $E \cap A'_1 \times \cdots \times A'_d$. We say that two hyper-graphs are *disjoint* if their corresponding vertex sets are disjoint. Let $K_{A_1,\ldots,A_d}$ denote the *complete d-partite hyper-graph* over colour classes $A_1,\ldots,A_d$ whose edge set is $A_1 \times \cdots \times A_d$. Henceforth, we denote the cardinality of A_i by n_i. If we identify A_i with $[n_i]$, then we write $K_{n_1,\ldots,n_d}$ instead of $K_{A_1,\ldots,A_d}$. If $n_1 = \cdots = n_d$, then we write $K^{(d)}_n$ instead of $K_{n_1,\ldots,n_d}$. Over the edge set of $K_{A_1,\ldots,A_d}$ we consider the natural partial order relation $\preceq$ defined by

$$(v_1,\ldots,v_d) \preceq (v'_1,\ldots,v'_d) \Longleftrightarrow v_i \leqslant v'_i \quad \text{for all } 1 \leqslant i \leqslant d.$$

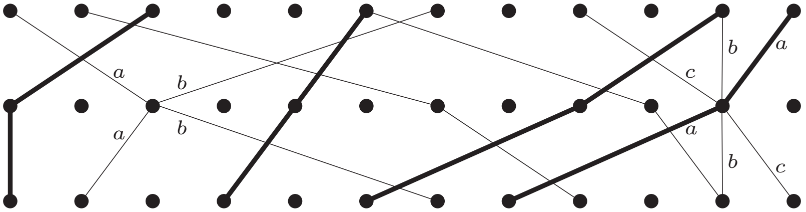


Figure 2. An example of a 3-partite hyper-graph H with 36 vertices and 10 hyper-edges. Rows depict partitions, nodes are depicted as circles, and hyper-edges are shown as lines (when hyper-edges share a vertex, disambiguation is done by labelling the line segments corresponding to the same hyper-edge). Note that $L(H) = 4$ (the edges belonging to a largest non-crossing matching are depicted as thicker lines).

We say that a collection of node-disjoint edges $M \subseteq E(H)$ is a *non-crossing hyper-matching* if for every pair of edges $e, f \in M$ it holds that $e \preceq f$ or $f \preceq e$. When $H = (A_1, \ldots, A_d; E)$ is such that $E(H)$ is a non-crossing hyper-matching, we say that H is a non-crossing d-partite hyper-graph, or simply a non-crossing hyper-matching. Furthermore, we let $L(H)$ denote the size of a largest non-crossing hyper-matching of H (for an example, see Figure 2) and let $L(\mathcal{F})$ be the random variable $L(H)$ when H is chosen according to a distribution $\mathcal{F}$ over d-partite hyper-graphs. When we want to stress that we are dealing with only two colour classes, we will speak of graphs and matchings instead of hyper-graphs and hyper-matchings.

Now, consider a family of distributions $\mathcal{D} = (\mathcal{D}(K_{A_1,\ldots,A_d}))$ where each $\mathcal{D}(K_{A_1,\ldots,A_d})$ is a probability distribution over subgraphs of $K_{A_1,\ldots,A_d}$. In this work we are interested in understanding what we refer to as the longest non-crossing matching problem, *i.e.*, the behaviour of the expectation of $L(H)$ when H is chosen according to various distinct families of distributions $\mathcal{D} = (\mathcal{D}(K_n^{(d)}))$ and n goes to infinity. Of course, in order to be able to derive some meaningful results we will need some assumptions on the distributions $\mathcal{D}(K_n^{(d)})$. Below, we encompass in a definition a minimal set of assumptions that are both easy to establish and general enough to capture several relevant scenarios.

Definition 1. Let $\mathcal{D} = (\mathcal{D}(K_{A_1,\ldots,A_d}))$ be a family of distributions where $\mathcal{D}(K_{A_1,\ldots,A_d})$ is a probability distribution over the collection of hyper-subgraphs of $K_{A_1,\ldots,A_d}$. We say that $\mathcal{D}$ is a random d-partite hyper-graph model if, for H chosen according to $\mathcal{D}(K_{A_1,\ldots,A_d})$, the following two conditions hold.

(1) *Monotonicity.* If $A'_i \subseteq A_i$ with $1 \leqslant i \leqslant d$ and $n'_i = |A'_i|$, then the distribution of $H|_{A'_1\times\cdots\times A'_d}$ is $\mathcal{D}(K_{n'_1,\ldots,n'_d})$.

(2) *Block independence.* If $A'_i, A''_i \subseteq A_i$ are disjoint with $1 \leqslant i \leqslant d$, then the hyper-graphs $H' = H|_{A'_1\times\cdots\times A'_d}$ and $H'' = H|_{A''_1\times\cdots\times A''_d}$ are independent (and so, $L(H')$ and $L(H'')$ are also independent).

For some of the results we will establish (in particular for the symmetric binomial random graph model), the following weaker notion will suffice.

Definition 2. Let $\mathcal{D} = (\mathcal{D}(K_n^{(d)}))$ be a family of distributions where each $\mathcal{D}(K_n^{(d)})$ is a probability distribution over the collection of hyper-subgraphs of $K_n^{(d)}$. We say that $\mathcal{D}$ is a weak random d-partite hyper-graph model if, for H chosen according to $\mathcal{D}(K_n^{(d)})$, the following two conditions hold.

(1) *Weak monotonicity.* If $A' \subseteq [n]$, $|A'| = n'$, then the distribution of $H|_{A'\times\cdots\times A'}$ is $\mathcal{D}(K_{n'}^{(d)})$.
(2) *Weak block independence.* If $A', A'' \subseteq [n]$ are disjoint sets, then $H' = H|_{A'\times\cdots\times A'}$ and $H'' = H|_{A''\times\cdots\times A''}$ are independent (and so $L(H')$ and $L(H'')$ are also independent).

The reader may easily verify that the following distributions (on which we will focus attention) give rise to random d-partite hyper-graph models.

- $\Sigma(K_{n_1,\dots,n_d}, k)$, the *random d-word model*: the distribution over the set of hyper-subgraphs obtained from $K_{n_1,\dots,n_d}$ when each element in the vertex set of $K_{n_1,\dots,n_d}$ is uniformly and independently randomly assigned one of k letters and where an edge is always discarded if its nodes are not assigned the same letters.
- $\mathcal{G}(K_{n_1,\dots,n_d}, p)$, the *d-dimensional binomial random hyper-graph model*: the distribution over the set of hyper-subgraphs H of $K_{n_1,\dots,n_d}$ where the events $\{H \mid e \in E(H)\}$ for $e \in E(K_{n_1,\dots,n_d})$ have probability p and are mutually independent.

The model $\Sigma(K_{n_1,\dots,n_d}, k)$ is referred to as the random word model because it arises when one considers the letters of d words $\omega_1, \dots, \omega_d$ of length $n_1, \dots, n_d$, respectively. The letters in each word are chosen uniformly and independently from a finite alphabet of size k. Then, each word is identified with a colour class of a hyper-subgraph H of $K_{n_1,\dots,n_d}$ whose hyper-edges are the $(v_1, \dots, v_d) \in V(H)$ for which $v_1, \dots, v_d$ have been assigned the same letter. It is easy to see that the longest common subsequence of $\omega_1, \dots, \omega_d$ equals ℓ if and only if $L(H) = \ell$. The random word model thus encompasses the longest common subsequence problem discussed above. Similarly, the attentive reader probably already noticed that the binomial random graph model also encompasses the already discussed point lattice process considered by Seppäläinen [29].

Inspired by the work of Baik and Rains [8] cited above, where symmetric variants of the longest increasing subsequence problem were considered, we will also study the following symmetric version of the binomial random graph model (see Figure 3).

- $\mathcal{S}(K_{n,n}, p)$, the *symmetric binomial random graph model*: the distribution over the set of subgraphs H of $K_{n,n}$ where $(i,j) \in E(H)$ if and only if $(j,i) \in E(H)$, for $1 \leqslant i < j \leqslant n$ and all the events $\{H \mid (i,j), (j,i) \in E(H)\}$ for $1 \leqslant i < j \leqslant n$, have probability p and are mutually independent.

Note that $(\mathcal{S}(K_{n,n}, p))_{n\in\mathbb{N}}$ is not a random model according to Definition 1, but it is a weak random model according to Definition 2.

Henceforth, given a random bipartite graph model $\mathcal{D} = (\mathcal{D}(\cdot))$, any value that is constant across the distributions $\mathcal{D}(\cdot)$ will be called the *internal parameter* of the model, *e.g.*, $1/p$ and k in $\mathcal{G}(\cdot, p)$ and $\Sigma(\cdot, k)$, respectively.

The main purpose of this work is to establish a general result, referred to as the Main Theorem, with a minimal set of easily verifiable hypothesis, that characterizes the (adequately normalized) limit behaviour of $\mathsf{E}[L(\mathcal{D}(K_n^{(d)}, p))]$ when d is fixed and both n and the internal parameter t go to infinity. We also show several applications of our Main Theorem. Specifically, we characterize aspects of the limiting behaviour for the three

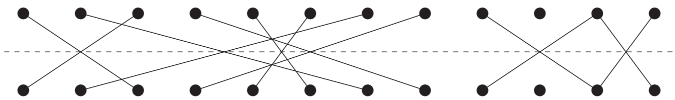


Figure 3. A bipartite graph in the support of $\mathcal{S}(K_{12,12},p)$.

previously introduced random hyper-graph models. In the following section we formally state our Main Theorem and the results of its application.

1.3. Main contributions

A straightforward application of Talagrand's inequality (as stated in [23, Theorem 2.29]) yields that both $L(\Sigma(K_{n,n},k))$ and $L(\mathcal{G}(K_{n,n},p))$ are concentrated around any one of their (potentially not unique) medians. As we shall see, the same is true for $L(\Sigma(K_n^{(d)},k))$ and $L(\mathcal{G}(K_n^{(d)},p))$. Somewhat equivalent statements hold for the symmetric binomial random graph model. The following general notion encompasses the concentration type requirement the random hyper-graph models will need to satisfy in order for our Main Theorem to be applicable.

Definition 3. Let $\mathcal{F}$ be a distribution over d-partite hyper-graphs. We say that $\mathcal{F}$ has *concentration constant* h if there exists a median Med of $L(\mathcal{F})$ such that, for all $s \geqslant 0$,

$$\mathsf{Pr}\big[L(\mathcal{F}) \leqslant (1-s)\mathsf{Med}\big] \leqslant 2\exp\big(-hs^2\mathsf{Med}\big),$$
$$\mathsf{Pr}\big[L(\mathcal{F}) \geqslant (1+s)\mathsf{Med}\big] \leqslant 2\exp\left(-h\frac{s^2}{1+s}\mathsf{Med}\right).$$

We say that the random d-partite hyper-graph model $\mathcal{D} = (\mathcal{D}(\cdot))$ has concentration constant h if each $\mathcal{D}(\cdot)$ has concentration constant h.

Note that if one can estimate a median of $L(\mathcal{F})$ for some distribution $\mathcal{F}$, show that the median and mean are close, and establish that $\mathcal{F}$ has a concentration constant, then one can derive a concentration (around its mean) result for $L(\mathcal{F})$. Unfortunately, it is generally not easy to estimate a median of $L(\mathcal{D}(K_{n_1,\dots,n_d}))$ for the distributions $\mathcal{D}(K_{n_1,\dots,n_d})$ we consider; however, we will be able to approximate them under some assumptions on $n_1,\dots,n_d$. In particular, we will show that there is a median that is proportional to the geometric mean of $n_1,\dots,n_d$. The following definition captures the aforementioned assumptions we will need, and the sort of approximation guarantee that we can establish.

Definition 4. Let $\mathcal{D} = (\mathcal{D}(K_{n_1,\dots,n_d}))$ be a random d-partite hyper-graph model with internal parameter t. Fix $n_1,\dots,n_d$ and let

$$N = \left(\prod_{i=1}^{d} n_i\right)^{1/d} \quad \text{and} \quad S = \sum_{i=1}^{d} n_i$$

denote the geometric mean and sum of $n_1,\dots,n_d$, respectively. We say that $\mathcal{D}$ admits a (c,λ,θ)-*approximate median* (or simply a (c,λ,θ)-*median*) for some $c > 0$ and $0 \leqslant \lambda \leqslant \theta$ if, for all $\delta > 0$, there are sufficiently large constants $a(\delta)$, $b(\delta)$, and $t'(\delta)$, such that for all $t \geqslant t'$, for which

$$N \geqslant at^{\lambda}, \qquad \text{(1.1: size lower bound condition)}$$

$$Sb \leqslant t^{\theta}, \qquad \text{(1.2: size upper bound condition)}$$

it holds that every median Med of $L(\mathcal{D}(K_{n_1,\ldots,n_d}))$ satisfies

$$(1-\delta)\frac{cN}{t^{\lambda}} \leqslant \mathsf{Med} \leqslant (1+\delta)\frac{cN}{t^{\lambda}}.$$

In other words, if $\mathcal{D} = (\mathcal{D}(K_{n_1,\ldots,n_d}))$ is a a random d-partite hyper-graph model with internal parameter t that admits a (c, λ, θ)-median, and N (resp. S) as in the preceding definition are such that $N = \Omega(t^{\lambda})$ (resp. $S = O(t^{\theta})$), then for sufficiently large t, every median of $L(\mathcal{D}(n_1, \ldots, n_d))$ will be close to $cNt^{-\lambda}$. Although the approximate median notion defined above might at first glance seem artificial, we will see that it is possible to obtain such types of approximations for the random hyper-graph models we are interested in.

Returning to our discussion, the relevance of the notion of approximate median, when the random hyper-graph model admits a concentration constant, is that it allows us to derive concentration bounds around an approximation of the median, which in turn will be closed to the mean. Endowed with such estimates of the mean, we can easily derive the sought-after limiting behaviour of these expected values. This, in essence, is the crux of our approach to tackling all variants of the largest non-crossing matching problem.

Unfortunately, the approximation of $\mathsf{Med}[L(\mathcal{D}(n_1, \ldots, n_d))]$ guaranteed by the existence of a (c, λ, θ)-median, as in Definition 4, holds for the rather restrictive condition (1.2):

$$b\sum_{i=1}^{d} n_i \leqslant t^{\theta}.$$

However, the monotonicity and block independence properties of random hyper-graph models allow us to relax the restriction and still obtain essentially the same conclusion. More precisely, it will be possible to obtain the same guarantee, but requiring only that the sum of the n_i is not too large in comparison with the geometric mean of the n_i. Moreover, and of crucial importance, under the same conditions one can show that the median and mean of $L(\mathcal{D}(n_1, \ldots, n_d))$ are close to each other. The following result, which is the main result of this work, precisely states the claims made in the preceding informal discussion.

Theorem 1.1 (Main Theorem). *Let $\mathcal{D} = (\mathcal{D}(K_{n_1,\ldots,n_d}))$ be a random hyper-graph model with internal parameter t and concentration constant h which admits a (c, λ, θ)-median. Fix $n_1, \ldots, n_d$ and let N and S denote the geometric mean and sum of $n_1, \ldots, n_d$, respectively. Let $0 \leqslant \eta \leqslant \lambda/(d-1)$ be such that $\eta < \theta - \lambda$, and $g : \mathbb{R}_+ \to \mathbb{R}_+$ be such that $g = O(t^{\eta})$.*

For all $\epsilon > 0$ there exist t_0 and A sufficiently large that if $t \geqslant t_0$ satisfies $N \geqslant At^{\lambda}$ (size constraint) and $S \leqslant g(t)N$ (balance constraint), then

$$(1-\epsilon)\frac{cN}{t^{\lambda}} \leqslant \mathsf{E}\left[L(\mathcal{D}(K_{n_1,\ldots,n_d}))\right] \leqslant (1+\epsilon)\frac{cN}{t^{\lambda}}, \tag{1.3}$$

and the following bounds hold:

- *If* Med *is a median of $L(\mathcal{D}(K_{n_1,\ldots,n_d}))$, then*

$$(1-\epsilon)\frac{cN}{t^{\lambda}} \leqslant \mathsf{Med} \leqslant (1+\epsilon)\frac{cN}{t^{\lambda}}. \tag{1.4}$$

- *There is an absolute constant $K > 0$ such that*

$$\mathsf{Pr}\left[L(\mathcal{D}(K_{n_1,\ldots,n_d})) \leqslant (1-\epsilon)\frac{cN}{t^{\lambda}}\right] \leqslant \exp\left(-Kh\epsilon^2\frac{cN}{t^{\lambda}}\right), \tag{1.5}$$

$$\mathsf{Pr}\left[L(\mathcal{D}(K_{n_1,\ldots,n_d})) \geqslant (1+\epsilon)\frac{cN}{t^{\lambda}}\right] \leqslant \exp\left(-Kh\frac{\epsilon^2}{1+\epsilon}\frac{cN}{t^{\lambda}}\right). \tag{1.6}$$

Moreover, if $n_1 = \cdots = n_d = n$ and $\mathcal{D} = (\mathcal{D}(K_n^{(d)}))$ is just a weak random hyper-graph model, then the the lower bounds in (1.3) *and* (1.4)*, and inequality* (1.5)*, still hold.*

As a consequence of the above-stated Main Theorem, with some additional work, we can derive several results concerning the asymptotic behaviour of the expected length of a largest non-crossing matching for all of the random models introduced above. Our first two applications of the Main Theorem concern the random binomial hyper-graph model $(\mathcal{G}(K_n^{(d)}, p))_{n\in\mathbb{N}}$ and the random word model $(\Sigma(K_n^{(d)}, k))_{n\in\mathbb{N}}$. The asymptotic behaviour of the length of a largest non-crossing hyper-matching for both of these models is (interestingly!) related to a constant c_d that arises in the work of Bollobás and Winkler [11] concerning the height of a largest chain among random points independently chosen in the d-dimensional unit cube $[0,1]^d$. Specifically, for the random binomial hyper-graph model, we show the following.

