Universität des Saarlandes
Max-Planck-Institut für Informatik

Multiple Choice Allocations with Small Maximum Loads

Dissertation

zur Erlangung des Grades des
Doktors der Ingenieurwissenschaften (Dr.-Ing.)
der Naturwissenschaftlich-Technischen Fakultäten
der Universität des Saarlandes

vorgelegt von

Megha Khosla

Saarbrücken
March 2014

Dekan der Naturwissenschaftlich-Technischen Fakultät I	Prof. Mark Groves
Vorsitzender	Prof. Dr. Matthias Hein
Berichterstatter	Prof. Dr. Kurt Mehlhorn
Berichterstatter	Prof. Dr. Konstantinos Panagiotou
Beisitzer	Dr. Jens M. Schmidt
Tag des Promotionskollquiums	04.03.2014

Abstract

The idea of using multiple choices to improve allocation schemes is now well understood and is often illustrated by the following example. Suppose n balls are allocated to n bins with each ball choosing a bin independently and uniformly at random. The *maximum load*, or the number of balls in the most loaded bin, will then be approximately $\frac{\log n}{\log \log n}$ with high probability. Suppose now the balls are allocated sequentially by placing a ball in the least loaded bin among the $k \geq 2$ bins chosen independently and uniformly at random. Azar, Broder, Karlin, and Upfal [1] showed that in this scenario, the maximum load drops to $\frac{\log \log n}{\log k} + \Theta(1)$, with high probability, which is an exponential improvement over the previous case.

In this thesis we investigate multiple choice allocations from a slightly different perspective. Instead of minimizing the maximum load, we fix the bin capacities and focus on maximizing the number of balls that can be allocated without overloading any bin. In the process that we consider we have $m = \lfloor cn \rfloor$ balls and n bins. Each ball chooses k bins independently and uniformly at random. *Is it possible to assign each ball to one of its choices such that the no bin receives more than ℓ balls?* For all $k \geq 3$ and $\ell \geq 2$ we give a critical value, $c^*_{k,\ell}$, such that when $c < c^*_{k,\ell}$ an allocation is possible with high probability and when $c > c^*_{k,\ell}$ this is not the case.

In case such an allocation exists, *how quickly can we find it?* Previous work on total allocation time for case $k \geq 3$ and $\ell = 1$ has analyzed a *breadth first strategy* which is shown to be linear only in expectation. We give a simple and efficient algorithm which we also call *local search allocation*(LSA) to find an allocation for all $k \geq 3$ and $\ell = 1$. Provided the number of balls are below (but arbitrarily close to) the theoretical achievable load threshold, we give a *linear* bound for the total allocation time that holds with high probability. We demonstrate, through simulations, an order of magnitude improvement for total and maximum allocation times when compared to the state of the art method.

Our results find applications in many areas including hashing, load balancing, data management, orientability of random hypergraphs and maximum matchings in a special class of bipartite graphs.

Zusammenfassung

Die Idee, mehrere Wahlmöglichkeiten zu benutzen, um Zuordnungsschemas zu verbessern, ist mittlerweile gut verstanden und wird oft mit Hilfe des folgenden Beispiels illustriert: Man nehme an, dass n Kugeln auf n Behälter verteilt werden und jede Kugel unabhängig und gleichverteilt per Zufall ihren Behälter wählt. Die *maximale Auslastung*, bzw. die Anzahl an Kugeln im meist befüllten Behälter, wird dann mit hoher Wahrscheinlichkeit schätzungsweise $\frac{\log n}{\log \log n}$ sein. Alternativ können die Kugeln sequenziell zugeordnet werden, indem jede Kugel $k \geq 2$ Behälter unabhängig und gleichverteilt zufällig auswählt und in dem am wenigsten befüllten dieser k Behälter platziert wird. Azar, Broder, Karlin, and Upfal [1] haben gezeigt, dass in diesem Szenario die maximale Auslastung mit hoher Wahrscheinlichkeit auf $\frac{\log \log n}{\log k} + \Theta(1)$ sinkt, was eine exponentielle Verbesserung des vorhergehenden Falls darstellt.

In dieser Doktorarbeit untersuchen wir solche Zuteilungschemas von einem etwas anderen Standpunkt. Statt die maximale Last zu minimieren, fixieren wir die Kapazitäten der Behälter und konzentrieren uns auf die Maximierung der Anzahl der Kugeln, die ohne Überlastung eines Behälters zugeteilt werden können. In dem von uns betrachteten Prozess haben wir $m = \lfloor cn \rfloor$ Kugeln und n Behälter. Jede Kugel wählt unabhängig und gleichverteilt zuällig k Behälter. *Ist es möglich, jeder Kugel einen Behälter ihrer Wahl zuzuordnen, so dass kein Behälter mehr als ℓ Kugeln erhält?* Für alle $k \geq 3$ und $\ell \geq 2$ geben wir einen kritischen Wert $c^*_{k,\ell}$ an, sodass für $c < c^*_{k,\ell}$ eine Zuordnung mit hoher Wahrscheinlich möglich ist und für $c > c^*_{k,\ell}$ nicht.

Im Falle, dass solch eine Zuordnung existiert, stellt sich die Frage, *wie schnell diese gefunden werden kann.* Die bisher durchgeführten Arbeiten zur Gesamtzuordnungszeit im Falle $k \geq 3$ and $\ell = 1$ haben eine *Breitensuchstrategie* analysiert, welche nur im Erwartungswert linear ist. Wir präsentieren einen einfachen und effizienten Algorithmus, welchen wir *local search allocation* (LSA) nennen und der Zuteilungen für alle $k \geq 3$ und $\ell = 1$ findet. Sofern die Anzahl der Kugeln unter (aber beliebig nahe an) der theoretisch erreichbaren Lastschwelle ist, zeigen wir eine *lineare* Schranke für die Gesamtzuordnungszeit, die mit hoher Wahrscheinlichkeit gilt. Anhand von Simulationen demonstrieren wir eine Verbesserung der Gesamt- und Maximalzuordnungszeiten um eine Größenordnung im Vergleich zu anderen aktuellen Methoden.

Unsere Ergebnisse finden Anwendung in vielen Bereichen einschließlich Hashing, Lastbalancierung, Datenmanagement, Orientierbarkeit von zufälligen Hypergraphen und maximale Paarungen in einer speziellen Klasse von bipartiten Graphen.

Acknowledgements

I would like to thank my supervisor Kurt Mehlhorn for providing me an opportunity to work in his group and for allowing me to pursue my own line of research. His guidance was of immense help. I thank Konstantinos Panagiotou who guided me in the initial phase of my PhD. I am thankful to Nikolaos Fountoulakis and Konstantinos Panagiotou for their collaboration on the first problem of this thesis. I have greatly benefitted from their experience and our interesting discussions.

I am also thankful to Alexander Hartmann for working with me on one of my favorite problems. I also thank Tim Byrnes for advising me during my internship in Tokyo. Though our works could not be placed into this thesis because of its structure I still consider them as an essential part of my research and PhD.

I thank my parents and my family for being a source of continued emotional support and teaching me to aim high. I thank the PhD gang of AG1 for starting very interesting discussions at times. I especially thank Ali for always being ready to review my work and providing valuable suggestions. I thank all my wonderful friends for pouring in my ears words of encouragement whenever I needed them.

I thank my loving husband Avishek for always supporting me. Without his constant encouragement and support I would have never finished this thesis.

Contents

List of Figures

Dedicated to Avishek

Chapter 1

Introduction

1.1 Multiple Choice Allocation

Balls-into-bins processes describe in an abstract setting several multiple-choice scenarios, and allow for a systematic and unified theoretical treatment. In general, the goal of these processes is to allocate a set of independent balls (representing tasks, jobs) to a set of bins (representing resources, servers) and, thereby, to minimize the *maximum load* (the number of balls in the most loaded bin). The idea of using multiple choices to improve allocation schemes is now well understood and often illustrated by the following example.

Suppose n balls are placed into n bins by allocating each ball to a bin chosen independently and uniformly at random. It is well known that, in this case, the maximum load will be approximately $\frac{\log n}{\log \log n}$ with high probability[1]. Azar, Broder, Karlin, and Upfal [1] improved this result by considering the following multiple choice scenario. Suppose that the balls are placed sequentially, such that for each ball we choose k bins independently and uniformly at random and place the ball into the less loaded bin (breaking ties arbitrarily). In this case, the maximum load drops to $\frac{\log \log n}{\log k} + \Theta(1)$, with high probability, which is an exponential improvement over the previous case. The above result clearly demonstrates the gain obtained by using more than one choice.

In this thesis we look at the multiple choice process in a slightly different manner. Instead of minimizing the maximum load we fix the bin capacities and then focus on strategies which can maximize the number of balls that can be placed without overloading any bin. We aim to answer the following question.

[1]Throughout this thesis we use *with high probability* to mean with probability $1 - n^{-\zeta}$ for some constant $\zeta > 0$. Also log refers to the natural logarithms.

Question 1 : *What is the maximum number of balls that can be allocated to n bins so that each ball is assigned to one of its k randomly chosen bins, and no bin has more than ℓ balls?*

The motivation behind answering such a question is manifold. For example consider *cuckoo hashing* [2], a technique used to build large hash tables. We consider here a slight variation of the original idea, see also the paper [3] 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 $k \geq 3$ 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?

In a data management setting we are given n hard disks (or any other means of storing large amounts of information), which can be accessed independently of each other. We want to store there a big data set redundantly, that gives us some degree of fault tolerance, and at the same time minimize the number of I/O steps needed to retreive the data (see [4] for more details). To accomplish this, we allocate k copies of each block randomly on n hard disks.

What is the maximum number of data blocks that can be read with at most ℓ parallel queries on each disk ?

As a last example consider *load balancing* in which balls represent the jobs and the bins are the machines. Assume that we have n machines each with capacity ℓ. Additionally each job chooses randomly k machines and need to be assigned to one of them.

What is the maximum number of jobs than can be allocated to n machines such that no machine receives more than ℓ jobs ?

We answer the above questions by giving a critical load threshold (dependent on the number of bins) such that when the number of balls is less than this threshold, an allocation is possible with high probability and otherwise this is not the case. Assuming that the number of balls are below the load threshold, the second question then is how quickly can one find an allocation.

Question 2 : *Suppose that there exists an allocation such that each of the m balls is allocated to one of its k random choices and no bin receives more than ℓ balls. How quickly can one find such an allocation ?*

We answer the above question for all $k \geq 3$ and $\ell = 1$ and provide a simple algorithm which run in linear time with high probability. We assume an *online* setting such that

each ball chooses k random bins on arrival and it has to be placed as and when it appears. Such a setting is quite useful in hashing in which items have to be placed when they appear and no knowledge of their choices is known to the algorithm prior to their arrival, or in online load balancing in which jobs have to be assigned as soon as they arrive.

1.2 Orientation of Hypergraphs

The first question addressed in this thesis can also be phrased in terms of *orientation of graphs* or more generally *orientations of k-uniform hypergraphs*. The n bins are represented as vertices and each of the m balls form an edge with its k-vertices representing the k random choices of the ball. In fact, this is a random (multi)hypergraph $H^*_{n,m,k}$ (or random (multi)graph $G^*_{n,m}$ for $k = 2$) with n vertices and m edges where each edge is drawn uniformly at random (with replacement) from the set of all k-multisubsets of the vertex set. An ℓ-orientation of a graph then amounts to a mapping of each edge to one of its vertices such that no vertex receives more than ℓ edges. Note that the properties of $H^*_{n,m,k}$ are essentially same as that of the simple random hypergraph denoted by $H_{n,m,k}$ (or $G_{n,m}$ for $k = 2$) where multiple edges are forbidden. So $H_{n,m,k}$ is a k-uniform hypergraph drawn uniformly at random from the set of all simple k-uniform hypergraphs with n vertices and m edges.

The case $k = 2$ and $\ell \geq 1$ is well-understood. This case corresponds to the classical random graph $G_{n,m}$ drawn uniformly from the set of all graphs with n vertices and m distinct edges. A result of Fernholz and Ramachandran [5] and Cain, Sanders and Wormald [6] implies that there is a constant $c^*_{2,\ell}$ such that as $n \to \infty$

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

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.

Similarly, the case $\ell = 1$ and $k \geq 3$ is well understood. The threshold for 1-orientability of random hypergraphs is known from the work of the Fountoulakis and Panagiotou [7, 8], and Frieze and Melsted [9]. In particular, there is a constant $c^*_{k,1}$ such that as $n \to \infty$

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

We consider the general case, i.e., k and ℓ arbitrary. Our result also 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 $\ell \geq 2$ let ξ^* be the unique solution of the equation*

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

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

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

A similar result by using completely different techniques was also shown in a slightly different context by Gao and Wormald [10], with the restriction that the 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. Further generalizations of the concept of orientability of hypergraphs have been considered after our work in [11] and [12], where tight results are also obtained.

Note: This work was done in collaboration with Nikolaos Fountoulakis and Konstantinos Panagiotou.

1.3 An Efficient Algorithm

We now focus on the second question addressed in this thesis, which is to develop an algorithm for allocating the given balls into one of their choices without overloading any bin. The typical performance measures for such an algorithm are (1) *total allocation time*, i.e., the total time to allocate all balls and (2) *maximum allocation time* which is the maximum time required to allocate any ball. These parameters are also the main topics in this work.

We start by giving an overview of the already existing algorithms. As we already mentioned the problem of finding an optimal allocation (in context of this work) with maximum load ℓ is equivalent to finding an ℓ orientation of a random graph or hypergraph. For the case $k = 2$, several allocation algorithms and their analysis are closely connected to the cores of the associated graph. The ℓ core of a graph is the maximum vertex induced subgraph with minimum degree at least ℓ. For example, Czumaj and Stemann [13] gave a linear time algorithm achieving maximum load $O(m/n)$ based on computation of

all cores. The main idea was to repeatedly choose a vertex v with minimum degree and remove it from the graph, and assigning all its incident edges (balls) to vertex (bin) v.

Cain, Sanders, and Wormald [6] used a variation of the above approach and gave a linear time algorithm for computing an optimal allocation (asymptotically almost surely). Their algorithm first guesses the optimal load among the two likely values values ($\lceil m/n \rceil$ or $\lceil m/n \rceil + 1$). The algorithm starts with load value say $\ell = \lceil m/n \rceil$. Each time a vertex with degree at most ℓ and its incident edges are assigned to the bin represented by v. The above rule also called the mindegree rule will first reduce the graph to its $(\ell + 1)$-core. After that some edge (u, v) is picked according to some priority rule and assigned to one of its vertices. Again the mindegree rule is applied with respect to some conditions. In case the algorithm fails it is repeated after incrementing the load value.

Fernholz and Ramachandran [5] used a different approach of dealing with the vertices with degree greater than the maximum load. Their algorithm also called *excess degree reduction* (EDR) always chooses a vertex with minimum degree, d. If $d < \ell$ then this vertex is assigned all its incident edges and is removed from the graph. In case $d > 2\ell$ the algorithm fails. Otherwise, EDR replaces $d - \ell$ paths of the form (u, v, w) by bypass edges (u, w) and then orients all remaining edges ($\leq \ell$) incident to v towards v.

Note that the above described algorithms requires the complete knowledge of the graph right from the beginning. In contrast we might need methods to assign balls in an online manner, i.e., balls make their random choices only on arrival and have to be assigned as and when they arrive. Such methods usually involve moving of balls among its chosen locations whenever required. For example, in cuckoo hashing, when an item i appears it is assigned to one of its free choices. In case all its k choices are occupied, then one of its chosen locations say *loc* is selected. One of the items already placed on *loc* is moved out and the item i is placed. The moved out item then looks for a free location among its other choices and the procedure is repeated till an empty location is found.

It is often useful to understand cuckoo hashing in a graph theoretic setting, where each item corresponds to a vertex on one side of a bipartite graph and locations correspond to vertices on the other side. There is an edge between each item and its chosen locations. Then the sequence of moves for assigning an item (described in the previous paragraph) defines an augmenting path in this graph.

For the online setting, the case $k = 2$ and $\ell = 1$ is well understood [2, 14]. Note that for each move (except the first one) there is exactly one choice for the algorithm. The case $k \geq 3$ is more interesting. For $k = 3$ and $\ell = 1$, Fotakis et. al [3] provides a *breadth first search* (BFS) approach. Essentially, if the k choices for the item i are full, one considers the other choices of the k items in those locations, and if all those locations are filled,

one considers the other choices of the items in those locations, and so on. The total allocation time with this approach is shown to be linear only in expectation.

For the same case, Frieze, Melsted and Mitzemmacher [9], and Fountoulakis, Panagiotou and Steger [15] analyzed the *random walk method*, in which one chooses a location randomly from among the k filled choices of an item. More precisely if the k choices of an item i are full, one chooses a random location, say *loc* from among the k locations. The already placed item, say i' is moved out to make room for i. The item i' then looks for an empty location from among its $k-1$ choices. If all its choices are full a location is again selected randomly and the above procedure is repeated. Both of the above mentioned works gave polylogarithmic bounds for maximum allocation time.

