The Multiple-orientability Thresholds for Random Hypergraphs

Nikolaos Fountoulakis Megha Khosla Konstantinos Panagiotou*

Max Planck Institute for Informatics
Saarbrücken, Germany

Abstract

A k-uniform hypergraph $H = (V, E)$ is called ℓ-orientable, if there is an assignment of each edge $e \in E$ to one of its vertices $v \in e$ such that no vertex is assigned more than ℓ edges. Let $H_{n,m,k}$ be a hypergraph, drawn uniformly at random from the set of all k-uniform hypergraphs with n vertices and m edges. In this paper we establish the threshold for the ℓ-orientability of $H_{n,m,k}$ for all $k \geq 3$ and $\ell \geq 1$, i.e., we determine a critical quantity $c^*_{k,\ell}$ such that with probability $1 - o(1)$ the graph $H_{n,cn,k}$ has an ℓ-orientation if $c < c^*_{k,\ell}$, but fails doing so if $c > c^*_{k,\ell}$.

Our result has various applications including sharp load thresholds for cuckoo hashing, load balancing with guaranteed maximum load, and massive parallel access to hard disk arrays.

1 Introduction

This paper studies the property of multiple orientability of random hypergraphs. For any integers $k \geq 2$ and $\ell \geq 1$, a k-uniform hypergraph is called *ℓ-orientable*, if for each edge we can select one of its vertices, so that all vertices are selected at most ℓ times. This definition generalizes the classical notion of orientability of graphs, where we want to orient the edges under the condition that no vertex has in-degree larger than ℓ. In this paper, we consider random k-uniform hypergraphs $H_{n,m,k}$, for $k \geq 3$, with n vertices and $m = \lfloor cn \rfloor$ edges. Our main result establishes the existence of a critical density $c^*_{k,\ell}$ (determined explicitly in Thorem 1.1), such that when c crosses this value the probability that the random hypergraph is ℓ-orientable drops abruptly from $1 - o(1)$ to $o(1)$, as the number of vertices n grows.

The case $k = 2$ and $\ell \geq 1$ arbitrary is well-understood. In fact, this case corresponds to the classical random graph $G_{n,m}$ drawn uniformly from the set of all graphs with n vertices and m edges. A result of Fernholz and Ramachandran [7] and Cain, Sanders and Wormald [3] implies that there is a constant $c^*_{2,\ell}$ such that

$$\mathbb{P}\left(G_{n,\lfloor cn \rfloor} \text{ is } \ell\text{-orientable}\right) \overset{(n\to\infty)}{=} \begin{cases} 0, & \text{if } c > c^*_{2,\ell} \\ 1, & \text{if } c < c^*_{2,\ell} \end{cases}.$$

In other words, there is a critical value such that when the average degree is below this, then with high probability an ℓ-orientation exists, and otherwise not. We want to remark at this point that the orientation can be found efficiently by solving a matching problem on a suitably defined bipartite graph, but we will not consider computational issues any further in this paper.

Similarly, the case $\ell = 1$ and $k \geq 3$ arbitrary is also well-understood. The threshold for the 1-orientabilty is known from the work of the first and the third author [9], and Frieze and Melsted [10], and also it follows from the work of Dietzfelbinger et al. [5]. In particular, there is a constant $c^*_{k,1}$ such that

$$\mathbb{P}\left(H_{n,\lfloor cn \rfloor,k} \text{ is 1-orientable}\right) \overset{(n\to\infty)}{=} \begin{cases} 0, & \text{if } c > c^*_{k,1} \\ 1, & \text{if } c < c^*_{k,1} \end{cases}.$$

In this paper we consider the general case, i.e., k and ℓ arbitrary. Our main result is summarized in the following theorem, and settles the threshold for the ℓ-orientability property of random hypergraphs for all k and ℓ.

Theorem 1.1. *For integers $k \geq 3$ and $\ell \geq 1$ let ξ^* be the unique solution of the equation*

(1.1)
$$k\ell = \frac{\xi^* Q(\xi^*, \ell)}{Q(\xi^*, \ell+1)}, \text{ where } Q(x,y) = 1 - e^{-x} \sum_{j<y} \frac{x^j}{j!}.$$

Let $c^*_{k,\ell} = \frac{\xi^*}{kQ(\xi^*,\ell)^{k-1}}$. Then

(1.2)
$$\mathbb{P}\left(H_{n,\lfloor cn \rfloor,k} \textit{ is } \ell\textit{-orientable}\right) \overset{(n\to\infty)}{=} \begin{cases} 0, & \textit{if } c > c^*_{k,\ell} \\ 1, & \textit{if } c < c^*_{k,\ell} \end{cases}.$$

A similar result by using completely different techniques was also shown recently in a slightly different context by Gao and Wormald [11], with the restriction that the

*This author is supported by the Humboldt Foundation.

product $k\ell$ is large. So, our result fills the remaining gap, and treats especially the cases of small k and arbitrary ℓ, which are most interesting in practical applications.

1.1 Some Applications

Cuckoo Hashing The paradigm of many choices has influenced significantly the design of efficient data structures and, most notably, hash tables. *Cuckoo hashing*, introduced by Pagh and Rodler [15], is a technique that extends this concept. We consider here a slight variation of the original idea, see also the paper [8] by Fotakis, Pagh, Sanders and Spirakis, where we are given a table with n locations, and we assume that each location can hold ℓ items. Each item to be inserted chooses randomly $k \geq 2$ locations and has to be placed in any one of them. How much load can cuckoo hashing handle before collisions make the successful assignment of the available items to the chosen locations impossible? Practical evaluations of this method have shown that one can allocate a number of elements that is a large proportion of the size of the table, being very close to 1 even for small values of $k\ell$ such as 4 or 6. Our main theorem provides the theoretical foundation for this empirical observation: if the number of items is less than $c^*_{k,\ell}n$, then it is highly likely that they can be allocated. However, if their number is larger, then most likely every allocation will have an overfull bin. Our result thus proves a conjecture about the threshold loads of cuckoo hashing made in [5].

Load Balancing In a typical load balancing problem we are given a set of $m = \lfloor cn \rfloor$ identical jobs, and n machines on which they can be executed. Suppose that each job may choose randomly among k different machines. Is there any upper bound for the maximum load that can be guaranteed with high probability? Our main result implies that as long as $c < c^*_{k,\ell}$, then there is an assignment of the jobs to their favorable machines such that no machine is assigned more than ℓ different tasks.

Parallel Access to Hard Disks In our final application we are given n hard disks (or any other means of storing large amounts of information), which can be accessed independently of each other. The main objective is to store there a big data set, such that by making ℓ parallel queries to the n machines we can retrieve in total $m = \lfloor cn \rfloor$ different data blocks. What is the largest possible value of c such that this can be achieved with high probability, if we store each block of data redundantly k times by using hash functions? Theorem 1.1 implies that with high probability we will select a duplicate of each data block, provided that $c < c^*_{k,\ell}$.

2 Proof Strategy & The Upper Bound

Our main result follows immediately from the two theorems below. The first statement says that $H_{n,m,k}$ has a subgraph of density $> \ell$ (i.e., the fraction of edges and vertices in this subgraph is greater than ℓ) if $c > c^*_{k,\ell}$. We denote by the $(\ell{+}1)$-core of a hypergraph its maximum subgraph that has minimum degree at least $\ell + 1$.

THEOREM 2.1. *Let $c^*_{k,\ell}$ be defined as in Theorem 1.1. If $c > c^*_{k,\ell}$, then with probability $1 - o(1)$ the $(\ell{+}1)$-core of $H_{n,cn,k}$ has density greater than ℓ.*

Note that this implies the statement in the first line of (1.2), as by the pigeonhole principle it is impossible to orient the edges of a hypergraph with density larger than ℓ so that each vertex has indegree at most ℓ.

The above theorem is not very difficult to prove, as the core of random hypergraphs and its structural characteristics have been studied quite extensively in recent years, see e.g. the results by Cooper [4], Molloy [14] and Kim [13]. However, it requires some technical work, which is accomplished in Section 2.1. The heart of this paper is devoted to the "subcritical" case, where we show that the above result is essentially tight.

THEOREM 2.2. *Let $c^*_{k,\ell}$ be defined as in Theorem 1.1. If $c < c^*_{k,\ell}$, then with probability $1 - o(1)$ all subgraphs of $H_{n,cn,k}$ have density smaller than ℓ.*

Proof. [Proof of Theorem 1.1.] Let us construct an auxiliary bipartite graph $B = (\mathcal{E}, \mathcal{V}; E)$, where $\mathcal{E}$ represents the m edges and $\mathcal{V} = \{1, \dots, n\} \times \{1, \dots, \ell\}$ represents the n vertices of $H_{n,m,k}$. Also, $\{e, (i,j)\} \in E$ if the eth edge contains vertex i, and $1 \leq j \leq \ell$. Note that $H_{n,m,k}$ is ℓ-orientable if and only if B has a left-perfect matching, and by Hall's theorem such a matching exists if and only if for all $\mathcal{I} \subseteq \mathcal{E}$ we have that $|\mathcal{I}| \leq |\Gamma(\mathcal{I})|$, where $\Gamma(\mathcal{I})$ denotes the set of neighbors of the vertices in $\mathcal{I}$ in $\mathcal{V}$.

Observe that $\Gamma(\mathcal{I})$ is precisely the set of ℓ copies of the vertices that are contained in the hyperedges corresponding to items in $\mathcal{I}$. So, if $c < c^*_{k,\ell}$, Theorem 2.2 guarantees that with high probability for all $\mathcal{I}$ we have $|\mathcal{I}| \leq |\Gamma(\mathcal{I})|$ and therefore B has a left-perfect matching. On the other hand, if $c > c^*_{k,\ell}$, then with high probability there is a set $\mathcal{I}$ such that $|\mathcal{I}| > |\Gamma(\mathcal{I})|$; choose for example $\mathcal{I}$ to be the set of items that correspond to the edges in the $(\ell{+}1)$-core of $H_{n,m,k}$. Hence a matching does not exist in this case, and the proof is completed.

2.1 Proof of Theorem 2.1 and the Value of $c^*_{k,\ell}$

The aim of this section is to determine the value $c^*_{k,\ell}$ and prove Theorem 2.1. Moreover, we will introduce some known facts and tools that will turn out to be very useful in the study of random hypergraphs, and will be used later on in the proof of Theorem 2.2 as well. In what follows we will be referring to a hyperedge of size k as a (k-)edge and we will be calling a hypergraph where all of its hyperedges are of size k a k-graph.

Models of Random Hypergraphs For the sake of convenience we will carry out our calculations in the $H_{n,p,k}$ model of random k-graphs. This is the "higher-dimensional" analogue of the well-studied $G_{n,p}$ model, where each possible (k-)edge is included independently with probability p. More precisely, given $n \geq k$ vertices we obtain $H_{n,p,k}$ by including each k-tuple of vertices with probability p, independently of every other k-tuple.

Standard arguments show that if we adjust p suitably, then the $H_{n,p,k}$ model is essentially equivalent to the $H_{n,cn,k}$ model. Let us be more precise. Suppose that $\mathcal{P}$ is a *convex* hypergraph property, that is, whenever we have three hypergraphs H_1, H_2, H_3 such that $H_1 \subseteq H_2 \subseteq H_3$ and $H_1, H_3 \in \mathcal{P}$, then also $H_2 \in \mathcal{P}$. We also assume that $\mathcal{P}$ is closed under automorphisms. Any monotone property is also convex and, therefore, the properties examined in Theorem 2.2. The following proposition is a generalization of Proposition 1.15 from [12, p.16] and its proof is very similar to the proof of that – so we omit it.

PROPOSITION 2.1. *Let $\mathcal{P}$ be a convex property of hypergraphs, and let $p = ck/\binom{n-1}{k-1}$, where $c > 0$. If $\mathbb{P}(H_{n,p,k} \in \mathcal{P}) \to 1$ as $n \to \infty$, then $\mathbb{P}(H_{n,cn,k} \in \mathcal{P}) \to 1$ as well.*

Working on the $(\ell{+}1)$-core of $H_{n,p,k}$ – the Cloning Model Recall that the $(\ell + 1)$-core of a hypergraph is its maximum subgraph that has minimum degree (at least) $\ell + 1$. At this point we introduce the main tool for our analysis. The *cloning model* with parameters (N, D, k), where N and D are probability distributions over the non-negative integers, is defined as follows. We generate a graph in three stages.