Theorem 1.2. *For $0 < p < 1$, and $d \in \mathbb{N}$, $d \geqslant 2$, there exists a constant $\delta_{p,d}$ such that*

$$\lim_{n\to\infty}\frac{1}{n}\mathsf{E}\left[L(\mathcal{G}(K_n^{(d)}, p))\right] = \inf_{n\in\mathbb{N}}\frac{1}{n}\mathsf{E}\left[L(\mathcal{G}(K_n^{(d)}, p))\right] = \delta_{p,d},$$

and $\delta_{p,d}/\sqrt[d]{p} \to c_d$ when $p \to 0$.

When the underlying model is the one that arises when considering the length of a longest common subsequence of d randomly chosen words over a finite alphabet, *i.e.*, the random d-word model, we establish the following.

Theorem 1.3. *For $k, d \in \mathbb{N}$, with $d \geqslant 2$, there exists a constant $\gamma_{k,d}$ such that*

$$\lim_{n\to\infty}\frac{1}{n}\mathsf{E}\left[L(\Sigma(K_n^{(d)}, k))\right] = \inf_{n\in\mathbb{N}}\frac{1}{n}\mathsf{E}\left[L(\Sigma(K_n^{(d)}, k))\right] = \gamma_{k,d},$$

and $k^{1-1/d}\gamma_{k,d} \to c_d$ when $k \to \infty$.

The $d = 2$ cases of Theorems 1.2 and 1.3 were established by Kiwi, Loebl and Matoušek [25]. This work generalizes and strengthens the arguments developed in [25] and elicits new connections with other previously studied problems (most notably in [11]).

Finally, we consider symmetric versions of random graph models introduced above and show how the Main Theorem, plus some additional observations, allows us to characterize some aspects of the asymptotic behaviour of the length of a longest non-crossing matching. Specifically, we prove the following result.

Theorem 1.4. *For* $0 < p < 1$, *there exists a constant* σ_p *such that*

$$\lim_{n\to\infty} \frac{1}{n}\mathsf{E}\left[L(\mathcal{S}(K_{n,n}, p))\right] = \inf_{n\in\mathbb{N}} \frac{1}{n}\mathsf{E}\left[L(\mathcal{S}(K_{n,n}, p))\right] = \sigma_p,$$

and $\sigma_p/\sqrt{p} \to 2$ *when* $p \to 0$.

1.4. Related problems

A natural question in connection to longest increasing subsequences of multi-row arrays is that of computing the length $\mathsf{lis}(\alpha)$ of a d-row array α, and more generally, finding a longest increasing subsequence of maximum length. One can easily show that these tasks can be performed by a simple dynamic program in time $O(n^d)$, where n is the number of columns of α. The latter holds under no assumption about the way in which instances are generated. A more interesting question arises if one maintains the consideration that the arrays are generated by a random process, but revealed one column at a time: How well can one sequentially choose (without clairvoyance) an increasing subsequence? For two-row arrays obtained from randomly and uniformly chosen permutations of $[n]$, the question was considered by Samuels and Steele [28], who demonstrated an asymptotically optimal policy which prescribes selection of a variable if and only if it exceeds the last variable selected so far by no more than a threshold parameter. The limiting upper bound $\sqrt{2n}$ was derived by careful analysis of the dynamic programming equation for computing the longest increasing subsequence (see Gnedin [21] for a simple proof along these lines). Coffman, Flatto and Weber [17] noted that the increasing subsequence problem can be regarded as a particular case of the sequential selection problem with a sum constraint. They used this connection to derive an elegant upper bound for non-clairvoyant policies based on sums of order statistics (see also Bruss and Robertson [15], Rhee and Talagrand [27], and Boshuizen and Kertz [12]). Baryshnikov and Gnedin [9] considered an extension to a scenario which, cast in the terminology of this work, can be thought of as relating to d-row arrays for $d > 2$ (see [13, 14] for other variants). For more recent work concerning sequential selection policies for longest increasing subsequence-type problems (*e.g.*, unimodal and alternating increasing subsequences), see Arlotto, Chen, Shepp and Steele [2] and Arlotto and Steele [3, 4].

There is a rich family of related problems in which a decision maker considers a (randomly generated) sequence of n objects and must decide whether to accept or reject each one at the time of presentation. Perhaps the most famous one is the classic *secretary problem* (see, *e.g.*, Dynkin [19] or Gilbert and Mosteller [20]), in which the objects are real values, the decision maker can only select a single one, and his goal is to maximize the probability of selecting the maximum. On more complex scenarios, the decision maker must select more than one object, in such a way that the selected subset satisfies a combinatorial constraint. The case mentioned above of selecting an increasing subsequence is an example of this setting. Many other interesting variants such as selecting a feasible set of a knapsack (see Babaioff, Immorlica, Kempe and Kleinberg [5]), or an independent set in a matroid (see Babaioff, Immorlica and Kleinberg [6] for the original formulation, and Dinitz [18] for a recent survey) have been proposed. The study of generalizations of the secretary problem is a very active research area, both in applied probability and in online algorithms.

1.5. Organization

For the sake of clarity of exposition and given that the arguments employed are different, we prove the lower and upper bounds (as well as lower and upper tail bounds) of the statement of the Main Theorem in separate sections. Specifically, in Section 2, we establish all the lower bounds and lower tail bounds claimed in the Main Theorem. In Section 3, we prove the upper bounds and upper tail bounds stated in the Main Theorem, thence completing its proof. Finally, in Section 4, we apply the Main Theorem to three distinct scenarios. Specifically, we consider the cases where the underlying random model is the binomial random hyper-graph model, the random word model, and the symmetric binomial random graph model.

2. Lower bounds

In this section we will establish the lower bounds claimed in the statement of the Main Theorem, *i.e.*, the lower bounds in (1.3) and (1.4), and inequality (1.5).

Let $\mathcal{D}$, c, λ, θ, η, and ϵ be as in the statement of the Main Theorem. Let $\delta > 0$ be sufficiently small that

$$(1-\delta)^2(1-2\delta) \geqslant 1 - \frac{\epsilon}{2}, \qquad \text{(2.1: definition of } \delta\text{)}$$

and let $a = a(\delta)$, $b = b(\delta)$ and $t' = t'(\delta)$ as guaranteed by the definition of a (c, λ, θ)-median.

Since $g = O(t^\eta)$, there are constants $C_g > 1$ and $t_g \geqslant 1$ such that $g(t) \leqslant C_g t^\eta$ for all $t \geqslant t_g$. Choose A sufficiently large that

$$A \geqslant \max\left\{ \frac{2a}{1-\delta}, \frac{2}{1-(1-\delta)^d} C_g^{d-1} t_g^{\eta(d-1)-\lambda}, \frac{2}{h\delta^2 c} \ln(2/\delta), \frac{16\ln(2)}{hc\epsilon^2} \right\}. \tag{2.2}$$

Since $\eta < \theta - \lambda$, we can choose $t_0 > \max\{t_g, t'(\delta)\}$ sufficiently large that, for all $t \geqslant t_0$,

$$g(t) \leqslant C_g t^\eta \quad \text{and} \quad C_g b A t^\eta \leqslant t^{\theta-\lambda}. \tag{2.3}$$

Now, assume $t > t_0$ and that the geometric mean N and sum S of $n_1, \ldots, n_d$ satisfy the size and balance constraints of the Main Theorem. These constraints guarantee that

$$N \geqslant A t^\lambda \quad \text{and} \quad S \leqslant g(t) N \leqslant C_g N t^\eta. \tag{2.4}$$

Finally, assume H is chosen according to $\mathcal{D}(K_{n_1,\ldots,n_d})$.

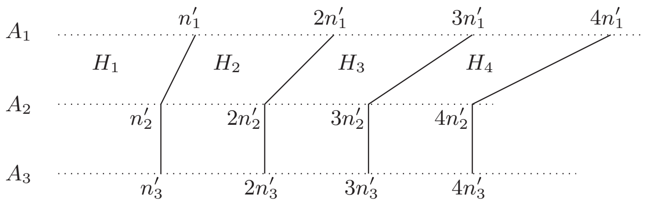


Figure 4. Illustration, for $d = 3$ and $q = 4$, of the construction of blocks $H_1, \dots, H_q$.

If the n_i satisfy the size conditions (1.1) and (1.2) of the definition of a (c, λ, θ)-median and since the model admits a concentration constant, we would then have a concentration bound around cNt^λ for $L(H)$. Unfortunately, when some of the n_i are large, then S will be large, and the size upper bound condition (1.2) need not be satisfied, leaving us without the desired concentration bound. To overcome this situation, we split H into hyper-subgraphs $H_1, \dots, H_q$ of roughly the same size, which we will refer to as *blocks*. The blocks will be vertex-disjoint, and the proportion between the sizes of the colour classes in each H_i will be roughly the same as the one in H. However, the crucial new aspect is that the size upper bound condition (1.2) will be satisfied in each block H_i allowing us to derive a concentration bound for $L(H_i)$. This will later allow us to obtain a concentration bound for $L(H)$; details follow.

Let $q = \lceil N/(At^\lambda) \rceil$. For each $1 \leqslant j \leqslant d$, let $n'_j = \lfloor n_j/q \rfloor$. Henceforth, let N' and S' denote the geometric mean and the sum of the n'_j. Denote the jth colour class of H by A_j. Recall that A_j is totally ordered. Let $A_{j,1}$ be the first n'_j elements of A_j, let $A_{j,2}$ be the following n'_j elements of A_j, and so on and so forth up to defining $A_{j,q}$. Clearly, the $A_{j,i}$ are disjoint but do not necessarily cover all of A_j. Now, for $1 \leqslant i \leqslant q$, define H_i as the hyper-subgraph of H restricted to $A_{1,i} \times \cdots \times A_{d,i}$ (for an illustration, see Figure 4). Observe that the proportion between the sizes of the colour classes of H_i is roughly the same as the one among the colour classes of H.

Note that by monotonicity, the distribution of H_i is $\mathcal{D}(K_{n'_1,\dots,n'_d})$. Moreover, since the H_i are disjoint, by block independence, their distributions are independent. Thus $L(H_1), \dots, L(H_q)$ are independent random variables. A crucial, though trivial, observation is that

$$L(H) \geqslant \sum_{i=1}^{q} L(H_i). \qquad (2.5)$$

On the other hand, by definition of q and (2.4), we obtain

$$\frac{N}{At^\lambda} \leqslant q \leqslant \frac{N}{At^\lambda} + 1 \leqslant \frac{2N}{At^\lambda}. \qquad \text{(2.6: estimate of } q\text{)}$$

An important but non-trivial observation is that each $n'_j \geqslant 1$, or equivalently, that $n_j \geqslant q$. In fact, we will show the stronger claim $n_j \geqslant 2N/At^\lambda$. We do this by contradiction. Suppose without loss of generality that $n_1 < 2N/At^\lambda$. Using that the arithmetic mean of

$\{n_2,\dots,n_d\}$ is larger than or equal to their geometric mean, we get

$$S \geqslant \sum_{i=2}^{d} n_i \geqslant (d-1)\prod_{i=2}^{d} n_i^{1/(d-1)} = (d-1)\frac{N^{d/(d-1)}}{n_1^{1/(d-1)}} > (d-1)N\left(\frac{At^{\lambda}}{2}\right)^{1/(d-1)}.$$

By (2.2),

$$A \geqslant \frac{2}{1-(1-\delta)^d} C_g^{d-1} t_g^{\eta(d-1)-\lambda} \geqslant 2C_g^{d-1}.$$

Combining this with the previous inequality, and using the hypothesis $\lambda/(d-1) \geqslant \eta$ of the Main Theorem, we get

$$S > (d-1)NC_g t^{\lambda/(d-1)} \geqslant NC_g t^{\eta},$$

which contradicts (2.4) and completes the proof of the claim.

In order to estimate the geometric mean N' of $n'_1,\dots,n'_d$, the next result will be useful.

Lemma 2.1. *If* $x_1,\dots,x_d$ *are real numbers greater than or equal to* 1, *then*

$$\prod_{j=1}^{d}(x_j-1) \geqslant \prod_{j=1}^{d} x_j - \left(\sum_{j=1}^{d} x_j\right)^{d-1}.$$

Proof. Clear for $d=1$. For $d \geqslant 2$, by the induction hypothesis, the fact that $x_d \geqslant 1$, and since

$$\max_{j=1,\dots,d-1} x_j \leqslant \sum_{j=1}^{d-1} x_j,$$

we obtain

$$\begin{aligned}\prod_{j=1}^{d} x_j - \prod_{j=1}^{d}(x_j-1) &= (x_d-1)\left(\prod_{j=1}^{d-1} x_j - \prod_{j=1}^{d-1}(x_j-1)\right) + \prod_{j=1}^{d-1} x_j \\ &\leqslant (x_d-1)\left(\sum_{j=1}^{d-1} x_j\right)^{d-2} + \prod_{j=1}^{d-1} x_j \leqslant (x_d-1)\left(\sum_{j=1}^{d-1} x_j\right)^{d-2} + \left(\sum_{j=1}^{d-1} x_j\right)^{d-1} \\ &= \left(\sum_{j=1}^{d-2} x_j\right)^{d-2}\left(x_d - 1 + \sum_{j=1}^{d-1} x_j\right) \leqslant \left(\sum_{j=1}^{d-2} x_j\right)^{d-2} \sum_{j=1}^{d} x_j.\end{aligned}$$

The desired conclusion follows immediately from the fact that the x_j are positive. □

Using the definition of

$$N' = \prod_{j=1}^{d}(\lfloor n_j/q \rfloor)^{1/d},$$

the preceding lemma, and our previous observation that each $n'_j \geqslant 1$, we conclude that

$$\frac{N}{q} = \prod_{j=1}^{d}\left(\frac{n_j}{q}\right)^{1/d} \geqslant N' \geqslant \left(\prod_{j=1}^{d}\left(\frac{n_j}{q}-1\right)\right)^{1/d}$$
$$\geqslant \left(\prod_{j=1}^{d}\frac{n_j}{q} - \left(\sum_{j=1}^{d}\frac{n_j}{q}\right)^{d-1}\right)^{1/d} = \frac{N}{q}\left(1-q\frac{S^{d-1}}{N^d}\right)^{1/d}.$$

By (2.4) and our estimate (2.6) of q,

$$q\frac{S^{d-1}}{N^d} \leqslant \frac{2}{A}C_g^{d-1}t^{\eta(d-1)-\lambda}.$$

Given the way we have chosen A, we have that $(1-qS^{d-1}/N^d)^{1/d} \geqslant 1-\delta$, and thus

$$\frac{N}{q} \geqslant N' \geqslant \frac{N}{q}(1-\delta). \qquad \text{(2.7: estimate of } N'\text{)}$$

Based on the preceding estimate of N' and the estimate for q, we will now show that $n'_1,\dots,n'_d$ satisfy the size conditions (1.1) and (1.2) required by the definition of a (c,λ,θ)-median. Indeed, by our estimates (2.7) and (2.6) of N' and q, and (2.2),

$$N' \geqslant \frac{N}{q}(1-\delta) \geqslant \frac{1}{2}At^{\lambda}(1-\delta) \geqslant at^{\lambda}.$$

Moreover, by the definition of S', (2.6), (2.4), and (2.3),

$$S'b \leqslant \frac{Sb}{q} \leqslant \frac{SbAt^{\lambda}}{N} \leqslant C_g bAt^{\lambda+\eta} \leqslant t^{\theta}.$$

Now, choose H' according to $\mathcal{D}(K_{n'_1,\dots,n'_d})$ and let Med' be a median of $L(\mathcal{D}(K_{n'_1,\dots,n'_d}))$. By the definition of a (c,λ,θ)-median, we get that $cN't^{-\lambda}(1-\delta) \leqslant \mathsf{Med}' \leqslant cN't^{-\lambda}(1+\delta)$. Moreover, using the definition of the constant of concentration, we get

$$\mathsf{Pr}\left[L(H') \geqslant (1-2\delta)\frac{cN'}{t^{\lambda}}\right] \geqslant \mathsf{Pr}\left[L(H') \geqslant \left(1-\frac{\delta}{1-\delta}\right)\mathsf{Med}'\right]$$
$$\geqslant 1-2\exp\left(-h\frac{\delta^2}{(1-\delta)^2}\mathsf{Med}'\right) \geqslant 1-2\exp\left(-h\frac{\delta^2}{1-\delta}\frac{cN'}{t^{\lambda}}\right).$$

Using Markov's inequality, we get

$$\mathsf{E}[L(H')] \geqslant (1-2\delta)\frac{cN'}{t^{\lambda}}\left(1-2\exp\left(-h\frac{\delta^2}{1-\delta}\frac{cN'}{t^{\lambda}}\right)\right).$$

As observed above, $N' \geqslant At^{\lambda}(1-\delta)/2$, so by the choice of A (see (2.2)) we get that

$$\mathsf{E}[L(H')] \geqslant (1-2\delta)(1-\delta)cN't^{-\lambda}.$$

Hence, using that $L(H) \geqslant \sum_{i=1}^{q} L(H_i)$, (2.7), (2.1) and elementary algebra,

$$\mathsf{E}[L(H)] \geqslant \sum_{i=1}^{q}\mathsf{E}[L(H_i)] \geqslant (1-2\delta)(1-\delta)q\frac{cN'}{t^{\lambda}}$$
$$\geqslant (1-2\delta)(1-\delta)^2\frac{cN}{t^{\lambda}} \geqslant (1-\epsilon/2)\frac{cN}{t^{\lambda}}.$$

We have thus established the lower bound claimed in (1.3).