Optimal allocations can also be computed in polynomial time using maximum flow computations and with high probability achieve a maximum load of $\lceil m/n \rceil$ or $\lceil m/n \rceil + 1$ [4].

In this thesis we propose a simple and efficient algorithm which we call *local search allocation* (LSA) to find an allocation for the case $k \geq 3$ and $\ell = 1$. Our algorithm runs in linear time with high probability.

Theorem 1.2. *Let $k \geq 3$. For any fixed $\varepsilon > 0$, set $m = (1-\varepsilon)c^*_{k,1}n$. Assume that each of the m balls chooses k random bins from a total of n bins. Then with high probability local search allocation finds an optimal allocation of these balls in time $O(n)$.*

A simple reduction suggests that to match the probability bounds given by our algorithm, BFS would require $O(n \log n)$ run time. The random walk method does not provide any guarantees for the total allocation time. In fact it might run for ever in some worst case. Our algorithm in contrast finds an allocation (with probability 1) whenever it exists. We also present experimental results comparing the performance of these two algorithms. The results reveal that local search allocation is 5 to 10 times faster when total allocation and maximum allocation times are compared. With a small change our algorithm can be extended to the case $\ell \geq 2$. Our simulations for this case predicts that LSA requires linear time for finding an optimal allocation.

One of the very important applications of our result is a faster algorithm for finding maximum cardinality matchings in a special class of sparse random bipartite graphs.

A Faster Matching Algorithm

Consider a bipartite graph $G = (L \cup R, E)$ where L represents the set of m balls and R is the set of n bins. Each $v \in L$ chooses k vertices (independently and uniformly at

random) in R as neighbors. The problem of allocating balls into bins now reduces to finding a perfect matching or a maximum cardinality matching in G.

In random sparse graphs Bast et al. [16] showed that the algorithm of Hopkraft and Karp requires $O(m \log n)$ time, with high probability, to find a maximum matching. Motwani [17] proved this result for random graphs when the average degree is at least $\log(n)$. Goel, Kapralov and Khanna [18] gave a linear time algorithm for finding perfect matchings in regular bipartite graphs. Various heuristics for finding maximum matchings can be found in [19, 20].

Our algorithm computes a perfect matching (whenever it exists) in a k-left regular random bipartite graph in time $O(n)$ with high probability. This is the most efficient algorithm (to the best of my knowledge) for this special class of bipartite graphs.

1.4 Organization

The thesis is organized as follows. Each chapter consists of one problem together with its detailed introduction and obtained results. The directions for future work are presented at the end of each chapter.

Chapter 2: This chapter deals with the multiple orientability thresholds for random hypergraphs. For integers k and ℓ we compute a threshold $c^*_{k,\ell}$ such that when the density of a random k-uniform hypergraph (ratio of number of edges to that of vertices) is below $c^*_{k,\ell}$, then the hypergraph is ℓ-orientable with high probability, otherwise this is not the case.

Indication of source : The contents of Chapter 2 has been previously published in SODA 2011 [21]. The full version of this work has been submitted to the journal Combinatorics, Probability and Computing.

Chapter 3: In this chapter we propose and analyze an efficient algorithm which we call *local search allocation* to find an optimal allocation of balls-into-bins for the case $k \geq 3, \ell = 1$. As a corollary we obtain an efficient algorithm for finding perfect matchings in a special class of random bipartite graphs.

Indication of source: The contents of Chapter 3 has been previously published in ESA 2013 [22].

Chapter 2

The Multiple-orientability Thresholds for Random Hypergraphs

2.1 Introduction

In this chapter we study 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. The main result of this chapter is the following theorem, which establishes the existence of a critical density $c^*_{k,\ell}$ 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.

Theorem 2.1. *For integers $k \geq 3$ and $\ell \geq 2$ let ξ^* be the unique solution of the equation*

$$k\ell = \frac{\xi^* Q(\xi^*, \ell)}{Q(\xi^*, \ell + 1)}, \text{ where } Q(x, y) = 1 - e^{-x} \sum_{j<y} \frac{x^j}{j!}. \tag{2.1}$$

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

$$\mathbb{P}\left(H_{n,\lfloor cn \rfloor,k} \text{ is } \ell\text{-orientable}\right) \overset{(n\to\infty)}{=} \begin{cases} 0, & \text{if } c > c^*_{k,\ell} \\ 1, & \text{if } c < c^*_{k,\ell} \end{cases}. \tag{2.2}$$

2.2 Proof Strategy

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 ratio of the number of edges to the number of 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 2.2. *Let $c^*_{k,\ell}$ be defined as in Theorem 2.1. 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 (2.2), 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 for example the results by Cooper [23], Molloy [24] and Kim [25]. However, it requires some technical work, which is accomplished in Section 2.4. The heart of this work is devoted to the "subcritical" case, where we show that the above result is essentially tight.

Theorem 2.3. *Let $c^*_{k,\ell}$ be defined as in Theorem 2.1. If $c < c^*_{k,\ell}$, then with probability $1-o(1)$ all subgraphs of $H_{n,cn,k}$ have density smaller than ℓ.*

Proof of Theorem 2.1. 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 $1 \le j \le \ell$. 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 $|\mathcal{I}| \le |\Gamma(\mathcal{I})|$, 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 2.3 guarantees that with high probability for all $\mathcal{I}$ we have $|\mathcal{I}| \le |\Gamma(\mathcal{I})|$ 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. □

2.3 Technical Preliminaries

2.3.1 Models of Random Hypergraphs

We refer to a hyperedge of size k as a (k-)edge and call a hypergraph with all its hyperedges of size k a *k-graph*. 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 2.3. The following proposition is a generalization of Proposition 1.15 from [26, p.16] and its proof is very similar to the proof of that.

Proposition 2.4. *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.*

Proof. Let m' and m'' maximizes $\mathbb{P}(H_{n,m,k} \in \mathcal{P})$ for $m \leq cn$ and $m \geq cn$ respectively. Let E_p denote the edge set in $H_{n,p,k}$. We then have

$$\mathbb{P}(H_{n,p,k} \in \mathcal{P}) \leq \mathbb{P}\left(H_{n,m',k} \in \mathcal{P}\right) \mathbb{P}(|E_p| \leq cn) + \mathbb{P}(|E_p| > cn).$$

By central limit theorem we have $\mathbb{P}(|E_p| \leq cn) \overset{n\to\infty}{=} 1/2$, and therefore

$$1 = \lim_{n\to\infty} \mathbb{P}(H_{n,p,k} \in \mathcal{P}) \leq \frac{1}{2} \lim_{n\to\infty} \mathbb{P}\left(H_{n,m',k} \in \mathcal{P}\right) + \frac{1}{2},$$

which implies that $\lim_{n\to\infty} \mathbb{P}\left(H_{n,m',k} \in \mathcal{P}\right) = 1$. Similarly $\lim_{n\to\infty} \mathbb{P}\left(H_{n,m'',k} \in \mathcal{P}\right) = 1$. The convexity of $\mathcal{P}$ then yields $\mathbb{P}(H_{n,cn,k} \in \mathcal{P}) \to 1$. □

2.3.2 The Poisson Cloning Model for the $(\ell+1)$ -core

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 integer valued random variables, is defined as follows. We generate a graph in three stages.

1. We expose the value of N;
2. if $N \geq 1$ 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 k-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* [27, 28] $H_{\mathbf{d},k}$ 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 particular case 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 [25]. 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 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 [25].

Theorem 2.5. *If* $\mathbb{P}\left(\widetilde{H}_{n,p,k} \in \mathcal{P}\right) \to 0$ *as* $n \to \infty$, *then* $\mathbb{P}\left(H_{n,p,k} \in \mathcal{P}\right) \to 0$ *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 [25] implies the following statement, where we write "$x \pm y$" for the interval of numbers $(x-y, x+y)$.

Theorem 2.6. *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* β *satisfying* $|\beta| \leq \delta$.

In what follows, we say that a random variable is an ℓ-truncated Poisson variable, if it is distributed like a Poisson variable, conditioned on being at least ℓ. The following theorem, which is a special case of Theorem II.4.I in [29] 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 2.3 and will be presented in the next section.

Theorem 2.7. *Let X be a random variable taking real values and set $c(t) = \log \mathbb{E}(e^{tX})$, 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* of the random variable X) 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 lemma applies Theorem 2.7 to $(\ell+1)$-truncated Poisson random variables.

Lemma 2.8. *Let $X_1, \dots, X_s$ be i.i.d. $(\ell+1)$-truncated Poisson random variables with parameter Λ. For any $z > \ell + 1$, let T_z be the unique solution of $z = T_z \cdot \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}$ and*

$$I_\Lambda(z) = z(\log T_z - \log \Lambda) - T_z + \Lambda - \log Q(T_z, \ell+1) + \log Q(\Lambda, \ell+1). \tag{2.3}$$

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

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

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

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

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

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

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

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

Setting this expression to zero and solving for T gives the value of T_z as in the statement of the lemma. 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 (cf. Claim 2.24) and, as x approaches 0, it tends to $\ell + 1$. Letting t_z be such that $T_z = \Lambda e^{t_z}$, we obtain

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

and

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

The function $-c(t)$ is concave with respect to t (cf. Proposition VII.1.1 in [29, 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 $t \in \mathbb{R}$. Combining the above we obtain $I_\Lambda(z)$ as stated in the lemma. For $z = \frac{\Lambda Q(\Lambda,\ell)}{Q(\Lambda,\ell+1)}$ we have $T_z = \Lambda$ which yields $I_\Lambda(\mu) = 0$. As far as $I_\Lambda(\ell+1)$ 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 $I_\Lambda(\ell+1)$ as a limit as $T \to 0$ from the right and show that

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

We define

$$I_\Lambda(\ell+1) := \lim_{T \to 0^+} ((\ell+1)\log T - T - \log Q(T, \ell+1)) - (\ell+1)\log\Lambda + \Lambda + \log Q(\Lambda, \ell+1).$$

But

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

and therefore

$$I_\Lambda(\ell+1) = \log(\ell+1)! - (\ell+1)\log\Lambda + \Lambda + \log Q(\Lambda, \ell+1).$$

On the other hand, the independence of the X_i's guarantees that

$$\mathbb{P}\left(\sum_{i=1}^{s} X_i \leq s(\ell+1)\right) = [\mathbb{P}(X_1 = \ell+1)]^s = \left(\frac{\frac{e^{-\Lambda}\Lambda^{\ell+1}}{(\ell+1)!}}{Q(\Lambda,\ell+1)}\right)^s = \exp(-sI_\Lambda(\ell+1)).$$

Also, according to Theorem 2.7 the function $I_\Lambda(z)$ is non-negative and convex on its domain. So if $z \leq \mu$, then $\inf\{I_\Lambda(x) : x \leq z\} = I_\Lambda(z)$ and the second part of the lemma follows. □

Theorem II.3.3 in [29] along with the above lemma then implies the following corollary.

Corollary 2.9. *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 $s \to \infty$*

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

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

Corollary 2.10. *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 $1 - n^{-\omega(1)}$,*

$$\tilde{N}_{\ell+1} = Q(\xi, \ell+1)n \pm \delta n, \tag{2.4}$$

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

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 2.6.

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 2.6. In particular, we can assume that 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 $\Lambda_{c,k,\ell} = \xi + \beta$ for $|\beta| < \delta^2/2$. Let D be the sum of the d_i's. Therefore, Corollary 2.9 yields for any $\varepsilon > 0$ and a constant $C(\varepsilon) > 0$

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

The claim then follows from the fact that

$$\mathbb{E}(D) = \frac{\Lambda_{c,k,\ell} Q(\Lambda_{c,k,\ell}, \ell)}{Q(\Lambda_{c,k,\ell}, \ell+1)}$$

and the continuity of the above expression by choosing ε sufficiently small. □

2.4 Proof of the Upper Bound and the Critical Density

The aim of this section is to determine the value $c^*_{k,\ell}$ and prove Theorem 2.2. We proceed with the proof of Theorem 2.2, 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 $\widetilde{H}_{n,p,k}$. Applying Corollary 2.10 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

$$\xi^* \frac{Q(\xi^*,\ell)}{kQ(\xi^*,\ell+1)} = \ell \implies k\ell = \xi^* \frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}. \tag{2.6}$$

This is precisely (3.1). 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

$$c^*_{k,\ell} = \frac{\xi^*}{\bar{x}k} = \frac{\xi^*}{kQ(\xi^*,\ell)^{k-1}}. \tag{2.7}$$

Proof of Theorem 2.2. 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 increasing in ξ (cf. Corollary 2.25). Recall that $\xi = \bar{x}ck$. We conclude the proof by showing the following claim.

Claim 2.11. The quantity $\xi = \bar{x}ck$ is increasing with respect to c. So, for some fixed c, with probability $1 - o(1)$

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

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

$$ck = \frac{\xi}{Q(\xi, \ell)^{k-1}}. \tag{2.8}$$

The derivative of the function $F(\xi) := \frac{\xi}{Q(\xi,\ell)^{k-1}}$ with respect to ξ is given by

$$Q(\xi, \ell)^{-k} \left(Q(\xi, \ell) - (k-1)\xi \cdot \mathbb{P}\left(\mathrm{Po}(\xi) = \ell - 1\right)\right).$$

An easy calculation shows that $F'(\xi)$ is positive when ξ satisfies the inequality

$$\sum_{i \geq \ell} \frac{\xi^{i-\ell}}{i!} > \frac{k}{(\ell-1)!},$$

and negative otherwise. We therefore conclude that $F(\xi)$ is a convex function. Moreover, by the assumption in Theorem 2.6 we have $ck > \min_{x>0}(x/Q(x,\ell)^{k-1})$. This implies the function $\xi \cdot Q(\xi, \ell)^{-(k-1)}$ is strictly increasing in the domain of interest. Note that by (2.8) the first derivative of ξ with respect to c is given by $k/F'(\xi)$ which is positive by the above discussion, thus proving our claim. □

2.5 Proof of the Lower Bound

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, we write $v_H = |V_H|$ and $e_H = |V_H|$. 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 $e_U/v_U \geq \ell$. By a *maximal* ℓ-dense subset we mean that whenever we add a vertex to such a set, then its density drops below ℓ.

In order to prove Theorem 2.3 we will to show that whenever $c < c^*_{k,\ell}$, the random graph $H_{n,\lfloor cn \rfloor,k}$ 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 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 property.

Proposition 2.12. *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. In particular, the following lemma argues about subsets of size at most $0.6n$.

Lemma 2.13. *Let $c < \ell$ and $k \geq 3$, $\ell \geq 2$. With probability $1 - o(1)$, $H_{n,\lfloor cn \rfloor,k}$ contains no maximal ℓ-dense subset with less than $0.6n$ vertices.*

Proof. We first prove the lemma for all $k \geq 3$ and $\ell \geq 2$ except for the case $(k, \ell) = (3, 2)$ by using a rough first moment argument. The probability that an edge of $H_{n,cn,k}$ is contained completely in a subset U of the vertex set is given by

$$\binom{|U|}{k} / \binom{n}{k} \leq \left(\frac{|U|}{n}\right)^k.$$

Let $k/n \leq u \leq 0.6$ and for $x \in (0, 1)$ let $H(x) = -x \log x - (1 - x) \log(1 - x)$ denote the entropy function. Then

$$\mathbb{P}\left(\exists \ell\text{-dense subset with } un \text{ vertices}\right) \leq \binom{n}{un} \cdot \binom{cn}{\ell un} (u^k)^{\ell un} \leq e^{n((\ell+1)H(u) + k\ell u \log u)}. \tag{2.9}$$

We first show that the exponent attains its maximum at $u = k/n$ or $u = 0.6$. Let $u_{max} = 1 - (\ell + 1)/k\ell$. We note that the second derivative of the exponent in (2.9) equals

$$(k\ell(1 - u) - (\ell + 1))/(u(1 - u)),$$

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 attains a global maximum at $u = k/n$ 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 $(\ell+1)\log((1-u)/u) + 3\ell(1+\log(u))$, which is at least $2.8\ell - 0.41 > 0$ for $u = 0.6$. Hence, the exponent is increasing at $u = 0.6$.