1. We expose the value of N.

2. We expose the degrees $\mathbf{d} = (d_1, \ldots, d_N)$, where the d_i's are independent samples from the distribution D.

3. For each $1 \leq v \leq N$ we generate d_v copies, which we call *v-clones* or simply *clones*. Then we choose uniformly at random a matching from all perfect k-matchings on the set of all clones, i.e., all partitions of the set of clones into sets of size k. Note that such a matching may not exist – in this case we choose a random matching that leaves less than k clones unmatched. Finally, we construct the graph $H_{\mathbf{d},k}$ by contracting the clones to vertices, i.e., by projecting the clones of v onto v itself for every $1 \leq v \leq N$.

Note that the last stage in the above procedure is equivalent to the *configuration model* [2, 1] $H_{\mathbf{d},k}$ for random hypergraphs with degree sequence $\mathbf{d} = (d_1, \ldots, d_n)$. In other words, $H_{\mathbf{d},k}$ is a random multigraph where the ith vertex has degree d_i.

One special instantiation of the cloning model is the so-called *Poisson cloning model* $\widetilde{H}_{n,p,k}$ for k-graphs with n vertices and parameter $p \in [0, 1]$, which was introduced by Kim [13]. There, we choose $N = n$ with probability 1, and the distribution D is the Poisson distribution with parameter $\lambda := p\binom{n-1}{k-1}$. Note that here D is essentially the vertex degree distribution in the binomial random graph $H_{n,p,k}$, so we would expect that the two models behave similarly. The following statement confirms this, and is implied by Theorem 1.1 in [13].

THEOREM 2.3. *If $\mathbb{P}\left(\widetilde{H}_{n,p,k} \in \mathcal{P}\right) \to 0$ as $n \to \infty$, then $\mathbb{P}(H_{n,p,k} \in \mathcal{P}) \to 0$ as well.*

One big advantage of the Poisson cloning model is that it provides a very precise description of the $(\ell + 1)$-core of $\widetilde{H}_{n,p,k}$. Particularly, Theorem 6.2 in [13] implies the following statement, where we write "$x \pm y$" for the interval of numbers $(x - y, x + y)$.

THEOREM 2.4. *Let $\lambda_{k,\ell+1} := \min_{x>0} \frac{x}{Q(x,\ell)^{k-1}}$. Assume that $ck = p\binom{n-1}{k-1} > \lambda_{k,\ell+1}$. Moreover, let $\bar{x}$ be the largest solution of the equation $x = Q(xck, \ell)^{k-1}$, and set $\xi := \bar{x}ck$. Then, for any $0 < \delta < 1$ the following is true with probability $1 - n^{-\omega(1)}$. If $\tilde{N}_{\ell+1}$ denotes the number of vertices in the $(\ell + 1)$-core of $\widetilde{H}_{n,p,k}$, then*

$$\tilde{N}_{\ell+1} = Q(\xi, \ell + 1)n \pm \delta n.$$

Furthermore, the $(\ell + 1)$-core itself is distributed like the cloning model with parameters $(\tilde{N}_{\ell+1}, \mathrm{Po}_{\geq \ell+1}(\Lambda_{c,k,\ell}), k)$, where $\mathrm{Po}_{\geq \ell+1}(\Lambda_{c,k,\ell})$ denotes a Poisson random variable conditioned on being at least $(\ell + 1)$ and parameter $\Lambda_{c,k,\ell}$, where $\Lambda_{c,k,\ell} = \xi + \beta$, for some $|\beta| \leq \delta$.

In what follows, we say that a random variable is an $(\ell + 1)$-truncated Poisson variable, if it is distributed like a Poisson variable, conditioned on being at least $\ell + 1$. The following theorem, which is a special case of Theorem II.4.I in [6] from large deviation theory, bounds the sum of i.i.d. random variables. We apply the result to the case of i.i.d. $(\ell+1)$-truncated Poisson random variables, which are nothing but the degrees of the vertices of the $(\ell+1)$-core. As an immediate corollary we obtain tight bounds on the number of edges in the $(\ell+1)$-core of $\widetilde{H}_{n,p,k}$. Moreover, it also serves as our main tool in counting the expected number of subsets (with some density constraints) of the $(\ell+1)$-core, assuming that the degree sequence has been exposed. Such estimates are required for the proof of Theorem 2.2 and will be presented in the next section.

THEOREM 2.5. *Let X be a random variable taking real values and set $c(t) = \ln \mathbb{E}(e^{tX})$, for any $t \in \mathbb{R}$. For any $z > 0$ we define $I(z) = \sup_{t\in\mathbb{R}}\{zt - c(t)\}$. If $X_1, \dots, X_s$ are i.i.d. random variables distributed as X, then for $s \to \infty$*

$$\mathbb{P}\left(\sum_{i=1}^{s} X_i \le sz\right) = \exp\left(-s\inf\{I(x) : x \le z\}(1+o(1))\right).$$

The function $I(z)$ is non-negative and convex.

The function $I(z)$ (also known as the rate function) in the above theorem measures the discrepancy between z and the expected value of the sum of the i.i.d. random variables in the sense that $I(z) \ge 0$ with equality if and only if z equals the expected value of X. The following theorem applies Theorem 2.5 to $(\ell+1)$-truncated Poisson random variables.

THEOREM 2.6. *Let $X_1, \dots, X_s$ be i.i.d. $(\ell+1)$-truncated Poisson random variables with parameter Λ. Let T_z be the unique solution of $z = T_z \cdot \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}$, where $z > \ell+1$. Let*

$$\begin{aligned} I(z) = {} & z(\ln T_z - \ln\Lambda) - T_z + \Lambda \\ & - \ln Q(T_z, \ell+1) + \ln Q(\Lambda, \ell+1). \end{aligned}$$

Then $I(z)$ is continuous for all $z > \ell+1$ and convex. It has a unique minimum at $\mu = \Lambda \cdot \frac{Q(\Lambda,\ell)}{Q(\Lambda,\ell+1)}$, where $I(\mu) = 0$. Moreover uniformly for any z such that $\ell+1 \le z \le \mu$, we have as $s \to \infty$

$$\mathbb{P}\left(\sum_{i=1}^{s} X_i \le sz\right) \le \exp(-sI(z)(1+o(1))).$$

Proof. We shall first calculate $c(t) = \ln \mathbb{E}(e^{tX})$, where X is an $(\ell+1)$-truncated Poisson random variable with parameter Λ. We note that

$$\begin{aligned} \exp\{c(t)\} &= \frac{\sum_{j\ge\ell+1} e^{tj} \cdot \frac{e^{-\Lambda}\Lambda^j}{j!}}{Q(\Lambda,\ell+1)} \\ &= e^{-\Lambda} \cdot e^{\Lambda e^t} \cdot \frac{\sum_{j\ge\ell+1} \frac{e^{-\Lambda e^t}(e^t\Lambda)^j}{j!}}{Q(\Lambda,\ell+1)} \\ &= e^{\Lambda e^t - \Lambda} \cdot \frac{Q(\Lambda e^t,\ell+1)}{Q(\Lambda,\ell+1)}. \end{aligned}$$

Differentiating $zt - c(t)$ with respect to t we obtain

$$\begin{aligned} (zt - c(t))' &= z - \ln\left(e^{\Lambda e^t - \Lambda} \cdot \frac{Q(\Lambda e^t,\ell+1)}{Q(\Lambda,\ell+1)}\right)' \\ &= z - \Lambda e^t - (\ln Q(\Lambda e^t,\ell+1))' \\ &= z - \Lambda e^t + \frac{\Lambda e^t \cdot (Q(\Lambda e^t,\ell+1) - Q(\Lambda e^t,\ell))}{Q(\Lambda e^t,\ell+1)}. \end{aligned}$$

Substituting $T = \Lambda e^t$ we get

$$\begin{aligned} (zt - c(t))' &= z - T + \frac{T \cdot (Q(T,\ell+1) - Q(T,\ell))}{Q(T,\ell+1)} \\ &= z - T \cdot \frac{Q(T,\ell)}{Q(T,\ell+1)}. \end{aligned}$$

Setting the derivate to 0, we obtain a unique T that solves the above and which we denote T_z. The uniqueness of the solution for $z > \ell+1$ follows from the fact that the function $x \cdot \frac{Q(x,\ell)}{Q(x,\ell+1)}$ is strictly increasing with respect to x and, as x approaches 0, it tends to $\ell+1$. In other words, T_z is the unique positive real number that satisfies

$$z = T_z \cdot \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}. \tag{2.3}$$

Letting t_z be such that $T_z = \Lambda e^{t_z}$, we obtain

$$-c(t_z) = -T_z - \ln Q(T_z,\ell+1) + \Lambda + \ln Q(\Lambda,\ell+1)$$

and

$$t_z z = z(\ln T_z - \ln\Lambda).$$

The function $-c(t)$ is concave with respect to t (cf. Proposition VII.1.1 in [6, p. 229]); also adding the linear term zt does preserve concavity. So t_z is the point where the unique maximum of $zt - c(t)$ is attained over $t \in \mathbb{R}$. Therefore,

$$\begin{aligned} I(z) = {} & z(\ln T_z - \ln\Lambda) - T_z + \Lambda \\ & - \ln Q(T_z,\ell+1) + \ln Q(\Lambda,\ell+1). \end{aligned} \tag{2.4}$$

For $z = \frac{\Lambda Q(\Lambda,\ell)}{Q(\Lambda,\ell+1)}$, we have $T_z = \Lambda$ and therefore the above equality yields $I(\mu) = 0$. As far as $I(\ell+1)$

is concerned, note that strictly speaking this is not defined, as there is no positive solution of the equation $\ell+1 = T \cdot \frac{Q(T,\ell)}{Q(T,\ell+1)}$. We will express $I(\ell+1)$ as a limit as $T \to 0$ from the right and show that

$$\mathbb{P}\left(\sum_{i=1}^{s} X_i \le s(\ell+1)\right) = \exp(-sI(\ell+1)).$$

We define

$$\begin{aligned} I(\ell+1) := \lim_{T\to 0^+} &\left((\ell+1)\ln T - T - \ln Q(T,\ell+1)\right) \\ &- (\ell+1)\ln\Lambda + \Lambda + \ln Q(\Lambda,\ell+1). \end{aligned}$$

But

$$\begin{aligned} &\lim_{T\to 0^+} \left((\ell+1)\ln T + T - \ln Q(T,\ell+1)\right) \\ &= \lim_{T\to 0^+} \ln \frac{T^{\ell+1}}{e^T Q(T,\ell+1)} \\ &= \lim_{T\to 0^+} \ln \frac{T^{\ell+1}}{\frac{T^{\ell+1}}{(\ell+1)!} + \frac{T^{\ell+2}}{(\ell+2)!} + \cdots} \\ &= \lim_{T\to 0^+} \ln \frac{1}{\frac{1}{(\ell+1)!} + \frac{T}{(\ell+2)!} + \cdots} = \ln(\ell+1)!\ , \end{aligned}$$

and therefore

$$I(\ell+1) = \ln(\ell+1)! - (\ell+1)\ln\Lambda + \Lambda + \ln Q(\Lambda,\ell+1).$$

In the following we compute the required probability for $z = \ell+1$.

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{s} X_i \le s(\ell+1)\right) &= \left(\mathbb{P}\left(\mathrm{Po}_{\ge \ell+1}(\Lambda) = \ell+1\right)\right)^s \\ &= \left(\frac{\frac{e^{-\Lambda}\Lambda^{\ell+1}}{(\ell+1)!}}{Q(\Lambda,\ell+1)}\right)^s \\ &= \exp(-sI(\ell+1)). \end{aligned}$$

Also, according to Theorem 2.5 the function $I(z)$ is non-negative and convex on its domain. So if $z \le \mu$, then $\inf\{I(x) : x \le z\} = I(z)$ and the second part of the lemma follows.

Theorem II.3.3 in [6] along with the above lemma then implies the following corollary.