Now, we proceed to show (1.5). Note that

$$\mathsf{Pr}\left[L(H) \leqslant (1-\epsilon)\frac{cN}{t^{\lambda}}\right] \leqslant \sum_{\substack{(s_1,\dots,s_q)\in\mathbb{N}^q \\ s_1+\cdots+s_q\leqslant(1-\epsilon)cNt^{-\lambda}}} \mathsf{Pr}[L(H_i)=s_i, i=1,\dots,q]. \tag{2.8}$$

Let $\mathcal{T}$ be the set of indices of the summation in the preceding displayed equation. Also, for $T=(s_1,\dots,s_q)$ belonging to $\mathcal{T}$, let P_T denote $\mathsf{Pr}[L(H_i)=s_i, i=1,\dots,q]$. We will show that P_T is exponentially small with respect to $cNt^{-\lambda}$. More precisely, we will show that $\ln P_T \leqslant q\ln(2) - cNt^{-\lambda}h\epsilon^2/8$. Recalling that the $L(H_i)$ are independent and distributed as $L(H')$ when H' is chosen according to $\mathcal{D}(K_{n'_1,\dots,n'_d})$,

$$P_T = \prod_{i=1}^{q} \mathsf{Pr}[L(H_i)=s_i] \leqslant \left(\mathsf{Pr}[L(H')\leqslant s_i]\right)^q.$$

Again, by the way in which H' is chosen, the definition of Med', and the definition of a (c,λ,θ)-median, for i such that $s_i \leqslant (1-\delta)cN't^{-\lambda} \leqslant \mathsf{Med}' \leqslant (1+\delta)cN't^{-\lambda} \leqslant 2cN't^{-\lambda}$, it holds that

$$\begin{aligned}\mathsf{Pr}[L(H')\leqslant s_i] &= \mathsf{Pr}\left[L(H') \leqslant \left(1-\frac{\mathsf{Med}'-s_i}{\mathsf{Med}'}\right)\mathsf{Med}'\right] \\ &\leqslant 2\exp\left(-h\frac{(\mathsf{Med}'-s_i)^2}{\mathsf{Med}'}\right) \leqslant 2\exp\left(-\frac{ht^{\lambda}}{2cN'}((1-\delta)cN't^{-\lambda}-s_i)^2\right).\end{aligned}$$

Hence, for all $1\leqslant i\leqslant q$,

$$\mathsf{Pr}[L(H')\leqslant s_i] \leqslant 2\exp\left(-\frac{ht^{\lambda}}{2cN'}\left(\max\{0,(1-\delta)cN't^{-\lambda}-s_i\}\right)^2\right),$$

and then

$$\begin{aligned}-\ln P_T &\geqslant -\sum_{i=1}^{q}\ln \mathsf{Pr}[L(H')\leqslant s_i] \\ &\geqslant -q\ln(2) + \frac{ht^{\lambda}}{2cN'}\sum_{i=1}^{q}\left(\max\{0,(1-\delta)cN't^{-\lambda}-s_i\}\right)^2.\end{aligned}$$

By the Cauchy–Schwarz inequality, our estimate (2.7) of N', the fact that

$$s_1+\cdots+s_q \leqslant (1-\epsilon)cNt^{-\lambda},$$

and since by the definition of δ (see (2.1)) we know that $(1-\delta)^2 \geqslant 1-\epsilon/2$,

$$\begin{aligned}\sqrt{q\sum_{i=1}^{q}\left(\max\{0,(1-\delta)cN't^{-\lambda}-s_i\}\right)^2} &\geqslant \sum_{i=1}^{q}\max\{0,(1-\delta)cN't^{-\lambda}-s_i\} \\ &\geqslant (1-\delta)cN'qt^{-\lambda} - \sum_{i=1}^{q}s_i \geqslant (1-\delta)^2cNt^{-\lambda} - (1-\epsilon)cNt^{-\lambda} \geqslant \frac{cN\epsilon}{2t^{\lambda}}.\end{aligned}$$

Combining the last two displayed inequalities and recalling our estimate (2.7) of N', we get the desired bound for P_T, since

$$-\ln P_T \geqslant -q\ln(2) + \frac{ht^{\lambda}}{2cN'q}\cdot\frac{c^2N^2\epsilon^2}{4t^{2\lambda}} \geqslant -q\ln(2) + \frac{hcN\epsilon^2}{8t^{\lambda}}.$$

By (2.8) and using the standard estimate $\binom{a}{b} \leqslant (ea/b)^b$, we have

$$\Pr\left[L(H) \leqslant (1-\epsilon)\frac{cN}{t^{\lambda}}\right] \leqslant \sum_{T\in\mathcal{T}} P_T \leqslant |\mathcal{T}|\cdot\max_{T\in\mathcal{T}} P_T \leqslant \binom{\lfloor(1-\epsilon)cNt^{-\lambda}\rfloor + q}{q}\cdot\max_{T\in\mathcal{T}} P_T$$
$$\leqslant \exp\left(q\ln\left(2e[1+(1-\epsilon)cNt^{-\lambda}/q]\right) - \frac{hcN\epsilon^2}{8t^{\lambda}}\right).$$

Now, by the estimate of q (2.6) we know that $N \leqslant qAt^{\lambda} \leqslant 2N$. Thus, if we require that A is sufficiently large that $\ln(2e[1+(1-\epsilon)cA]) \leqslant Ahc\epsilon^2/32$, we get that

$$\Pr\left[L(H) \leqslant (1-\epsilon)\frac{cN}{t^{\lambda}}\right] \leqslant \exp\left(\frac{2N}{At^{\lambda}}\ln(2e[1+(1-\epsilon)cA]) - \frac{hcN\epsilon^2}{8t^{\lambda}}\right)$$
$$\leqslant \exp\left(-\frac{h\epsilon^2cN}{16t^{\lambda}}\right).$$

This proves the lower bound claimed in (1.5).

What remains is to show the lower bound in (1.4). By the estimate of q (see (2.6)) we have $N \geqslant At^{\lambda}$, which together with our choice of A (see (2.2)) implies that

$$\exp\left(-\frac{h\epsilon^2cN}{16t^{\lambda}}\right) \leqslant \exp\left(-\frac{h\epsilon^2c}{16}A\right) \leqslant \frac{1}{2}.$$

Combining the last two displayed equations, it follows that

$$\Pr\left[L(H) \leqslant (1-\epsilon)cNt^{-\lambda}\right] \leqslant \frac{1}{2},$$

implying that any median of $L(H)$ must be at least $(1-\epsilon)cNt^{-\lambda}$.

Remark. The reader may check that all claims proved in this section would still hold if, instead of $\mathcal{D} = (\mathcal{D}(K_{n_1,\ldots,n_d}))$, we had worked with a weak random hyper-graph model $\mathcal{D} = (\mathcal{D}(K_n^{(d)}))$. Indeed, if this were the case, then for H chosen according to $\mathcal{D}(K_n^{(d)})$, the hyper-graphs $H_1,\ldots,H_q$ obtained above from H would have all their colour classes of equal size, and the weak random hyper-graph model assumption is all that would be needed to carry forth the arguments laid out in this section.

3. Upper bounds

In this section we will establish the upper bounds claimed in the statement of the Main Theorem, *i.e.*, the upper bounds in (1.3) and (1.4), and inequality (1.6). The proof of the latter of these bounds, the upper tail bound, is rather long. For sake of clarity of exposition, we have divided its proof into three parts. First, in Section 3.1, we introduce some useful variables. In Section 3.2, we establish (1.6) for not too large values of the geometric mean N. Then, in Section 3.3, we consider the case where N is large. Finally, in Section 3.4, we conclude the proof of the bounds claimed in the Main Theorem.

3.1. Basic variable definitions

For the rest of this section, let $\mathcal{D}$, c, λ, θ, η, and ϵ be as in the statement of the Main Theorem. Define

$$\delta = \min\left\{1, \frac{\epsilon^2}{1+\epsilon}, \frac{\epsilon}{6}\right\}. \qquad \text{(3.1: definition of } \delta\text{)}$$

Let $a = a(\delta)$, $b = b(\delta)$ and $t' = t'(\delta)$ as guaranteed by the definition of a (c, λ, θ)-median. Choose A so that

$$A = \max\left\{\frac{a}{\delta}, \frac{8\ln(2)}{h\delta c}\right\}. \tag{3.2}$$

For technical reasons, it will be convenient to fix constants α and β such that

$$\lambda < \alpha < \beta < \theta - \eta.$$

We shall also encounter two constants K_1 and K_2, depending solely on d. Since $g = O(t^\eta)$, there are constants $C_g > 1$ and $t_g \geqslant t'$ such that for all $t \geqslant t_g$ it holds that $g(t) \leqslant C_g t^\eta$, and

$$\max\{9At^\lambda, e\} \leqslant t^\alpha \leqslant t^\beta \leqslant \frac{1}{bdC_g} t^{\theta-\eta}, \tag{3.3}$$

$$\frac{2\alpha K_1}{\delta c K_2 h} t^\lambda \leqslant \frac{t^\alpha}{\ln(t)}. \tag{3.4}$$

Now consider $t \geqslant t_g$ and positive integers $n_1, n_2, \ldots, n_d$ with geometric mean N, summing up to S, and satisfying both the size ($N \geqslant At^\lambda$) and balance constraints ($S \leqslant g(t)N$) of the Main Theorem. Furthermore, define $M = cNt^{-\lambda}$ and choose H according to $\mathcal{D}(K_{n_1,\ldots,n_d})$. In the following two sections, we separately consider the case where N is less than or (respectively) is at least t^β.

3.2. Upper tail bound for not too large values of N

Throughout this section, we assume $N < t^\beta$. We will show that H satisfies the size bound restrictions (1.1) and (1.2) in the definition of a (c, λ, θ)-median. The fact that $\mathcal{D}$ admits a concentration constant h will allow us to obtain a bound on the upper tail of $L(H)$.

Let $t \geqslant t_g$. Since N satisfies both the size and balance constraints of the Main Theorem, by (3.2), (3.3), and the definition (3.1) of δ,

$$N \geqslant At^\lambda \geqslant \frac{at^\lambda}{\delta} \geqslant at^\lambda,$$

$$Sb \leqslant g(t)bN < C_g bt^{\eta+\beta} \leqslant \frac{t^\theta}{d} \leqslant t^\theta.$$

Thus, $n_1, \ldots, n_d$ satisfy both the size lower and upper bound conditions (1.1) and (1.2) of the definition of a (c, λ, θ)-median. Hence, if H is chosen according to $\mathcal{D}(K_{n_1,\ldots,n_d})$, then every median Med of $L(H)$ is at a distance of at most δM from M. By simple algebra, the definitions of the concentration constant and the (c, λ, θ)-median, and given that by the definition of δ we know that $\delta < \epsilon$, we have

$$\begin{aligned}\mathsf{Pr}\big[L(H) \geqslant (1+\epsilon)M\big] &= \mathsf{Pr}\left[L(H) \geqslant \left(1 + \frac{(1+\epsilon)M - \mathsf{Med}}{\mathsf{Med}}\right)\mathsf{Med}\right] \\ &\leqslant 2\exp\left(-h\frac{((1+\epsilon)M - \mathsf{Med})^2}{(1+\epsilon)M}\right) \leqslant \exp\left(\ln(2) - h\frac{(\epsilon-\delta)^2}{1+\epsilon}M\right).\end{aligned}$$

Again by the way in which δ is defined we have that $\delta \leqslant \epsilon/2$ and $\delta \leqslant \epsilon^2/(1+\varepsilon)$. Using this, (3.2) and recalling that $N \geqslant At^{\lambda}$ and $M = cNt^{-\lambda}$, we have

$$\begin{aligned}\mathsf{Pr}\big[L(H) \geqslant (1+\epsilon)M\big] &\leqslant \exp\left(\frac{Ah\delta c}{8} - \frac{h\epsilon^2}{4(1+\epsilon)}M\right) \\ &\leqslant \exp\left(\frac{h\delta Nc}{8t^{\lambda}} - \frac{h\epsilon^2}{4(1+\epsilon)}M\right) \leqslant \exp\left(-\frac{h\epsilon^2}{8(1+\epsilon)}M\right).\end{aligned}$$

We have thus established (1.6) for $N < t^{\beta}$.

3.3. Upper tail bound for large values of N

We now consider the case where $N \geqslant t^{\beta}$. The magnitude of N is such that we cannot directly apply the definition of a (c, λ, θ)-median to a hyper-graph generated according to $\mathcal{D}(K_{n_1,\ldots,n_d})$, and thus derive the sought-after exponentially small tail bound. We again resort to the block partitioning technique introduced in the proof of the lower bound. However, both the block partitioning and the analysis are more delicate and involved in the case of the upper bound.

3.3.1. Block partition. Let $l = t^{\alpha}$, $L = C_g t^{\eta+\alpha}$ and

$$m_{\max} = \lceil (1+\epsilon)M \rceil. \tag{3.5}$$

In what follows, we shall upper-bound the probability that H chosen according to $\mathcal{D}(K_{n_1,\ldots,n_d})$ has a non-crossing hyper-matching of size at least $m_{\max}$, *i.e.*, the probability that $L(H) \geqslant m_{\max}$.

We begin with a simple observation: since different edges of a non-crossing hyper-matching of H cannot have vertices in common, $L(H) \leqslant n_i$ for all i. It immediately follows that $L(H)$ is upper-bounded by the geometric mean of the n_i, that is, $L(H) \leqslant N$. Thus, if $m_{\max} > N$, then $\mathsf{Pr}\big[L(H) \geqslant m_{\max}\big] = 0$. This justifies why, in the ensuing discussion, we assume that $m_{\max} \leqslant N$.

Let J be a non-crossing hyper-subgraph of $K_{n_1,\ldots,n_d}$ such that the number of edges of J is (exactly equal to) $m_{\max}$. We shall partition the edge set of J into consecutive sets of edges which we will refer to as *blocks*. The partition will be such that for any colour class, the set of vertices appearing in a block are 'not too far apart', the precise meaning to be clarified shortly. The maximum number of edges in any block will be $s_{\max}$, where

$$s_{\max} = \left\lfloor \frac{l}{N} m_{\max} \right\rfloor. \tag{3.6}$$

Given two edges e and $\widetilde{e}$ of J, such that $e \preceq \widetilde{e}$, we let $[e, \widetilde{e}]$ denote the collection of edges f of J such that $e \preceq f \preceq \widetilde{e}$. We now define a partition into blocks of the edge set of J,

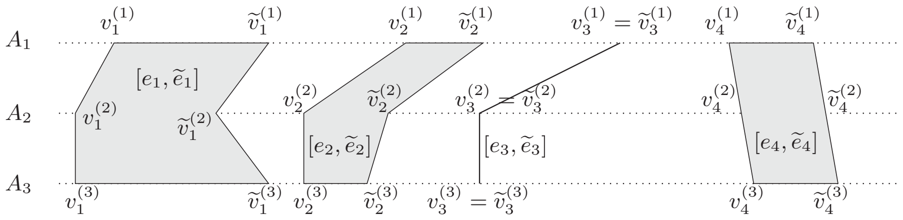


Figure 5. Partition into blocks of a hyper-graph. Each block $[e_i, \widetilde{e}_i]$ (shown in light grey) contains at most $s_{\max}$ edges and at most L vertices from each colour class.

denoted by $\mathcal{P}(J)$, as follows: $\mathcal{P}(J) = \{[e_i, \widetilde{e}_i] \mid 1 \leqslant i \leqslant q\}$ where the e_i, the $\widetilde{e}_i$, and q are determined via the following inductive process.