We can now infer that for $k = 3$, $\ell \geq 3$ and $k \geq 4$, $\ell \geq 2$, the exponent is either maximized at $u = k/n$ or at $u = 0.6$. Note that

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

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

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

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

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

So, the maximum is obtained at $u = k/n$ for n sufficiently large, and we conclude the case in which $(k, \ell) \neq (3, 2)$ with

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

For the case $(k, \ell) = (3, 2)$ a counting argument as above involving the 2-dense sets does not work, and we will use the property that the considered set are *maximal* 2-dense. By (2.7) we obtain $c^*_{3,2} < 1.97$. Let $p = c'/\binom{n-1}{2}$, where $c' = 3 \cdot c \leq 3 \cdot c^*_{3,2} \leq 5.91$. A simple application of Stirling's formula reveals

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

Let U be a maximal 2-dense subset of $H_{n,cn,3}$. 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

$$\begin{aligned}&\mathbb{P}\left(H_{n,cn,3} \text{ contains a maximal 2-dense subset } U \text{ with at most } 0.6n \text{ vertices}\right) =\\ &O(\sqrt{n}) \cdot \mathbb{P}\left(H_{n,p,3} \text{ contains a maximal 2-dense subset } U \text{ with at most } 0.6n \text{ vertices}\right).\end{aligned}$$

To complete the proof it is therefore sufficient to show that the latter probability is $o(n^{-1/2})$. By Proposition 2.12 the event that $H_{n,p,3}$ contains a maximal 2-dense subset U implies that there exists 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 show that the expected number of such sets with at most $0.6n$ vertices is $o(1)$. We accomplish this in two steps. Note that if a subset U is maximal 2-dense, then certainly $|U| \geq 5$. Let us begin with the case $s := |U| \leq n^{1/3}$. 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, for large n the probability that $H_{n,p,3}$ contains such a subset with at most $\lfloor n^{1/3} \rfloor$ vertices is bounded by

$$\sum_{s=5}^{\lfloor n^{1/3} \rfloor} \sum_{\theta=0}^{1} n^s s^{6s+3\theta} p^{2s+\theta} < \sum_{s=5}^{\lfloor n^{1/3} \rfloor} 2n^s s^{6s+3} p^{2s} = \sum_{s=5}^{\lfloor n^{1/3} \rfloor} 2 \left(n s^6 \left(\frac{c'}{\binom{n-1}{2}} \right)^2 \right)^s \cdot s^3$$
$$\leq n \sum_{s=5}^{\lfloor n^{1/3} \rfloor} 2 \left(c'^2 n^{(1+6/3)-4} \right)^s \leq n \sum_{s=5}^{\lfloor n^{1/3} \rfloor} \left(n^{-1+o(1)} \right)^s = n^{-4+o(1)}.$$

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

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

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

$$\binom{\binom{un}{3}}{2un+\theta} \leq \left(\frac{e(un)^3}{6(2un+\theta)} \right)^{2un} = \exp\left\{ 2un \log \left(\frac{e(un)^2}{12} \right) + O(1) \right\}.$$

Finally, the probability that a vertex outside of U has a degree less than $2-\theta$ in $|U|$ is at most

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

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

$$P_u \leq \sum_{\theta=0}^{1} \binom{n}{un} \binom{\binom{un}{3}}{2un+\theta} p^{2un+\theta} (1-p)^{\binom{un}{3}-2un-\theta} \cdot \left(e^{-u^2 c'} (1 + u^2 c')(1 + O(1/n)) \right)^{(1-u)n}$$
$$\leq \exp\left\{ n \left(H(u) + 2u \log \left(\frac{e u^2 n^2}{12} \right) + 2u \log p \right) - p \left(\binom{un}{3} - 2un - 1 \right) \right.$$
$$\left. + (1-u)n(-u^2 c' + \log(1 + u^2 c')) + O(1/n) \right\}$$

$$\leq \exp\left\{n\left(H(u) + 2u\log\left(\frac{ec'u^2}{6}\right) - \frac{u^3c'}{3} + (1-u)(-u^2c' + \log(1+u^2c'))\right) + O(1/n)\right\}.$$

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

$$\begin{aligned}
\frac{2u}{c'} - \frac{u^3}{3} + (1-u)\left(-u^2 + \frac{u^2}{1+u^2c'}\right) &\overset{c'\leq 5.91}{\geq} \frac{2u}{6} - \frac{u^3}{3} + (1-u)\left(-u^2 + \frac{u^2}{1+6u^2}\right) \\
&= \frac{u}{3} - \frac{u^3}{3} - u^2 + \frac{u^2}{1+6u^2} + u^3 - \frac{u^3}{1+6u^2} \\
&= \frac{u}{3} + \frac{2u^3}{3} - \frac{6u^4}{1+6u^2} - \frac{u^3}{1+6u^2} \\
&= \frac{u}{3} + \frac{4u^5 - 6u^4}{1+6u^2} - \frac{u^3}{3(1+6u^2)} \\
&= u\left(\frac{1}{3} - \frac{u^2/3 + 6u^3 - 4u^4}{1+6u^2}\right) \overset{u\leq 0.6}{\geq} u\left(\frac{1}{3} - 0.29\right) \overset{u>0}{>} 0,
\end{aligned}$$

thus implying that for all $u \in (0, 0.6]$ the exponent is increasing with respect to c'. Therefore, it is sufficient to consider only the case when $c' = 5.91$.

The derivative of the exponent with respect to u equals $\log(c'^2u^3(1-u)) + 6 - \log 6 - \log(1+u^2c') - ((1-u)2u^3c'^2/(1+u^2c'))$. As the function $\log(c'u^3) + (2u^4c'^3/(1+u^2c'))$ is increasing and $\log\big((1-u)/(1+u^2c')\big) - (2u^3c'^2/(1+u^2c'))$ is decreasing in u, there is at most one $n^{-2/3} \leq u_0 \leq 0.6$ 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 $n^{-2/3} \leq u \leq 0.6$.

□

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 2.5 and Proposition 2.4 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.

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 (2.6). 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 2.11), 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 $\delta \to 0$ by continuity of the largest solution of $x = Q(xck, \ell)^{k-1}$.

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

$$n_{\ell+1} = Q(\xi, \ell+1)n \quad \text{and} \quad m_{\ell+1} = \frac{\xi Q(\xi,\ell)}{kQ(\xi,\ell+1)} n_{\ell+1}. \tag{2.10}$$

By applying Corollary 2.10 (and using δ^3 instead of δ) we obtain that with probability $1 - n^{-\omega(1)}$

$$\tilde{N}_{\ell+1} = n_{\ell+1} \pm \delta^3 n \quad \text{and} \quad \tilde{M}_{\ell+1} = m_{\ell+1} \pm \delta^3 n. \tag{2.11}$$

Moreover, by applying Theorem 2.6 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

$$\Lambda_{c,k,\ell} = \xi \pm \delta^3 = \xi^* - \delta \pm \delta^3, \tag{2.12}$$

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 $x = \xi^*$. By applying Taylor's Theorem to $\frac{xQ(x,\ell)}{Q(x,\ell+1)}$ around $x = \xi^*$ we obtain

$$m_{\ell+1} = (1 - e_{k,\ell} \cdot \delta + \Theta(\delta^2))\ell \cdot n_{\ell+1}, \text{ where } \frac{\xi Q(\xi,\ell)}{Q(\xi,\ell+1)} = k\ell(1 - e_{k,\ell} \cdot \delta + \Theta(\delta^2)). \tag{2.13}$$

Recall that $H_{\mathbf{d},k}$ is a random hypergraph where the ith vertex has degree d_i. 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 2.14. *Let $k \geq 3, \ell \geq 2$ and $\mathbf{d} = (d_1, \ldots, d_N)$ 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 $\mathbb{P}_{\mathbf{d},k}$ denotes the probability measure on the space of k-uniform hypergraphs with degree sequence given by $\mathbf{d}$, $\mathcal{B}(\beta, q)$ denotes the event that U is a maximal ℓ-dense set in $H_{\mathbf{d},k}$, and $H(x) = -x\log x - (1-x)\log(1-x)$*

denotes the entropy function, then

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

Proof. Recall that $H_{\mathbf{d},k}$ is obtained by beginning with d_i clones for each $1 \leq i \leq N$ 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 $0 \leq \theta < \ell$. So, by applying Proposition 2.12 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 $1,\ldots,k$. Let $\mathcal{K} = \{1,\ldots,k\}$, and let $X_1,\ldots,X_M$ be independent random sets such that for $1 \leq i \leq M$

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

Note that $|X_i|$ follows 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 as $\mathrm{Bin}(kM,q)$. 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

$$\begin{aligned}\mathbb{P}_{\mathbf{d},k}(\mathcal{B}(\beta,q)) &\leq \sum_{\theta=0}^{\ell-1}\mathbb{P}(R_k = B_\theta | X = kqM) = \sum_{\theta=0}^{\ell-1}\frac{\mathbb{P}(X = kqM \wedge R_k = B_\theta)}{\mathbb{P}(X = kqM)} \\ &= O\left(\sqrt{M}\right)\sum_{\theta=0}^{\ell-1}\mathbb{P}(X = kqM \wedge R_k = B_\theta).\end{aligned} \tag{2.14}$$

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

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

Then

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

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,

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

Thus, we have

$$\begin{aligned}
\mathbb{P}\left(X = kqM \wedge R_k = B_\theta\right) &\le \sum_{\theta=0}^{\ell-1} \binom{M}{B_\theta} q^{kqM}(1-q)^{k(1-q)M}(2^k-1)^{M-B_\theta} \\
&\le \sum_{\theta=0}^{\ell-1} M^\theta \binom{M}{\ell|U|} e^{-kMH(q)}(2^k-1)^{M-\ell|U|}\cdot(2^k-1)^{-\theta} \\
&\le \ell M^\ell \binom{M}{\ell|U|}(2^k-1)^{M-\ell|U|}e^{-kMH(q)}.
\end{aligned}$$

The claim then follows by combining the above facts and (2.14). □

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. In particular, 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 $X^{(\ell)}_{q,\beta}$ 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 2.15. *Let $\mathcal{B}(q,\beta)$ be defined as in Lemma 2.14, and let $\mathbf{d}$ be the degree sequence of the core of $\widetilde{H}_{n,p,k}$. Then*

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

By applying Lemma 2.13 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 prove Theorem 2.3 for sets of size bigger than $0.6n \ge 0.6\tilde{N}_{\ell+1}$. We also 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 (2.13) implies that for small δ all larger subsets have density smaller than ℓ. Moreover, the total degree

D of any ℓ-dense subset with $\beta\tilde{N}_{\ell+1}$ vertices is at least $k\ell \cdot \beta\tilde{N}_{\ell+1}$, i.e.,

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

By (2.11) and (2.13), we infer $\tilde{M}_{\ell+1} = \ell(1-\Theta(\delta))$ which combined with above inequality implies that $q \geq (1+\Theta(\delta))\beta$. Note that as each of the vertices in C has degree at least $\ell+1$, 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

$$\begin{aligned} k\tilde{M}_{\ell+1} &\geq q \cdot k\tilde{M}_{\ell+1} + (\ell+1)(\tilde{N}_{\ell+1} - \beta\tilde{N}_{\ell+1}) \\ &\Rightarrow q \leq 1 - \frac{(\ell+1)(1-\beta)\tilde{N}_{\ell+1}}{k\tilde{M}_{\ell+1}} \overset{(2.11),(2.13)}{\leq} 1 - \frac{(\ell+1)(1-\beta)}{k\ell}, \end{aligned}$$

where the last inequality holds for any small enough δ. Therefore, we fix β and q as follows.

$$0.6 < \beta < 1 - e_{k,\ell}\delta/2 \quad \text{and} \quad (1+\Theta(\delta))\beta \leq q \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}. \tag{2.15}$$

With Lemma 2.14 and Corollary 2.15 in hand we are ready to show the following.

Lemma 2.16. *Let $m_{\ell+1}$ and $n_{\ell+1}$ be as defined in (2.10) and $\mathcal{E}$ be the event that (2.11) holds. Then*

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

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

$$\begin{aligned} \mathbb{P}\left(X^{(\ell)}_{q,\beta} > 0\right) &\leq \mathbb{P}\left(X^{(\ell)}_{q,\beta} > 0 \mid \mathcal{E}_1 \cap \mathcal{E}\right) + \mathbb{P}\left(\overline{\mathcal{E}_1}\right) + \mathbb{P}\left(\overline{\mathcal{E}}\right) \\ &= \sum_{\vec{d} \in \{\mathcal{E}\cap\mathcal{E}_1\}} \mathbb{P}\left(X^{(\ell)}_{q,\beta} > 0 \mid \mathcal{E}_1 \cap \mathcal{E} \text{ and } \mathbf{d} = \vec{d}\right) \cdot \mathbb{P}\left(\mathbf{d} = \vec{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\ &= \sum_{\vec{d} \in \{\mathcal{E}\cap\mathcal{E}_1\}} \mathbb{P}\left(X^{(\ell)}_{q,\beta} > 0 \mid \mathbf{d} = \vec{d}\right) \cdot \mathbb{P}\left(\mathbf{d} = \vec{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\ &\overset{\text{Cor. 2.15}}{=} \sum_{\vec{d} \in \{\mathcal{E}\cap\mathcal{E}_1\}} X_{q,\beta}(\vec{d})\mathbb{P}_{\vec{d},k}(\mathcal{B}(q,\beta)) \cdot \mathbb{P}\left(\mathbf{d} = \vec{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}) \\ &= n^3\ \mathbb{E}\left(X_{q,\beta} \mid \mathcal{E}\right) \cdot \sum_{\vec{d} \in \{\mathcal{E}\cap\mathcal{E}_1\}} \mathbb{P}_{\vec{d},k}(\mathcal{B}(q,\beta))\mathbb{P}\left(\mathbf{d} = \vec{d} \mid \mathcal{E}_1 \cap \mathcal{E}\right) + O(n^{-3}). \end{aligned}$$

Note that the assumption $\vec{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}$ implies that the number of vertices $\tilde{N}_{\ell+1}$ of $\vec{d}$ is $n_{\ell+1} \pm \delta^3 n$ and the number of edges $\tilde{M}_{\ell+1}$ is $m_{\ell+1} \pm \delta^3 n$, by $\mathcal{E}$. Further note that for

small enough δ

$$\tilde{M}_{\ell+1} \leq m_{\ell+1} + \delta^3 n \leq (1 - \Theta(\delta))\ell n_{\ell+1} + \delta^3 n \leq \ell \tilde{N}_{\ell+1} - \Theta(\delta)n$$

Using Stirling's formula we obtain

$$\binom{\tilde{M}_{\ell+1}}{\ell\beta\tilde{N}_{\ell+1}} < \binom{\ell\tilde{N}_{\ell+1}}{\ell\beta\tilde{N}_{\ell+1}} = \exp(\ell n_{\ell+1} H(\beta) + O(\delta^3 n)).$$

Thus, applying Lemma 2.14 we obtain uniformly for all $\vec{d} \in \{\mathcal{E} \cap \mathcal{E}_1\}$ that

$$\mathbb{P}_{\vec{d},k}(\mathcal{B}(q,\beta)) = (2^k - 1)^{m_{\ell+1} - \beta n_{\ell+1}} \cdot e^{\ell n_{\ell+1} H(\beta) - k m_{\ell+1} H(q) + O(\delta^3 n)}.$$

The claim follows. □

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

Lemma 2.17. *There exists $\delta_0 > 0$ such that whenever $\delta < \delta_0$*

$$\mathbb{E}\left(X_{q,\beta} | \mathcal{E}\right) < \exp\left(n_{\ell+1} H(\beta) - n_{\ell+1}(1-\beta) I_{\xi^*}\left(\frac{k\ell(1-q)}{1-\beta}\right) + 0.4 \cdot \frac{k\ell}{\xi^*} \cdot n_{\ell+1}\delta + O(\delta^2 n)\right),$$

where $I_{\xi^*}\left(\frac{k\ell(1-q)}{1-\beta}\right)$ is the rate function as defined in (2.3).