Corollary 2.1. *Let $X_1, \ldots, X_s$ be i.i.d. $(\ell+1)-$truncated Poisson random variables with parameter Λ and set $\mu = \mathbb{E}(X_1)$. For any $\varepsilon > 0$ there exists a constant $C = C(\varepsilon) > 0$ such that as $s \to \infty$*

$$\mathbb{P}\left(|\sum_{i=1}^{s} X_i - s\mu| \ge s\varepsilon\right) \le e^{-Cs}.$$

With the above results in hand we are ready to prove the following corollary about the density of the $(\ell+1)$-core.

Corollary 2.2. *Let $\tilde{N}_{\ell+1}$ and $\tilde{M}_{\ell+1}$ denote the number of vertices and edges in the $(\ell+1)$-core of $\widetilde{H}_{n,p,k}$. Also let $ck = p\binom{n-1}{k-1}$. Then, for any $0 < \delta < 1$, with probability $1 - n^{-\omega(1)}$,*

$$\tilde{N}_{\ell+1} = Q(\xi,\ell+1)n \pm \delta n, \tag{2.5}$$

$$\tilde{M}_{\ell+1} = \frac{\xi Q(\xi,\ell)}{kQ(\xi,\ell+1)}\tilde{N}_{\ell+1} \pm \delta n, \tag{2.6}$$

where $\xi := \bar{x}ck$ and $\bar{x}$ is the largest solution of the equation $x = Q(xck,\ell)^{k-1}$.

Proof. The statement about $\tilde{N}_{\ell+1}$ follows immediately from the first part of Theorem 2.4.

To see the second statement, we condition on certain values of $\tilde{N}_{\ell+1}$ and $\Lambda_{c,k,\ell}$ that lie in the intervals stated in Theorem 2.4. Then the total degree of the core of $\widetilde{H}_{n,p,k}$ is the sum of independent $(\ell+1)$-truncated Poisson random variables $d_1, \ldots, d_{\tilde{N}_{\ell+1}}$ with parameter $\Lambda_{c,k,\ell} \in \xi \pm \delta$. Let D be the sum of the d_i's. Therefore, Corollary 2.1 yields for any $\varepsilon > 0$ and a constant $C > 0$

$$\mathbb{P}\left(|D - \mathbb{E}(D)| \ge \varepsilon\right) \le e^{-C\tilde{N}_{\ell+1}}.$$

Also, $\mathbb{E}(D) = \frac{\Lambda_{c,k,\ell}Q(\Lambda_{c,k,\ell},\ell)}{Q(\Lambda_{c,k,\ell},\ell+1)} \cdot \tilde{N}_{\ell+1} = (1 \pm s\delta) \cdot \frac{\xi Q(\xi,\ell)}{Q(\xi,\ell+1)} \cdot \tilde{N}_{\ell+1}$, for some appropriate $s > 0$. The proof completes by choosing the initial δ as δ/s.

We proceed with the proof of Theorem 2.1, i.e., we will show that the $(\ell+1)$-core of $\widetilde{H}_{n,p,k}$ has density at least ℓ if $p = ck/\binom{n-1}{k-1}$ and $c > c^*_{k,\ell}$. Let $0 < \delta < 1$, and denote by $\tilde{N}_{\ell+1}$ and $\tilde{M}_{\ell+1}$ the number of vertices and edges in the $(\ell+1)$-core of $\widetilde{H}_{n,p,k}$. By applying Corollary 2.2 we obtain that with probability $1 - n^{-\omega(1)}$

$$\begin{aligned} \tilde{N}_{\ell+1} &= Q(\xi,\ell+1)n \pm \delta n \quad \text{and} \\ \tilde{M}_{\ell+1} &= \frac{\xi Q(\xi,\ell)}{kQ(\xi,\ell+1)}\tilde{N}_{\ell+1} \pm \delta n, \end{aligned}$$

where $\xi = \bar{x}ck$ and $\bar{x}$ is the largest solution of the equation $x = Q(xck,\ell)^{k-1}$. The value of $c^*_{k,\ell}$ is then obtained by taking $\tilde{M}_{\ell+1} = \ell\tilde{N}_{\ell+1}$, and ignoring the additive error terms. The above values imply that the critical ξ^* is given by the equation

$$\xi^* \frac{Q(\xi^*,\ell)}{kQ(\xi^*,\ell+1)} = \ell \implies k\ell = \xi^* \frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}. \tag{2.7}$$

This is precisely (1.1). So, the product $k\ell$ determines ξ^* and $\bar{x}$ satisfies $\bar{x} = Q(\bar{x}ck,\ell)^{k-1} = Q(\xi^*,\ell)^{k-1}$. Therefore, the critical density is

$$c^*_{k,\ell} = \frac{\xi^*}{\bar{x}k} = \frac{\xi^*}{kQ(\xi^*,\ell)^{k-1}}. \tag{2.8}$$

Proof. [Proof of Theorem 2.1] The above calculations imply that uniformly for any $0 < \delta < 1$, with probability $1 - o(1)$

$$\frac{\tilde{M}_{\ell+1}}{\tilde{N}_{\ell+1}} = \frac{1}{k}\,\frac{\xi Q(\xi,\ell)}{Q(\xi,\ell+1)} \pm \Theta(\delta).$$

In particular, if $c = c^*_{k,\ell}$, then $\tilde{M}_{\ell+1}/\tilde{N}_{\ell+1} = \ell \pm \Theta(\delta)$. To complete the proof it is therefore sufficient to show that the ratio $\frac{\xi Q(\xi,\ell)}{Q(\xi,\ell+1)}$ is an increasing function of c. Note that this is the expected value of an $(\ell+1)$-truncated Poisson random variable with parameter ξ, which is known and can be easily verified to be increasing in ξ. Recall that $\xi = \bar{x}ck$. We conclude the proof by showing the following claim.

Claim 1. *The quantity $\xi = \bar{x}ck$ is increasing with respect to c. So, for some fixed c, with probability* $1-o(1)$

$$\frac{\tilde{M}_{\ell+1}}{\tilde{N}_{\ell+1}} < \ell \quad , \textit{if } c < c^*_{k,\ell} \qquad \textit{and}$$
$$\frac{\tilde{M}_{\ell+1}}{\tilde{N}_{\ell+1}} > \ell \quad , \textit{if } c > c^*_{k,\ell}.$$

Indeed, recall that $\bar{x}$ satisfies $\bar{x} = Q(\bar{x}ck,\ell)^{k-1}$. Equivalently, $\bar{x}ck = ck \cdot Q(\bar{x}ck,\ell)^{k-1}$. We have

$$ck = \frac{\xi}{Q(\xi,\ell)^{k-1}}. \tag{2.9}$$

An easy calculation shows that the function on the right-hand side has a unique minimum for $\xi > 0$. Now by the assumption in Theorem 2.4 we have $ck > \min_{x>0} \frac{x}{Q(x,\ell)^{k-1}}$. This implies the function $\frac{\xi}{Q(\xi,\ell)^{k-1}}$ is strictly increasing in the domain of interest and, therefore, when ck increases, then the root of (2.9), that is, the product $\bar{x}ck$, increases as well.

3 Proof of Theorem 2.2

Let us begin with introducing some notation. For a hypergraph H we will denote by V_H its vertex set and by E_H its set of edges. Additionally, v_H and e_H shall denote the number of elements in the corresponding sets. For $U \subset V_H$ we denote by v_U, e_U the number of vertices in U and the number of edges joining vertices only in U. Finally, d_U is the total degree in U, i.e., the sum of the degrees in H of all vertices in U. We say that a subset U of the vertex set of a hypergraph is *ℓ-dense*, if $e_U/v_U \geq \ell$.

In order to prove Theorem 2.2, we need to show that whenever $c < c^*_{k,\ell}$, the random graph $H_{n,cn,k}$ does not contain any ℓ-dense subset with probability $1-o(1)$. We will accomplish this by proving that such a hypergraph does not contain any *maximal* ℓ-dense subset (The term "maximal" here means that whenever we add a vertex to such a set, then its density drops below ℓ.) with probability $1-o(1)$. Note that this is sufficient as any ℓ-dense subset will be contained in some maximal ℓ-dense subset. We shall use the following characterization of a maximal ℓ-dense subset.

Proposition 3.1. *Let H be a k uniform hypergraph with density less than ℓ and let U be a maximal ℓ-dense subset of V_H. Then there is a $0 \leq \theta < \ell$ such that $e_U = \ell \cdot v_U + \theta$. Also, for each vertex $v \in V_H \setminus U$ the corresponding degree d in U, i.e., the number of edges in H that contain v and all other vertices only from U, is less than $\ell - \theta$.*

Proof. If $\theta \geq \ell$, then we have $e_U \geq \ell \cdot (v_U + 1)$. Let $U' = U \cup \{v\}$, where v is any vertex in $V_H \setminus U$. Note that such a vertex always exists, as $U \neq V_H$. Let d be the degree of v in U. Then

$$\frac{e_{U'}}{v_{U'}} = \frac{e_U + d}{v_U + 1} \geq \frac{e_U}{v_U + 1} \geq \ell,$$

which contradicts the maximality of U in H. Similarly, if there exists a vertex $v \in V_H \setminus U$ with degree $d \geq \ell - \theta$ in U, then we could obtain a larger ℓ-dense subset of V_H by adding v to U.

We begin with showing that whenever $c < \ell$, the random graph $H_{n,cn,k}$ does not contain small maximal ℓ-dense subsets. Note that this is sufficient since $c^*_{k,\ell} < \ell$. In particular, the following lemma argues about subsets of size at most $0.6n$. The proof uses a first moment argument that is based on rough counting and will be presented in the next subsection.

Lemma 3.1. *Let $c < \ell$ and $k \geq 3$, $\ell \geq 2$. Then, $H_{n,cn,k}$ contains no maximal ℓ-dense subset with less than $0.6n$ vertices with probability $1 - o(1)$.*

In order to deal with larger subsets we switch to the Poisson cloning model. Let C denote the $(\ell+1)$-core of $\widetilde{H}_{n,p,k}$, where $p = ck/\binom{n-1}{k-1}$, and note that Theorem 2.3 and Proposition 2.1 guarantee that $\widetilde{H}_{n,p,k}$ and $H_{n,cn,k}$ are sufficiently similar. Observe that any minimal ℓ-dense set in $\widetilde{H}_{n,p,k}$ is always a subset of C, as otherwise, by removing vertices of degree at most ℓ the density would not decrease. In other words, C contains all minimal ℓ-dense subsets, and so it is enough to show that the core does not contain any ℓ-dense subset. Therefore, from now on we will restrict our attention to the study of C, and we want to remark that the conclusion of Lemma 3.1 is also true for C.

Assume that the degree sequence of C is given by $\mathbf{d} = (d_1, \ldots, d_{\tilde{N}_{\ell+1}})$, where we denote by $\tilde{N}_{\ell+1}$ the number of vertices in C. Thus, the number of edges in C is $\tilde{M}_{\ell+1} = k^{-1}\sum_{i=1}^{\tilde{N}_{\ell+1}} d_i$. For $q, \beta \in [0,1]$ let $X_{q,\beta} = X_{q,\beta}(C) = X_{q,\beta}(\mathbf{d})$ denote the number of subsets of C with $\lfloor \beta \tilde{N}_{\ell+1} \rfloor$ vertices and total degree $\lfloor qk\tilde{M}_{\ell+1} \rfloor$.

Let $\xi^* = \bar{x}^* c^*_{k,\ell} k$, where $\bar{x}^*$ is the largest solution of the equation $x = Q(xc^*_{k,\ell}k, \ell)^{k-1}$, and note that ξ^* satisfies (2.7). Moreover, let ξ be given by $\xi = \bar{x}ck$, where $\bar{x}$ is the largest solution of the equation $x = Q(xck, \ell)^{k-1}$. As ξ is increasing with respect to c (cf. Claim 1), there exists a $\delta > 0$ and a $\gamma = \gamma(\delta) > 0$ such that $c = c^*_{k,\ell} - \gamma$ and $\xi = \xi^* - \delta$. Also $\gamma \to 0$ as $\delta \to 0$.