- e_1 is the first (smallest according to $\preceq$) edge of J.
- Assuming $e_i = (v_1^{(i)}, v_2^{(i)}, \dots, v_d^{(i)})$ has already been defined, $\widetilde{e}_i = (\widetilde{v}_1^{(i)}, \widetilde{v}_2^{(i)}, \dots, \widetilde{v}_d^{(i)})$ is the last edge of J satisfying the following two conditions (see Figure 5 for an illustration):
  - $[e_i, \widetilde{e}_i]$ has at most $s_{\max}$ elements,
  - $\widetilde{v}_j^{(i)} - v_j^{(i)} \leqslant L$ for all $1 \leqslant j \leqslant d$ (where we have relied on the abuse of notation entailed by our identification of the jth colour class of $K_{n_1,\dots,n_d}$ with the set $\{1, 2, \dots, n_j\}$ endowed with the natural order).
- Assuming $\widetilde{e}_i$ has already been defined and provided there are edges e of J strictly larger than $\widetilde{e}_i$, we define e_{i+1} to be the smallest such e.

Clearly, the value taken by q above depends on J. Nevertheless, we will show that the following estimate of $q = |\mathcal{P}(J)|$ holds for all J non-crossing hyper-subgraphs of $K_{n_1,\dots,n_d}$:

$$\frac{N}{l} \leqslant |\mathcal{P}(J)| \leqslant \frac{3N}{l}. \qquad \text{(3.7: estimate of } q\text{)}$$

Note that each block has at most $s_{\max}$ edges and recall that $|E(J)| = m_{\max}$. Thus,

$$q \geqslant m_{\max}/s_{\max} \geqslant N/l.$$

Now, we say that a block is *short* if it is either $[e_q, \widetilde{e}_q]$ or a block with exactly $s_{\max}$ edges. Let I_0 be the collection of indices of short blocks. It follows that $m_{\max} \geqslant (|I_0| - 1)s_{\max}$. Furthermore, since $m_{\max} \geqslant M = cNt^{-\lambda}$, we have

$$\frac{m_{\max}}{s_{\max}} \leqslant \frac{N}{l} \cdot \frac{1}{1 - N/(lm_{\max})} \leqslant \frac{N}{l} \cdot \frac{1}{1 - t^{\lambda-\alpha}/c}.$$

Since (3.4) implies that $t^{\lambda-\alpha} < c/2$, we conclude that $|I_0| \leqslant 2N/l$.

Say a block is *regular* if it is not short, and let $I_1 = [q] \setminus I_0$ be the set of indices of such blocks. We shall use the term *block cover* for the collection of all nodes between the first edge of the block (inclusive) and the first edge of the next block (exclusive). By the definition of block partition, if the ith block is regular, then for some colour class j we must have $v_j^{(i+1)} - v_j^{(i)} > L$. Hence,

$$\sum_{j=1}^{d} (v_j^{(i+1)} - v_j^{(i)}) > L.$$

In other words, a regular block gives rise to a block cover of cardinality at least L. Since every node belongs to at most one block cover, $|I_1| \leqslant S/L$. Recalling that $L = C_g t^{\eta} l$ and that S satisfies the balance condition (hence, $S \leqslant C_g t^{\eta} N$ for $t \geqslant t_g$), we conclude that $|I_1| \leqslant N/l$.

Putting together the conclusions reached in the last two paragraphs, we see that $q = |I_0| + |I_1| \leqslant 3N/l$, which establishes the claimed estimate of q.

3.3.2. Partition types. Let s_i be the number of edges of J in the ith block $[e_i, \widetilde{e}_i]$ of the partition $\mathcal{P}(J)$. Let q be the number of blocks of $\mathcal{P}(J)$. We refer to the $(3q)$-tuple

$$T = (e_1, \widetilde{e}_1, s_1, \dots, e_q, \widetilde{e}_q, s_q)$$

as the *type* of partition $\mathcal{P}(J)$, and denote it by $T(\mathcal{P}(J))$. Furthermore, let $\mathcal{T}$ be the collection of all possible partition types of non-crossing hyper-subgraphs of $K_{n_1,\dots,n_d}$ with exactly $m_{\max}$ edges.

Lemma 3.1. *There is a constant K_1, depending only on d, such that*

$$|\mathcal{T}| \leqslant \exp\left(K_1 \frac{N}{l} \ln(l)\right).$$

Proof. Observe that each e_i is completely determined by specifying its vertices. Hence, the number of ways of choosing $e_1, \dots, e_q$ is at most the number of ways of choosing q elements from each node colour class, *i.e.*, at most $\prod_{i=1}^{d} \binom{n_i}{q}$. The number of choices for $\widetilde{e}_1, \dots, \widetilde{e}_q$ is bounded by the same quantity. On the other hand, since J has exactly $m_{\max}$ edges, the number of choices for $s_1, \dots, s_q$ is at most the number of ways of summing up to $m_{\max}$ with q positive integer summands. Since we are assuming that $m_{\max} \leqslant N$ (see comment in the second paragraph of Section 3.3.1), the aforementioned quantity can be bounded by $\binom{N}{q}$. Using that $\binom{a}{b} \leqslant (ea/b)^b$ we obtain, for fixed q, that the number of types is at most

$$\binom{N}{q}\left(\prod_{i=1}^{d}\binom{n_i}{q}\right)^2 \leqslant \left(\frac{eN}{q}\right)^q \left(\prod_{i=1}^{d}(en_i/q)\right)^{2q} = \left(\frac{eN}{q}\right)^{q+2qd}.$$

Recalling our estimate (3.7) for q, we get that

$$|\mathcal{T}| \leqslant \sum_{q=\lceil N/l \rceil}^{\lfloor 3N/l \rfloor} \left(\frac{eN}{q}\right)^{q(1+2d)} \leqslant \frac{3N}{l}(el)^{3(1+2d)N/l}.$$

Since $\ln(x) \leqslant x$ for all $x > 0$ and by (3.3), we know that $l = t^{\alpha} \geqslant e$, and

$$\ln|\mathcal{T}| \leqslant \ln\left(\frac{3N}{l}\right) + (1+2d)\frac{3N}{l}(1+\ln(l)) \leqslant (2+2d)\frac{3N}{l}(1+\ln(l)).$$

Since $1 + \ln l \leqslant 2\ln l$, the desired conclusion follows choosing $K_1 = 12(1+d)$. □

3.3.3. Probability of a block partition occurring. The purpose of this section is to show that for a given fixed type T, the probability that a hyper-graph chosen according to

$\mathcal{D}(K_{n_1,\dots,n_d})$ contains a hyper-subgraph of type T with $m_{\max}$ edges is exponentially small in M. Specifically, we will prove the following result.

Lemma 3.2. *For $T \in \mathcal{T}$, let P_T denote the probability that a hyper-subgraph randomly chosen according to $\mathcal{D}(K_{n_1,\dots,n_d})$ contains a non-crossing hyper-subgraph J with $m_{\max}$ edges such that $T(\mathcal{P}(J)) = T$. Then, for some absolute constant $K_2 > 0$,*

$$P_T \leqslant \exp\left(-K_2 h \frac{\epsilon^2}{1+\epsilon} M\right).$$

We now proceed with the proof of the preceding result. Let $T = (e_1, \widetilde{e}_1, s_1, \dots, e_q, \widetilde{e}_q, s_q)$. As before, for all i, let $e_i = (v_1^{(i)}, v_2^{(i)}, \dots, v_q^{(i)})$ and $e_i = (\widetilde{v}_1^{(i)}, \widetilde{v}_2^{(i)}, \dots, \widetilde{v}_q^{(i)})$. Let H be chosen according to $\mathcal{D}(K_{n_1,\dots,n_d})$, and let H_i be the hyper-subgraph of H induced by the nodes between e_i and $\widetilde{e}_i$, that is,

$$v_1^{(i)}, v_1^{(i)}+1, \dots, \widetilde{v}_1^{(i)}, v_2^{(i)}, v_2^{(i)}+1, \dots, \widetilde{v}_2^{(i)}, \dots, v_d^{(i)}, v_d^{(i)}+1, \dots, \widetilde{v}_d^{(i)}.$$

Note that H_i is distributed according to $\mathcal{D}(K_{n_1^{(i)}, n_2^{(i)}, \dots, n_d^{(i)}})$, where $n_j^{(i)} = \widetilde{v}_j^{(i)} - v_j^{(i)} + 1$ is the size of the jth colour class of H_i. Moreover, if there is a non-crossing hyper-subgraph J of H such that $T(J) = T$, then it must hold that $L(H_i) \geqslant s_i$, for all $i = 1, \dots, q$. Since by hypothesis, $\mathcal{D}$ satisfies the block independence property, the events $L(H_i) \geqslant s_i$, $i = 1, \dots, q$, are independent, so

$$P_T \leqslant \prod_{i=1}^{q} \Pr\left[L\left(\mathcal{D}(K_{n_1^{(i)}, n_2^{(i)}, \dots, n_d^{(i)}})\right) \geqslant s_i\right].$$

Now, let N_i and S_i denote the geometric mean and sum of $n_1^{(i)}, \dots, n_d^{(i)}$, respectively. The ith term in the product of the last displayed equation will be small provided the sizes of the colour classes of H_i, *i.e.*, the $n_j^{(i)}$, satisfy the size conditions (1.1) and (1.2) of the definition of a (c, λ, θ)-median. Unfortunately, this may not occur for every i, somewhat complicating the analysis. Below we see how to handle this situation.

Since $T(\mathcal{P}(J)) = T$, we know that $n_1^{(i)}, n_2^{(i)}, \dots, n_d^{(i)} \leqslant L$. Recalling that $\alpha < \beta$ and applying (3.3), we conclude that $S_i b \leqslant dbL = C_g dbt^{\eta+\alpha} \leqslant C_g dbt^{\eta+\beta} \leqslant t^{\theta}$, so the size upper bound condition (1.2) of the definition of a (c, λ, θ)-median holds. However, the same might not be true regarding the size lower bound condition $N_i \geqslant at^{\lambda}$. In order to handle this situation, we artificially augment the size of the blocks where the condition fails. Specifically, for all $i = 1, \dots, q$ and $j = 1, \dots, d$ we define

$$\overline{n}_j^{(i)} = \max\{\delta n_j A t^{\lambda} / N, n_j^{(i)}\}.$$

Unsurprisingly, we let $\overline{N}_i$ and $\overline{S}_i$ denote the geometric mean and sum of the $\overline{n}_j^{(i)}$, respectively. Observe that when we augment the sizes of the colour classes of the hyper-graphs chosen, by the monotonicity property of random hyper-graph models, the probability of finding a non-crossing hyper-subgraph of size at least s_i increases. Hence,

$$P_T \leqslant \prod_{i=1}^{q} \Pr\left[L\left(\mathcal{D}\left(K_{n_i^{(1)}, \dots, n_i^{(d)}}\right)\right) \geqslant s_i\right] \leqslant \prod_{i=1}^{q} \Pr\left[L\left(\mathcal{D}\left(K_{\overline{n}_i^{(1)}, \dots, \overline{n}_i^{(d)}}\right)\right) \geqslant s_i\right]. \tag{3.8}$$

We claim that the $\overline{N}_i$ and $\overline{S}_i$ satisfy the size conditions (1.1) and (1.2) in the definition of a (c, λ, θ)-median. Indeed since $n_j^{(i)} \leqslant L$, and

$$\delta n_j \frac{At^\lambda}{N} \leqslant \delta n_j \frac{t^\alpha}{N} \leqslant \delta \frac{St^\alpha}{N} \leqslant \delta g(t) t^\alpha \leqslant \delta C_g t^{\alpha+\eta} = \delta L \leqslant L,$$

it follows that $\overline{n}_j^{(i)} \leqslant L$. Thence, by the same argument used to bound S_i before augmenting the block sizes, $\overline{S}_i b \leqslant t^\theta$. On the other hand, by (3.2) we have $A \geqslant a/\delta$; therefore

$$\overline{N}_i = \left(\prod_{j=1}^d \overline{n}_j^{(i)}\right)^{1/d} \geqslant \frac{\delta A t^\lambda}{N} \left(\prod_{j=1}^d n_j\right)^{1/d} = \delta A t^\lambda \geqslant a t^\lambda.$$

This concludes the proof of the stated claim.

Now, let $\overline{\mathsf{Med}}_i$ be a median of

$$L\Big(\mathcal{D}\Big(K_{\overline{n}_i^{(1)},\dots,\overline{n}_i^{(d)}}\Big)\Big).$$

By the definition of a (c, λ, θ)-median,

$$(1-\delta) c \overline{N}_i t^{-\lambda} \leqslant \overline{\mathsf{Med}}_i \leqslant (1+\delta) c \overline{N}_i t^{-\lambda}.$$

Hence, for all i such that $s_i \geqslant (1+\delta) c \overline{N}_i t^{-\alpha} \geqslant \overline{\mathsf{Med}}_i$, and using that h is a concentration constant for the random model $\mathcal{D}$, we get

$$\begin{aligned}\Pr\Big[L\Big(\mathcal{D}\Big(K_{\overline{n}_i^{(1)},\dots,\overline{n}_i^{(d)}}\Big)\Big) \geqslant s_i\Big] &\leqslant 2\exp\left(-h\frac{(s_i - \overline{\mathsf{Med}}_i)^2}{s_i}\right) \\ &\leqslant 2\exp\left(-h\frac{(s_i - (1+\delta)c\overline{N}_i t^{-\lambda})^2}{s_i}\right).\end{aligned}$$

Since $s_i \leqslant s_{\max}$, we get that for all i,

$$\begin{aligned}\Pr\Big[L\Big(\mathcal{D}\Big(K_{\overline{n}_i^{(1)},\dots,\overline{n}_i^{(d)}}\Big)\Big) \geqslant s_i\Big] &\leqslant 2\exp\left(-h\frac{(\max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\})^2}{s_i}\right) \\ &\leqslant 2\exp\left(-h\frac{(\max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\})^2}{s_{\max}}\right).\end{aligned}$$

Using the last bound and (3.8),

$$\begin{aligned}-\ln P_T &\geqslant -\ln\left(\prod_{i=1}^q \Pr\Big[L\Big(\mathcal{D}\Big(K_{\overline{n}_i,\dots,\overline{n}_i^{(d)}}\Big)\Big) \geqslant s_i\Big]\right) \\ &\geqslant -q\ln(2) + \frac{h}{s_{\max}} \sum_{i=1}^q \big(\max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\}\big)^2. \qquad (3.9)\end{aligned}$$

We now focus on the summation in the last term in the preceding displayed equation. We bound it from below by the following generalization of Hölder's inequality (see for example [22, Theorem 11]): For any collection of positive real numbers $x_{i,j}$, $1 \leqslant i \leqslant q$, $1 \leqslant j \leqslant d$,

$$\sum_{i=1}^q \prod_{j=1}^d x_{i,j} \leqslant \prod_{j=1}^d \left(\sum_{i=1}^q x_{i,j}^d\right)^{1/d}. \qquad (3.10)$$

Setting $x_{i,j} = (\overline{n}_j^{(i)})^{1/d}$ in the aforementioned inequality, observing that by the definition of $\overline{n}_j^{(i)}$ we have $\overline{n}_j^{(i)} \leqslant n_j^{(i)} + \delta n_j At^{\lambda}/N$, and recalling that the sum of $n_j^{(1)}, \dots, n_j^{(q)}$ is at most n_j,

$$\sum_{i=1}^{q} \overline{N}_i \leqslant \left(\prod_{j=1}^{d} \sum_{i=1}^{q} \overline{n}_j^{(i)}\right)^{1/d} \leqslant \left(\prod_{j=1}^{d} \sum_{i=1}^{q} (n_j^{(i)} + \delta n_j At^{\lambda}/N)\right)^{1/d} \leqslant N(1 + \delta q At^{\lambda}/N).$$

Because of our estimate (3.7) for q and (3.3), we conclude that

$$\sum_{i=1}^{q} \overline{N}_i \leqslant N(1 + 3\delta At^{\lambda-\alpha}) \leqslant N(1 + \delta/3) \leqslant N(1 + \delta).$$

By the Cauchy–Schwarz inequality and recalling that the sum of the s_i is $m_{\max} = \lceil (1 + \epsilon)M \rceil$,

$$\sqrt{q \sum_{i=1}^{q} \left(\max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\}\right)^2} \geqslant \sum_{i=1}^{q} \max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\}$$
$$\geqslant m_{\max} - (1+\delta)ct^{-\lambda} \sum_{i=1}^{q} \overline{N}_i \geqslant M(1+\epsilon) - M(1+\delta)^2.$$

Let us now see that the just derived lower bound is actually positive. Recall that by the definition of δ (see (3.1)) we know that $\delta \leqslant \epsilon/6$ and $\delta \leqslant 1$, so

$$(1+\epsilon) - (1+\delta)^2 = \epsilon - 2\delta - \delta^2 \geqslant \epsilon - 3\delta \geqslant \epsilon/2.$$

We then have

$$\sqrt{q \sum_{i=1}^{q} \left(\max\{0, s_i - (1+\delta)c\overline{N}_i t^{-\lambda}\}\right)^2} \geqslant \frac{\epsilon M}{2}.$$

Combining the last inequality with (3.9), and since $s_{\max} \leqslant (l/N)(1+\epsilon)M$, we find that

$$-\ln P_T \geqslant -q\ln(2) + \frac{h}{qs_{\max}} \cdot \frac{\epsilon^2 M^2}{4} \geqslant -q\ln(2) + \frac{hN\epsilon^2 M}{4q(1+\epsilon)l}.$$

Our estimate (3.7) for q states that $q \leqslant 3N/l$. Moreover, $l = t^{\alpha} \geqslant 9At^{\lambda}$ by (3.3); therefore

$$-\ln P_T \geqslant -\frac{N\ln(2)}{3At^{\lambda}} + \frac{h\epsilon^2 M}{12(1+\epsilon)} = \frac{cN}{t^{\lambda}}\left(\frac{h\epsilon^2}{12(1+\epsilon)} - \frac{\ln(2)}{3Ac}\right).$$

By (3.2) and definition (3.1) of δ we know that $A \geqslant 8\ln(2)/(hc\delta)$ and $\delta \leqslant \epsilon^2/(1+\epsilon)$, implying that

$$-\ln P_T \geqslant \frac{cN}{t^{\lambda}}\left(\frac{h\epsilon^2}{12(1+\epsilon)} - \frac{h\delta}{24}\right) \geqslant \frac{\epsilon^2}{1+\epsilon} \cdot \frac{hM}{24}.$$

We have thus shown that Lemma 3.2 holds taking $K_2 = 1/24$.