Proof. Let $t = \lfloor \beta \tilde{N}_{\ell+1} \rfloor$. Conditional on $\mathcal{E}$ there are $\binom{\tilde{N}_{\ell+1}}{t} = e^{n_{\ell+1}H(\beta) + O(\delta^3 n)}$ 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. We label the vertices as $1, \ldots, \tilde{N}_{\ell+1}$ so that the vertices which are not in U are indexed from $t+1$ to $\tilde{N}_{\ell+1}$. Let the random variable d_i denote the degree of vertex i. We recall that $d_1, d_2, \ldots, d_{\tilde{N}_{\ell+1}}$ are i.i.d. $(\ell+1)$-truncated Poisson variables with parameter $\Lambda = \Lambda_{c,k,\ell} = \xi \pm \delta^3$ and mean $\mu_\Lambda = \Lambda \frac{Q(\Lambda,\ell)}{Q(\Lambda,\ell+1)}$. By Taylor's expansion of μ_λ around ξ we obtain

$$\mu_\Lambda = \xi \frac{Q(\xi,\ell)}{Q(\xi,\ell+1)} \pm \Theta(\delta^3).$$

We will calculate the probability of the event $\sum_{i=1}^{t} d_i = qk\tilde{M}_{\ell+1}$ conditional on $\mathcal{E}$. This is equivalent to calculating the probability of the event $\sum_{i=t}^{\tilde{N}_{\ell+1}} d_i = k(1-q)\tilde{M}_{\ell+1}$ conditional on $\mathcal{E}$ which by using (2.10) is same as the event

$$\sum_{i=t+1}^{\tilde{N}_{\ell+1}} \frac{d_i}{\tilde{N}_{\ell+1} - t} = \xi \frac{Q(\xi,\ell)}{Q(\xi,\ell+1)} \cdot \frac{1-q}{1-\beta} \pm \Theta(\delta^3).$$

Let us abbreviate $z = \xi\frac{Q(\xi,\ell)}{Q(\xi,\ell+1)} \cdot \frac{1-q}{1-\beta} \pm \Theta(\delta^3)$. Using the lower bound of q from (2.15) we obtain

$$\mu_\Lambda - z = \xi\frac{Q(\xi,\ell)}{Q(\xi,\ell+1)} \cdot \frac{\beta}{1-\beta}\Theta(\delta) \pm \Theta(\delta^3) > 0.$$

As $I_\Lambda(x)$ is a non-negative convex function and $I_\Lambda(\mu_\Lambda) = 0$, $I_\Lambda(x)$ is a decreasing function for $x < \mu_\Lambda$. Therefore, by Lemma 2.8

$$\mathbb{P}\left(\sum_{i=t+1}^{\tilde{N}_{\ell+1}} d_i = z(\tilde{N}_{\ell+1} - t) \mid \mathcal{E}\right) = \exp\left(-n_{\ell+1}(1-\beta)\cdot I_\Lambda(z)(1+o(1))\right)$$

and

$$I_\Lambda(z) = z(\log T_z - \log\Lambda) - T_z + \Lambda - \log Q(T_z,\ell+1) + \log Q(\Lambda,\ell+1),$$

where T_z is the unique solution of $z = T_z \cdot \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}$. Note that

$$\frac{\partial I_\Lambda(z)}{\partial\Lambda} = -\frac{z}{\Lambda} + 1 + \frac{\frac{e^{-\Lambda}\Lambda^\ell}{\ell!}}{Q(\Lambda,\ell+1)} = -\frac{z}{\Lambda} + \frac{Q(\Lambda,\ell)}{Q(\Lambda,\ell+1)} = \frac{\mu_\Lambda - z}{\Lambda}.$$

But recall that $\Lambda = \xi \pm \delta^3 = \xi^* - \delta \pm \delta^3$. So using Taylor's expansion around ξ^* to write $I_\Lambda(z)$ in terms of $I_{\xi^*}(z)$ we obtain

$$I_\Lambda(z) = I_{\xi^*}(z) - \left(\frac{\mu_{\xi^*} - z}{\xi^*}\right)(\delta \pm \delta^3) \pm O(\delta^2) = I_{\xi^*}(z) - \frac{\mu_{\xi^*}}{\xi^*} \cdot \frac{q-\beta}{1-\beta}\,\delta \pm O(\delta^2).$$

The last equality holds as $z = \mu_{\xi^*}\frac{1-q}{1-\beta}(1 - e_{k,\ell}\delta + \Theta(\delta^2))$. Since $\beta > 0.6$ we have $q - \beta < 0.4$. Also $\mu_{\xi^*} = k\ell$. Therefore,

$$I_\Lambda(z) \geq I_{\xi^*}(z) - \frac{k\ell}{\xi^*} \cdot \frac{0.4}{1-\beta}\,\delta - \pm O(\delta^2). \tag{2.16}$$

We will now approximate $I_{\xi^*}(z)$ in terms of $I_{\xi^*}\left(k\ell\frac{1-q}{1-\beta}\right)$. Note that

$$\frac{\partial I_{\xi^*}(z)}{\partial z} = \log T_z - \log\xi^*.$$

By Taylor's expansion of $I^*_\xi(z)$ around $z_0 := k\ell\frac{1-q}{1-\beta}$ we obtain

$$I_{\xi^*}(z) = I^*_\xi\left(k\ell\frac{1-q}{1-\beta}\right) + \delta \cdot e_{k,\ell}\left(k\ell\frac{1-q}{1-\beta}\right)\left(\log\frac{\xi^*}{T_{z_0}}\right) \pm O(\delta^2). \tag{2.17}$$

By Claim 2.24 the function μ_t is increasing with respect to t. This implies that $T_{z_0} < \xi^*$ as $z_0 < k\ell$, whereby $\log\frac{\xi^*}{T_{z_0}} > 0$. Also recall that $e_{k,\ell}$ denotes the value of the partial derivative of $\frac{1}{k\ell}\cdot\frac{tQ(t,\ell)}{Q(t,\ell+1)}$ with respect to t at $t = \xi^*$. Again, Claim 2.24 implies that this

is positive. We therefore obtain

$$I_{\xi^*}(z) > I^*_\xi\left(k\ell\frac{1-q}{1-\beta}\right) - \Theta(\delta^2) \tag{2.18}$$

Combining (2.16), (2.17) and (2.18) we obtain

$$I_\Lambda(z) > I^*_\xi\left(k\ell\frac{1-q}{1-\beta}\right) - \frac{k\ell}{\xi^*}\cdot\frac{0.4}{1-\beta}\ \delta - O(\delta^2).$$

The proof is then completed by using the fact that $\mathbb{P}(\mathcal{E}) = 1 - n^{-\omega(1)}$. □

Lemma 2.16 along with Lemmas 2.14 and 2.17 yield the following estimate.

Lemma 2.18. *There exists $\delta_0 > 0$ such that whenever $\delta < \delta_0$*

$$\mathbb{P}\left(X^{(\ell)}_{q,\beta} > 0\right) < \ O(n^{-3}) + F(\beta, q; \ell),$$

where

$$\begin{aligned} F(\beta,q;\ell) = \ & (2^k-1)^{m_{\ell+1}-\ell\beta n_{\ell+1}} \\ & \cdot \exp\left((\ell+1)n_{\ell+1}H(\beta) - km_{\ell+1}H(q) - n_{\ell+1}(1-\beta)I_{\xi^*}\left(\frac{k\ell(1-q)}{1-\beta}\right)\right. \\ & \left. + 0.4\cdot\frac{k\ell}{\xi^*}\cdot n_{\ell+1}\cdot\delta + O(\delta^2 n)\right), \end{aligned}$$

We can now complete the proof of Lemma 2.19 by showing the above probability is $o(1)$. We proceed as follows. Let us abbreviate

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

By using Lemma 2.18 we infer that

$$\frac{1}{n_{\ell+1}}\log F(\beta,q;\ell) \le f(\beta,q) + e_{k,\ell}\cdot\delta\cdot k\ell\left(H(q) - \frac{\log(2^k-1)}{k} + \frac{0.4}{e_{k,\ell}\cdot\xi^*}\right) + O(\delta^2).$$

By Claim 2.26 $e_{k,\ell} > 0.77/\xi^*$. So

$$\frac{1}{n_{\ell+1}}\log F(\beta,q;\ell) \le f(\beta,q) + e_{k,\ell}\cdot\delta\cdot k\ell\left(H(q) - \frac{\log(2^k-1)}{k} + 0.52\right) + O(\delta^2). \tag{2.19}$$

We will now prove the main tool for the proof of Theorem 2.3.

Lemma 2.19. *There exists $\hat{\delta} = \hat{\delta}(k,\ell) > 0$ such that if $\delta < \hat{\delta}$ the following holds. With probability $1 - n^{-\omega(1)}$, for any $0.6 < \beta \leq 1 - e_{k,\ell}\delta/2$ and $\beta < q \leq 1 - \frac{(\ell+1)(1-\beta)}{k\ell}$, we have $X^{(\ell)}_{q,\beta} = 0$.*

Proof. To deduce this lemma, we first bound $f(\beta, q)$.

Claim 2.20. For any $k \geq 3$ and $\ell \geq 2$, there exist $\varepsilon_0, C > 0$ such that for any $\varepsilon < \varepsilon_0$ the following holds. For any $0.6 < \beta \leq 1 - \varepsilon$, and q as in Lemma 2.19, we have

$$f(\beta, q) \leq -C\varepsilon.$$

The proof of Lemma 2.19 will be complete as long as we show that for δ small enough the rest of the right-hand side of (2.19) is negative. Firstly, let $\delta_1 = \delta_1(k,\ell)$ be such that for any $\delta < \delta_1$ we have $1 - e_{k,\ell}\delta/2 > 0.999$. We will consider a case distinction according to the value of q.

If $q < 0.99$, then $\beta < 0.99$ as well, and Claim 2.20 implies that $f(\beta, q) \leq -0.01 \cdot C$, where $C > 0$ depends on k and ℓ. Then let $\delta_2 = \delta_2(k,\ell) > 0$ be such that for $\delta < \delta_2$, we have

$$e_{k,\ell} \cdot \delta \cdot k\ell \left(H(0.6) - \frac{\log(2^k - 1)}{k} + 0.52 \right) + O(\delta^2) < 0.005 \cdot C.$$

Here recall that $\beta \geq 0.6$. So for any $\delta < \min\{\delta_0, \delta_1, \delta_2\}$, (2.19) implies that

$$\frac{1}{n_{\ell+1}} \log F(\beta, q; \ell) \leq -0.005 \cdot C.$$

Assume now that $q \geq 0.99$. The monotonicity of the entropy function implies that

$$H(q) - \frac{\log(2^k - 1)}{k} + 0.52 \leq H(0.99) - \frac{\log(2^k - 1)}{k} + 0.52 \overset{k\geq 3}{<} -0.072.$$

Now with $0.6 \leq \beta \leq 1 - e_{k,\ell} \cdot \delta/2$ as in Lemma 2.19, the bound of Claim 2.20 substituted in (2.19) yields

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

In turn, this is at most $-Ce_{k,\ell} \cdot \delta/4$, if $\delta < \delta_3 = \delta_3(k,\ell)$. The above cases imply that if $\delta < \min\{\delta_0, \delta_1, \delta_2, \delta_3\} =: \hat{\delta}$, then with probability $1 - e^{-\Omega(n_{\ell+1})} - O(n^{-3})$ we have $X^{(\ell)}_{q,\beta} = 0$, for all β and q as in Lemma 2.19. □

With the above result at hand we can finally complete the proof of Theorem 2.3.

Proof of Theorem 2.3. 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 2.13, which we do not repeat here. Moreover, by Theorem 2.5 and Proposition 2.4, it is enough to consider the $(\ell+1)$-core C of $\widetilde{H}_{n,p,k}$, where $p = ck/\binom{n-1}{k-1}$. The proof is completed by applying Lemma 2.19, as we can choose $\delta > 0$ as small as we please. □

The rest of the chapter is devoted to the proof of Claim 2.20 and contains a detailed analysis of the function f. 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 where the partial derivative of f with respect to β vanishes, then $f(\beta, q_0) \leq -C_1\varepsilon$. 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 2.8, $I_{\xi^*}(z) = z\left(\log T_z - \log \xi^*\right) - \log Q(T_z, \ell+1) - T_z + \log Q(\xi^*, \ell+1) + \xi^*$, where T_z is the unique solution of $z = T_z \cdot \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}$. Differentiating this with respect to z we obtain

$$\begin{aligned} I'_{\xi^*}(z) &= \log T_z - \log \xi^* + \frac{z}{T_z}\frac{dT_z}{dz} - \frac{dT_z}{dz} - \frac{Q(T_z,\ell) - Q(T_z,\ell+1)}{Q(T_z,\ell+1)}\frac{dT_z}{dz} \\ &= \log T_z - \log \xi^* + \frac{z}{T_z}\frac{dT_z}{dz} - \frac{Q(T_z,\ell)}{Q(T_z,\ell+1)}\frac{dT_z}{dz} \\ &= \log T_z - \log \xi^*. \end{aligned} \tag{2.20}$$

However, in the differentiation of f we need to differentiate $I_{\xi^*}(k\ell(1-q)/(1-\beta))$ with respect to q. Using (2.20), we obtain

$$\frac{\partial I_{\xi^*}\left(\frac{k\ell(1-q)}{1-\beta}\right)}{\partial q} = -\frac{k\ell}{1-\beta}\left(\log H_q - \log \xi^*\right),$$

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 the derivative of the entropy function, $H'(q)$ is given by $\log\left(\frac{1-q}{q}\right)$. All the above

facts together yield the derivative of $f(\beta, q)$ with respect to q

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

Therefore, if q_0 is a critical point, that is, if $\frac{\partial f(\beta,q)}{\partial q}\Big|_{q=q_0} = 0$, then with $T_0 = H_{q_0}$, q_0 satisfies

$$T_0 = \xi^*\frac{1-q_0}{q_0} \quad \textit{and} \quad \frac{k\ell(1-q_0)}{1-\beta} = \frac{T_0 Q(T_0,\ell)}{Q(T_0,\ell+1)}. \tag{2.21}$$

At this point, we have the main tool that will allow us to evaluate $f(\beta, q_0)$. We will use (2.21) in order to eliminate T_0 and express $f(\beta, q_0)$ solely as a function of q_0.

Claim 2.21. For any given $\beta \in (0.6, 1)$, if $q_0 = q_0(\beta)$ satisfies (2.21), then

$$f(\beta,q_0) = \log\left(e^{(\ell+1)H(\beta)}q_0^{k\ell}\left(\frac{(2^k-1)(1-q_0)}{q_0}\right)^{\ell(1-\beta)}\cdot\left(\frac{(1-\beta)(k\ell-\xi^*)}{k\ell q_0 - \xi^*(1-\beta)}\right)^{1-\beta}\right). \tag{2.22}$$

Proof. Note that

$$\begin{aligned} I\left(\frac{k\ell(1-q_0)}{1-\beta}\right) &= \frac{k\ell(1-q_0)}{1-\beta}\log\frac{T_0}{\xi^*} + \log\left(\frac{e^{\xi^*}Q(\xi^*,\ell+1)}{e^{T_0}Q(T_0,\ell+1)}\right) \\ &\overset{(2.21)}{=} \frac{k\ell(1-q_0)}{1-\beta}\log\left(\frac{1-q_0}{q_0}\right) + \log\left(\frac{e^{\xi^*}Q(\xi^*,\ell+1)}{e^{T_0}Q(T_0,\ell+1)}\right). \end{aligned}$$

Therefore,

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

Also, the definition of the entropy function implies that

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

Thus

$$-(1-\beta)I\left(\frac{k\ell(1-q_0)}{1-\beta}\right) - k\ell H(q_0) = \log\left(q_0^{k\ell}\left(\frac{e^{T_0}Q(T_0,\ell+1)}{e^{\xi}Q(\xi^*,\ell+1)}\right)^{1-\beta}\right). \tag{2.23}$$

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 (2.21) with respect to $e^{T_0}Q(T_0, \ell+1)$ yields

$$e^{T_0}Q(T_0,\ell+1) = e^{T_0}\frac{T_0Q(T_0,\ell)}{z_0} = \frac{e^{T_0}T_0}{z_0}\left(Q(T_0,\ell+1) + e^{-T_0}\frac{{T_0}^\ell}{\ell!}\right).$$

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

$$z_0 - T_0 = \frac{k\ell(1-q_0)}{1-\beta} - \frac{\xi^*(1-q_0)}{q_0} = \frac{(1-q_0)(k\ell q_0 - \xi^*(1-\beta))}{(1-\beta)q_0}.$$

Thus we obtain

$$\begin{aligned}\log(e^{T_0}Q(T_0,\ell+1)) &= \log\left(\frac{{T_0}^{\ell+1}}{(z-T_0)\ell!}\right)\\ &\overset{(2.21)}{=} \log\left(\left(\frac{\xi^*(1-q_0)}{q_0}\right)^{\ell+1}\cdot\frac{(1-\beta)q_0}{(1-q_0)(k\ell q_0-\xi^*(1-\beta))\ell!}\right)\\ &= \log\left(\frac{(\xi^*)^{\ell+1}}{\ell!}\left(\frac{1-q_0}{q_0}\right)^{\ell}\cdot\frac{1-\beta}{k\ell q_0-\xi^*(1-\beta)}\right).\end{aligned}$$

Also, by definition of ξ^* we have $k = \frac{\xi^*Q(\xi^*,\ell)}{\ell Q(\xi^*,\ell+1)}$ which is equivalent to $k\ell = \xi^*\left(1+\frac{e^{-\xi^*}(\xi^*)^\ell/\ell!}{Q(\xi^*,\ell+1)}\right)$ and implies $e^{\xi^*}Q(\xi^*,\ell+1) = \frac{(\xi^*)^{\ell+1}/\ell!}{k\ell-\xi^*}$. Substituting this into (2.23) and adding the remaining terms, we obtain (2.22). □

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 2.22. 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

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

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 $\beta > (k\ell+\ell+1)/(2k\ell+\ell+1)$. Elementary algebra then yields that $(k\ell+\ell+1)/(2k\ell+\ell+1)$ is a decreasing function in k and ℓ. In particular its maximum is 0.6 for $k = 3$ and $\ell = 2$. Since $\beta > 0.6$ 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

$$h(\beta) := \log\left(\left(\left(\frac{(2^k-1)(\ell+1)}{k\ell-(\ell+1)(1-\beta)}\right)^{\ell}\frac{k\ell-\xi^*}{k\ell-(1+\ell+\xi^*)(1-\beta)}\right)^{1-\beta}\left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}\right) + \log(\beta^{-(\ell+1)\beta}).$$

Bounding $f(\beta, q)$ globally

To conclude the proof of Claim 2.20 it suffices to show that there exist ε_0 and $C > 0$ such that for any $\varepsilon < \varepsilon_0$ the following bounds hold

$$h(\beta), f(\beta, 1-(\ell+1)(1-\beta)/k\ell), f(\beta,\beta) \leq -C\varepsilon, \tag{2.24}$$

for all $0.6 \leq \beta \leq 1-\varepsilon$. These three inequalities will be shown in Claims 2.28, 2.29 and 2.30, respectively.