In the sequel we will assume that $\delta > 0$ is fixed (and sufficiently small for all our estimates to hold), and we will choose $c < c^*_{k,\ell}$ such that $c = c^*_{k,\ell} - \gamma$ and $\xi = \xi^* - \delta$. Set

$$(3.10) \qquad \begin{aligned} n_{\ell+1} &= Q(\xi, \ell+1)n \quad \text{and} \\ m_{\ell+1} &= \frac{\xi Q(\xi,\ell)}{kQ(\xi,\ell+1)} n_{\ell+1}. \end{aligned}$$

By applying Corollary 2.2 we obtain that with probability $1 - n^{-\omega(1)}$

$$(3.11) \quad \tilde{N}_{\ell+1} = n_{\ell+1} \pm \delta^3 n \quad \text{and} \quad \tilde{M}_{\ell+1} = m_{\ell+1} \pm \delta^3 n.$$

Moreover, by applying Theorem 2.4 we infer that C is distributed like the cloning model with parameters $\tilde{N}_{\ell+1}$ and vertex degree distribution $\mathrm{Po}_{\geq \ell+1}(\Lambda_{c,k,\ell})$, where

$$(3.12) \qquad \Lambda_{c,k,\ell} = \xi \pm \delta^3 = \xi^* - \delta \pm \delta^3,$$

Recall that the definition of ξ^* implies that $k\ell = \frac{\xi^* Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}$. Let $e_{k,\ell}$ denote the value of the first derivative of $\frac{xQ(x,\ell)}{k\ell Q(x,\ell+1)}$ with respect to x at $x = \xi^*$. Taylor's Theorem, applied to $\frac{xQ(x,\ell)}{Q(x,\ell+1)}$ around $x = \xi^*$ implies that

$$(3.13) \qquad \begin{aligned} m_{\ell+1} &= (1 - e_{k,\ell}\cdot\delta + \Theta(\delta^2))\ell \cdot n_{\ell+1}, \text{ where} \\ \frac{\xi Q(\xi,\ell)}{Q(\xi,\ell+1)} &= k\ell(1 - e_{k,\ell}\cdot\delta + \Theta(\delta^2)). \end{aligned}$$

We will now state the main tool for the proof of Theorem 2.2.

Lemma 3.2. *Let $\delta > 0$ be sufficiently small. Then the following holds with probability $1 - n^{-\omega(1)}$. For any $0.6 \leq \beta \leq 1 - e_{k,\ell}\delta/2$ and $\beta \leq q \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$ we have $X^{(\ell)}_{q,\beta} = 0$.*

With the above result at hand we can finally complete the proof of Theorem 2.2.

Proof. [Proof of Theorem 2.2] Firstly, note that it is enough to argue that with probability $1-o(1)$ the $(\ell+1)$-core does not contain any maximal ℓ-dense subset; this follows from the discussion after Lemma 3.1, which we do not repeat here. Moreover, by Theorem 2.3 and Proposition 2.1, it is enough to consider the $(\ell+1)$-core C of $\widetilde{H}_{n,p,k}$, where $p = ck/\binom{n-1}{k-1}$.

By applying Lemma 3.1 we obtain that $H_{n,cn,k}$ does not obtain any ℓ-dense set with less that $0.6n$ vertices. This is particularly also true for C, and so it remains to show the claim for sets of size at least $0.6n \geq 0.6\tilde{N}_{\ell+1}$, where $\tilde{N}_{\ell+1}$ is the number of vertices in C.

We observe that it is sufficient to argue about subsets of size up to, say, $(1 - e_{k,\ell}\delta/2)\tilde{N}_{\ell+1}$, as (3.11) implies that for small δ all larger subsets have density smaller than ℓ. Moreover, the lower bound for q is implied by the fact that the total degree D of any ℓ-dense subset with $\beta\tilde{N}_{\ell+1}$ vertices is at least $k\ell \cdot \beta\tilde{N}_{\ell+1}$. Also

$$D = k \cdot q\tilde{M}_{\ell+1} \Rightarrow k\ell \cdot \beta\tilde{N}_{\ell+1} \leq k \cdot q\tilde{M}_{\ell+1} \Rightarrow q \geq \beta.$$

To see the upper bound in the range of q, recall that each of the vertices in C has degree at least $\ell+1$. More precisely, the total degree of the $(\ell+1)$-core with a ℓ-dense subset with $\beta\tilde{N}_{\ell+1}$ vertices and degree $q \cdot k\tilde{M}_{\ell+1}$ satisfies

$$\begin{aligned} k\tilde{M}_{\ell+1} &\geq q \cdot k\tilde{M}_{\ell+1} + (\ell+1)(\tilde{N}_{\ell+1} - \beta\tilde{N}_{\ell+1}) \\ &\Rightarrow q \leq 1 - \frac{(\ell+1)(1-\beta)\tilde{N}_{\ell+1}}{k\tilde{M}_{\ell+1}} \\ &\overset{(3.11),(3.13)}{\leq} 1 - \frac{(\ell+1)(1-\beta)}{k\ell}. \end{aligned}$$

The proof is then completed by applying Lemma 3.2, as we can choose $\delta > 0$ as small as we please.

References

[1] E.A. Bender and E.R. Canfield. The asymptotic number of labelled graphs with given degree sequence. *J. Combin. Theory Ser. A*, 24(3):296 – 307, 1978.

[2] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *Europ. J. Combin*, 1:311–316, 1980.

[3] J. A. Cain, P. Sanders, and N. Wormald. The random graph threshold for k-orientiability and a fast algorithm for optimal multiple-choice allocation. In *SODA '07*, pages 469–476, 2007.

[4] C. Cooper. The cores of random hypergraphs with a given degree sequence. *Random Structures & Algorithms*, 25(4):353–375, 2004.

[5] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink. Tight thresholds for cuckoo hashing via XORSAT. In *ICALP'10*, volume 6198 of *Lecture Notes in Computer Science*, pages 213–225. 2010.

[6] R. Ellis. *Entropy, large deviations, and statistical mechanics.* Classics in Mathematics. Springer-Verlag, Berlin, 2006.

[7] D. Fernholz and V. Ramachandran. The k-orientability thresholds for $G_{n,p}$. In *SODA '07*, pages 459–468, 2007.

[8] D. Fotakis, R. Pagh, P. Sanders, and P. Spirakis. Space efficient hash tables with worst case constant access time. In *STACS '03*, volume 2607 of *Lecture Notes in Computer Science*, pages 271–282. 2003.

[9] N. Fountoulakis and K. Panagiotou. Orientability of random hypergraphs and the power of multiple choices. In *ICALP'10*, volume 6198 of *Lecture Notes in Computer Science*, pages 348–359. 2010.

[10] A. Frieze and P. Melsted. Maximum matchings in random bipartite graphs and the space utilization of cuckoo hashtables. arXiv:0910.5535v3 [cs.DS] (29 Oct 2009), 2009.

[11] P. Gao and N. C. Wormald. Load balancing and orientability thresholds for random hypergraphs. In *STOC '10*, pages 97–104, 2010.

[12] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[13] J. H. Kim. Poisson cloning model for random graphs. Manuscript, 2006.

[14] M. Molloy. Cores in random hypergraphs and boolean formulas. *Random Structures & Algorithms*, 27(1):124–135, 2005.

[15] R. Pagh and F. F. Rodler. Cuckoo hashing. In *ESA '01*, pages 121–133, 2001.

A Missing Proofs from Section 3

Proof of Lemma 3.1 We will divide the proof in 2 steps. First, we prove it for all k and ℓ except for $k = 3,\ \ell = 2$ by showing the following.

LEMMA A.1. *Let $c < \ell$. Then for any $k \geq 3, \ell \geq 2$ except for the case $k = 3$ and $\ell = 2$, $H_{n,cn,k}$ contains no ℓ-dense subset with less than $0.6n$ vertices with probability $1 - o(1)$.*

Proof. The probability that an edge of $H_{n,cn,k}$ is contained completely in a subset U of the vertex set is $\binom{|U|}{k}/\binom{n}{k} \leq \left(\frac{|U|}{n}\right)^k$. Let $\frac{k}{n} \leq u \leq 0.6$ and $H(x) = -x \ln x - (1-x)\ln(1-x)$ denote the entropy function. Then

$$\begin{aligned} &\mathbb{P}\,(\exists\ \ell\text{-dense subset with } un \text{ vertices}) \\ &\leq \binom{n}{un} \cdot \binom{cn}{\ell un} (u^k)^{\ell un} \qquad \text{(A.1)} \\ &\leq e^{n((\ell+1)H(u) + k\ell u \ln u)}. \end{aligned}$$

We first show that the exponent attains its maximum at $u = k/n$ or $u = 0.6$. Let $u_{max} = 1 - \frac{\ell+1}{k\ell}$. We note that the second derivative of the exponent in the expression above equals $\frac{k\ell(1-u)-(\ell+1)}{u(1-u)}$, which is positive for $k \geq 3, \ell \geq 2$ and $u \in (0, u_{max}]$. Hence the exponent is convex for $u \leq u_{max}$, implying that it is maximized at $u = \frac{k}{n}$ or at $u = (k\ell - (\ell+1))/k\ell$. Moreover, for any $k \geq 4, \ell \geq 2$ we have $u_{max} > 0.6$.

The case $k = 3$ and $\ell \geq 3$ is slightly more involved. Note that $u_{max} \geq 5/9$ in this case. The second derivative of the exponent is negative for $u \in (u_{max}, 1)$, implying that the function is concave in the specified range. But the first derivative of the exponent is $(\ell+1)\ln((1-u)/u) + 3\ell(1 + \ln(u))$, which is at least $2.8\ell - 0.41$ for $u = 0.6$. Hence, the exponent is increasing at $u = 0.6$.

We can now infer that for all combinations of k and ℓ under consideration, the exponent is either maximized at $u = k/n$ or at $u = 0.6$. Note that

$$(\ell+1)H\left(\frac{k}{n}\right) + \frac{k^2\ell}{n}\ln\left(\frac{k}{n}\right) = -\frac{(k^2\ell - (\ell+1)k)\ln n}{n} + O\left(\frac{1}{n}\right). \qquad \text{(A.2)}$$

Also for $k \geq 4$ and $\ell \geq 2$ we obtain

$$\begin{aligned} (\ell+1)H(0.6) + k\ell \cdot 0.6\ln(0.6) &\leq (\ell+1)H(0.6) + 4\ell \cdot 0.6\ln(0.6) \\ &\leq H(0.6) - 0.56\ell \leq -0.44, \end{aligned}$$

and for $k = 3$ and $\ell \geq 3$

$$\begin{aligned} (\ell+1)H(0.6) + k\ell \cdot 0.6\ln(0.6) &\leq (\ell+1)H(0.6) + 3\ell \cdot 0.6\ln(0.6) \\ &\leq H(0.6) - 0.24\ell \leq -0.04. \end{aligned}$$

So, the maximum is obtained at $u = k/n$, and we conclude the proof with

$$\begin{aligned} &\mathbb{P}\,(\exists\ \ell - \text{dense subset with} \leq 0.6n \text{ vertices}) \\ &= \sum_{u=k/n}^{0.6} O(n^{-k^2\ell+(\ell+1)k}) = O(n^{-8}). \end{aligned}$$

We shall now complete the proof of Lemma 3.1 by bounding the probability of existence of a maximal 2-dense subset in a 3-uniform random hypergraph.

Lemma A.2. *For any $c \le c^*_{3,2}$, $H_{n,cn,3}$ contains no maximal 2-dense subset with at most $0.6n$ vertices with probability $1-o(1)$.*

Proof. Let U be a maximal 2–dense subset of $H_{n,cn,3}$. By Proposition 3.1 we infer that there is a $\theta \in \{0,1\}$, such that $e_U = 2\cdot v_U + \theta$ and all vertices in $V_H \setminus U$ have degree less than $2-\theta$ in U . We will now show that the expected number of such sets with at most $0.6n$ vertices is $o(1)$.