3.3.4. Upper tail bound. We are now ready to finally prove (1.6) for $N \geqslant t^{\beta}$. First, note that

$$\mathsf{Pr}\left[L\left(\mathcal{D}\left(K_{n_1,\ldots,n_d}\right)\right) \geqslant m_{\max}\right] \leqslant \sum_{T\in\mathcal{T}} P_T \leqslant |\mathcal{T}| \cdot \max_{T\in\mathcal{T}} P_T.$$

By Lemmas 3.1 and 3.2, using the fact that $l = t^{\alpha}$ by our choice of t_g, so (3.4) would hold, and recalling that by definition (3.1) of δ we have $\delta \leqslant \epsilon^2/(1+\epsilon)$, and given that $M = cNt^{-\lambda}$, we have

$$\begin{aligned}
&\mathsf{Pr}\left[L\left(\mathcal{D}\left(K_{n_1,\ldots,n_d}\right)\right) \geqslant m_{\max}\right]\\
&\quad\leqslant \exp\left(K_1\frac{N}{l}\ln(l) - K_2h\frac{\epsilon^2}{1+\epsilon}M\right) = \exp\left(K_1\alpha\frac{N}{t^{\alpha}}\ln(t) - K_2h\frac{\epsilon^2}{1+\epsilon}M\right)\\
&\quad\leqslant \exp\left(\frac{\delta K_2 h}{2}\frac{cN}{t^{\lambda}} - K_2h\frac{\epsilon^2}{1+\epsilon}M\right) \leqslant \exp\left(-\frac{K_2h\epsilon^2}{2(1+\epsilon)}M\right).
\end{aligned}$$

We thus conclude that (1.6) holds for any constant $K \leqslant K_2/2$ (since $K_2 = 1/24$, any $K \leqslant 1/48$ would do).

3.4. Upper bounds for the mean and median

We will now establish the two remaining unproved bounds claimed in the Main Theorem, *i.e.*, (1.3) and (1.4).

Fix $\epsilon = \epsilon_0 > 0$ and choose δ, A, α, β, C_g, t_g, K_1 and K_2 as in Section 3.1. We can view δ as a function of ϵ, henceforth denoted by $\delta(\epsilon)$. Similarly, we can view A and t_g as functions of δ, denoted by $A(\delta)$ and $t_g(\delta)$ respectively. Let A' be a sufficiently large constant such that

$$\exp\left(-Kh\frac{\epsilon_0^2}{4(1+\epsilon_0/2)}cA'\right) \leqslant \frac{\epsilon_0}{28}, \quad \text{and} \quad \frac{7}{6Kh} \leqslant \frac{\epsilon_0}{4}cA'. \qquad \text{(3.11: definition of } A')$$

Observe that by definition, $\delta(\epsilon) = 1$ for every $\epsilon \geqslant 6$. Moreover, let

$$\delta_0 = \delta(\epsilon_0/2), \quad \widetilde{A} = \max\{A(\delta_0), A(1), A'\}, \quad \text{and} \quad \widetilde{t}_g = \max\{t_g(\delta_0), t_g(1)\}.$$

Let $t \geqslant \widetilde{t}_g$ and consider the positive integers $n_1,\ldots,n_d$ with geometric mean N and summing S satisfying the size and balance constraints in the statement of the Main Theorem, *i.e.*,

$$N \geqslant \widetilde{A}t^{\lambda}, \quad \text{and} \quad Sb \leqslant g(t)N.$$

The choice of $\widetilde{t}_g$ and $\widetilde{A}$ guarantee that (1.6) holds for $\epsilon = \epsilon_0/2$ and for all $\epsilon \geqslant 6$.

As usual, let H be chosen according to $\mathcal{D}(K_{n_1,\ldots,n_d})$ and let $M = cNt^{-\lambda}$. We follow Iverson and use the bracket notation $[\![X]\!] = 1$ if event X is true, and $[\![X]\!] = 0$ otherwise. Observe that

$$\begin{aligned}
\mathsf{E}[L(H)] = {}&\mathsf{E}\left[L(H)\cdot[\![0 \leqslant L(H) < (1+\epsilon_0/2)M)]\!]\right]\\
&+ \mathsf{E}\left[L(H)\cdot[\![(1+\epsilon_0/2)M \leqslant L(H) < 7M]\!]\right] + \mathsf{E}\left[L(H)\cdot[\![L(H) \geqslant 7M]\!]\right].
\end{aligned}$$

Let us now upper-bound separately each of the terms in the right-hand side of the preceding displayed equation. The first one is trivially upper-bounded by $(1+\epsilon_0/2)M$.

Using (1.6), the fact that $N \geqslant \widetilde{A}t^{\lambda} \geqslant A't^{\lambda}$, and the definition (3.11) of A', we obtain

$$\begin{aligned}\mathsf{E}\big[L(H)\cdot [\![(1+\epsilon_0/2)M \leqslant L(H) < 7M]\!]\big] &\leqslant 7M\mathsf{Pr}\big[L(H) > (1+\epsilon_0/2)M\big]\\ &\leqslant 7M\exp\Big(-Kh\frac{\epsilon_0^2}{4(1+\epsilon_0/2)}\cdot\frac{cN}{t^{\lambda}}\Big) \leqslant 7M\exp\Big(-Kh\frac{\epsilon_0^2}{4(1+\epsilon_0/2)}\cdot cA'\Big) \leqslant \frac{M\epsilon_0}{4}.\end{aligned}$$

Now let us consider the third term. For $\epsilon \geqslant 6$, (1.6) holds and $\epsilon/(1+\epsilon) \geqslant 6/7$. Furthermore, using that $M = cNt^{-\lambda} \geqslant c\widetilde{A} \geqslant cA'$, and the definition (3.11) of A',

$$\begin{aligned}\mathsf{E}\big[L(H)\cdot [\![L(H) \geqslant 7M]\!]\big] &= \int_{7M}^{\infty}\mathsf{Pr}\big[L(H) > t\big]dt = M\int_{6}^{\infty}\mathsf{Pr}\big[L(H) > (1+\epsilon)M\big]d\epsilon\\ &\leqslant M\int_{6}^{\infty}\exp\Big(-Kh\frac{\epsilon^2}{1+\epsilon}\cdot M\Big)d\epsilon \leqslant M\int_{6}^{\infty}\exp\Big(-\frac{6Kh}{7}\cdot M\epsilon\Big)d\epsilon\\ &= M\Big(\frac{6Kh}{7}M\Big)^{-1}\exp\Big(-\frac{36Kh}{7}\cdot M\Big) \leqslant M\Big(\frac{6Kh}{7}\cdot cA'\Big)^{-1} \leqslant \frac{M\epsilon_0}{4}.\end{aligned}$$

Summarizing, we have that $\mathsf{E}[L(H)] \leqslant (1+\epsilon_0)M$, which proves (1.3).

Finally, we establish (1.4). Again, let $\epsilon > 0$ and choose δ, A, α, β, C_g, t_g, K_1 and K_2 as in Section 3.1. Let

$$A' = \max\Big\{A, \frac{(1+\epsilon)\ln(2)}{Khc\epsilon^2}\Big\}. \qquad \text{(3.12: definition of } A')$$

Now, let $t \geqslant t_g$ and $n_1, \ldots, n_d$ be positive integers with geometric mean N and summing up to S, satisfying the size and balance constraints with respect to the just defined constant A', that is,

$$N \geqslant A't^{\lambda}, \quad \text{and} \quad Sb \leqslant g(t)N.$$

By (1.6) and (3.12), it follows that

$$\mathsf{Pr}\big[L(H) \geqslant (1+\epsilon)M\big] \leqslant \exp\Big(-Kh\frac{\epsilon^2}{1+\epsilon}\cdot\frac{cN}{t^{\lambda}}\Big) \leqslant \exp\Big(-Kh\frac{\epsilon^2}{1+\epsilon}\cdot cA'\Big) \leqslant \frac{1}{2}.$$

Hence, every median of $L(H)$ is at most $(1+\epsilon)M$, thus establishing (1.4) and completing the proof of the Main Theorem. □

4. Applications

4.1. Preliminaries

For future reference we determine below concentration constants for the binomial and word models.

Proposition 4.1. *The d-dimensional binomial random hyper-graph model admits a concentration constant of* $1/4$. *The random d-word model admits a concentration constant of* $1/(4d)$.

Proof. Let H be chosen according to $\mathcal{G}(K_{n_1,\ldots,n_d}, p)$. Since $L(H)$ depends exclusively on whether or not an edge appears in H (and by independence among these events), it follows that $L(H)$ is 1-Lipschitz, *i.e.*, $|L(H) - L(H\triangle\{e\})| \leqslant 1$. Moreover, if $L(H) \geqslant r$, then there is a set of r edges witnessing the fact that $L(H) \geqslant r$, for every H containing such a set of r edges. A direct application of Talagrand's inequality (as stated in [23, Theorem 2.29]) proves the claim about the concentration constant for the d-dimensional binomial random hyper-graph model. The case of the random d-word model is similar and left to the reader to verify. □

Also, for the sake of conciseness and in order to avoid irritating reiteration of by now default notation, throughout this section, whenever not explicitly defined, N and S always denote the geometric mean and sum of the positive integers $n_1, \ldots, n_d$, respectively.

4.2. Random binomial hyper-graph model

In this section we apply the Main Theorem to the d-dimensional binomial random hyper-graph model.

We will show that the constant c in the definition of a (c, λ, θ)-median for this model is related to a constant that arises in the study of the asymptotic behaviour of the length of a longest increasing subsequence of $d-1$ randomly chosen permutations of $[n]$, when n goes to infinity. We first recall some known facts about this problem. Given positive integers d and n, consider d permutations $\pi_1, \ldots, \pi_d$ of $[n]$. We say that $\{(i_j, \pi_1(i_j), \ldots, \pi_d(i_j)) \mid 1 \leqslant j \leqslant \ell\}$ is an increasing sequence of $(\pi_1, \ldots, \pi_d)$ of length ℓ if $i_1 < i_2 < \cdots < i_\ell$ and $\pi_t(i_1) < \pi_t(i_2) < \cdots < \pi_t(i_\ell)$ for $1 \leqslant t \leqslant d$. We let $\mathsf{lis}_{d+1}(n)$ denote the random variable corresponding to the length of a longest increasing subsequence of $(\pi_1, \ldots, \pi_d)$ when $\pi_1, \ldots, \pi_d$ are randomly and uniformly chosen. The study of the asymptotic characteristics of the distribution of $\mathsf{lis}_d(n)$ will be henceforth referred to as Ulam's problem in d dimensions (note that the $d = 2$ case corresponds precisely to the setting discussed in the first paragraph of the Introduction).

Ulam's problem in d dimensions can be restated geometrically as follows. Consider choosing $\vec{x}(1), \ldots, \vec{x}(n)$ uniformly and independently in the d-dimensional unit cube $[0,1]^d$ endowed with the natural componentwise partial order. Let $H_d(n)$ be the length of a largest chain $C \subseteq \{\vec{x}(1), \ldots, \vec{x}(n)\}$. It is not hard to see that $H_d(n)$ and $\mathsf{lis}_d(n)$ follow the same distribution. Bollobás and Winkler [11] have shown that for every d there exists a constant c_d such that $H_d(n)/\sqrt[d]{n}$ (and thus also $\mathsf{lis}_d(n)/\sqrt[d]{n}$) goes to c_d as $n \to \infty$. Only the values $c_1 = 1$ and $c_2 = 2$ are known for these constants. However, in [11] it is shown that $c_i \leqslant c_{i+1}$ and $c_i < e$ for all i, and that $\lim_{d \to \infty} c_d = e$.

Now, let us go back to our discussion concerning the random binomial hyper-graph model. Our immediate goal is to estimate a median of $L(\mathcal{G}(K_{n_1,\ldots,n_d}, p))$. Consider H chosen according to $\mathcal{G}(K_{n_1,\ldots,n_d}, p)$ and let H' be the hyper-subgraph of H obtained from H after removal of all edges incident to nodes of degree at least 2. Let $E = E(H)$ and $E' = E(H')$. In order to approximate a median of $L(\mathcal{G}(K_{n_1,\ldots,n_d}, p))$ it will be useful to estimate first the expected value of $L(H')$. We now come to a crucial observation: $L(H')$ is precisely the length of a largest chain (for the natural order among edges) contained in E', or equivalently the length of a longest increasing subsequence of $d-1$ permutations of $\{1, \ldots, |E'|\}$. The preceding observation will enable us to build on the known results concerning Ulam's problem and use them in the analysis of the longest non-crossing matching problem for the random binomial hyper-graph model. In particular, the following concentration result due to Bollobás and Brightwell [10] for the length of a d-dimensional longest increasing subsequence will be useful for our purposes.

Theorem 4.2 (Bollobás and Brightwell [10, Theorem 8]). *For every $d \geqslant 2$, there is a constant D_d such that, for m sufficiently large and $2 < \lambda < m^{1/2d}/\log\log m$,*

$$\mathsf{Pr}\left[|\mathsf{lis}_d(m) - \mathsf{E}\left[\mathsf{lis}_d(m)\right]| > \frac{\lambda D_d m^{1/2d}\log(m)}{\log\log(m)}\right] \leqslant 80\lambda^2 e^{-\lambda^2}.$$

We will not directly apply the preceding result. Instead, we rely on the following.

Corollary 4.3. *For every $d \geqslant 2$, $t > 0$ and $\alpha > 0$, there exists $m_0 = m_0(t, \alpha, d)$ sufficiently large that if $m \geqslant m_0$, then*

$$\mathsf{Pr}\left[|\mathsf{lis}_d(m) - c_d m^{1/d}| > t c_d m^{1/d}\right] \leqslant \alpha.$$

Proof. Let D_d be the constant in the statement of Theorem 4.2. By definition of Ulam's constant, we know that $\lim_{n\to\infty} \mathsf{E}[\mathsf{lis}_d(m)]/\sqrt[d]{m} = c_d$. Hence, we can choose $m_0 = m_0(t, \alpha, d)$ sufficiently large that for all $m \geqslant m_0$, Theorem 4.2 holds, and in addition the following conditions are satisfied:

$$|\mathsf{E}\left[\mathsf{lis}_d(m)\right] - c_d m^{1/d}| < \frac{1}{2} t c_d m^{1/d}, \tag{4.1}$$

$$\lambda = \lambda(m) \overset{\text{def}}{=} \frac{t c_d}{2D_d} \cdot \frac{m^{1/2d}\log\log(m)}{\log(m)} \leqslant \frac{m^{1/2d}}{\log\log(m)} \quad \text{and} \quad 80\lambda^2 e^{-\lambda^2} \leqslant \alpha. \tag{4.2}$$

(Both conditions can be satisfied since $(\log\log(m))^2 = o(\log(m))$ and $\lambda(m) \to \infty$ when $m \to \infty$.) It follows that for all $m > m_0$,

$$\begin{aligned}
&\mathsf{Pr}\left[|\mathsf{lis}_d(m) - c_d m^{1/d}| > t c_d m^{1/d}\right] \\
&\quad\leqslant \mathsf{Pr}\left[|\mathsf{lis}_d(m) - \mathsf{E}\left[\mathsf{lis}_d(m)\right]| + |\mathsf{E}\left[\mathsf{lis}_d(m)\right] - c_d m^{1/d}| > t c_d m^{1/d}\right] \\
&\quad\leqslant \mathsf{Pr}\left[|\mathsf{lis}_d(m) - \mathsf{E}\left[\mathsf{lis}_d(m)\right]| > \frac{1}{2} t c_d m^{1/d}\right] \\
&\quad= \mathsf{Pr}\left[|\mathsf{lis}_d(m) - \mathsf{E}\left[\mathsf{lis}_d(m)\right]| > \frac{\lambda D_d m^{1/2d}\log(m)}{\log\log(m)}\right] \\
&\quad\leqslant 80\lambda^2 e^{-\lambda^2}.
\end{aligned}$$

□

For future reference, we recall a well-known variant of Chebyshev's inequality.