We will first bound $k\ell - \xi^*$ which we will require to bound the above functions.

Claim 2.23. Let $k \geq 3$, $\ell \geq 2$ and ξ^* satisfies (2.6). Then $\xi^* > k\ell - 0.36$. Moreover, $k\ell - \xi^* < 0.19$ for $k = 3, \ell \geq 4$ and $k \geq 4, \ell \geq 2$.

Proof. Recall that $k\ell = \frac{\xi^* Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}$. By definition we have

$$\frac{k\ell}{\xi^*} = \frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)} = 1 + \frac{\mathbb{P}\left(\mathrm{Po}(\xi^*) = \ell\right)}{Q(\xi^*,\ell+1)} = 1 + \frac{1}{\sum_{i\geq 1} \frac{(\xi^*)^i}{(\ell+1)\ldots(\ell+i)}}. \tag{2.25}$$

Let

$$\mathcal{S} := \sum_{i\geq 1} \frac{(\xi^*)^i}{(\ell+1)\ldots(\ell+i)} \quad \text{and} \quad \mathcal{S}_i := \frac{(\xi^*)^i}{(\ell+1)\ldots(\ell+i)}.$$

Substituting $\xi^* = \frac{k\ell}{1+1/\mathcal{S}}$ we obtain

$$\mathcal{S}_i = \frac{\left(\frac{1}{1+1/\mathcal{S}}\right)^i}{\left(\frac{1}{k} + \frac{1}{k\ell}\right)\ldots\left(\frac{1}{k} + \frac{i}{k\ell}\right)}. \tag{2.26}$$

By (2.26) we have

$$\mathcal{S} > \mathcal{S}_1 = \frac{k\ell \cdot \frac{\mathcal{S}}{\mathcal{S}+1}}{\ell+1} \implies \mathcal{S} > \frac{k\ell}{\ell+1} - 1 \geq 1. \tag{2.27}$$

So $\xi^* = \frac{k\ell}{1+1/\mathcal{S}} > \frac{k\ell}{2}$ and thus $\xi^* \geq 3\ell/2$. Therefore we obtain

$$\mathcal{S} > \frac{k\ell/2}{\ell+1} + \frac{(k\ell/2)^2}{(\ell+1)(\ell+2)} + \frac{(k\ell/2)^3}{(\ell+1)(\ell+2)(\ell+3)}.$$

The right-hand side is clearly increasing in k and ℓ. Therefore, substituting $k = 3$ and $\ell = 2$ we obtain $\mathcal{S} > 2.2$, implying that

$$\xi^* > (11/16)k\ell \geq (33/16)\ell. \tag{2.28}$$

In order to improve the bound upon $k\ell - \xi^*$ we use the fact that $k\ell - \xi^* = \xi^*/\mathcal{S}$ and show that $\frac{\mathcal{S}}{\xi^*} > 1$.

$$\begin{aligned}\frac{\mathcal{S}}{\xi^*} &= \sum_{i\geq 1} \frac{(\xi^*)^{i-1}}{(\ell+1)\ldots(\ell+i)} = \frac{1}{\ell+1}\left(\sum_{i\leq \ell} \frac{(\xi^*)^{i-1}}{(\ell+2)\ldots(\ell+i)} + \sum_{i\geq \ell+1} \frac{(\xi^*)^{i-1}}{(\ell+2)\ldots(\ell+i)}\right) \\ &\overset{(2.28)}{>} \frac{1}{\ell+1}\left(\ell + \sum_{i\geq \ell+1} \frac{(2\ell)^{i-1}}{(\ell+2)\ldots(\ell+i)}\right).\end{aligned}$$

For $\ell \geq 3$ observe that the term for $i = \ell + 1$ is

$$\frac{(2\ell)^{i-1}}{(\ell+2)(\ell+3)\dots(2\ell+1)} > \frac{2\ell \cdot 2\ell}{(2\ell-1)(2\ell+1)} > 1$$

For $\ell = 2$ we have $\sum_{i\geq \ell+1} \frac{(2\ell)^{i-1}}{(\ell+2)\dots(\ell+i)} > \sum_{i=3}^{5} \frac{4^3}{(2+i)(2+i-1)\dots 5} > 1$. By (2.25), we have $k\ell - \xi^* = \frac{1}{\sum_{i\geq 1} \frac{(\xi^*)^{i-1}}{(\ell+1)\dots(\ell+i)}}$, and so

$$\frac{1}{k\ell - \xi^*} > \sum_{i\geq \ell+1} \frac{(\xi^*)^{i-1}}{(\ell+1)\dots(\ell+i)} > \sum_{i\geq \ell+1} \frac{(k\ell-1)^{i-1}}{(\ell+1)\dots(\ell+i)}.$$

Let $S_i(k,\ell) = \frac{(k\ell-1)^{i-1}}{(\ell+1)\dots(\ell+i)}$. Clearly $S_i(k,\ell)$ is increasing with respect to k. Taking the derivative with respect to ℓ we obtain that

$$\begin{aligned}
\frac{\partial}{\partial \ell} S_i(k,\ell) &= S_i(k,\ell)\left(\frac{k(i-1)}{k\ell-1} - \frac{1}{\ell+1} - \frac{1}{\ell+2} - \dots - \frac{1}{\ell+i}\right) \\
&> S_i(k,\ell)\left(\frac{i-1}{\ell} - \frac{1}{\ell+1} - \frac{1}{\ell+2} - \dots - \frac{1}{\ell+i}\right) \\
&= \frac{S_i(k,\ell)}{\ell}\left(\frac{1}{\ell+1} + \frac{2}{\ell+2} + \dots + \frac{i}{\ell+i} - 1\right) \\
&> \frac{S_i(k,\ell)}{\ell}\left(\frac{i-1}{\ell+i-1} + \frac{i}{\ell+i} - 1\right) \overset{i\geq \ell+1}{>} \frac{S_i(k,\ell)}{\ell}\left(\frac{1}{2} + \frac{\ell+1}{2\ell+1} - 1\right) > 0.
\end{aligned}$$

Therefore, for all $i \geq \ell+1$, $S_i(k,\ell)$ increases with respect to ℓ. Numerical computations show that

$$\left(\sum_{i\geq \ell+1} S_i(3,3)\right)^{-1} < 0.34\ ,\ \left(\sum_{i\geq \ell+1} S_i(3,4)\right)^{-1} < 0.15 \text{ and } \left(\sum_{i\geq \ell+1} S_i(4,2)\right)^{-1} < 0.19.$$

For the case $(k,\ell) = (3,2)$ we obtain $k\ell - \xi^* < 0.36$ by direct computation. □

Claim 2.24. For every $t \geq 1$, the function $x \to xQ(x,t-1)/Q(x,t)$ is increasing for $x > 0$.

Proof. Set

$$g_t(x) := \frac{1}{(t-1)!} \cdot \frac{1}{\frac{1}{t!} + \frac{x}{(t+1)!} + \frac{x^2}{(t+2)!} + \dots}.$$

Then

$$\frac{xQ(x,t-1)}{Q(x,t)} = \frac{x(Q(x,t) + \mathbb{P}(\mathrm{Po}(x) = t-1))}{Q(x,t)} = x + g_t(x).$$

To see the claim it thus suffices to show that

$$-g_t'(x) < 1.$$

But

$$-g_t'(x) = \frac{1}{(t-1)!}\,\frac{\frac{1}{(t+1)!}+\frac{2x}{(t+2)!}+\frac{3x^2}{(t+3)!}+\cdots}{\left(\frac{1}{t!}+\frac{x}{(t+1)!}+\frac{x^2}{(t+2)!}+\cdots\right)^2}.$$

We, therefore, need to prove that

$$\frac{1}{(t-1)!}\left(\frac{1}{(t+1)!}+\frac{2x}{(t+2)!}+\frac{3x^2}{(t+3)!}+\cdots\right) < \left(\frac{1}{t!}+\frac{x}{(t+1)!}+\frac{x^2}{(t+2)!}+\cdots\right)^2. \tag{2.29}$$

We compare the coefficients on both sides one by one. Note that

$$\frac{1}{(t-1)!(t+1)!} < \frac{1}{t!^2} \Leftrightarrow t < t+1.$$

Moreover,

$$\frac{2}{(t-1)!(t+2)!} < \frac{2}{t!(t+1)!} \Leftrightarrow t < t+2.$$

Next, the coefficient of x^s for $s \geq 2$ on the right-hand side is

$$\begin{cases} 2\sum_{i=0}^{\lfloor\frac{s-1}{2}\rfloor}\frac{1}{(t+i)!(t+s-i)!}+\frac{1}{\left(t+\lceil\frac{s-1}{2}\rceil\right)!^2}, & \text{if } s \text{ is even,} \\ 2\sum_{i=0}^{\lfloor\frac{s-1}{2}\rfloor}\frac{1}{(t+i)!(t+s-i)!}, & \text{if } s \text{ is odd} \end{cases}.$$

Note that in any case we have (essentially) $s+1$ summands. So it suffices to show that each one of them is larger than the $1/(s+1)$th of the coefficient of x^s on the left-hand side, that is, $\frac{1}{(t-1)!(t+s+1)!}$. But this is the case, as for any $0 \leq i \leq s$.

$$\frac{1}{(t-1)!(t+s+1)!} < \frac{1}{(t+i)!(t+s-i)!} \Leftrightarrow (t+i)\cdots t < (t+s+1)\cdots(t+s-i+1).$$

This now concludes the proof of the claim. □

We immediately obtain the following corollary.

Corollary 2.25. *Let $k \geq 3$, $\ell \geq 2$ and ξ^* satisfies (2.6). Then $\xi^*\frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}$ is increasing with respect to ξ^*.*

Claim 2.26. Let $e_{k,\ell}$ be the value of derivative of $\frac{xQ(x,\ell)}{k\ell\cdot Q(x,\ell+1)}$ with respect to x at $x=\xi^*$. Then $e_{k,\ell} > \frac{0.77}{\xi^*}$.

Proof. We write

$$\frac{xQ(x,\ell)}{Q(x,\ell+1)} = \frac{x(Q(x,\ell+1)+\mathbb{P}\left(\mathrm{Po}(x)=\ell\right))}{Q(x,\ell+1)} = x + \frac{\frac{1}{\ell!}}{\frac{1}{(\ell+1)!}+\frac{x}{(\ell+2)!}+\frac{x^2}{(\ell+3)!}+\cdots}.$$

By definition

$$\begin{aligned}
e_{k,\ell}\cdot k\ell &= 1-\frac{1}{\ell!}\,\frac{\frac{1}{(\ell+2)!}+\frac{2\xi^*}{(\ell+3)!}+\frac{3\xi^{*2}}{(\ell+4)!}+\cdots}{\left(\frac{1}{(\ell+1)!}+\frac{\xi^*}{(\ell+2)!}+\frac{\xi^{*2}}{(\ell+3)!}+\cdots\right)^2} = 1-(k\ell-\xi^*)\cdot\frac{\frac{1}{(\ell+2)!}+\frac{2\xi^*}{(\ell+3)!}+\frac{3\xi^{*2}}{(\ell+4)!}+\cdots}{\frac{1}{(\ell+1)!}+\frac{\xi^*}{(\ell+2)!}+\frac{\xi^{*2}}{(\ell+3)!}+\cdots}\\
&= 1-(k\ell-\xi^*)\cdot\left(1-\frac{\frac{\ell+1}{(\ell+2)!}+\frac{(\ell+1)\xi^*}{(\ell+3)!}+\frac{(\ell+1)\xi^{*2}}{(\ell+4)!}+\cdots}{\frac{1}{(\ell+1)!}+\frac{\xi^*}{(\ell+2)!}+\frac{\xi^{*2}}{(\ell+3)!}+\cdots}\right)\\
&= 1-(k\ell-\xi^*)\cdot\left(1-\frac{\ell+1}{\xi^*}\cdot\left(1-\frac{\frac{1}{(\ell+1)!}}{\frac{1}{(\ell+1)!}+\frac{\xi^*}{(\ell+2)!}+\frac{\xi^{*2}}{(\ell+3)!}+\cdots}\right)\right)\\
&= 1-(k\ell-\xi^*)\left(1-\frac{\ell+1}{\xi^*}+\frac{k\ell-\xi^*}{\xi^*}\right) = 1-(k\ell-\xi^*)\left(-\frac{\ell+1}{\xi^*}+\frac{k\ell}{\xi^*}\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
e_{k,\ell} &= \frac{1}{k\ell}-\frac{k\ell-\xi^*}{k\ell}\left(-\frac{\ell+1}{\xi^*}+\frac{k\ell}{\xi^*}\right) = \frac{1}{k\ell}+\frac{\ell+1}{\xi^*}-\frac{\ell+1}{k\ell}-\frac{k\ell-\xi^*}{\xi^*}\\
&= \frac{1}{\xi^*}-\frac{k\ell-\xi^*}{\xi^*}+\frac{k\ell-\xi^*}{\xi^* k}.
\end{aligned}$$

One can check that for $(k,\ell)=(3,2)$, $e_{k,\ell}>\frac{0.77}{\xi^*}$ and for $(k,\ell)=(3,3)$, $e_{k,\ell}>\frac{0.89}{\xi^*}$. For other values we use

$$e_{k,\ell}\cdot\xi^* > 1-(k\ell-\xi^*).$$

which by second part of Claim 2.23 is at least 0.81. □

Claim 2.27. For any $k\geq 3$ and $\ell\geq 2$ we have $\xi^*<k\ell$ and

$$\xi^* > k\ell-\frac{e^{-k\ell}(k\ell)\cdot(k\ell-0.36)^\ell}{\ell!}\left(1-\exp\left(\frac{-(k\ell-\ell+0.64)^2}{2k\ell-0.72}\right)\right)^{-1}.$$

Proof. We have $k\cdot\ell=\xi^*\cdot\frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}$. As $\frac{Q(\xi^*,\ell)}{Q(\xi^*,\ell+1)}>1$ for all ξ^* and ℓ, we deduce that $\xi^*<k\ell$. By Claim 2.23 we know that for all $k\geq 3$ and $\ell\geq 2$, $\xi^*>k\ell-0.36$. In order to improve upon the above bound, note first that

$$\xi^* = k\ell\cdot\frac{Q(\xi^*,\ell+1)}{Q(\xi^*,\ell)} = k\ell-k\ell\frac{\mathbb{P}\left(\mathrm{Po}(\xi^*)=\ell\right)}{Q(\xi^*,\ell)} \geq \quad k\ell-k\ell\frac{\mathbb{P}\left(\mathrm{Po}(k\ell-0.36)=\ell\right)}{Q(k\ell-0.36,\ell)}. \tag{2.30}$$

Let X be a Poisson random variable with parameter $\mu=k\ell-0.36$. Thus, $Q(k\ell-0.36,\ell)=1-\mathbb{P}\left(X\leq\ell-1\right)$. We define $\delta=1-(\ell-1)/\mu$. Now, for any $t<0$ we have

$$\begin{aligned}
\mathbb{P}\left(X\leq\ell-1\right) =& \mathbb{P}\left(X\leq(1-\delta)\mu\right) = \mathbb{P}\left(e^{tX}\geq e^{t(1-\delta)\mu}\right)\\
&\leq \frac{\mathbb{E}\left(e^{tX}\right)}{e^{t(1-\delta)\mu}} = \frac{\exp(-\mu+\mu\cdot e^t)}{\exp(t(1-\delta)\mu)}.
\end{aligned}$$

Setting $t = \log(\ell - 1) - \log(\mu)$ we have

$$\mathbb{P}\left(X \leq \ell - 1\right) < \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu} < \exp\left(\frac{-(\mu - \ell + 1)^2}{2\mu}\right). \tag{2.31}$$

The combination of (2.30) and (2.31) lead us to the stated lower bound. □

In what follows we use the following definition

$$t(k,\ell) := \left(1 - \frac{0.36}{k\ell}\right)^{\ell} \left(1 - \exp\left(\frac{-(k\ell - \ell + 0.64)^2}{2k\ell - 0.72}\right)\right)^{-1}.$$

We are now ready to deduce the inequalities in (2.24), starting with a bound on $h(\beta)$.