Numerical calculations show that $c^*_{3,2} \approx 1.97$. Let $p = c'/\binom{n-1}{2}$, where $c' \le 3\cdot c \le 5.91$. A simple application of Stirling's formula reveals

$$\mathbb{P}\left(H_{n,p,3} \text{ has exactly } cn \text{ edges}\right) = (1+o(1))(2\pi cn)^{-1/2}.$$

As the distribution of $H_{n,cn,3}$ is the same as the distribution of $H_{n,p,3}$ conditioned on the number of edges being precisely cn we infer that

$$\mathbb{P}\left(H_{n,cn,3} \text{ contains a subset } U \text{ with } \le 0.6n \text{ vertices}\right) = O(\sqrt{n})\cdot\mathbb{P}\left(H_{n,p,3} \text{ contains a subset } U \text{ with } \le 0.6n \text{ vertices}\right).$$

To complete the proof it is therefore sufficient to show that the latter probability is $o(n^{-1/2})$.

We accomplish this in two steps. Note that if a subset U is maximal 2-dense, then certainly $|U| \ge 5$. Let us begin with the case $s := |U| \le n^{1/3}$. There are at most n^s ways to choose the vertices in U, and at most $s^{3(2s+\theta)}$ ways to choose the edges that are contained in U. Hence, the probability that $H_{n,p,3}$ contains a bad subset with at most $\lfloor n^{1/3}\rfloor$ vertices is bounded for large n from above by

$$\begin{aligned}\sum_{s=5}^{\lfloor n^{1/3}\rfloor} n^s s^{6s+3\theta} p^{2s+\theta} &\le \sum_{s=5}^{\lfloor n^{1/3}\rfloor} n^s s^{6s+3} p^{2s}\\ &= \sum_{s=5}^{\lfloor n^{1/3}\rfloor}\left(ns^6\left(\frac{c'}{\binom{n-1}{2}}\right)^2\right)^s\cdot s^3\\ &\le n\sum_{s=5}^{\lfloor n^{1/3}\rfloor}\left(n^{(1+6/3)-4}\cdot O(1)\right)^s\\ &\le n\sum_{s=5}^{\lfloor n^{1/3}\rfloor}\left(n^{-1+o(1)}\right)^s = n^{-4+o(1)}.\end{aligned}$$

Let us now consider the case $n^{1/3} \le |U| \le 0.6n$. We note that

$$\ln p = \ln\left(\frac{c'}{\binom{n-1}{2}}\right) = \ln\frac{2c'}{n^2} + \Theta\left(\frac{1}{n}\right).$$

Also, there are $\binom{n}{un} \le e^{nH(u)}$ ways to select U containing un vertices. Moreover, the number of ways to choose the $2un+\theta$ edges that are completely contained in U is

$$\begin{aligned}\binom{\binom{un}{3}}{2un+\theta} &\le \binom{un}{3}\binom{\binom{un}{3}}{2un}\\ &\le \frac{(un)^3}{6}\left(\frac{e(un)^3}{12un}\right)^{2un}\\ &= \exp\left\{2un\ln\left(\frac{e(un)^2}{12}\right)+O(\ln n)\right\}.\end{aligned}$$

Finally, the probability that a vertex in $V_H\setminus U$ has a degree of at most $1-\theta$ in $|U|$ is at most

$$(1-p)^{\binom{un}{2}} + \binom{un}{2}p(1-p)^{\binom{un}{2}-1} = e^{-u^2c'}(1+u^2c')(1+O(1/n)).$$

Combining the above facts we obtain that the probability P_u that $H_{n,p,3}$ contains a maximal 2-dense subset U with $2un$ vertices is at most

$$\begin{aligned}P_u \le&\binom{n}{un}\binom{\binom{un}{3}}{2un+\theta}p^{2un+\theta}(1-p)^{\binom{un}{3}-2un-\theta}\\ &\cdot\left(e^{-u^2c'}(1+u^2c')(1+O(1/n))\right)^{(1-u)n}\\ \le&\exp\left\{n\left(H(u)+2u\ln\left(\frac{eu^2n^2}{12}\right)+2u\ln p\right)\right.\\ &-p\left(\binom{un}{3}-2un-1\right)\\ &\left.+(1-u)n(-u^2c'+\ln(1+u^2c'))+O(\ln n)\right\}\\ \le&\exp\left\{n\left(H(u)+2u\ln\left(\frac{ec'u^2}{6}\right)-\frac{u^3c'}{3}\right.\right.\\ &\left.\left.+(1-u)(-u^2c'+\ln(1+u^2c'))\right)+O(\ln n)\right\}.\end{aligned}$$

If we fix u, the derivative of the exponent with respect to c' is given by

$$\frac{2u}{c'} - \frac{u^3}{3} + (1-u)\left(-u^2+\frac{u^2}{1+u^2c'}\right) \overset{c'\le 5.91}{\ge} \frac{2u}{6} - \frac{u^3}{3} + (1-u)\left(-u^2+\frac{u^2}{1+6u^2}\right).$$

A numerical calculation shows that the latter is positive for all $u \le 0.6$, thus implying that for all such u the exponent is increasing with respect to c'. Therefore, it is sufficient to consider only the case when $c' = 5.91$.

Now, the derivative of the exponent with respect to u equals $\ln(c'^2u^3(1-u))+6-\ln 6-\ln(1+u^2c')-\frac{(1-u)2u^3c'^2}{1+u^2c'}$. As the function $\ln(c'u^3)+\frac{2u^4c'^3}{1+u^2c'}$ is increasing and $\ln\left(\frac{1-u}{1+u^2c'}\right)-\frac{2u^3c'^2}{1+u^2c'}$ is decreasing in u, there is at most one $n^{-2/3} \le u_0 \le 0.6$ where the derivative of the exponent vanishes. Moreover the derivative of the exponent at $u = 0.6$ is positive. Therefore, u_0 is a global minimum, and the bound on P_u is maximized at either at $u = n^{-2/3}$ or at $u = 0.6$.

Elementary algebra then yields that the left point is the right choice, giving the estimate $P_u = o(2^{-n^{1/3}})$, and the proof concludes by adding up this expression for all admissible $n^{-2/3} \le u \le 0.6$.

Proof of Lemma 3.2 We will accomplish the proof in a number of steps. We start by bounding the probability that a given subset of the vertices in $H_{\mathbf{d},k}$ is maximal ℓ–dense. In particular, we will work on the Stage 3 of the exposure process, i.e., when the number of vertices and degree sequence of the core have already been exposed. We will show the following.

LEMMA A.3. *Let $k \ge 3, \ell \ge 2$ and $\mathbf{d} = (d_1, \dots, d_N)$ be a degree sequence and $U \subseteq \{1, ..., N\}$ such that $|U| = \lfloor \beta N \rfloor$. Moreover, set $M = k^{-1}\sum_{i=1}^N d_i$ and $q = (kM)^{-1}\sum_{i\in U} d_i$. Assume that $M < \ell \cdot N$. If $\mathcal{B}(\beta, q)$ denotes the event that U is a maximal ℓ-dense set of $H_{\mathbf{d},k}$, then*

$$P_{\mathbf{d},k}(\mathcal{B}(\beta, q)) \le O(M^{\ell+0.5})\binom{M}{\ell|U|}e^{-kMH(q)}(2^k-1)^{M-\ell|U|},$$

where $H(x) = -x\ln x - (1-x)\ln(1-x)$ denotes the entropy function, and $\mathbb{P}_{\mathbf{d},k}$ denotes the probability measure on the space of Stage 3 of the exposure process, given the outcomes of the first two stages.

Proof. Recall that $H_{\mathbf{d},k}$ is obtained by beginning with d_i clones for each $1 \le i \le N$ and by choosing uniformly at random a perfect k-matching on this set of clones. This is equivalent to throwing kM balls into M bins such that every bin contains k balls. In order to estimate the probability for $\mathcal{B}(\beta, q)$ assume that we color the kqM clones of the vertices in U with red, and the remaining $k(1-q)M$ clones with blue. Let θ be an integer such that $0 \le \theta < \ell$. So, by applying Proposition 3.1 we are interested in the probability for the event that there are exactly $B_\theta = \ell|U| + \theta$ bins with k red balls. We estimate the above probability as follows. We begin by putting into each bin k *black* balls, labeled with the numbers $1, \dots, k$. Let $\mathcal{K} = \{1, \dots, k\}$, and let $X_1, \dots, X_M$ be independent random sets such that for $1 \le i \le M$

$$\forall \mathcal{K}' \subseteq \mathcal{K} \ : \ \mathbb{P}(X_i = \mathcal{K}') = q^{|\mathcal{K}'|}(1-q)^{k-|\mathcal{K}'|}.$$

Note that $|X_i|$ is the binomial distribution $\mathrm{Bin}(k, q)$. We then recolor the balls in the ith bin that are in X_i with red, and all others with blue. So, the total number of red balls is $X = \sum_{i=1}^M |X_i|$. Note that $\mathbb{E}(X) = kqM$, and that X is distributed like $\mathrm{Bin}(kM, q)$. A straightforward application of Stirling's formula then gives

$$\mathbb{P}(X = kqM) = \mathbb{P}(X = \mathbb{E}(X)) = (1+o(1))(2\pi q(1-q)kM)^{-1/2}.$$

Let R_j be the number of X_i's that contain j elements. Then

$$\begin{aligned} P_{\mathbf{d},k}(\mathcal{B}(\beta, q)) &\le \mathbb{P}(R_k = B_\theta | X = kqM) \\ &= \frac{\mathbb{P}(X = kqM \wedge R_k = B_\theta)}{\mathbb{P}(X = kqM)} \qquad \text{(A.3)} \\ &= O\left(\sqrt{M}\right)\mathbb{P}(X = kqM \wedge R_k = B_\theta). \end{aligned}$$

Let $p_j = \mathbb{P}(|X_i| = j) = \binom{k}{j}q^j(1-q)^{k-j}$. Moreover, define the set of integer sequences

$$\mathcal{A} = \left\{(b_0, \dots, b_{k-1}) \in \mathbb{N}^k \right. \\ \left. : \sum_{j=0}^{k-1} b_j = M - B_\theta \text{ and } \sum_{j=0}^{k-1} jb_j = kqM - kB_\theta\right\}.$$

Then

$$\mathbb{P}(X = kqM \wedge R_k = B_\theta) \\ \le \sum_{(b_0,\dots,b_{k-1})\in\mathcal{A}} \binom{M}{b_0, \dots, b_{k-1}, B_\theta} \cdot \left(\prod_{j=0}^{k-1} p_j^{b_j}\right) \cdot p_k^{B_\theta}.$$

Now observe that the summand can be rewritten as

$$\binom{M}{B_\theta} q^{kqM}(1-q)^{k(1-q)M} \cdot \binom{M - B_\theta}{b_0, \dots, b_{k-1}} \prod_{j=0}^{k-1}\binom{k}{j}^{b_j}.$$

Also,

$$\begin{aligned} \sum_{(b_0,\dots,b_{k-1})\in\mathcal{A}} \binom{M - B_\theta}{b_0, \dots, b_{k-1}} \prod_{j=0}^{k-1}\binom{k}{j}^{b_j} &\le \left(\textstyle\sum_{j=0}^{k-1}\binom{k}{j}\right)^{M-B_\theta} \\ &= (2^k-1)^{M-B_\theta}. \end{aligned}$$

Thus, we have

$$\begin{aligned} &\mathbb{P}(X = kqM \wedge R_k = B_\theta) \\ &\le \binom{M}{B_\theta} q^{kqM}(1-q)^{k(1-q)M}(2^k-1)^{M-B_\theta} \\ &\le M^\theta \binom{M}{\ell|U|} e^{-kMH(q)}(2^k-1)^{M-\ell|U|} \cdot (2^k-1)^{-\theta} \\ &\le M^\ell \binom{M}{\ell|U|}(2^k-1)^{M-\ell|U|}e^{-kMH(q)}. \end{aligned}$$

Thus, by using (A.3) and the above facts we infer that

$$P_{\mathbf{d},k}(\mathcal{B}(\beta, q)) \le O(M^{\ell+0.5})\binom{M}{\ell|U|}(2^k-1)^{M-\ell|U|}e^{-kMH(q)}.$$

As already mentioned, the above lemma gives us a bound on the probability that a subset of the $(\ell+1)$-core with a given number of vertices and total degree is maximal ℓ-dense, assuming that the degree sequence is given. Particularly, we work on the probability space of stage 3 of the exposure process. In order to show that the $(\ell+1)$-core contains no ℓ-dense subset, we will estimate the number of such subsets. Recall that $X_{q,\beta}(\mathbf{d})$ denotes the number of subsets of $H_{\mathbf{d},k}$ with $\lfloor \beta\tilde{N}_{\ell+1}\rfloor$ vertices and total degree $\lfloor q \cdot k\tilde{M}_{\ell+1}\rfloor$. Let also $X_{q,\beta}^{(\ell)}$ denote the number of these sets that are maximal ℓ-dense. As an immediate consequence of Markov's inequality we obtain the following corollary.