Proposition 4.4 (Chebyshev's inequality for indicator random variables). *Let $X_1, \ldots, X_m$ be random variables taking values in $\{0, 1\}$ and let X denote $X_1 + \cdots + X_m$. Also, let*

$$\Delta = \sum_{i,j:i\neq j} \mathsf{E}[X_i X_j].$$

Then, for all $t \geqslant 0$,

$$\mathsf{Pr}\left[|X - \mathsf{E}[X]| \geqslant t\right] \leqslant \frac{1}{t^2}\left(\mathsf{E}[X](1 - \mathsf{E}[X]) + \Delta\right).$$

Moreover, if $X_1, \ldots, X_m$ are independent, then

$$\mathsf{Pr}\left[|X - \mathsf{E}[X]| \geqslant t\right] \leqslant \frac{\mathsf{E}[X]}{t^2}.$$

Proof. Observe that since X_i is an indicator variable, then $\mathsf{E}[X_i^2] = \mathsf{E}[X_i]$. Thus, if we let $\mathsf{V}[X]$ denote the variance of X,

$$\mathsf{V}[X] = \mathsf{E}[X^2] - (\mathsf{E}[X])^2 = \sum_{i=1}^{m} \mathsf{E}[X_i^2] + \Delta - (\mathsf{E}[X])^2 = \mathsf{E}[X](1 - \mathsf{E}[X]) + \Delta.$$

A direct application of Chebyshev's inequality yields the first bound claimed. The second stated bound follows from the first one, and the fact that if $X_1, \ldots, X_m$ are independent, then $\Delta \leqslant (\mathsf{E}[X])^2$. □

We will also need the following two lemmas.

Lemma 4.5. *Let N and S denote the geometric mean and sum of $n_1, \ldots, n_d$. If*

$$\widetilde{N} = \left(\prod_{j=1}^{d}(n_j - 1)\right)^{d},$$

then $N^d - \widetilde{N}^d \leqslant S^{d-1}$.

Proof. Direct application of (3.10). □

Lemma 4.6. *Let N and S denote the geometric mean and sum of $n_1, \ldots, n_d$. If*

$$\widetilde{N} = \left(\prod_{j=1}^{d}(n_j - 1)\right)^{d},$$

then the following hold:

$$\mathsf{E}[|E|] = N^d p, \tag{4.3}$$

$$\mathsf{E}[|E'|] = N^d p(1-p)^{N^d - \widetilde{N}^d} \geqslant N^d p(1 - S^{d-1}p), \tag{4.4}$$

$$\mathsf{E}[|E \setminus E'|] \leqslant N^d S^{d-1} p^2. \tag{4.5}$$

Moreover, for all $\eta > 0$,

$$\mathsf{Pr}\left[|E| - \mathsf{E}[|E|] \geqslant \eta \mathsf{E}[|E|]\right] \leqslant \frac{1}{\eta^2 \mathsf{E}[|E|]}. \tag{4.6}$$

Proof. Let $K = K_{n_1,\dots,n_k}$, and for each $e \in E(K)$ let X_e and Y_e denote the indicators of the events $e \in E$ and $e \in E'$, respectively. Note that

$$|E| = \sum_{e \in E(K)} X_e \quad \text{and} \quad |E'| = \sum_{e \in E(K)} Y_e.$$

Clearly, $\mathsf{E}[X_e] = p$ for all $e \in E(K)$. Moreover, $e \in E'$ if and only if $e \in E$ and no edge $f \in E \setminus \{e\}$ intersects e. Since the number of edges in $E(K)$ that intersect any given $e \in E(K)$ is exactly $N^d - \widetilde{N}^d$, we have that

$$\mathsf{E}[Y_e] = p(1-p)^{N^d - \widetilde{N}^d}.$$

Observing that $|E(K)| = N^d$, we obtain (4.3) and the first equation in (4.4). On the other hand, since $(1-p)^m \geqslant 1 - pm$ and by Lemma 4.5, we can finish the proof of (4.4) by noting that

$$\mathsf{E}[|E'|] = N^d p(1-p)^{N^d - \widetilde{N}^d} \geqslant N^d p(1 - (N^d - \widetilde{N}^d)p) \geqslant N^d p(1 - S^{d-1}p).$$

Inequality (4.5) is a consequence of (4.3), (4.4), and the fact that $E' \subseteq E$, as follows:

$$\mathsf{E}[|E \setminus E'|] = \mathsf{E}[|E| - |E'|] \leqslant N^d S^{d-1} p^2.$$

Applying Chebyshev's inequality for independent indicator random variables to the collection $\{X_e \mid e \in E(K)\}$ yields (4.6). □

We are now ready to exploit the fact, already mentioned, that $L(H')$ equals the length of a longest increasing subsequence of $d-1$ permutations of $\{1,\dots,|E'|\}$, and then apply Corollary 4.3 in order to estimate its value. Formally, we prove the following claim.

Proposition 4.7. *Let $\delta > 0$, $d \geqslant 2$, and N and S be the geometric mean and sum of positive integers $n_1,\dots,n_d$, respectively. Moreover, let $M = c_d N p^{1/d}$, where c_d is the d-dimensional Ulam constant. Then, there is a constant $C = C(\delta, d)$ sufficiently large that the following hold.*

- *If $Np^{1/d} \geqslant C$ and $12S^{2d-2}p^{2-1/d} \leqslant d^{d-1}\delta c_d$, then every median of $L(\mathcal{G}(K_{n_1,\dots,n_d}, p))$ is at most $(1+\delta)M$.*
- *If $Np^{1/d} \geqslant C$ and $12S^{d-1}p \leqslant \delta$, then every median of $L(\mathcal{G}(K_{n_1,\dots,n_d}, p))$ is at least $(1-\delta)M$.*

Proof. To prove that every median of $L(\mathcal{G}(K_{n_1,\dots,n_d}, p))$ is at most $(1+\delta)M$, it suffices to show that $\mathsf{Pr}[L(H) \geqslant (1+\delta)M]$ is at most $1/2$. To establish the latter, note that $L(H) \leqslant L(H') + |E \setminus E'|$, and hence

$$\begin{aligned}\mathsf{Pr}\big[L(H) \geqslant (1+\delta)M\big] &\leqslant \mathsf{Pr}\left[|E \setminus E'| \geqslant \frac{M\delta}{2}\right] + \mathsf{Pr}\big[L(H') \geqslant (1+\delta/2)M\big] \\ &\leqslant \mathsf{Pr}\left[|E \setminus E'| \geqslant \frac{M\delta}{2}\right] + \mathsf{Pr}\left[|E'| \geqslant (1+\delta/2)\frac{M^d}{c_d^d}\right] \\ &\quad + \mathsf{Pr}\left[L(H') \geqslant (1+\delta/2)M, |E'| < (1+\delta/2)\frac{M^d}{c_d^d}\right].\end{aligned}$$

We now separately upper-bound each of the latter three terms. For the first one, we rely on Markov's inequality, inequality (4.4) of Lemma 4.6, the fact that $N \leqslant S/d$, and our hypothesis, to conclude that

$$\begin{aligned}\mathsf{Pr}\left[|E \setminus E'| \geqslant \frac{M\delta}{2}\right] &\leqslant \frac{2}{M\delta}\mathsf{E}\left[|E \setminus E'|\right] \leqslant \frac{2N^d S^{d-1} p^2}{\delta c_d N p^{1/d}} \\ &= \frac{2N^{d-1} S^{d-1} p^{2-1/d}}{\delta c_d} \leqslant \frac{2S^{2d-2} p^{2-1/d}}{d^{d-1}\delta c_d} \leqslant \frac{1}{6}.\end{aligned}$$

To bound the second term, note that $|E| \geqslant |E'|$, and recall (4.3) and (4.6) of Lemma 4.6, so

$$\mathsf{Pr}\left[|E'| \geqslant (1+\delta/2)\frac{M^d}{c_d^d}\right] = \mathsf{Pr}\left[|E| \geqslant (1+\delta/2)\mathsf{E}[|E|]\right] \leqslant \frac{4}{\delta^2 \mathsf{E}[|E|]} = \frac{4}{\delta^2 N^d p}.$$

Since by assumption $N^d p \geqslant C^d$, it suffices to take $C^d \geqslant 24/\delta^2$ in order to derive an upper bound of $1/6$ for the second term.

Finally, we focus on the third term. Let $m = \lfloor (1+\delta/2)M^d/c_d^d \rfloor$. Recall that conditioned on $|E'| = n'$, the random variable $L(H')$ follows the same distribution as $\mathsf{lis}(n')$. Thus, since $n \geqslant n'$ implies that $\mathsf{lis}(n)$ dominates $\mathsf{lis}(n')$, and given that $(1+x)^a \leqslant 1 + ax$ for $x \geqslant -1$ and $0 < a < 1$,

$$\begin{aligned}\mathsf{Pr}\left[L(H') \geqslant (1+\delta/2)M, |E'| < (1+\delta/2)\frac{M^d}{c_d^d}\right] &\leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \geqslant \frac{1+\delta/2}{(1+\delta/2)^{1/d}} c_d m^{1/d}\right] \\ \leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \geqslant \frac{1+\delta/2}{1+\delta/(2d)} c_d m^{1/d}\right] &= \mathsf{Pr}\left[\mathsf{lis}_d(m) \geqslant \left(1 + \frac{(d-1)\delta}{2d+\delta}\right) c_d m^{1/d}\right].\end{aligned}$$

Setting $t = (d-1)\delta/(2d+\delta)$ and requiring that $C^d \geqslant m_0 + 1$ with $m_0 = m_0(t, 1/6, d)$ as in Corollary 4.3, and since by assumption $N^d p \geqslant C^d$, we have

$$m = \lfloor (1+\delta/2)M^d/c_d^d \rfloor = \lfloor (1+\delta/2)N^d p \rfloor \geqslant \lfloor C^d \rfloor \geqslant m_0.$$

Thus, we can apply Corollary 4.3 and conclude that

$$\mathsf{Pr}\left[L(H') \geqslant (1+\delta/2)M, |E'| < (1+\delta/2)M^d/c_d^d\right] \leqslant \frac{1}{6}.$$

In summary, $\mathsf{Pr}\left[L(H) \geqslant (1+\delta)M\right] \leqslant 3(1/6) = 1/2$, as we wanted to show.

Now, to prove that every median of $L(\mathcal{G}(K_{n_1,\dots,n_d}, p))$ is at least $(1-\delta)M$, it suffices to show that $\mathsf{Pr}\left[L(H) \leqslant (1-\delta)M\right]$ is at most $1/2$. Note that $L(\cdot)$ is non-negative, so we can always assume that $\delta \leqslant 1$. Since $L(H') \leqslant L(H)$,

$$\begin{aligned}&\mathsf{Pr}\left[L(H) \leqslant (1-\delta)M\right] \\ &\quad\leqslant \mathsf{Pr}\left[|E| \leqslant (1-\delta)\frac{M^d}{c_d^d}\right] + \mathsf{Pr}\left[L(H') \leqslant (1-\delta)M, |E| > (1-\delta)\frac{M^d}{c_d^d}\right] \\ &\quad\leqslant \mathsf{Pr}\left[|E| \leqslant (1-\delta)\frac{M^d}{c_d^d}\right] + \mathsf{Pr}\left[|E \setminus E'| \geqslant (\delta/2)\frac{M^d}{c_d^d}\right] \\ &\qquad + \mathsf{Pr}\left[L(H') \leqslant (1-\delta)M, |E'| > (1-\delta/2)\frac{M^d}{c_d^d}\right].\end{aligned}$$

As above, we separately bound each of the three latter terms. In the case of the first term, by (4.3) and (4.6) of Lemma 4.6,

$$\mathsf{Pr}\left[|E| \leqslant (1-\delta)\frac{M^d}{c_d^d}\right] = \mathsf{Pr}\left[|E| \leqslant (1-\delta)\mathsf{E}[|E|]\right] \leqslant \frac{1}{\delta^2\mathsf{E}[|E|]} = \frac{1}{\delta^2 N^d p}.$$

Since by assumption $N^d p \geqslant C^d$, it suffices to take $C^d \geqslant 6/\delta^2$ in order to establish an upper bound of $1/6$ for the term under consideration.

To bound the second term, simply apply Markov's inequality, use (4.5) of Lemma 4.6, and recall that by assumption $12pS^{d-1} \leqslant \delta$. An upper bound of $1/6$ follows for the term under consideration.

Now, to bound the third term, let $m = \lceil (1-\delta/2)M^d/c_d^d \rceil$. Recall that conditioned on $|E'| = n'$, the random variable $L(H')$ follows the same distribution as $\mathsf{lis}(n')$. Thus, since $n' \geqslant n$ implies that $\mathsf{lis}(n')$ dominates $\mathsf{lis}(n)$, some basic arithmetic and given that $(1+x)^a \leqslant 1 + ax$ for $x \geqslant -1$ and $0 < a < 1$,

$$\begin{aligned}
&\mathsf{Pr}\left[L(H') \leqslant (1-\delta)M, |E'| > (1-\delta/2)\frac{M^d}{c_d^d}\right] \leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \leqslant (1-\delta)M\right] \\
&\leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \leqslant \frac{1-\delta}{(1-\delta/2)^{1/d}} c_d m^{1/d}\right] \leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \leqslant (1-\delta/2)^{1-1/d} c_d m^{1/d}\right] \\
&\leqslant \mathsf{Pr}\left[\mathsf{lis}_d(m) \leqslant \left(1 - \frac{\delta}{2}\left(1-\frac{1}{d}\right)\right) c_d m^{1/d}\right].
\end{aligned}$$

Taking $t = (\delta/2)(1-1/d)$, requiring that $C \geqslant (m_0/(1-\delta/2))^{1/d}$ with $m_0 = m_0(t, 1/6, d)$ as in Corollary 4.3, and since by assumption $N^d p \geqslant C^d$, we get

$$m \geqslant (1-\delta/2)\frac{M^d}{c_d^d} = (1-\delta/2)N^d p \geqslant (1-\delta/2)C^d \geqslant m_0.$$

Thus, we can apply Corollary 4.3 and conclude that the third term is also upper-bounded by $1/6$.

Summarizing, $\mathsf{Pr}\left[L(H) \leqslant (1-\delta)M\right] \leqslant 3(1/6) = 1/2$, as we wanted to show. □

Corollary 4.8. *Let $d \geqslant 2$. If $t = 1/p$, then the model $(\mathcal{G}(K_{n_1,\ldots,n_d}, p))$ of internal parameter t admits a (c, λ, θ)-median where*

$$(c, \lambda, \theta) = \left(c_d, \frac{1}{d}, \frac{2d-1}{2d(d-1)}\right).$$

Proof. Let H be chosen according to $\mathcal{G}(K_{n_1,\ldots,n_d}, p)$, $M = cN/t^\lambda = c_d N p^{1/d}$, $\delta > 0$, and let $C(\delta)$ be as in Proposition 4.7. Define $a(\delta) = C(\delta)$, $b(\delta) = (12/(\delta d^{d-1} c_d))^{1/(2d-2)}$ and $t'(\delta)$ sufficiently large that $t > t'(\delta)$ and $t^{1-1/(2d)} < (\delta/12)t(b(\delta))^{d-1}$. Note that if $t > t'(\delta)$, $N \geqslant a(\delta)t^{1/d}$, and $Sb(\delta) \leqslant t^{(2d-1)/(2d(d-1))}$, then the hypothesis of Proposition 4.7 will be satisfied, and thence every median of $L(H)$ will be between $(1-\delta)M$ and $(1+\delta)M$. □

Recalling that by Proposition 4.1 we know that $h = 1/4$ is a concentration constant for the d-dimensional binomial random hyper-graph model, by Corollary 4.8 and the Main Theorem, we obtain the following.

Theorem 4.9. *Let $\epsilon > 0$ and $g : \mathbb{R} \to \mathbb{R}$ be such that, for a given $0 \leqslant \eta < 1/(2d(d-1))$, we obtain $g(t) = O(t^{\eta})$. Fix $n_1, \dots, n_d$ and let N and S denote their geometric mean and sum, respectively. There exists a sufficiently small p_0 and sufficiently large A such that if $p \leqslant p_0$, $Np^{1/d} \geqslant A$ and $S \leqslant g(1/p)N$, then for $M = c_d N p^{1/d}$ where c_d is the d-dimensional Ulam constant,*

$$(1-\epsilon)M \leqslant \mathsf{E}\left[L(\mathcal{G}(K_{n_1,\dots,n_d}, p))\right] \leqslant (1+\epsilon)M,$$

and the following hold.

- *If $\mathsf{Med}\left[L(\mathcal{G}(K_{n_1,\dots,n_d}, p))\right]$ is a median of $L(\mathcal{G}(K_{n_1,\dots,n_d}, p))$,*

$$(1-\epsilon)M \leqslant \mathsf{Med}\left[L(\mathcal{G}(K_{n_1,\dots,n_d}, p))\right] \leqslant (1+\epsilon)M.$$

- *There is an absolute constant $C > 0$ such that*

$$\mathsf{Pr}\left[L(\mathcal{G}(K_{n_1,\dots,n_d}, p)) \leqslant (1-\epsilon)M\right] \leqslant \exp(-C\epsilon^2 M),$$
$$\mathsf{Pr}\left[L(\mathcal{G}(K_{n_1,\dots,n_d}, p)) \geqslant (1+\epsilon)M\right] \leqslant \exp\left(-C\frac{\epsilon^2}{1+\epsilon}M\right).$$

We are now ready to prove Theorem 1.2, which is the main result of this section, and which was stated in the main contributions section.