Claim 2.28. For any $k \geq 3$ and $\ell \geq 2$ there is a $C_1 > 0$ such that for any $0 < \varepsilon < 1$ and any $0.6 \leq \beta \leq 1 - \varepsilon$ we have $h(\beta) \leq -C_1\varepsilon$.

Proof. By Claim 2.27, we have $k\ell - t(k,\ell) \cdot \frac{e^{-k\ell}(k\ell)^{\ell+1}}{\ell!} < \xi^* < k\ell$. Using these bounds for ξ^* we obtain

$$\begin{aligned} e^{h(\beta)} &< \beta^{-(\ell+1)\beta} \left(\frac{(2^k-1)(\ell+1)}{k\ell - (\ell+1)(1-\beta)}\right)^{\ell(1-\beta)} \left(\frac{t(k,\ell) \cdot \frac{e^{-k\ell}(k\ell)^{\ell+1}}{\ell!}}{k\ell - (\ell + k\ell + 1)(1-\beta)}\right)^{1-\beta} \\ &\quad \times \left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell} \\ &= \left(\frac{2^k - 1}{e^k \cdot \beta^{\frac{\beta}{(1-\beta)}}}\right)^{\ell(1-\beta)} \left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{-\ell(1-\beta)} \cdot \left(1 - \frac{(\ell + k\ell + 1)(1-\beta)}{k\ell}\right)^{-(1-\beta)} \\ &\quad \times \left(\frac{(\ell+1)^{\ell} \cdot t(k,\ell)}{\beta^{\frac{\beta}{(1-\beta)}}\ell!}\right)^{1-\beta} \left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}. \end{aligned} \tag{2.32}$$

Using the inequality $(1-x)^{-1} \leq \exp\left(x + \frac{x^2}{1.4}\right)$ for $x \leq 0.4$ we can deduce

$$\beta^{\frac{-\beta}{1-\beta}} = (1 - (1-\beta))^{\frac{-\beta}{1-\beta}} \leq e^{\beta + \frac{(1-\beta)\beta}{1.4}}. \tag{2.33}$$

Also,

$$\begin{aligned} &\left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{-1} \leq \exp\left\{\frac{(\ell+1)(1-\beta)}{k\ell} + \frac{(\ell+1)^2(1-\beta)^2}{1.4(k\ell)^2}\right\}, \\ &\left(1 - \frac{(1+\ell+k\ell)(1-\beta)}{k\ell}\right)^{-1/\ell} \leq \exp\left\{\frac{(1-\beta)(1+\ell+k\ell)}{k\ell^2} + \frac{(1-\beta)^2(1+\ell+k\ell)^2}{k^2\ell^3}\right\}, \\ &\left(1 - \frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell} < \exp\left(-(\ell+1)(1-\beta) - \frac{(\ell+1)^2(1-\beta)^2}{2k\ell}\right). \end{aligned}$$

By Stirling's formula and (2.33) we have

$$\frac{(\ell+1)^\ell}{\ell!\cdot\beta^{\frac{\beta}{1-\beta}}} < \frac{(1+1/\ell)^\ell \exp(\ell)}{\sqrt{2\pi\ell}} \exp\left(\beta + \frac{\beta(1-\beta)}{1.4}\right).$$

Now combining the last two terms in (2.32) we obtain

$$\begin{aligned}
&\left(\frac{(\ell+1)^\ell\cdot t(k,\ell)}{\beta^{\frac{\beta}{(1-\beta)}}\ell!}\right)^{1-\beta}\left(1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)^{k\ell}\\
&<\left(\frac{(1+1/\ell)^\ell\cdot t(k,\ell)}{\sqrt{2\pi\ell}}\right)^{1-\beta}\exp\left(\beta(1-\beta)+\frac{\beta(1-\beta)^2}{1.4}-(1-\beta)-\frac{(\ell+1)^2(1-\beta)^2}{2k\ell}\right)\\
&=\left(\frac{(1+1/\ell)^\ell\cdot t(k,\ell)}{\sqrt{2\pi\ell}}\right)^{1-\beta}\exp\left(\beta(1-\beta)+\frac{\beta(1-\beta)^2}{1.4}-(1-\beta)-\left(1+\frac{1}{\ell}\right)\frac{(\ell+1)(1-\beta)^2}{2k}\right).
\end{aligned}$$

Also recall that

$$t(k,\ell)=\left(1-\frac{0.36}{k\ell}\right)^\ell\left(1-\exp\left(\frac{-(k\ell-\ell+0.64)^2}{2k\ell-0.72}\right)\right)^{-1}.$$

Substituting these bounds in (2.32) we obtain

$$e^{h(\beta)} < \left(\left(\frac{2^k-1}{\exp\left(k-\Delta_{k,\ell,\beta}\right)}\right)^\ell\cdot\frac{(1+1/\ell)^\ell\exp\left(\beta+\frac{\beta(1-\beta)}{1.4}-1\right)}{\sqrt{2\pi\ell}\cdot\left(1-\exp\left(\frac{-(k\ell-\ell+0.64)^2}{2k\ell-0.72}\right)\right)}\right)^{1-\beta}, \tag{2.34}$$

where

$$\begin{aligned}
\Delta_{k,\ell,\beta} :=&\beta+\frac{(1-\beta)\beta}{1.4}+\frac{(\ell+1)(1-\beta)}{k\ell}+\frac{(\ell+1)^2(1-\beta)^2}{1.4(k\ell)^2}+\frac{(1-\beta)(1+k\ell+\ell)}{k\ell^2}\\
&+\frac{(1-\beta)^2(1+k\ell+\ell)^2}{k^2\ell^3}-\left(1+\frac{1}{\ell}\right)\frac{(\ell+1)(1-\beta)}{2k\ell}\\
=&\beta+\frac{(1-\beta)\beta}{1.4}+\frac{(\ell+1)(1-\beta)}{2k\ell}+\frac{(\ell+1)^2(1-\beta)^2}{1.4(k\ell)^2}+\frac{(1-\beta)(1+k\ell+\ell)}{k\ell^2}\\
&+\frac{(1-\beta)^2(1+k\ell+\ell)^2}{k^2\ell^3}-\frac{1}{\ell}\frac{(\ell+1)(1-\beta)}{2k\ell}\\
=&\beta+\frac{(1-\beta)\beta}{1.4}+\frac{(1+1/\ell)(1-\beta)}{2k}+\frac{(1+1/\ell)^2(1-\beta)^2}{1.4\ k^2}+\frac{(1-\beta)(1/2k\ell+1+1/2k)}{\ell}\\
&+\frac{(1-\beta)^2(1/k\ell+1+1/k)^2}{\ell}.
\end{aligned}$$

We note that $\Delta_{k,\ell,\beta}$ is decreasing in k and ℓ. The partial derivative of $\Delta_{k,\ell,\beta}$ with respect to β is given by

$$\begin{aligned}
\Delta'_{k,\ell,\beta} := \frac{\partial\Delta_{k,\ell,\beta}}{\partial\beta} =&\frac{12}{7}-\frac{10}{7}\beta-\frac{1+1/\ell}{2k}-\frac{(1+1/\ell)^2(1-\beta)}{(0.7)k^2}-\frac{1/2k\ell+1+1/2k}{\ell}\\
&-\frac{2(1-\beta)(1/k\ell+1+1/k)^2}{\ell}.
\end{aligned}$$

Observe that $\frac{\partial \Delta_{k,\ell,\beta}}{\partial \beta}$ is increasing with k and ℓ. Let

$$p(k,\ell,\beta) := \left(\frac{2^k-1}{\exp\left(k-\Delta_{k,\ell,\beta}\right)}\right) \quad \textit{and} \quad g(k,\ell) := \frac{\exp(1)}{\sqrt{2\pi\ell}\cdot\left(1-\exp\left(\frac{-(k\ell-\ell+0.64)^2}{2k\ell-0.72}\right)\right)}.$$

One can check that

$$e^{h(\beta)} < ((p(k,\ell,\beta))^{\ell} g(k,\ell))^{1-\beta}.$$

We start with the case $k \geq 4$. Firstly note that $\Delta'_{4,2,\beta} = -519/448 + (297/448)\beta$ which is negative for all $\beta < 1$. Also, as $(2^k-1)\cdot\exp(-k)$ is decreasing in k and $\Delta_{k,\ell,\beta}$ is decreasing in k and ℓ we infer that for $k \geq 4, \ell \geq 2$, thus the maximum value of $p(k,\ell,\beta)$ is $p(4,2,0.6)$. Numerical computations show that $p(4,2,0.6) < 0.97$. Now, clearly $g(k,\ell)$ is decreasing in k and ℓ. Moreover, one can check that $g(3,2) < 0.91$, which completes the proof for $k \geq 4, \ell \geq 2$.

For the case $k = 3$, firstly note that $\Delta'_{3,5,\beta} = 229/875 - (52/125)\beta$, which implies that $\Delta_{3,5,\beta}$ is maximized at $\beta = \beta_{max} = 229/364$. Therefore, for $\ell \geq 5$, $p(3,\ell,\beta)$ is maximized at $p(3,5,\beta_{max})$. Numerical computations show that $p(3,5,\beta_{max}) < 0.98$.

For the cases $\ell \leq 4$, firstly note that $\Delta'_{3,4,\beta} = -1/21 - 17\beta/96 \overset{\beta>0}{<} 0$. Now let

$$m(k,\ell,\beta) := p(k,\ell,\beta)^{\ell} g(k,\ell).$$

Recall that $\Delta'_{k,\ell,\beta}$ is increasing in k and ℓ. Also, $\Delta_{3,4,\beta}$ is decreasing in β. We can therefore conclude that for all $\beta \geq 0.6$ and $\ell \leq 4$, $m(3,\ell,\beta) \leq m(3,\ell,0.6)$. One can check that $m(3,3,0.6) < 0.93$ and $m(3,4,0.6) < 0.62$. The case $\ell = 2$ is more tedious. We substitute $k = 3, \ell = 2$ in (2.34).

$$\begin{aligned} e^{\frac{h(\beta)}{1-\beta}} &< \left(\frac{7}{\exp\left(3-\Delta_{3,2,\beta}\right)}\right)^2 \cdot \frac{(1+1/2)^2\exp\left(\beta+\frac{\beta(1-\beta)}{1.4}-1\right)}{\sqrt{4\pi}\cdot\left(1-\exp\left(\frac{-(4.64)^2}{11.28}\right)\right)} \\ &< \left(\frac{7}{\exp\left(3-\Delta_{3,2,\beta}-\frac{\beta}{2}-\frac{\beta(1-\beta)}{2.8}\right)}\right)^2 \cdot \frac{2.25\cdot\exp\left(-1\right)}{\sqrt{4\pi}\cdot\left(1-\exp\left(\frac{-(4.64)^2}{11.28}\right)\right)} \end{aligned} \tag{2.35}$$

Now we check that the partial derivative of $\Delta_{3,2,\beta} + \frac{\beta}{2} + \frac{\beta(1-\beta)}{2.8}$ with respect to β is less than $-0.91 + 0.47\beta$, which implies that the right-hand side is decreasing with respect to β for $\beta \leq 1$. We complete the proof by calculating the above expression for $\beta = 0.6$ which gives $e^{h(\beta)} < (0.91)^{1-\beta}$.

□

Claim 2.29. For any $k \geq 3$ and $\ell \geq 2$ there exist $\varepsilon_0 > 0$ and $C_2 > 0$ such that the following holds. For any $\varepsilon < \varepsilon_0$, if $0.6 < \beta \leq 1 - \varepsilon$ we have

$$f(\beta, \beta) < -C_2\varepsilon.$$

Proof. By Lemma 2.8, it follows that substituting $q = \beta$ in $\frac{k\ell(1-q)}{1-\beta}$ we have

$$I_{\xi^*}\left(\frac{k\ell(1-\beta)}{1-\beta}\right) = 0.$$

So,

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

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 preserves its convexity. Note that $-H(1-\varepsilon) < -\varepsilon\log(1/\varepsilon)$, whereby it follows that there exists a constant $C_2 = C_2(k, \ell) > 0$ such that for any $0 < \varepsilon < 1/e$ we have

$$f(1-\varepsilon, 1-\varepsilon) < -C_2\varepsilon\log(1/\varepsilon) < -C_2\varepsilon.$$

Since $H(0.6) > 0.6$, we have

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

The derivative of this function with respect to k is $-0.6\ell + \ell \cdot 0.4\frac{2^k \log 2}{2^k - 1}$. 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 $-0.6k + 0.6 + 0.4\log(2^k - 1)$ which is again a decreasing function in k and less than -0.42 at $k = 3$. So, we may set $k = 3$ and $\ell = 2$, thus obtaining $f(0.6, 0.6) < -1.8 + 0.8\log 7 < -0.24$. The above analysis along with the convexity of $f(\beta, \beta)$ imply the claimed statement. □

Claim 2.30. For all $k \geq 3$ and $\ell \geq 2$ there is a $C_3 > 0$ such that for all ε and 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:

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

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

We get

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

Substituting for $I(\ell+1)$ the value given in Lemma 2.8 and since $e^{\xi}Q(\xi,\ell+1) = \xi^{\ell+1}/\ell!(k\ell-\xi)$ we obtain for $\beta < 1$

$$\frac{\partial f\left(\beta, 1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)}{\partial \beta} = \log\left(\left(\frac{k\ell-(\ell+1)(1-\beta)}{(\ell+1)\beta}\right)^{\ell+1}(2^k-1)^{-\ell}\cdot\frac{\ell+1}{k\ell-\xi}\right).$$

We will show that the fraction inside the logarithm is greater than 1. Note first that

$$\frac{k\ell-(\ell+1)(1-\beta)}{(\ell+1)\beta} = \frac{1}{\beta}\left(\frac{k\ell-(\ell+1)}{\ell+1}\right)+1 = \frac{1}{\beta}\left(\frac{(k-1)\ell-1}{\ell+1}\right)+1$$

is decreasing with respect to β – so we obtain a lower bound by setting $\beta = 1$. Substituting $\beta = 1$ we obtain

$$\frac{\partial f\left(\beta, 1-\frac{(\ell+1)(1-\beta)}{k\ell}\right)}{\partial \beta} > \log\left(\left(\frac{k\ell}{\ell+1}\right)^{\ell+1}(2^k-1)^{-\ell}\cdot\frac{\ell+1}{k\ell-\xi}\right).$$

By Claim 2.27, for all $k \geq 3$ and $\ell \geq 2$ we have $k\ell - \xi \leq \frac{e^{-k\ell}(k\ell)^{\ell+1}}{\ell!(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}$ which yields

$$\begin{aligned}\left(\frac{k\ell}{\ell+1}\right)^{\ell+1}(2^k-1)^{-\ell}\cdot\frac{(\ell+1)}{k\ell-\xi} \geq& \frac{e^{k\ell}\ell!(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}{(2^k-1)^{\ell}(\ell+1)^{\ell}} \\ =& \frac{e^{k\ell}\ell!(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}{\ell^{\ell}(2^k-1)^{\ell}(1+1/\ell)^{\ell}} \\ \overset{1+x\leq e^x}{>}& \frac{\ell!}{e\cdot\ell^{\ell}}\cdot\frac{e^{k\ell}(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}{(2^k-1)^{\ell}}\end{aligned} \tag{2.36}$$

Using the bounds $\ell! \geq \sqrt{2\pi\ell}(\ell/e)^{\ell}$ and $1+x \leq e^x$ we can further bound the right-hand side of (2.36) as follows:

$$\frac{\ell!}{e\cdot\ell^{\ell}}\cdot\frac{e^{k\ell}(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}{(2^k-1)^{\ell}} \geq \frac{\sqrt{2\pi\ell}}{e^{\ell+1}}\cdot\frac{e^{k\ell}(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})}{(2^k-1)^{\ell}}. \tag{2.37}$$

It is easy to verify that $\sqrt{2\pi\ell}(1-e^{-(k\ell-\ell+0.64)^2/2k\ell-0.72})$ is increasing in k and ℓ. Also the first derivative of the function $e^k/(2^k-1)$ with respect to k is $e^k(2^k(1-\log(2))-1)/(2^k-1)^2$ 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 $e^{k\ell-\ell-1}(2^k-1)^{-\ell}(k-\log(2^k-1)-1)$ 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 greater than 1.2 for $k = 3, \ell = 2$. 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 $\ell \geq 2$. □

2.6 Conclusion and Future Directions

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. In this chapter we computed tight density thresholds for multiple orientability of random hypergraphs. 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. 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 $c > c^*_{k,\ell}$.