COROLLARY A.1. *Let $\mathcal{B}(q, \beta)$ be defined as in Lemma A.3, and let $\mathbf{d}$ be the degree sequence of the core of $\widetilde{H}_{n,p,k}$. Then*

$$\mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0 \mid \mathbf{d}\right) \le X_{q,\beta}(\mathbf{d})\mathbb{P}_{\mathbf{d},k}(\mathcal{B}(q, \beta)).$$

Let $\mathcal{E}$ be the event that

$$\mathcal{E} : \tilde{N}_{\ell+1} = n_{\ell+1} \pm \delta^3 n \text{ and } \tilde{M}_{\ell+1} = m_{\ell+1} \pm \delta^3 n. \tag{A.4}$$

where $m_{\ell+1}$ and $n_{\ell+1}$ are given by (3.10). Note that by Corollary 2.2 we have $\mathbb{P}(\mathcal{E}) = 1 - n^{-\omega(1)}$. With Lemma A.3 and Corollary A.1 in hand we are ready to show the following.

LEMMA A.4. *Let $q, \beta \in [0, 1)$. Then*

$$\mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0\right) = \mathbb{E}\left(X_{q,\beta}|\mathcal{E}\right)(2^k - 1)^{m_{\ell+1} - \ell\beta n_{\ell+1}} \cdot e^{\ell n_{\ell+1} H(\beta) - k m_{\ell+1} H(q) + O(\delta^2 n)} + O\left(n^{-3}\right).$$

Proof. Let $\mathcal{E}_1$ be the event that conditional on $\mathcal{E}$ we have $X_{q,\beta} \leq n^3 \mathbb{E}(X_{q,\beta} \mid \mathcal{E})$. Markov's inequality immediately implies that $\mathbb{P}(\mathcal{E}_1 \mid \mathcal{E}) \geq 1 - n^{-3}$. If $\bar{d}$ is a vector, we write $\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}$ to denote that $\bar{d}$ is a possible degree sequence of C if the events $\mathcal{E}$ and $\mathcal{E}_1$ are realized. We have

$$\begin{aligned}
\mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0\right) &\leq \mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0 \mid \mathcal{E}_1 \cap \mathcal{E}\right) + \mathbb{P}\left(\overline{\mathcal{E}_1}\right) + \mathbb{P}\left(\overline{\mathcal{E}}\right) \\
&= \sum_{\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}} \mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0 \mid \mathcal{E}_1 \cap \mathcal{E} \text{ and } \mathbf{d} = \bar{d}\right) \\
&\qquad \cdot \mathbb{P}\left(\mathbf{d} = \bar{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\
&= \sum_{\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}} \mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0 \mid \mathbf{d} = \bar{d}\right) \\
&\qquad \cdot \mathbb{P}\left(\mathbf{d} = \bar{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\
&\overset{(\text{Cor. } A.1)}{\leq} \sum_{\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}} X_{q,\beta}(\bar{d}) \mathbb{P}_{\bar{d},k}(\mathcal{B}(q,\beta)) \\
&\qquad \cdot \mathbb{P}\left(\mathbf{d} = \bar{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\
&\leq n^3\, \mathbb{E}\left(X_{q,\beta} \mid \mathcal{E}\right) \\
&\quad \cdot \sum_{\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}} \mathbb{P}_{\bar{d},k}(\mathcal{B}(q,\beta)) \mathbb{P}\left(\mathbf{d} = \bar{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) \\
&\quad + O(n^{-3}).
\end{aligned}$$

Note that the assumption $\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}$ implies that the number of vertices N induced by $\bar{d}$ is $n_{\ell+1} \pm \delta^3 n$ and the number of edges M is $m_{\ell+1} \pm \delta^3 n$, by (A.4). To apply Lemma A.3 we estimate

$$\binom{M}{\ell\beta N} \overset{(\mathcal{E})}{\leq} \binom{\ell N}{\ell \beta N} = e^{\ell n_{\ell+1} H(\beta) + O(\delta^2 n)}.$$

Thus, applying Lemma A.3 we obtain uniformly for all $\bar{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}$ that

$$\mathbb{P}_{\bar{d},k}(\mathcal{B}(q,\beta)) = (2^k - 1)^{m_{\ell+1} - \beta n_{\ell+1}} \cdot e^{\ell n_{\ell+1} H(\beta) - k m_{\ell+1} H(q) + O(\delta^2 n)}.$$

The claim follows.

The following lemma bounds the expected value of $X_{q,\beta}$ conditional on $\mathcal{E}$.

LEMMA A.5. *Let $\delta > 0$ be sufficiently small. Then, for $\beta \in [0, 1)$ and $\beta \leq q \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$*

$$\mathbb{E}\left(X_{q,\beta}|\mathcal{E}\right) = \exp\left(n_{\ell+1} H(\beta) - n_{\ell+1}(1-\beta) I\left(\tfrac{k\ell(1-q)}{1-\beta}\right) \cdot (1 + o(1)) + O(\delta^2 n)\right)$$

where $I(z)$ is given in (2.4).

Proof. Let $t = \lfloor \beta \tilde{N}_{\ell+1} \rfloor$. Conditioned on $\mathcal{E}$ there are $\binom{\tilde{N}_{\ell+1}}{t} = e^{n_{\ell+1} H(\beta) + O(\delta^2 n)}$ ways to select a set with t vertices. We shall next calculate the probability that one of them has the claimed property, and the statement will follow from the linearity of expectation. Let U be a fixed subset of the vertex set of C that has size t and let $d_1, \ldots, d_t$ denote the random variables that are the degrees of the vertices in U. Thus, we want to estimate the probability of the event $\sum_{i=1}^t d_i = qk\tilde{M}_{\ell+1}$ conditional on $\mathcal{E}$. Let $d_{t+1}, \ldots, d_{\tilde{N}_{\ell+1}}$ denote the random variables that are the degrees of the vertices which do not belong to U; (without conditioning on $\mathcal{E}$) these are i.i.d. $(\ell+1)$-truncated Poisson variables with parameter $\Lambda_{c,k,\ell} = \xi \pm \delta^2$. If $\sum_{i=1}^t d_i = qk\tilde{M}_{\ell+1}$ and the event $\mathcal{E}$ is realized, then recall (A.4) and (3.13), and note that

$$\frac{\sum_{i=t+1}^{\tilde{N}_{\ell+1}} d_i}{\tilde{N}_{\ell+1} - t} = \frac{k(1-q)\tilde{M}_{\ell+1}}{(1-\beta)\tilde{N}_{\ell+1}} \leq \frac{k\ell(1-q)}{1-\beta}(1 - e_{k,\ell}\delta + \Theta(\delta^2)).$$

The last expression is at most $k\ell(1 - e_{k,\ell}\delta + \Theta(\delta^2)) = \xi \frac{Q(\xi,\ell)}{Q(\xi,\ell+1)} + \Theta(\delta^2) = \mu + \Theta(\delta^2)$.

Let $z = \frac{k\ell(1-q)}{1-\beta}(1 - e_{k,\ell}\delta + \Theta(\delta^2))$. A straightforward argument by using Taylor's theorem shows that $I(z) = I\left(\frac{k\ell(1-q)}{1-\beta}\right) + O(\delta^2)$. So, by applying Theorem 2.6 and the fact $\mathbb{P}(\mathcal{E}) = 1 - n^{-\omega(1)}$ the proof is completed.

Lemma A.4 along with Lemmas A.3 and A.5 yield the following estimate.

LEMMA A.6. *Let $\beta \in [0, 1)$ and $\beta \leq q \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$. Then*

$$\begin{aligned}
\mathbb{P}\left(X_{q,\beta}^{(\ell)} > 0\right) =& O(n^{-3}) + (2^k - 1)^{m_{\ell+1} - \ell\beta n_{\ell+1}} \\
&\cdot \exp\left((\ell+1) n_{\ell+1} H(\beta) - k m_{\ell+1} H(q) \right.\\
&\left. - n_{\ell+1}(1-\beta) I\left(\frac{k\ell(1-q)}{1-\beta}\right) + O(\delta^2 n)\right).
\end{aligned}$$

We can now complete the proof of Lemma 3.2 by showing the above probability is $o(1)$. We proceed as follows.

$$\begin{aligned}
F(\beta, q; \ell) :=& (2^k - 1)^{m_{\ell+1} - \ell\beta n_{\ell+1}} \\
&\cdot \exp\left((\ell+1) n_{\ell+1} H(\beta) - k m_{\ell+1} H(q) \right.\\
&\left. - n_{\ell+1}(1-\beta) I\left(\frac{k\ell(1-q)}{1-\beta}\right) + O(\delta^2 n)\right),
\end{aligned}$$

and

$$f(\beta, q) := (\ell+1)H(\beta) + \ell \cdot (1-\beta)\ln(2^k-1) - k\ell \cdot H(q) - (1-\beta) I\left(\frac{k\ell(1-q)}{1-\beta}\right).$$

By using Lemma A.6 we infer that

$$\frac{1}{n_{\ell+1}} \ln F(\beta, q; \ell) \le f(\beta, q) + e_{k,\ell} \cdot \delta \cdot \ell \left(kH(q) - \ln(2^k-1)\right) + O(\delta^2).$$

We will show the following.

CLAIM 2. *There exists a $C > 0$ such that for any small enough $\varepsilon > 0$ the following is true. Let $0.6 \le \beta \le 1-\varepsilon$, and q as in Lemma 3.2. Then*

$$f(\beta, q) \le -C\varepsilon + O(\delta^2).$$

This completes the proof of Lemma 3.2 as follows. First, note that $q \ge \beta \ge 0.6$. Then by the monotonicity of the entropy function for $q \ge 0.6$ we have

$$kH(q) - \ln(2^k-1) \le kH(0.99) - \ln(2^k-1).$$

A simple calculation and the fact $H(0.99) < 0.06$ show that the above expression is negative for all $k \ge 3$. Now if $0.6 \le \beta \le 1 - e_{k,\ell} \cdot \delta/2$ as in Lemma 3.2, then the above claim yields for sufficiently small $\delta > 0$

$$\frac{1}{n_{\ell+1}} \ln F(\beta, q; \ell) \le -Ce_{k,\ell} \cdot \delta/2 + O(\delta^2).$$

This implies that with probability $1 - e^{-\Omega(\delta n_{\ell+1})} - O(n^{-3})$ we have $X^{(\ell)}_{q,\beta} = 0$ for all β and q as in Lemma 3.2.

The rest of the paper is devoted to the proof of Claim 2. We proceed as follows. We will fix arbitrarily a β and we will consider $f(\beta, q)$ solely as a function of q. Then we will show that if $q_0 = q_0(\beta)$ is a point inside the domain where $\partial f/\partial q = 0$, then $f(\beta, q_0) \le -C_1\varepsilon$. Additionally, we will show that this holds for $f(\beta, \beta)$ and $f\left(\beta, 1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)$.