Proof of Theorem 1.2. Let n, n', n'' be positive integers such that $n = n' + n''$. Clearly,

$$\mathsf{E}\left[L(\mathcal{G}(K_n^{(d)}, p))\right] \geqslant \mathsf{E}\left[L(\mathcal{G}(K_{n'}^{(d)}, p))\right] + \mathsf{E}\left[L(\mathcal{G}(K_{n''}^{(d)}, p))\right].$$

By subadditivity, it follows that the limit of $\mathsf{E}\left[L(\mathcal{G}(K_n^{(d)}, p))\right]$ when normalized by n exists and equals $\delta_{p,d} = \inf_{n\in\mathbb{N}} \mathsf{E}\left[L(\mathcal{G}(K_n^{(d)}, p))/n\right]$. A direct application of Theorem 4.9 yields that $\delta_{p,d}/\sqrt[d]{p} \to c_d$ when $p \to 0$. □

4.3. Random word model

In this section, we consider the random d-word model. The structure, arguments and type of derived results are similar to those obtained in the preceding section. However, the intermediate calculations are somewhat longer and more involved.

As in the preceding section, we first show that the random model under consideration admits a (c, λ, θ)-median. Now consider H chosen according to $\Sigma(K_{n_1,\dots,n_d}, k)$ and let H' be the hyper-subgraph of H obtained from H as in the preceding section (*i.e.*, by removal of all edges incident to nodes of degree at least 2). Let $E = E(H)$ and $E' = E(H')$. For the random word model, the analogue of Lemma 4.6 is as follows.

Lemma 4.10. *For positive integers $n_1,\dots,n_d$, let N and S be the geometric mean and sum, respectively. Then,*

$$\mathsf{E}[|E|] = \frac{N^d}{k^{d-1}}, \tag{4.7}$$

$$\mathsf{E}[|E'|] = \frac{N^d}{k^{d-1}}\left(\frac{k-1}{k}\right)^{S-d} \geqslant \frac{N^d}{k^{d-1}}\left(1-\frac{S}{k}\right), \tag{4.8}$$

$$\mathsf{E}[|E \setminus E'|] \leqslant \frac{N^d S}{k^d}. \tag{4.9}$$

Moreover, for all $\eta > 0$,

$$\mathsf{Pr}[|E'| - \mathsf{E}[|E'|] \geqslant \eta \mathsf{E}[|E'|]] \leqslant \frac{1}{\eta^2 \mathsf{E}[|E'|]} + \frac{1}{\eta^2}\left(\left(\frac{k-1}{k-2}\right)^{2d-1} - 1\right). \tag{4.10}$$

Proof sketch. Consider K, X_e, and Y_e exactly as defined in the proof of Lemma 4.6. Note that now

$$\mathsf{E}\left[X_e\right] = 1/k^{d-1} \quad \text{and} \quad \mathsf{E}\left[Y_e\right] = (1-1/k)^{S-d}/k^{d-1} \geqslant (1-S/k)/k^{d-1}.$$

Recalling that $|E(K)| = N^d$, both (4.7) and the first equation of (4.8) follow. To conclude (4.8) and (4.9), just note that

$$\mathsf{E}[|E \setminus E'|] = \sum_{e \in E(K)} \left(\mathsf{E}[X_e] - \mathsf{E}[Y_e]\right) \leqslant \frac{N^d S}{k^d}.$$

The last inequality in the statement of the lemma follows from Proposition 4.4, and noting that if $e \cap f \neq \emptyset$, then $\mathsf{E}[Y_e Y_f] = 0$, while if $e \cap f = \emptyset$, then

$$\mathsf{E}[Y_e Y_f] = \mathsf{E}[Y_e]\mathsf{E}[Y_f]\left(\frac{k(k-2)}{(k-1)^2}\right)^{S-1}\left(\frac{k-1}{k-2}\right)^{2d-1} \leqslant \mathsf{E}[Y_e]\mathsf{E}[Y_f]\left(\frac{k-1}{k-2}\right)^{2d-1}. \qquad \square$$

We can now estimate the median of $L(\Sigma(K_{n_1,\dots,n_d},k))$.

Proposition 4.11. *Let $\delta > 0$, $d \geqslant 2$, and N and S be the geometric mean and sum of positive integers $n_1,\dots,n_d$, respectively. Moreover, let $M = c_d N/k^{1-1/d}$, where c_d is the d-dimensional Ulam constant. Then, there are sufficiently large constants $C = C(\delta)$ and $K = K(\delta)$ such that the following hold.*

- *If $k \geqslant K$, $N \geqslant Ck^{1-1/d}$, $12S^d \leqslant \delta c_d k^{d-1+1/d}$, and $S \leqslant k/2$, then every median of $L(\Sigma(K_{n_1,\dots,n_d},k))$ is upper-bounded by $(1+\delta)M$.*
- *If $k \geqslant K$, $N \geqslant Ck^{1-1/d}$, and $S \leqslant \delta k/2$, then every median of $L(\Sigma(K_{n_1,\dots,n_d},k))$ is at least $(1-\delta)M$.*

Proof sketch. We proceed as in the proof of Proposition 4.7. To establish the first stated claim it suffices to show that $\mathsf{Pr}\left[L(H) \geqslant (1+\delta)M\right]$ is at most 1/2. Thus, it is enough to show that under the hypothesis of the first item, each of the following three terms can be bounded by 1/6:

$$\Pr\left[|E \setminus E'| \geqslant \frac{M\delta}{2}\right], \qquad \Pr\left[|E'| \geqslant (1+\delta/2)\frac{M^d}{c_d^d}\right],$$

$$\Pr\left[L(H') \geqslant (1+\delta/2)M, |E'| < (1+\delta/2)\frac{M^d}{c_d^d}\right].$$

To bound the first term, we rely on Markov's inequality, the fact that $N < S$, inequality (4.9) of Lemma 4.10, and the hypothesis, to obtain

$$\Pr\left[|E \setminus E'| \geqslant \frac{M\delta}{2}\right] \leqslant \frac{2SN^{d-1}}{\delta c_d k^{d-1+1/d}} \leqslant \frac{2S^d}{\delta c_d k^{d-1+1/d}} \leqslant \frac{1}{6}.$$

To bound the second term, we note that by (4.8) of Lemma 4.10 we have that

$$M^d/c_d^d \geqslant \mathsf{E}\left[|E'|\right] \geqslant (1 - S/k)N^d/k^{d-1},$$

and we then apply (4.10) of Lemma 4.10, to obtain

$$\Pr\left[|E'| \geqslant (1+\delta/2)\frac{M^d}{c_d^d}\right] \leqslant \frac{4k^{d-1}}{\delta^2 N^d(1-S/k)} + \frac{4}{\delta^2}(48\delta^2).$$

By hypothesis, $S \leqslant k/2$, so if $K = K(\delta)$ is sufficiently large, then the right-hand side of the last displayed inequality is at most $8/(\delta^2 C^d) + 1/12$, which is at most 1/6 provided $C^d \geqslant 96/\delta^2$.

Now, to bound the third term, we consider $m = \lfloor (1+\delta/2)M^d/c_d^d \rfloor$ and proceed as in Proposition 4.7, and similarly obtain

$$\Pr\left[L(H') \geqslant (1+\delta/2)M, |E'| < (1+\delta/2)\frac{M^d}{c_d^d}\right] \leqslant \Pr\left[\mathsf{lis}_d(m) \leqslant \left(1 + \frac{(d-1)\delta}{2d+\delta}\right)c_d m^{1/d}\right].$$

Setting $t = (d-1)\delta/(2d+\delta)$ and choosing C so $C^d \geqslant m_0 + 1$, where $m_0 = m_0(t, 1/6, d)$ is as in Corollary 4.3, some basic algebra yields that $m = \lfloor (1+\delta/2)N^d/k^{d-1} \rfloor \geqslant \lfloor C^d \rfloor \geqslant m_0$, so the hypothesis of Corollary 4.3 is satisfied, and hence its conclusion gives the 1/6 sought-after upper bound.

The proof of the second claimed item follows the same argument as the analogous item of Proposition 4.7. We leave the details to the interested reader. □

Corollary 4.12. *The model $(\Sigma(K_{n_1,\ldots,n_d}, k))$ of internal parameter k admits a (c, λ, θ)-median where*

$$(c, \lambda, \theta) = \left(c_d, 1 - \frac{1}{d}, 1 - \frac{1}{d} + \frac{1}{d^2}\right).$$

Proof. Choose H according to $\Sigma(K_{n_1,\ldots,n_d}, k)$. Let $M = cN/k^\lambda = c_d N/k^{1-1/d}$, $\delta > 0$, and $C(\delta)$ and $K(\delta)$ be as in Proposition 4.11. Define $a(\delta) = C(\delta)$, $b(\delta) = (12/(\delta c_d))^{1/d}$ and $k'(\delta) > K(\delta)$ sufficiently large that $k > k'(\delta)$ and $k^{1-1/d+1/d^2} < \min\{\delta, 1\} b(\delta) k/2$. Observe that if $k > k'(\delta)$, $N \geqslant a(\delta) k^{1-1/d}$, and $Sb(\delta) \leqslant k^{1-1/d+1/d^2}$, then the hypothesis of

Proposition 4.11 will be satisfied, and thence every median of $L(H)$ will be between $(1-\delta)M$ and $(1+\delta)M$. □

Recalling that by Proposition 4.1 we have that $h = 1/(4d)$ is a concentration constant for the random d-word model, by the preceding corollary and the Main Theorem, we obtain the following.

Theorem 4.13. *Let $\epsilon > 0$ and $g : \mathbb{R} \to \mathbb{R}$ be such that $g(k) = O(k^{\eta})$ for a given $0 \leqslant \eta < 1/d^2$. Fix $n_1, \ldots, n_d$ and let N and S denote their geometric mean and sum, respectively. There exists sufficiently large constants k_0 and A such that if $k \geqslant k_0$, $N \geqslant Ak^{1-1/d}$ and $S \leqslant g(k)N$, then for $M = c_d N/k^{1-1/d}$, where c_d is the d-dimensional Ulam constant,*

$$(1-\epsilon)M \leqslant \mathsf{E}\big[L(\Sigma(K_{n_1,\ldots,n_d}, k))\big] \leqslant (1+\epsilon)M,$$

and the following hold.

- *If $\mathsf{Med}\big[L(\Sigma(K_{n_1,\ldots,n_d}, k))\big]$ is a median of $L(\Sigma(K_{n_1,\ldots,n_d}, k))$,*

$$(1-\epsilon)M \leqslant \mathsf{Med}\big[L(\Sigma(K_{n_1,\ldots,n_d}, k))\big] \leqslant (1+\epsilon)M.$$

- *There is an absolute constant $C > 0$ such that*

$$\mathsf{Pr}\big[L(\Sigma(K_{n_1,\ldots,n_d}, k)) \leqslant (1-\epsilon)M\big] \leqslant \exp\left(-\frac{C}{d}\epsilon^2 M\right),$$
$$\mathsf{Pr}\big[L(\Sigma(K_{n_1,\ldots,n_d}, k)) \geqslant (1+\epsilon)M\big] \leqslant \exp\left(-\frac{C}{d}\frac{\epsilon^2}{1+\epsilon}M\right).$$

We are now ready to prove Theorem 1.3, which is this section's main result, and which was stated in the main contributions section.

Proof of Theorem 1.3. Let n, n', n'' be positive integers such that $n = n' + n''$. Clearly,

$$\mathsf{E}\big[L(\Sigma(K^{(d)}_n, k))\big] \geqslant \mathsf{E}\big[L(\Sigma(K^{(d)}_{n'}, k))\big] + \mathsf{E}\big[L(\Sigma(K^{(d)}_{n''}, k))\big].$$

By subadditivity, it follows that the limit of $\mathsf{E}\big[L(\Sigma(K^{(d)}_n, k))\big]$ when normalized by n exists and equals

$$\gamma_{k,d} = \inf_{n\in\mathbb{N}} \mathsf{E}\big[L(\Sigma(K^{(d)}_n, k))/n\big].$$

A direct application of Theorem 4.13 yields that $k^{1-1/d}\gamma_{k,d} \to c_d$ when $k \to \infty$. □

4.4. Symmetric binomial random graph model

Throughout this section we focus on the study of $L(\mathcal{D})$ when $\mathcal{D}$ is $S(K_{n,n}, p)$, as defined in the Introduction.

First, we study the behaviour of $L(G)$ when G is chosen according $\mathcal{S}(K_{n,n}, p)$. Recall that in this case, the collection of events $\{(x,y),(y,x)\} \subseteq E(G)$ are independent, and each one occurs with probability p. Also note that $(x,y) \in E(G)$ if and only if $(y,x) \in E(G)$. Any graph for which this equivalence holds will be said to be *symmetric*, thus motivating the use of the word 'symmetric' in naming the random graph model. As usual, we begin our study with the determination of the concentration constant for the random model under study.

Lemma 4.14. *The symmetric binomial random model* $(\mathcal{S}(K_{n,n}, p))_{n\in\mathbb{N}}$ *admits a concentration constant of* $1/4$.

Proof. Direct application of Talagrand's inequality (see [23, Theorem 2.29]). □

As in the study of the binomial model (Section 4.2) and the word model (Section 4.3), given a graph G chosen according to $\mathcal{S}(K_{n,n}, p)$ we will consider a reduced graph G' obtained from G by removal of all edges incident to nodes of degree at least 2. An important observation is that the graph G' thus obtained is also symmetric. Since G' is symmetric, the number of vertices of degree 1 in each of the two colour classes of G' must be even, say $2m$. Thus, the arcs between nodes of degree 1 in G' can be thought of as an involution of $[2m]$ without fixed points. In fact, given that the distribution of G' is invariant under permutation of its nodes, the distribution of G' is also invariant under this permutation, and the resulting associated involution is distributed as a random involution of $[2m]$ without fixed points. We shall see that under proper assumptions $L(G)$ and $L(G')$ are essentially equal. Thus, $L(G)$ behaves (approximately) like the length of a longest increasing subsequence of a randomly chosen involution of $[2m]$ without fixed points. This partly explains our recollection below of some results about the length of a longest increasing subsequence of randomly chosen involutions.

Let $\mathcal{I}_{2m}$ be the distribution of a uniformly chosen involution of $[2m]$ without fixed points. Let $L(\mathcal{I}_{2m})$ denote the length of the longest increasing subsequence of an involution chosen according to $\mathcal{I}_{2m}$. Baik and Rains [8] showed that the expected value of $L(\mathcal{I}_{2m})$ is roughly $2\sqrt{2m}$, for m large. Moreover, Kiwi [24, Theorem 5] established the following concentration result for $L(I_{2m})$ (we state the result in a weaker form).

Theorem 4.15. *For m sufficiently large and every* $0 \leqslant s \leqslant 2\sqrt{2m}$,

$$\mathsf{Pr}\left[|L(\mathcal{I}_{2m}) - \mathsf{E}\left[L(\mathcal{I}_{2m})\right]| \geqslant s + 32(2m)^{1/4}\right] \leqslant 4\exp\left(-\frac{s^2}{16e^{3/2}\sqrt{2m}}\right).$$

Corollary 4.16. *For every* $0 \leqslant t \leqslant 1$ *and* $\alpha > 0$ *there exists a* $m_0 = m_0(t, \alpha)$ *sufficiently large that, for all* $m \geqslant m_0$,

$$\mathsf{Pr}\left[|L(\mathcal{I}_{2m}) - 2\sqrt{2m}| \geqslant 2t\sqrt{2m}\right] \leqslant \alpha.$$

Proof. Let $m_0 = m_0(t, \alpha)$ be sufficiently large that Theorem 4.15 and the following conditions hold for all $m > m_0$:

- $|\mathsf{E}\left[L(\mathcal{I}_{2m})\right] - 2\sqrt{2m}| + 32(2m)^{1/4} \leqslant t\sqrt{2m}$,
- $4e^{-t^2\sqrt{2m}/16e^{3/2}} \leqslant \alpha$.

It follows that

$$\begin{aligned}
&\Pr\left[|L(\mathcal{I}_{2m}) - 2\sqrt{2m}| \geqslant 2t\sqrt{2m}\right] \\
&\quad\leqslant \Pr\left[|L(\mathcal{I}_{2m}) - \mathsf{E}[L(\mathcal{I}_{2m})]| \geqslant 2t\sqrt{2m} - |\mathsf{E}[L(\mathcal{I}_{2m})] - 2\sqrt{2m}|\right] \\
&\quad\leqslant \Pr\left[|L(\mathcal{I}_{2m}) - \mathsf{E}[L(\mathcal{I}_{2m})]| \geqslant t\sqrt{2m} + 32(2m)^{1/4}\right] \\
&\quad\leqslant 4e^{-t^2\sqrt{2m}/16e^{3/2}}.
\end{aligned}$$

□

We now proceed to show that the symmetric random model $\mathcal{S}(K_{n,n}, p)$ admits a (c, λ, θ)-median, where the constant c is related to a constant that arises in the study of the asymptotic behaviour of $L(\mathcal{I}_{2m})$. We will need the following analogues of Lemmas 4.6 and 4.10.