An important future direction is to extend our result for random inhomogeneous k-uniform hypergraphs. Consider the case where the n vertices are partitioned equally into n/q sets. Each edge chooses 2 of these sets randomly. From one of these sets, say the primary set, it draws $k-1$ vertices randomly and one vertex randomly from the other set. What can we say about the orientability thresholds for such a hypergraph? This question is motivated by study of cuckoo hashing with pages [30]. To be more precise Dietzfelbinger, Mitzenmacher and Rink studied cuckoo hashing in a setting where memory is organized in large pages. Then each item (edge) can choose several locations (vertices) on a single page with some additional choices on a back up page. They showed experimentally that with $k-1$ choices on one page and a single backup location choice, one can achieve nearly the same loads as when each key has k random locations to choose from. It would be interesting to obtain provable performance bounds in this setting.

Chapter 3

Local Search Allocation

3.1 Introduction

In this chapter we consider the following process. There are n bins, initially empty, and $m = \lfloor cn \rfloor$ balls. Each ball chooses independently and uniformly at random $k \geq 3$ bins. We are looking for an allocation such that each ball is placed into one of its chosen bins and no bin has load greater than 1. How quickly can we find such an allocation?

We present a simple and novel algorithm that finds such an allocation (if it exists) and runs in *linear time* with high probability. We provide that each ball, in addition to having $k \geq 3$ choices, can also be moved among its choices on demand. An important example of an allocation strategy in this direction is *cuckoo hashing* [2, 3] which is a collision resolution scheme used in building hash tables. Here bins are the locations on the hash table and balls represent the items. In this scheme when a ball arrives, it chooses its k random bins (chosen using k random hash functions) and is allocated to one of them. In case the bin is full, the previously allocated ball is moved out and placed in one of its other $k - 1$ choices. This process may be repeated indefinitely or until a free bin is found. We give a simple algorithm that builds on the idea of cuckoo hashing and runs in linear time with high probability. Roughly speaking we propose an efficient strategy to choose the bin in case all the choices of the incoming ball are full.

We model the k-choice balls-into-bins game by a directed graph $G = (V, E)$ such that the set of vertices V corresponds to bins. We say a vertex is *occupied* if there is a ball assigned to the corresponding bin, otherwise it is *free*. Let $\mathcal{I}$ be the set of m balls. We represent each ball $x \in \mathcal{I}$ as a tuple of its k chosen bins, so we say a vertex $v \in x$ if v corresponds to one of the chosen bins of ball x. For vertices $u, v \in V$, a directed edge $e = (u, v) \in E$ if and only if there exists a ball $y \in \mathcal{I}$ so that the following two conditions

hold, (i) $u, v \in y$, and (ii) u is occupied by y. Note that a vertex with outdegree 0 is a free vertex. We denote the set of free vertices by F and the minimum of the distance of vertices in F from v by $d(v, F)$. Since G represents an allocation we call G an *allocation graph*.

Assume that at some instance a ball z arrives such that all of its k choices are occupied. Let $v \in z$ be the vertex chosen to place z. The following are the main observations.

1. The necessary condition for ball z to be successfully assigned to v is the existence of a path from v to F. This condition remains satisfied as long as some allocation is possible.
2. A free location will be found in the minimum number of steps if for all $u \in z$ the distance $d(v, F) \leq d(u, F)$.

With respect to our first observation, a natural question would be the following. Is it possible to place each of the $m = \lfloor cn \rfloor$ balls into one of their chosen bins such that each bin holds at most one ball? This has already been answered by [8, 9] in the following theorem.

Theorem 3.1. *For integers $k \geq 3$ let ξ^* be the unique solution of the equation*

$$k = \frac{\xi(1 - e^{-\xi})}{1 - e^{-\xi} - \xi e^{-\xi}}. \tag{3.1}$$

Let $c_k^* = \frac{\xi^*}{k(1-e^{-\xi^*})^{k-1}}$. Then

$$\mathbb{P}\left(\textit{allocation of } m = \lfloor cn \rfloor \textit{ balls to } n \textit{ bins is possible}\right) \overset{(n\to\infty)}{=} \begin{cases} 0, & \textit{if } c > c_k^* \\ 1, & \textit{if } c < c_k^* \end{cases}. \tag{3.2}$$

The proof of the above theorem is non-constructive, i.e., it does not give us an algorithm to find such an allocation. We propose a novel allocation algorithm called *local search allocation* (LSA) which runs in linear time with high probability. Moreover it is guaranteed to find an allocation if it exists. We state the main result of this chapter in the following theorem.

Theorem 3.2. *Let $k \geq 3$. For any fixed $\varepsilon > 0$, set $m = (1 - \varepsilon)c_k^* n$. Assume that each of the m balls chooses k random bins from a total of n bins. Then with high probability local search allocation finds an optimal allocation of these balls in time $O(n)$.*

Through simulations we demonstrate that the our allocation method requires drastically less number of selections (to place or replace an item) when compared to the *random*

walk method(which to our knowledge is also the state of art method for the process under consideration and is described latter). For instance the number of selections in the worst case is reduced by a factor of 10 when using our method.

Our second observation suggests that the allocation time depends on the selection of the bin, which we make for each assignment, from among the k possible bins. One can in principle use breadth first search (BFS) to always make assignments over the shortest path (in the allocation graph). BFS is analyzed in [3] and is shown to run in linear time only in *expectation*. One can also select uniformly at random a bin from the available bins. This resembles a random walk on the vertices of the allocation graph and is called the random walk insertion. In [9, 15] the authors analyzed the random walk insertion method and gave a polylogarithmic bound (with high probability) on the maximum allocation time, i.e., the maximum time it can take to allocate a single ball. The random walk method does not provide any guarantees for the total allocation time. In fact it might run for ever in some worst case.

Notation

Throughout this chapter we use n to denote the number of bins, m for the number of balls and k denotes the number of random choices of any ball. For an allocation graph $G = (V, E)$ and any two vertices $u, v \in V$, the shortest distance from u to v is denoted by $d(u, v)$. We denote the set of free vertices by F. We denote the shortest distance from a vertex $v \in V$ to any set of vertices say S by $d(v, S)$ which is defined as

$$d(v, S) := \min_{u \in S} d(v, u).$$

We use R to denote the set of vertices furthest to F, i.e.,

$$R := \{v \in V | d(v, F) \geq \max_{u \in V} d(u, F)\}.$$

For an integer $t \in \{0, 1, \ldots, n\}$ and a subset of vertex set $V' \subseteq V$ we use $N_t(u)$ and $N_t(V')$ to denote the set of vertices at distance at most t from the vertex $u \in V$ and the set V' respectively. Mathematically,

$$N_t(u) := \{v \in V \mid d(u, v) \leq t\} \quad \text{and} \quad N_t(V') := \{v \in V \mid d(v, V') \leq t\}.$$

In the next section we first prove the correctness of the algorithm, i.e., it finds an allocation in a finite number of steps whenever an allocation exists. We show that the algorithm takes a maximum of $O(n^2)$ time before it obtains a bin for each ball. We then proceed to give a stronger bound on the running time.

3.2 Algorithm Outline and Proof Strategy

In a nutshell LSA provides a deterministic strategy of how to select a vertex (bin) for placing a ball when all of its choices are occupied. We assign to each vertex $v \in V$ an integer label, $L(v)$. Initially all vertices have 0 as their labels. Note that at this stage, for all $v \in V$, $L(v) = d(v, F)$, i.e., the labels on the vertices represent their shortest distances from F. When a ball x appears, it chooses the vertex with the least label from among its k choices. If the vertex is free, the ball is placed on it. Otherwise, the previous ball is kicked out. The label of the vertex is then updated and set to one more than the minimum label of the remaining $k-1$ choices of the ball x. The kicked out ball chooses the bin with minimum label from its k choices and the above procedure is repeated till an empty bin is found. Note that to maintain the labels of the vertices as their shortest distances to F we would require to update labels of the neighbors of the selected vertex and the labels of their neighbors and so on. This corresponds to performing a breadth first search starting from the selected vertex. We avoid the BFS and perform only local updates and therefore the name local search allocation.

We prove the optimality and efficiency of LSA in two steps. First we show that the algorithm is correct and finds an allocation in polynomial time. To this end we prove that, at any instance, the label of a vertex is at most its shortest distance to the set of free vertices. Therefore, no vertex can have a label greater than $n-1$. This would imply that the algorithm could not run indefinitely and would stop after making at most n changes at each location. We then show that the local search allocation method will find an allocation in a time proportional to the sum of distances of the n vertices to F (in the resulting allocation graph). We complete the proof by showing that (i) if for $\varepsilon > 0$, $m = (1-\varepsilon)c_k^*$ balls are placed in n bins using k random choices for each ball then the corresponding allocation graph has two special structural properties with high probability, and (ii) if the allocation graph has these two properties, then the sum of distances of its vertices to F is linear in n.

Recall from Chapter 1 that the k-choice balls-into-bins process with m balls and n bins can be represented by a k-uniform hypergraph on n vertices and m edges. In the following section we give some structural results about random hypergraphs. We will use these results in Section 3.3 to argue about the above mentioned two special structural properties of the allocation graph.

Balls-into-Bins and Random Hypergraphs

As already mentioned we can model the balls into bins game as a hypergraph. Each bin can be viewed as a vertex and each ball as an edge. The k vertices of each edge represent the k-random choices of each ball. In fact, this is a random hypergraph with n vertices and m edges where each edge is drawn uniformly at random (with replacement) from the set of all k-multisubsets of the vertex set. Therefore, a proper allocation of balls is possible if and only if the corresponding hypergraph is 1-*orientable*, i.e., if there is an assignment of each edge $e \in E$ to one of its vertices $v \in e$ such that each vertex is assigned at most one edge. We denote a random (multi)hypergraph with n vertices and m edges by $H^*_{n,m,k}$. We define the *density* of a hypergraph as the ratio of the number of edges to the number to its vertices.

We will need the following results from [15] about the expansion properties of a random hypergraph. In the analysis of LSA we would see that these properties help us to infer that the allocation graph expands considerably and the maximum label of any vertex there is $O(\log n)$.

Theorem 3.3. *Let for any fixed $\varepsilon > 0$, $m = (1-\varepsilon)c^*_k n$. Then there exists a $\delta = \delta(\varepsilon, k)$ such that any subhypergraph of $H^*_{n,m,k}$ has density at most $(1-\delta)$ with probability $1 - O(1/n)$.*

The proof of the following lemma is similar to that in [15]. The parameters here are adjusted to our requirements; so we present the proof for completeness.

Lemma 3.4. *Let $m < c^*_k n$ and $\alpha < 1/(k-1)$. Then for every integer s such that $1 \le s \le \alpha n$, there exists a constant $\zeta > 0$ such that the following holds with probability $1 - n^{-\zeta}$. The number of vertices spanned by any set of edges of size s in $H^*_{n,m,k}$ is greater than $\left(k - 1 - \frac{\log(k-1)e^k}{\log \frac{1}{\alpha(k-1)}}\right) s$.*

Proof. Recall that each edge in $H_{n,m,k}$ is a multiset of size k. Therefore, the probability that an edge of $H_{n,m,k}$ is contained completely in a subset of size t of the vertex set is given by $\frac{t^k}{n^k}$. Thus the expected number of sets of edges of size s that span at most t vertices is at most $\binom{m}{s}\binom{n}{t}\left(\frac{t^k}{n^k}\right)^s$. Note that by the following approximation for factorials for positive integer a

$$\left(\frac{a}{e}\right)^a \sqrt{2\pi a} \le a! \le \left(\frac{a}{e}\right)^a e\sqrt{a},$$

we obtain for $0 < b < a$

$$\begin{aligned}\binom{a}{b} = \frac{a!}{b!(a-b)!} \leq & \frac{\left(\frac{a}{e}\right)^a e\sqrt{a}}{\left(\frac{b}{e}\right)^b \left(\frac{a-b}{e}\right)^{a-b} \sqrt{2\pi b}\sqrt{2\pi(a-b)}} = \frac{e}{2\pi} \cdot \left(1 - \frac{b}{a}\right)^{-(a-b+1/2)} \left(\frac{a}{b}\right)^b \\ < & \frac{\exp\left(1 + b + \frac{b}{2a} + \frac{b^2}{2a^2} - \frac{b^3}{a^2}\right)}{2\pi} \left(\frac{a}{b}\right)^b < \frac{\exp\left(1 + \frac{b}{2a} - \frac{b^3}{2a^2}\right)}{2\pi} \left(\frac{ae}{b}\right)^b \\ < & \frac{\exp(1.5)}{2\pi} \left(\frac{ae}{b}\right)^b < \left(\frac{ae}{b}\right)^b.\end{aligned}$$

Using the above bounds for $m < c_k^* n$ and setting $t = (k - 1 - \delta_s)s$ we obtain

$$\begin{aligned}\binom{m}{s}\binom{n}{t}\left(\frac{t}{n}\right)^{ks} < & \left(\frac{nc_k^* e}{s}\right)^s \left(\frac{ne}{t}\right)^t \cdot \left(\frac{t}{n}\right)^{ks} < \left(\frac{nc_k^* e}{s}\right)^s \left(\frac{ne}{t}\right)^t \cdot \left(\frac{t}{n}\right)^{ks} \\ = & \left(\frac{nc_k^*}{s}\right)^s \left(\frac{n}{t}\right)^{t-ks} e^{t+s} = \left(\frac{nc_k^* e^{k-\delta_s}}{s}\right)^s \left(\frac{n}{(k-1-\delta_s)s}\right)^{-(1+\delta_s)s} \\ < & \left(\frac{nc_k^*}{s}\right)^s \left(\frac{n}{(k-1)s}\right)^{-(1+\delta_s)s} e^{ks} \\ = & \left(\left(\frac{n}{(k-1)s}\right)^{-\delta_s} \cdot (k-1)e^k c_k^*\right)^s.\end{aligned}$$

Moreover from [8] we know that $c_k^* < 1$. Let β be such that $(1+\beta)c_k^* = 1$. Setting $\delta_s = \log(k-1)e^k / \log \frac{n}{s(k-1)}$ we obtain

$$\left(\left(\frac{n}{(k-1)s}\right)^{-\delta_s} \cdot (k-1)e^k c_k^*\right)^s = (1+\beta)^s.$$

Therefore, for $\delta_s = 1 + \ln_{k-1} e^k / \ln_{k-1} \frac{n}{s} - 1$ and $\alpha < 1/(k-1)$, the probability that there exists a set of edges of size s, where $\log n \leq s \leq \alpha n$, spanning at most $(k-1-\delta_s)s$ vertices is $\mathrm{O}((1+\beta)^{-\log n}) = O(1/n^{\log(1+\beta)})$.

Note that for $\log n \leq s \leq \alpha n$, $\delta_s < 1 + \log_{k-1} e^k / (\log_{k-1} \frac{1}{\alpha} - 1)$. For the case $1 \leq s < \log n$, we substitute $\delta_s = \log(k-1)e^k / (\log \frac{1}{\alpha(k-1)} - 1)$. Then the expected number of sets of edges of size s spanning at most $(k-1-\delta_s)s$ vertices is at most

$$\left(\left(\frac{(k-1)\log n}{n}\right)^{\frac{\log(k-1)e^k}{\log \frac{1}{\alpha(k-1)} - 1}} \cdot ((k-1)e^k)\right)^s.$$

Therefore for large n the probability that there exists a set of edges of size $1 \leq s < \log n$ spanning at most $\left(k - 1 - \frac{\log(k-1)e^k}{\log \frac{1}{\alpha(k-1)}}\right) s$ vertices is at most $o(n^{-1/2})$, which completes the proof . □

3.3 Local Search Allocation and its Analysis

3.3.1 The Algorithm

Assume that we are given balls in an online fashion, i.e., each balls chooses its k random bins whenever it appears. Moreover, balls appear in an arbitrary order. The allocation using local search method goes as follows. For each vertex $v \in V$ we maintain a label. Initially each vertex is assigned a label 0. To assign a ball x at time t we select one of its chosen vertices v such that its label is minimum (among the k choices) and assign x to v. We assign a new label to v which is one more than the minimum label of the remaining $k-1$ choices of x. However, v might have already been occupied by a previously assigned ball y. In that case we kick out y and repeat the above procedure. Let $\mathbf{L} = \{L(v) \mid v \in V\}$ and $\mathbf{T} = \{T(v) \mid v \in V\}$ where $L(v)$ denotes the label of vertex v and $T(v)$ denotes the ball assigned to vertex v. We initialize $\mathbf{L}$ with all 0s and $\mathbf{T}$ with $\emptyset$, i.e., all vertices are free. We then use Algorithm 1 to assign an arbitrary ball when it appears.