Bounding $f(\beta, q)$ at its critical points Let β be fixed. We will evaluate $f(\beta, q)$ at a point where the partial derivative with respect to q vanishes. To calculate the partial derivative with respect to q, we first need to determine the derivative of $I(z)$ with respect to z. According to Lemma 2.6, $I(z) = z(\ln T_z - \ln \xi) - \ln Q(T_z, \ell+1) - T_z + \ln Q(\xi, \ell+1) + \xi$. So differentiating this with respect to z we obtain:

$$\begin{aligned} I'(z) &= \ln T_z - \ln \xi + z\frac{1}{T_z}\frac{dT_z}{dz} - \frac{dT_z}{dz} \\ &\quad - \frac{Q(T_z,\ell) - Q(T_z,\ell+1)}{Q(T_z,\ell+1)}\frac{dT_z}{dz} \\ &= \ln T_z - \ln \xi + z\frac{1}{T_z}\frac{dT_z}{dz} - \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}\frac{dT_z}{dz} \\ &\overset{(2.3)}{=} \ln T_z - \ln \xi. \end{aligned} \tag{A.5}$$

However, in the differentiation of f we need to differentiate $I(k\ell(1-q)/(1-\beta))$ with respect to q. Using (A.5), we obtain

$$\frac{\partial I\left(\frac{k\ell(1-q)}{1-\beta}\right)}{\partial q} = -\frac{k\ell}{1-\beta}\left(\ln H_q - \ln \xi\right),$$

where H_q is the unique solution of the equation

$$\frac{k\ell(1-q)}{1-\beta} = \frac{H_q \cdot Q(H_q,\ell)}{Q(H_q,\ell+1)}.$$

(Observe that the choice of the range of q is such that the left-hand side of the above equation is at least $\ell+1$. So, H_q is well-defined.) Also, an elementary calculation shows that $H'(q) = \ln\left(\frac{1-q}{q}\right)$. All the above facts together yield the derivative of $f(\beta, q)$ with respect to q:

$$\frac{\partial f(\beta,q)}{\partial q} = k\ell\left(-\ln\left(\frac{1-q}{q}\right) + \ln\frac{H_q}{\xi}\right).$$

Therefore, if q_0 is a critical point, that is, if $\left.\frac{\partial f(\beta,q)}{\partial q}\right|_{q=q_0} = 0$, then q_0 satisfies

$$T_0 = \xi\frac{1-q_0}{q_0}, \quad \text{where} \quad \frac{k\ell(1-q_0)}{1-\beta} = \frac{T_0 Q(T_0,\ell)}{Q(T_0,\ell+1)}. \tag{A.6}$$

At this point, we have the main tool that will allow us to evaluate $f(\beta, q_0)$. We will use (A.6) in order to eliminate T_0 and express $f(\beta, q_0)$ solely as a function of q_0.

CLAIM 3. *For any given β, if $q_0 = q_0(\beta)$ is a critical point of $f(\beta, q)$ with respect to q, then*

$$f(\beta, q_0) = \ln\left(e^{(\ell+1)H(\beta)} q_0^{k\ell}\left(\frac{(2^k-1)(1-q_0)}{q_0}\right)^{\ell(1-\beta)} \cdot \left(\frac{k\ell - \xi}{k\ell q_0 - \xi(1-\beta)}\right)^{1-\beta}\right). \tag{A.7}$$

Proof. Firstly note that

$$\begin{aligned} I\left(\frac{k\ell(1-q_0)}{1-\beta}\right) &= \frac{k\ell(1-q_0)}{1-\beta}\ln\frac{T_0}{\xi} + \ln\left(\frac{e^{\xi}Q(\xi,\ell+1)}{e^{T_0}Q(T_0,\ell+1)}\right) \\ &\overset{(A.6)}{=} \frac{k\ell(1-q_0)}{1-\beta}\ln\left(\frac{1-q_0}{q_0}\right) \\ &\quad + \ln\left(\frac{e^{\xi}Q(\xi,\ell+1)}{e^{T_0}Q(T_0,\ell+1)}\right), \end{aligned}$$

hence

$$\begin{aligned} -(1-\beta)I\left(\frac{k\ell(1-q_0)}{1-\beta}\right) &= -k\ell(1-q_0)\ln\left(\frac{1-q_0}{q_0}\right) \\ &\quad + (1-\beta)\ln\left(\frac{e^{T_0}Q(T_0,\ell+1)}{e^{\xi}Q(\xi,\ell+1)}\right) \\ &= -k\ell(1-q_0)\ln(1-q_0) \\ &\quad + k\ell\ln(q_0) - k\ell q_0\ln(q_0) \\ &\quad + (1-\beta)\ln\left(\frac{e^{T_0}Q(T_0,\ell+1)}{e^{\xi}Q(\xi,\ell+1)}\right). \end{aligned}$$

Also, the definition of the entropy function implies that

$$-k\ell H(q_0) = k\ell q_0 \ln(q_0) + k\ell(1-q_0)\ln(1-q_0).$$

Thus

$$\text{(A.8)}\quad \begin{aligned} &-(1-\beta)I\left(\frac{k\ell(1-q_0)}{1-\beta}\right) - k\ell H(q_0) \\ &= \ln\left(q_0^{k\ell}\left(\frac{e^{T_0}Q(T_0,\ell+1)}{e^{\xi}Q(\xi,\ell+1)}\right)^{1-\beta}\right). \end{aligned}$$

Let $z_0 := \frac{k\ell(1-q_0)}{1-\beta}$. Now we will express $e^{T_0}Q(T_0,\ell+1)$ as a rational function of T_0 and z_0. Solving (A.6) with respect to $e^{T_0}Q(T_0,\ell+1)$ yields

$$\begin{aligned} e^{T_0}Q(T_0,\ell+1) &= e^{T_0}\frac{T_0 Q(T_0,\ell)}{z_0} \\ &= \frac{e^{T_0}T_0}{z_0}\left(Q(T_0,\ell+1) + e^{-T_0}\frac{{T_0}^{\ell}}{\ell!}\right). \end{aligned}$$

Therefore,

$$e^{T_0}Q(T_0,\ell+1) = \frac{{T_0}^{\ell}}{\ell!}\left(\frac{z_0}{T_0}-1\right)^{-1}.$$

Note that

$$\begin{aligned} z_0 - T_0 &= \frac{k\ell(1-q_0)}{1-\beta} - \frac{\xi(1-q_0)}{q_0} \\ &= \frac{(1-q_0)(k\ell q_0 - \xi(1-\beta))}{(1-\beta)q_0}. \end{aligned}$$

Thus we obtain

$$\begin{aligned} \ln(e^{T_0}Q(T_0,\ell+1)) &= \ln\left(\frac{{T_0}^{\ell+1}}{(z-T_0)\ell!}\right) \\ &\overset{(A.6)}{=} \ln\left(\left(\frac{\xi(1-q_0)}{q_0}\right)^{\ell+1}\right. \\ &\qquad \left.\cdot \frac{(1-\beta)q_0}{(1-q_0)(k\ell q_0 - \xi(1-\beta))\ell!}\right) \\ &= \ln\left(\frac{\xi^{\ell+1}}{\ell!}\left(\frac{1-q_0}{q_0}\right)^{\ell}\right. \\ &\qquad \left.\cdot \frac{1-\beta}{k\ell q_0 - \xi(1-\beta)}\right). \end{aligned}$$

Also, by definition of ξ we have $k = \frac{\xi Q(\xi,\ell)}{\ell Q(\xi,\ell+1)}$ which is equivalent to $k\ell = \xi\left(1 + \frac{e^{-\xi}\xi^{\ell}/\ell!}{Q(\xi,\ell+1)}\right)$ which implies $e^{\xi}Q(\xi,\ell+1) = \frac{\xi^{\ell+1}/\ell!}{k\ell-\xi}$. Substituting this into (A.8) and adding the remaining terms, we obtain (A.7).

We will now treat q_0 as a free variable lying in the interval where q lies into, and we will study $f(\beta,q_0)$ for a fixed β as a function of q_0. In particular, we will show that for any fixed β in the domain of interest $f(\beta,q_0)$ is increasing. Thereafter, we will evaluate $f(\beta,q_0)$ at the largest possible value that q_0 can take, which is $1 - \frac{(\ell+1)(1-\beta)}{k\ell}$, and show that this value is negative.

Claim 4. *For any $k \geq 3, \ell \geq 2$ and for any $\beta > 0.6$ we have*

$$\frac{\partial f(\beta,q_0)}{\partial q_0} > 0.$$

Proof. The partial derivative of $f(\beta,q_0)$ with respect to q_0 is

$$\frac{\partial f(\beta,q_0)}{\partial q_0} = \frac{k\ell}{q_0} - \ell\frac{1-\beta}{1-q_0} - \ell\frac{1-\beta}{q_0} - \frac{k\ell(1-\beta)}{k\ell q_0 - \xi(1-\beta)}.$$

Since $q_0 \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$, we obtain

$$1 - q_0 \geq \frac{(\ell+1)(1-\beta)}{k\ell} \Rightarrow -\frac{1-\beta}{1-q_0} \geq -\frac{k\ell}{\ell+1}.$$

Also $q_0 \geq \beta$ and $\xi < k\ell$. Therefore,

$$k\ell q_0 - \xi(1-\beta) > k\ell\beta - k\ell(1-\beta) = 2\beta k\ell - k\ell = k\ell(2\beta-1).$$

Substituting these bounds into $\frac{\partial f(\beta,q_0)}{\partial q_0}$ yields

$$\begin{aligned} \frac{\partial f(\beta,q_0)}{\partial q_0} &> \frac{k\ell}{q_0} - \frac{k\ell^2}{\ell+1} - \frac{\ell(1-\beta)}{q_0} - \frac{1-\beta}{2\beta-1} \\ &= \frac{k\ell - \ell(1-\beta)}{q_0} - \frac{k\ell^2}{\ell+1} - \frac{1-\beta}{2\beta-1} \\ &\geq k\ell\frac{k\ell - l(1-\beta)}{k\ell - (\ell+1)(1-\beta)} - \frac{k\ell^2}{\ell+1} - \frac{1-\beta}{2\beta-1} \\ &\geq k\left(\ell - \frac{\ell^2}{\ell+1} - \frac{1-\beta}{k(2\beta-1)}\right) \\ &= k\left(\frac{\ell}{\ell+1} - \frac{1-\beta}{k(2\beta-1)}\right). \end{aligned}$$

But

$$\frac{\ell}{\ell+1} > \frac{1-\beta}{k(2\beta-1)},$$

as $k\ell(2\beta-1) > (\ell+1)(1-\beta)$, which is equivalent to $\beta > \frac{k\ell+\ell+1}{2k\ell+\ell+1}$. Elementary algebra then yields that $\frac{k\ell+\ell+1}{2k\ell+\ell+1}$ is a decreasing function in k and ℓ. In particular its maximum is 0.6 for $k = 3$ and $\ell = 2$. Since $\beta \geq 0.6$ the above holds.

We begin with setting $q_0 := 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$ into $f(\beta,q_0)$ and obtain a function which depends only on β, namely

$$\begin{aligned} h(\beta) := \ln\Bigg(&\beta^{-(\ell+1)\beta}\cdot \\ &\left(\left(\frac{(2^k-1)(\ell+1)}{k\ell-(\ell+1)(1-\beta)}\right)^{\ell}\frac{k\ell-\xi}{k\ell-(1+\ell+\xi)(1-\beta)}\right)^{1-\beta} \\ &\left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}\Bigg). \end{aligned}$$

Bounding $f(\beta,q)$ globally To conclude the proof of the lemma it suffices due to above arguments to show that for some $C > 0$

$$f(\beta,\beta), f(\beta, 1-(\ell+1)(1-\beta)/k\ell), h(\beta) \leq -C\varepsilon + O(\delta^2)$$

for all $0.6 \leq \beta \leq 1-\varepsilon$. We begin with $h(\beta)$.