Lemma 4.17. *Let n be a positive integer. Let G be chosen according to $\mathcal{S}(K_{n,n}, p)$. If E and E' denote $E(G)$ and $E(G')$, respectively, then*

$$\mathsf{E}[|E|] = pn(n-1), \tag{4.11}$$

$$\mathsf{E}[|E'|] = pn(n-1)(1-p)^{2n-4}, \tag{4.12}$$

$$\mathsf{E}[|E \setminus E'|] \leqslant 2p^2 n(n-1)(n-2). \tag{4.13}$$

Moreover, for $\eta > 0$,

$$\Pr[||E| - \mathsf{E}[|E|]| \geqslant \eta\mathsf{E}[|E|]] \leqslant \frac{2}{\eta^2\mathsf{E}[|E|]}. \tag{4.14}$$

Proof sketch. For $i \neq j$, let $X_{i,j}$ and $Y_{i,j}$ denote the indicator of the event $(i,j) \in E$ and $(i,j) \in E'$, respectively. Observing that

$$\mathsf{E}[X_{i,j}] = p, \quad \mathsf{E}[Y_{i,j}] = p(1-p)^{2n-4}, \quad |E| = \sum_{i,j:i\neq j} X_{i,j} \quad \text{and} \quad |E'| = \sum_{i,j:i\neq j} Y_{i,j},$$

equations (4.11) and (4.12) follow. Since $E' \subseteq E$, it follows that $|E \setminus E'| = |E| - |E'|$. Inequality (4.13) follows from (4.11) and (4.12) observing that $(1-p)^{2n-4} \geqslant 1 - (2n-4)p$.

To establish (4.14) we observe that $|E|$ can also be expressed as $2\sum_{i<j} X_{i,j}$ and that $\{X_{i,j} \mid i < j\}$ is a collection of independent random variables. To conclude, note that

$$\Delta \stackrel{\text{def}}{=} \sum_{\substack{(i,j),(k,l):i<j,k<l \\ (i,j)\neq(k,l)}} \mathsf{E}[X_{i,j}X_{k,l}] = \binom{n}{2}\left(\binom{n}{2} - 1\right)p^2 \leqslant \frac{\mathsf{E}[|E|]^2}{4},$$

and apply Chebyshev's inequality for indicator random variables to conclude (4.14). □

Proposition 4.18. *Let $\delta > 0$, $0 < p \leqslant 1$ and n be a positive integer. There is a sufficiently large constant $C_1 = C_1(\delta)$, and sufficiently small constants C_2 and C_3, such that the following hold.*

- *If $C_1/p \leqslant n^2 \leqslant C_2\delta/p^{3/2}$, then every median of $L(\mathcal{S}(K_{n,n}, p))$ is at most $2(1+\delta)n\sqrt{p}$.*
- *If $C_1/p \leqslant n^2 \leqslant C_3\delta^2/p^2$, then every median of $L(\mathcal{S}(K_{n,n}, p))$ is at least $2(1-\delta)n\sqrt{p}$.*

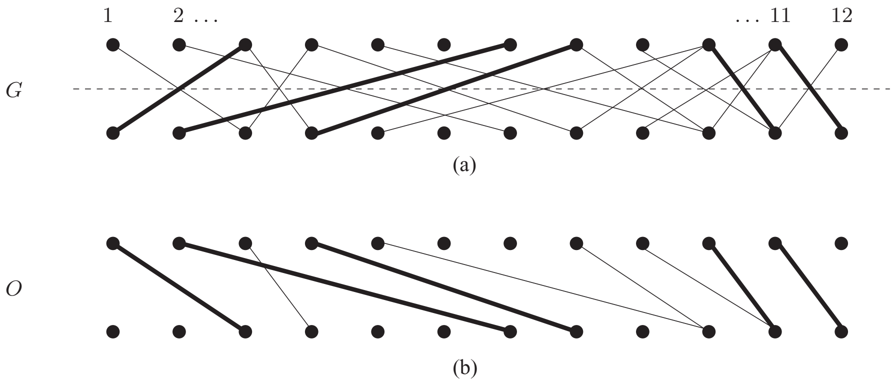


Figure 6. Illustration of (a) a graph G in the support of $\mathcal{S}(K_{12,12}, p)$ and (b) the graph O in the support of $\mathcal{O}(K_{12,12}, p)$ obtained from G by removal of all edges (x, y) such that $x \geqslant y$. Thicker edges represent a non-crossing matching M of G, and the associated non-crossing matching N of O with edge set $\{(\min\{x, y\}, \max\{x, y\}) \mid (x, y) \in E(M)\}$, respectively.

Proof. Similar to the proof of Proposition 4.7. □

We immediately have the following.

Corollary 4.19. *The model* $(\mathcal{S}(K_{n,n}, p))_{n\in\mathbb{N}}$ *of internal parameter* $t = 1/p$ *admits a* $(2, 1/2, 3/4)$-*median.*

We now define an auxiliary distribution which will be useful for our study.

- $\mathcal{O}(K_{n,n}, p)$, the oriented symmetric binomial random graph model: the distribution over the set of subgraphs H of $K_{n,n}$ where the events $\{H \mid (i, j) \in E(H)\}$ for $1 \leqslant i < j \leqslant n$, have probability p and are mutually independent, and the events $\{H \mid (i, j) \in E(H)\}$, $1 \leqslant j \leqslant i \leqslant n$, have probability 0.

(For an illustration of the distinction between distributions $\mathcal{S}(K_{n,n}, p)$ and $\mathcal{O}(K_{n,n}, p)$, see Figure 6.)

The following result justifies why we can henceforth work with either $L(\mathcal{S}(K_{n,n}, p))$ or $L(\mathcal{O}(K_{n,n}, p))$.

Lemma 4.20. *The random variables* $L(\mathcal{S}(K_{n,n}, p))$ *and* $L(\mathcal{O}(K_{n,n}, p))$ *are identically distributed.*

Proof. Let O be a graph in the support of $\mathcal{O}(K_{n,n}, p)$. We can associate with O a graph G over the same collection of vertices and having edge set $\{(x, y) \mid (x, y) \in E(O) \text{ or } (y, x) \in E(O)\}$. Clearly, G is a symmetric subgraph of $K_{n,n}$ and hence it belongs to the support of $\mathcal{S}(K_{n,n}, p)$. It is easy to see that the mapping from O to G is one-to-one. Moreover, the probability of G being chosen under $\mathcal{S}(K_{n,n}, p)$ is exactly equal to the probability of occurrence of O under $\mathcal{O}(K_{n,n}, p)$.

On the other hand, if M is a non-crossing subgraph of G, then there is a non-crossing subgraph of O (and hence of G), say N, whose size is the same as that of M. Indeed, it suffices to take as the collection of edges of N the set $\{(\min\{x,y\},\max\{x,y\}) \mid (x,y) \in E(M)\}$. (See Figure 6 for an illustration of the relation between M and N.) We get that $L(G) = L(O)$, which concludes the proof. □

We are now ready to prove the main result of this section.

Theorem 4.21. *For every $\epsilon > 0$ there is a sufficiently small constant p_0 and a sufficiently large constant A such that, for all $p \leqslant p_0$ and $n \geqslant A/\sqrt{p}$,*

$$(1-\epsilon)2n\sqrt{p} \leqslant \mathsf{E}\big[L(\mathcal{S}(K_{n,n},p))\big] \leqslant (1+\epsilon)2n\sqrt{p}, \tag{4.15}$$

and the following hold.

- *If* $\mathsf{Med}\big[L(\mathcal{S}(K_{n,n},p))\big]$ *is a median of* $L(\mathcal{S}(K_{n,n},p))$,

$$(1-\epsilon)2n\sqrt{p} \leqslant \mathsf{Med}\big[L(\mathcal{S}(K_{n,n},p))\big] \leqslant (1+\epsilon)2n\sqrt{p}. \tag{4.16}$$

- *There is an absolute constant $C > 0$ such that*

$$\mathsf{Pr}\big[L(\mathcal{S}(K_{n,n},p)) \leqslant (1-\epsilon)2n\sqrt{p}\big] \leqslant \exp\big(-C\epsilon^2 n\sqrt{p}\big), \tag{4.17}$$

$$\mathsf{Pr}\big[L(\mathcal{S}(K_{n,n},p)) \geqslant (1+\epsilon)2n\sqrt{p}\big] \leqslant \exp\bigg(-C\frac{\epsilon^2}{1+\epsilon}n\sqrt{p}\bigg). \tag{4.18}$$

Proof. Unfortunately, $(\mathcal{S}(K_{n,n},p))_{n\in\mathbb{N}}$ is not a random hyper-graph model, so we cannot immediately apply the Main Theorem. However, it is a weak random hyper-graph model. Hence, we can still use the Main Theorem to prove the lower bounds in (4.15) and (4.16), and inequality (4.17), using the fact that the model $\mathcal{S}(K_{n,n},p)$ with internal parameter $t = 1/p$ has a concentration constant $h = 1/4$ (Lemma 4.14) and admits a $(2,1/2,3/4)$-median (Corollary 4.19).

To prove the remaining bounds, consider a bipartite graph H chosen according to $\mathcal{G}(K_{n,n},p)$, and let O be the graph obtained from H by deletion of all its edges (x,y) such that $x \geqslant y$. Since O is a subgraph of H, it immediately follows that $L(O) \leqslant L(H)$. Note that O follows the distribution $\mathcal{O}(K_{n,n},p)$. By Lemma 4.20, $L(O)$ has the same distribution as $L(\mathcal{S}(K_{n,n},p))$. Hence, if n and p satisfy the hypothesis of Theorem 4.9,

$$\begin{aligned}\mathsf{E}\big[L(\mathcal{S}(K_{n,n},p))\big] &= \mathsf{E}[L(O)] \leqslant \mathsf{E}[L(H)] \leqslant (1+\epsilon)2n\sqrt{p},\\ \mathsf{Med}\big[L(\mathcal{S}(K_{n,n},p))\big] &= \mathsf{Med}[L(O)] \leqslant \mathsf{Med}[L(H)] \leqslant (1+\epsilon)2n\sqrt{p},\end{aligned}$$

and provided C is as in Theorem 4.9,

$$\begin{aligned}\mathsf{Pr}\big[L(\mathcal{S}(K_{n,n},p)) \geqslant (1+\epsilon)2n\sqrt{p}\big] &= \mathsf{Pr}\big[L(O) \geqslant (1+\epsilon)2n\sqrt{p}\big]\\ &\leqslant \mathsf{Pr}\big[L(H) \geqslant (1+\epsilon)2n\sqrt{p}\big] \leqslant \exp\bigg(-C\frac{\epsilon^2}{1+\epsilon}2n\sqrt{p}\bigg).\end{aligned}$$

This conclusion of the proof of the stated result. □

We can now establish Theorem 1.4.

Proof of Theorem 1.4. Let n, n', n'' be positive integers such that $n = n' + n''$. Clearly,

$$\mathsf{E}\left[L(\mathcal{S}(K_{n,n}, p))\right] \geqslant \mathsf{E}\left[L(\mathcal{S}(K_{n',n'}, p))\right] + \mathsf{E}\left[L(\mathcal{S}(K_{n'',n''}, k))\right].$$

By subadditivity, the limit of $\mathsf{E}[L(\mathcal{S}(K_{n,n}, p))]$ when normalized by n exists and it is equal to

$$\sigma_p = \inf_{n\in\mathbb{N}} \mathsf{E}\left[L(\mathcal{S}(K_{n,n}, p))/n\right].$$

A direct application of Theorem 4.21 yields that $\sigma_p/\sqrt{p} \to 2$ when $p \to 0$. □

Acknowledgements

We would like to thank the anonymous reviewers for their timely and thorough revision, many suggestions that helped us improve the presentation, as well as corrections and comments concerning technical matters.

References

[1] Aldous, D. and Diaconis, P. (1999) Longest increasing subsequences: From patience sorting to the Baik–Deift–Johansson theorem. *Bull. Amer. Math. Soc.* **36** 413–432.

[2] Arlotto, A., Chen, R. W., Shepp, L. A. and Steele, J. M. (2011) Online selection of alternating subsequences from a random sample. *J. Appl. Prob.* **48** 1114–1132.

[3] Arlotto, A. and Steele, J. M. (2011) Optimal sequential selection of a unimodal subsequence of a random sequence. *Combin. Probab. Comput.* **20** 799–814.

[4] Arlotto, A. and Steele, J. M. (2014) Optimal sequential selection of an alternating subsequence: A central limit theorem. *Adv. Appl. Probab.* **46** 536–559.

[5] Babaioff, M., Immorlica, N., Kempe, D. and Kleinberg, R. (2007) A knapsack secretary problem with applications. In *Proc. 10th APPROX and 11th RANDOM*, pp. 16–28.

[6] Babaioff, M., Immorlica, N. and Kleinberg, R. (2007) Matroids, secretary problems, and online mechanisms. In *Proc. 18th SODA*, pp. 434–443.

[7] 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.

[8] Baik, J. and Rains, E. (2001) Symmetrized random permutations. In *Random Matrix Models and their Applications* (P. M. Bleher and A. R. Its, eds), Vol. 40 of *Mathematical Sciences Research Institute Publications*, Cambridge University Press, pp. 1–19.

[9] Baryshnikov, Y. and Gnedin, A. V. (2000) Sequential selection of an increasing sequence from a multidimensional random sample. *Ann. Appl. Probab.* **10** 258–267.

[10] Bollobás, B. and Brightwell, B. (1992) The height of a random partial order: Concentration of measure. *Ann. Probab.* **2** 1009–1018.

[11] Bollobás, B. and Winkler, P. (1988) The longest chain among random points in Euclidean space. *Proc. Amer. Math. Soc.* **103** 347–353.

[12] Boshuizen, F. A. and Kertz, R. P. (1999) Smallest-fit selection of random sizes under a sum constraint: Weak convergence and moment comparisons. *Adv. Appl. Probab.* **31** 178–198.

[13] Bruss, F. T. and Delbaen, F. (2001) Optimal rules for the sequential selection of monotone subsequences of maximum expected length. *Stoch. Proc. Appl.* **96** 313–342.

[14] Bruss, F. T. and Delbaen, F. (2004) A central limit theorem for the optimal selection process for monotone subsequences of maximum expected length. *Stoch. Proc. Appl.* **114** 287–311.

[15] Bruss, F. T. and Robertson, J. B. (1991) 'Wald's Lemma' for sums of order statistics of i.i.d. random variables. *Adv. Appl. Probab.* **23** 612–623.

[16] Chvátal, V. and Sankoff, D. (1975) Longest common subsequences of two random sequences. *J. Appl. Probab.* **12** 306–315.

[17] Coffman, E. G., Flatto, L. and Weber, R. R. (1987) Optimal selection of stochastic intervals under a sum constraint. *Adv. Appl. Probab.* **19** 454–473.

[18] Dinitz, M. (2013) Recent advances on the matroid secretary problem. *SIGACT News* **44** 126–142.

[19] Dynkin, E. B. (1963) The optimum choice of the instant for stopping a Markov process. *Sov. Math. Doklady* **4** 627–629.

[20] Gilbert, J. P. and Mosteller, F. (1966) Recognizing the maximum of a sequence. *J. Amer. Statist. Assoc.* **61** 35–73.

[21] Gnedin, A. V. (2000) A note on sequential selection of permutations. *Combin. Probab. Comput.* **9** 13–17.

[22] Hardy, G., Littlewood, J. E. and Pólya, G. (1952) *Inequalities*, second edition, Cambridge University Press.

[23] Janson, S., Łuczak, T. and Rucinski, A. (2000) *Random Graphs*, Wiley.

[24] Kiwi, M. (2006) A concentration bound for the longest increasing subsequence of a randomly chosen involution. *Discrete Appl. Math.* **154** 1816–1823.

[25] Kiwi, M., Loebl, M. and Matoušek, J. (2005) Expected length of the longest common subsequence for large alphabets. *Adv. Math.* **197** 480–498.

[26] Odlyzko, A. and Rains, E. (1998) On longest increasing subsequences in random permutations. Technical report, AT&T Labs.

[27] Rhee, W. and Talagrand, M. (1991) A note on the selection of random variables under a sum constraint. *J. Appl. Probab.* **28** 919–923.

[28] Samuels, S. M. and Steele, J. M. (1981) Optimal sequential selection of a monotone sequence from a random sample. *Ann. Appl. Probab.* **9** 937–947.

[29] Seppäläinen, T. (1997) Increasing sequences of independent points on the planar lattice. *Ann. Appl. Probab.* **7** 886–898.

[30] Stanley, R. (2002) Recent progress in algebraic combinatorics. *Bull. Amer. Math. Soc.* **40** 55–68.