CLAIM 5. *For any $k \geq 3$ and $\ell \geq 2$ there is a $C_1 > 0$ such that for any $0.6 \leq \beta \leq 1-\varepsilon$ we have $h(\beta) \leq -C_1\varepsilon$.*

Proof. One can show that for all k and ℓ, we have $k\ell - \frac{e^{-k\ell}(k\ell)^{\ell+1}(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!} \leq \xi \leq k\ell$. Using these bounds for ξ we obtain

(A.9)

$$e^{h(\beta)} \leq \beta^{-(\ell+1)\beta}\left(\frac{(2^k-1)(\ell+1)}{k\ell-(\ell+1)(1-\beta)}\right)^{\ell(1-\beta)} \cdot \left(\frac{\frac{e^{-k\ell}(k\ell)^{\ell+1}(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!}}{k\ell-(\ell+k\ell+1)(1-\beta)}\right)^{1-\beta} \cdot \left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}$$
$$= \left(\frac{2^k-1}{e^k\cdot\beta^{\frac{\beta}{1-\beta}}}\right)^{\ell(1-\beta)}\left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{-\ell(1-\beta)} \cdot \left(1-\frac{(\ell+k\ell+1)(1-\beta)}{k\ell}\right)^{-(1-\beta)} \cdot \left(\frac{(\ell+1)^\ell(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta} \cdot \left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}.$$

By using the inequality $(1-x)^{-1} \leq \exp\{x+\frac{x^2}{1.4}\}$ for $x \leq 0.4$ we can infer $\beta^{\frac{-\beta}{1-\beta}} = (1-(1-\beta))^{\frac{-\beta}{1-\beta}} \leq e^{\beta+\frac{(1-\beta)\beta}{1.4}}$. Also,

$$\left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{-1} \leq \exp\left\{\frac{(\ell+1)(1-\beta)}{k\ell}+\frac{(\ell+1)^2(1-\beta)^2}{(1.4)(k\ell)^2}\right\} \text{ and}$$

$$\left(1-\frac{(1+\ell+k\ell)(1-\beta)}{k\ell}\right)^{-1/\ell} \leq \exp\left\{\frac{(1-\beta)(1+\ell+k\ell)}{k\ell^2}+\frac{(1-\beta)^2(1+\ell+k\ell)^2}{k^2\ell^3}\right\}.$$

Substituting these bounds we combine the first three terms of the right hand side of (A.9) and obtain

$$e^{h(\beta)} \leq \left(\frac{2^k-1}{\exp\left(k-\delta_{k,\ell}\right)}\right)^{\ell(1-\beta)} \cdot \left(\frac{(\ell+1)^\ell(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta} \cdot \left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell},$$

where $\delta_{k,\ell} = \beta+\frac{(1-\beta)\beta}{1.4}+\frac{(\ell+1)(1-\beta)}{k\ell}+\frac{(\ell+1)^2(1-\beta)^2}{(1.4)(k\ell)^2}+\frac{(1-\beta)(1+k\ell+\ell)}{k\ell^2}+\frac{(1-\beta)^2(1+k\ell+\ell)^2}{k^2\ell^3}$. Numerical calculations show that $\delta_{3,8} \leq 1.05$, $\delta_{4,3} \leq 1.2$ and $\delta_{5,2} \leq 1.5$ and it is easy to verify that $\delta_{k,\ell}$ is a decreasing function in k and ℓ. Also, it can be verified numerically that the first term in the product, i.e., $\frac{2^k-1}{\exp\left(k-\delta_{k,\ell}\right)} < 1$ for $\delta_{3,\ell} \leq 1.05$, $\delta_{4,\ell} \leq 1.2$ and $\delta_{5,\ell} \leq 1.5$.

Then it follows that

$$e^{h(\beta)} \leq \left(\frac{(\ell+1)^\ell(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta} \cdot \left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}$$
$$\overset{1+x\leq e^x}{\leq} \left(\frac{\ell^\ell\cdot e(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta} e^{-(\ell+1)(1-\beta)}$$
$$\leq \left(\frac{\ell^\ell\cdot e(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell^\ell e^{-\ell}\sqrt{2\pi\ell}\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta} e^{-(\ell+1)(1-\beta)}$$
$$= \left(\frac{1+e^{-(k\ell-\ell-1)^2/3k\ell}}{\sqrt{2\pi\ell}\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta}.$$

Let

$$g(k,\ell) := \left(\frac{1+e^{-(k\ell-\ell-1)^2/3k\ell}}{\sqrt{2\pi\ell}\cdot\beta^{\frac{\beta}{(1-\beta)}}}\right)^{1-\beta}.$$

A simple calculation yields that the first derivative of the function $(k\ell-\ell-1)^2/3k\ell$ with respect to k is $\frac{(k\ell)^2-(\ell+1)^2}{3k^2\ell}$ which is clearly positive in our setting for k and ℓ. Also the first derivative with respect to ℓ turns out to be $\frac{(k\ell-\ell)^2-1}{3k\ell^2}$ which is also positive.

So we can conclude that $g(k,\ell)$ is decreasing with respect to both k and ℓ. Therefore, it suffices to evaluate $g(k,\ell)$ at $(k=3,\ell=8)$, $(k=4,\ell=3)$ and $(k=5,\ell=2)$ respectively. We have $g(3,8) < 0.15$, $g(4,3) < 0.27$, $g(5,2) < 0.34$.

Also the function $x^{x/(1-x)}$ is equivalent to $(1-(1-x))^{x/(1-x)}$ which is at least $e^{-x-x(1-x)}$. The last expression is clearly a decreasing function in x. Thus we obtain $x^{x/(1-x)} \geq e^{-1} > 0.36$ for any $x \in (0,1)$. Therefore, there exists a constant $C_2 < 1$ such that $e^{h(\beta)} \leq (C_2)^\varepsilon$ which implies $h(\beta) \leq -C_1\varepsilon$ where $C_1 = -\ln(C_2)$. Hence our claim holds.

Details for the remaining cases ($k=3, \ell \leq 7$ and $k=4, \ell=2$) will be provided in the full version of the paper.

CLAIM 6. *For any $k \geq 3$ and $\ell \geq 2$ there is a $C_2 > 0$ such that for any $0.6 < \beta \leq 1-\varepsilon$ we have $f(\beta,\beta) < -C_2\varepsilon+O(\delta^2)$.*

Proof. By Lemma 2.6 we have $I(\mu) = I'(\mu) = 0$ and then we observe that

$$I\left(\frac{k\ell(1-\beta)}{1-\beta}\right) = I(k\ell) = I\left(\mu(1+O(\delta))\right) = I(\mu)+I'(\mu)O(\delta)+I''(\mu)O(\delta^2) = O(\delta^2).$$

So,

$$f(\beta,\beta) = -(k\ell-\ell-1)H(\beta)+\ell(1-\beta)\ln\left(2^k-1\right)+O(\delta^2).$$

Note that for any $k \geq 3$ and $\ell \geq 2$ this function is convex with respect to β, as $-H(\beta)$ is convex and the linear term that is added does not affect its convexity. Moreover, for $\beta = 1$, we have $f(1,1) = 0$. Since $H(0.6) > 0.6$, we have

$$f(0.6, 0.6) < -(k\ell - \ell - 1) \cdot 0.6 + 0.4\ell \ln \left(2^k - 1\right).$$

The derivative of this function with respect to k is $-0.6\ell + \ell \cdot 0.4 \frac{2^k \ln 2}{2^k - 1}$. A simple calculation shows that the second summand is less than 0.32ℓ for all $k \geq 3$. The derivative with respect to ℓ is $-0.6k + 0.6 + 0.4\ln(2^k - 1)$ which is again a decreasing function in k and is < -0.42 at $k = 3$. Thus, the function $f(0.6, 0.6)$ is decreasing with respect to k and ℓ. So, we may set $k = 3$ and $\ell = 2$, thus obtaining $f(0.6, 0.6) < -1.8 + 0.8\ln 7 < -0.24$. The above analysis along with the convexity of $f(\beta, \beta)$ finally imply with Taylor's Theorem the claimed statement.

Claim 7. *For all $k \geq 3$ and $\ell \geq 2$ there is a $C_3 > 0$ such that for all $\beta \leq 1 - \varepsilon$*

$$f(\beta, 1 - (\ell + 1)(1 - \beta)/k\ell) \leq -C_3\varepsilon.$$

Proof. Substituting $1-(\ell+1)(1-\beta)/k\ell$ for q into the formula of f we obtain:

$$\begin{aligned} f &\left(\beta, 1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right) \\ &= (\ell+1)H(\beta) + \ell(1-\beta)\ln(2^k - 1) \\ &\quad -kH\left(\frac{k\ell - (\ell+1)(1-\beta)}{k\ell}\right) - (1-\beta)I(\ell+1). \end{aligned}$$

Note that for $\beta = 1$ the expression is equal to 0. To deduce the bound we are aiming to, we will show that in fact $f(\beta, 1 - (\ell+1)(1-\beta)/k\ell)$ is an increasing function with respect to β. That is, we will show that its first derivative with respect to β is positive. Note that by Taylor's Theorem, this implies the claim.

We get

$$\begin{aligned} \frac{\partial f\left(\beta, 1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)}{\partial \beta} &= (\ell+1)\ln\left(\frac{1-\beta}{\beta}\right) - \ell\ln(2^k - 1) \\ &\quad -(\ell+1)\ln\left(\frac{(\ell+1)(1-\beta)}{k\ell - (\ell+1)(1-\beta)}\right) + I(\ell+1). \end{aligned}$$

Substituting for $I(\ell + 1)$ the value given in Lemma 2.6 and for $e^{\xi}Q(\xi, \ell+1) = \xi^{\ell+1}/\ell!(k\ell - \xi)$ we obtain

$$\begin{aligned} &\frac{\partial f\left(\beta, 1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)}{\partial \beta} \\ &\quad = \ln\left(\left(\frac{k\ell - (\ell+1)(1-\beta)}{(\ell+1)\beta}\right)^{\ell+1}(2^k - 1)^{-\ell} \cdot \frac{(\ell+1)}{k\ell - \xi}\right). \end{aligned}$$

We will show that the fraction inside the logarithm is greater than 1. Note first that the function $(k\ell-(\ell+1)(1-\beta))/(\ell+1)\beta$ is decreasing with respect to β – so we obtain a lower bound by setting $\beta = 1$. Also, for any $k \geq 3$ and $\ell \geq 2$ we can show that $k\ell - \xi \leq \frac{e^{-k\ell}(k\ell)^{\ell+1}(1+e^{-(k\ell-\ell-1)^2/3k\ell})}{\ell!}$. Moreover, we have $\ell! \geq \sqrt{2\pi\ell}(\ell/e)^{\ell}$ and $(1+x) \leq e^x$. All these bounds together yield that

(A.10)
$$\begin{aligned} &\left(\frac{k\ell - (\ell+1)(1-\beta)}{(\ell+1)\beta}\right)^{\ell+1}(2^k - 1)^{-\ell} \cdot \frac{(\ell+1)}{k\ell - \xi} \\ &\qquad \geq \frac{e^{k\ell}\ell!}{(2^k - 1)^{\ell}(\ell+1)^{\ell}(1 + e^{-(k\ell-\ell-1)^2/3k\ell})}. \end{aligned}$$

Solving for $k = 3$ and $\ell = 2$, we obtain the value for the right hand side of the above inequality > 1.13. Using the bounds $\ell! \geq \sqrt{2\pi\ell}(\ell/e)^{\ell}$ and $(1 + x) \leq e^x$ we can further simplify (A.10) as

(A.11)
$$\begin{aligned} &\frac{e^{k\ell}\ell!}{(2^k - 1)^{\ell}(\ell+1)^{\ell}(1 + e^{-(k\ell-\ell-1)^2/3k\ell})} \\ &\qquad \geq \frac{e^{k\ell}\sqrt{2\pi\ell}}{(2^k - 1)^{\ell}e^{\ell+1}(1 + e^{-(k\ell-\ell-1)^2/3k\ell})}. \end{aligned}$$

It is easy to verify that $\sqrt{2\pi\ell}(1 + e^{-(k\ell-\ell-1)^2/3k\ell})^{-1}$ is increasing in k and ℓ. Also the first derivative of the function $e^k/2^k - 1$ with respect to k is $e^k(2^k(1-\ln(2)) - 1)/(2^k - 1)^2$ which is positive for any $k \geq 3$. Moreover the first derivative of the function $e^{k\ell-\ell-1}/(2^k - 1)^{\ell}$ with respect to ℓ is $e^{k\ell-\ell-1}(2^k - 1)^{-\ell}(k - \ln(2^k - 1) - 1)$ which is positive for any $k \geq 3$ and $\ell \geq 2$. So we infer that the right hand side of the above inequality is increasing in both k and ℓ. Numerical calculations show that the right hand side of the above inequality is > 1.3 for $k = 3, \ell = 3$ and > 1.7 for $k = 4, \ell = 2$. The above arguments establish the fact that the derivative of $f(\beta, 1 - (\ell+1)(1-\beta)/k\ell)$ with respect to β is positive, for all $k \geq 3$ and $\ell \geq 2$.