Algorithm 1 AssignBall $(x, \mathbf{L}, \mathbf{T})$

```
1: Choose a bin v among the k choices of x with minimum label L(v).
2: if (L(v) >= n − 1) then
3:    EXIT                          ▷Allocation does not exist
4: else
5:    L(v) ← 1 + min (L(u)|u ≠ v and u ∈ x)
6:    if (T(v) ≠ ∅) then
7:       y ← T(v)                   ▷Move that replaces a ball
8:       T(v) ← x
9:       CALL AssignBall(y, L, T)
10:   else
11:      T(v) ← x                   ▷Move that places a ball
```

3.3.2 Labels and the Shortest Distances

We need some additional notation. In what follows a *move* denotes either placing a ball in a free bin or replacing a previously allocated ball. Let M be the total number of moves performed by the algorithm. For $p \in [M]$ we use $L_p(v)$ to denote the label of vertex v at the end of the pth move. Similarly we use F_p to denote the set of free vertices at the end of pth move. The corresponding allocation graph is denoted as $G_p = (V, E_p)$. We need the following proposition.

Proposition 3.5. *For all $p \in [M]$ and all $v \in V$, the shortest distance of v to F_p is at least the label of v, i.e., $d(v, F_p) \geq L_p(v)$.*

Proof. We first note that the label of a free vertex always remain 0, i.e.,

$$\forall p \in [M], \forall w \in F_p, \qquad L_p(w) = 0. \tag{3.3}$$

We will now show that throughout the algorithm the label of a vertex is at most one more than the label of any of its immediate neighbors (neighbors at distance 1). More precisely,

$$\forall p \in [M], \forall (u,v) \in E_p, \qquad L_p(u) \leq L_p(v) + 1. \tag{3.4}$$

We prove (3.4) by induction on the number of moves performed by the algorithm. Initially when no ball has appeared all vertices have 0 as their labels. When the first ball is assigned, i.e., there is a single vertex say u such that $L_1(u) = 1$. Clearly, (3.4) holds after the first move. Assume that (3.4) holds after p moves.

For the $(p+1)$th move let $w \in V$ be some vertex which is assigned a ball x. Consider an edge $(u,v) \in E_p$ such that $u \neq w$ and $v \neq w$. Note that the labels of all vertices $v \in V \setminus w$ remain unchanged in the $(p+1)$th move. Therefore by induction hypothesis, (3.4) is true for all edges which does not contain w. By Step 2 of Algorithm 1 the new label of w is one more than the minimum of the labels of its $k-1$ neighbors, i.e,

$$L_{p+1}(w) = \min_{w' \in x \setminus w} L_{p+1}(w') + 1.$$

Therefore (3.4) holds for all edges originating from w. Now consider a vertex $u \in V$ such that $(u,w) \in E_p$. Now by induction hypothesis we have $L_{p+1}(u) = L_p(u) \leq L_p(w) + 1$. Note that the vertex w was chosen because it had the minimum label among the k possible choices for the ball x, i.e.,

$$L_p(w) \leq \min_{w' \in x} L_p(w') = \min_{w' \in x \setminus w} L_{p+1}(w') < L_{p+1}(w).$$

We therefore obtain $L_{p+1}(u) \leq L_p(w) + 1 < L_{p+1}(w) + 1$, thereby completing the induction step. We can now combine (3.3) and (3.4) to obtain the desired result. To see this, consider a vertex v at distance $s < n$ to a free vertex $f \in F_p$ such that s is also the shortest distance from v to F_p. By iteratively applying (3.4) we obtain $L_p(v) \leq s + L_p(f) = d(v, F_p)$, which completes the proof. □

We know that whenever the algorithm visits a vertex, it increases its label by at least 1. Trivially the maximum distance of a vertex from a free vertex is $n-1$ (if an allocation exists), and so is the maximum label. Therefore the algorithm will stop in at most $n(n-1)$ steps, i.e., after visiting each vertex at most $n-1$ times, which implies that

the algorithm is correct and finds an allocation in $O(n^2)$ time. In the following we show that the total running time is proportional to the sum of labels of the n vertices.

Lemma 3.6. *Let $\mathbf{L}^*$ be the array of labels of the vertices after all balls have been allocated using Algorithm 1. Then the total time required to find an allocation is $O(\sum_{v\in V} L^*(v))$.*

Proof. Now each invocation of Algorithm 1 increases the label of the chosen vertex by at least 1. Therefore, if a vertex has a label ℓ at the end of the algorithm then it has been selected (for any move during the allocation process) at most ℓ times. Now the given number of balls can be allocated in a time proportional to the number of steps required to obtain the array $\mathbf{L}^*$ (when the initial set consisted of all zeros) and hence is $O(\sum_{v\in V} L^*(v))$. □

For notational convenience let $F := F_M$ and $G := G_M$ denote the set of free vertices and the allocation graph (respectively) at the end of the algorithm. By Proposition 3.5 we know that for each $v \in V$, $L^*(v) \leq d(v, F)$. Moreover, by Step 2 of Algorithm 1 the maximum value of a label is n. Thus the total sum of labels of all vertices is bounded as follows.

$$\sum_{v_i\in V} L^*(v_i)) \leq \min\left(\sum_{v_i\in V} d(v, F), n^2\right).$$

So our aim now is to bound the shortest distances such that the sum of these is linear in the size of G. We accomplish this in the following section.

3.3.3 Bounding the Distances

To compute the desired sum, i.e., $\sum_{v_i\in V} d(v, F)$, we study the structure of the allocation graph. The following lemma states that, with probability $1 - o(1)$, a fraction of the vertices in the allocation graph are at a constant distance to the set of free vertices, F. This would imply that the contribution for the above sum made by these vertices is $O(n)$.

Lemma 3.7. *For any fixed $\varepsilon > 0$, let $m = (1 - \varepsilon)c_k^* n$ balls are assigned to n bins using k random choices for each ball. Then the corresponding allocation graph $G = (V, E)$ satisfies the following with probability $1 - O(1/n)$: for every $\alpha > 0$ there exist $C = C(\alpha, \varepsilon) > 0$ and a set $S \subseteq V$ of size at least $(1 - \alpha)n$ such that every vertex $v \in S$ satisfies $d(v, F) \leq C$.*

Proof. We perform the following stripping process on G. We start with G and in each step remove all its free vertices and the edges they are contained in. Note that by

removing the edges, we have removed the balls placed on the corresponding vertices, thereby creating a new set of free vertices. For step i of the stripping process, we denote the set of vertices by V_i and the set of free vertices by F_i and let G_i be the corresponding allocation graph. The number of occupied vertices in G_i is then equal to $|V_i| - |F_i|$. As each vertex holds at most one ball, the number of remaining balls is $|V_i| - |F_i|$.

Let $H = (V, E')$ be a k-uniform hypergraph with n vertices representing the bins and m edges representing the balls. Each edge consists of k vertices or k choices of the ball. Note that the number of occupied vertices in G is equal to the number of edges in H. Similarly G_i corresponds to a subgraph in H induced on the vertex set V_i. Let us denote it by $H[V_i]$. The number of occupied vertices in G_i, i.e. $|V_i| - |F_i|$, then is the number of edges in $H[V_i]$. By Theorem 3.3, with probability $1 - o(1)$ we have $|F_i| \geq \delta |V_i|$. Also by the stripping process we obtain $|V_{i+1}| = |V_i| - |F_i|$. We can therefore conclude that, with probability $1 - o(1)$, $|V_{i+1}| \leq (1 - \delta)|V_i|$. Therefore, after $t \geq 1$ iterations of removing free vertices we obtain $|V_t| \leq (1 - \delta)^t |V|$. We can now choose $t = \lceil \ln_{(1-\delta)} \alpha \rceil$ to deduce that $|V_t| < \alpha |V|$. We complete the proof by substituting $S = V \setminus V_t \geq (1 - \alpha)n$ and $C = \lceil \log_{(1-\delta)} \alpha \rceil$. □

We remark that the above lemma has already been proved in [15] (in the hypergraph setting). A similar result has also been proved in [9] (in the bipartite matching setting) for $k \geq 8$. With respect to an allocation graph recall that we denote the set of vertices furthest from F by R. Also for an integer s, $N_s(R)$ denotes the set of vertices at distance at most s from R. The next lemma states that the neighborhood of R expands suitably with high probability. We remark that the estimate, for expansion factor, presented here is not the best possible but nevertheless suffices for our analysis.

Lemma 3.8. *For any fixed $\varepsilon > 0$, let $m = (1 - \varepsilon)c_k^* n$ balls are assigned to n bins using k random choices for each ball and $G = (V, E)$ be the corresponding allocation graph. Then for any $0 < \alpha < \frac{1}{k-1}$ and every integer s such that $1 \leq |N_s(R)| \leq \alpha n$, there exists a constant $\zeta > 0$ such that G satisfies the following with probability $1 - n^{-\zeta}$.*

$$|N_s(R)| > \left(k - 1 - \frac{\log e^k (k-1)}{\log \frac{1}{\alpha(k-1)}} \right) |N_{s-1}(R)|.$$

Proof. Recall that in the allocation graph G, R is the set of vertices furthest from the set of free vertices. The set of vertices at distance at most s from R is denoted by $N_s(R)$. Note that each occupied vertex in G holds one ball. By construction of the allocation graph $N_s(R)$ is the set of vertices representing the choices of balls placed on vertices in $N_{s-1}(R)$. In the hypergraph setting where each ball corresponds to an edge, $|N_s(R)|$ is

the number of vertices spanned by the set of edges of size $|N_{s-1}(R)|$. We can now use Lemma 3.4 to obtain the desired result. □

We define $\mu := \log e^k(k-1)/\log(-\alpha(k-1))$. For some fixed $\gamma > 0$ we set

$$\alpha := \exp\left(\frac{-k}{k-2-\gamma}\right)(k-1)^{-1-\frac{-1}{k-2-\gamma}}, \tag{3.5}$$

which implies that $\mu = k - 2 - \gamma$.

The following corollary follows from the above two lemmas.

Corollary 3.9. *With high probability, the maximum label of any vertex in the allocation graph is* $O(\log n)$.

Proof. Set α as in (3.5). Let d be the shortest distance of vertices in R to S. Then by Lemma 3.8 with high probability,

$$|N_d(R)| > \left(k - 1 - \frac{\log e^k(k-1)}{\log \frac{1}{\alpha(k-1)}}\right)|N_{d-1}(R)| = (1+\gamma)^d|R|,$$

which implies that $d < \log_{1+\gamma} \alpha n$. Note that the shortest distance of vertices in S to F is a constant $C(\alpha, \delta)$ for δ defined in Lemma 3.7. Moreover, by Proposition 3.5 the label of any vertex is upper bounded by its distance to the set of free vertices. Therefore, the label of any vertex v is such that $L(v) = O(\log_{1+\gamma} \alpha n)$. □

We now prove our main theorem.

Proof of Theorem 3.2. Set α as in (3.5). Then by Lemma 3.7, with probability $1 - O(1/n)$, there exists a $C = C(\alpha, \varepsilon)$ and a set S such that $|S| \geq (1-\alpha)n$ and every vertex $v \in S$ satisfies $d(v, F) \leq C$. Let $T + 1$ be the maximum of the distances of vertices in R to S, i.e.,

$$T = \max_{v \in R} d(v, S) - 1.$$

Clearly the number of vertices at distance at most T from R is at most αn, i.e., $|N_T(R)| \leq \alpha n$. Moreover for all $t < T$, $|N_t(R)| < |N_T(R)|$. Then by Lemma 3.8, for all $t \leq T$ the following holds with high probability,

$$|N_{t+1}(R)| > (k - 1 - \delta)\,|N_t(R)|.$$

One can check that for $\gamma > 0$ and α as chosen above, $\delta < k - 2 - \gamma$. The total distance of all vertices from F is then given by

$$D = \sum_{v \in N_T(R)} d(v, F) + \sum_{v \in S} d(v, F).$$

As every vertex in S is at a constant distance from F, we obtain $\sum_{v \in S} d(v, F) = O(n)$. Note that for every $i > 0$, $|N_i(R)| - |N_{i-1}(R)|$ is the number of vertices at distance i from R. Therefore,

$$\begin{aligned}
&\sum_{v \in N_T(R)} d(v, F) = (T + C)|N_0(R)| + \sum_{i=1}^{T} (T + C - i)(|N_i(R)| - |N_{i-1}(R)|) \\
&= (T + C)|N_0(R)| + \sum_{i=1}^{T} (T - i)(|N_i(R)| - |N_{i-1}(R)|) + C \sum_{i=1}^{T} (|N_i(R)| - |N_{i-1}(R)|) \\
&= (T + C)|N_0(R)| + \sum_{i=1}^{T} (T - i)(|N_i(R)| - |N_{i-1}(R)|) + C(|N_T(R)| - |N_0(R)|) \\
&= \sum_{i=1}^{T} \Big((T - i)(|N_i(R)| - |N_{i-1}(R)|) + |N_0(R)| \Big) + C \cdot |N_T(R)| = \sum_{i=0}^{T-1} |N_i(R)| + O(n).
\end{aligned}$$

Now with high probability, we have $|N_{T-j}(R)| < \frac{|N_T(R)|}{(k-1-\delta)^j}$. Therefore,

$$\sum_{i=0}^{T-1} |N_i(R))| < |N_T(R)| \sum_{j=1}^{T} \frac{1}{(k - 1 - \delta)^j} < |N_T(R)| \sum_{j=1}^{T} \frac{1}{(1 + \gamma)^j} = O(n),$$

which completes the proof of Theorem 3.2. □

We obtain the following corollary about maximum matchings in left regular random bipartite graphs. Recall that a bipartite graph $G = (L \cup R; E)$ is k-left regular if each vertex $v \in L$ has exactly k neighbors in R.

Corollary 3.10. *For $k \geq 3$ and c_k^* as defined in Theorem 3.1, let $G = (L \cup R; E)$ be a random k-left regular bipartite graph such that $|L|/|R| < c_k^*$. The local search allocation method obtains a maximum cardinality matching in G in time $O(|R|)$ with probability $1 - o(1)$.*

Proof. We assign label 0 to each of the vertices in R initially. Each vertex in L can be considered as a ball and let R be the set of bins. The k random choices for $v \in L$ (ball) are the k random neighbors of v. We can now find a matching for each $v \in L$ by using Algorithm 1. □

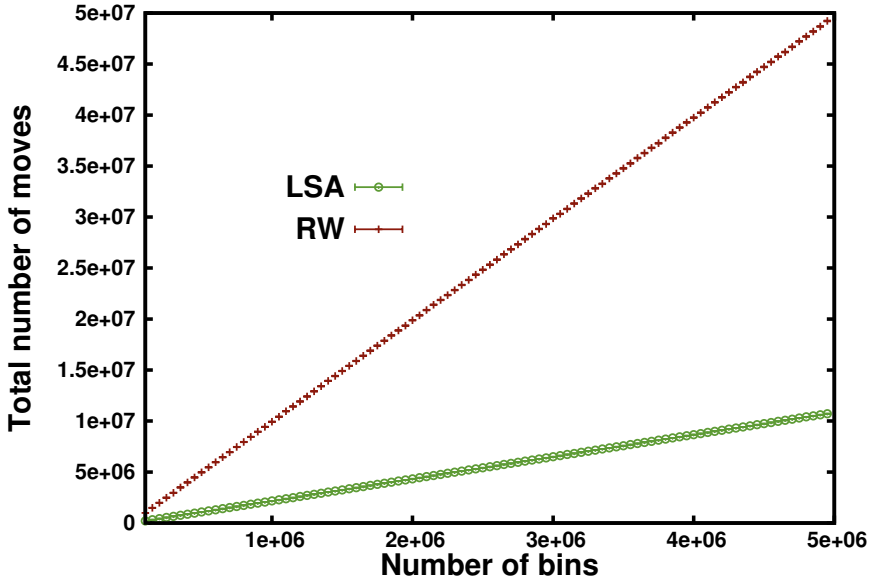
3.3.4 Experimental Results and Discussion

We present some simulations to compare the performance of local search allocation with the random walk method which (to the best of our knowledge) is currently the state-of-art method and so far considered to be the fastest algorithm for the case $k \geq 3$. We recall that in the random walk method we choose a bin at random from among the k possible bins to place the ball. If the bin is not free, the previous ball is moved out. The moved out ball again chooses a random bin from among its choices and the procedure goes on till an empty bin is found. In our experiments we consider $n \in [10^5, 5 \times 10^6]$ balls and $\lfloor cn \rfloor$ bins. The k random bins are chosen when the ball appears. All random numbers in our simulations are generated by *MT19937* generator of GNU Scientific Library [31].

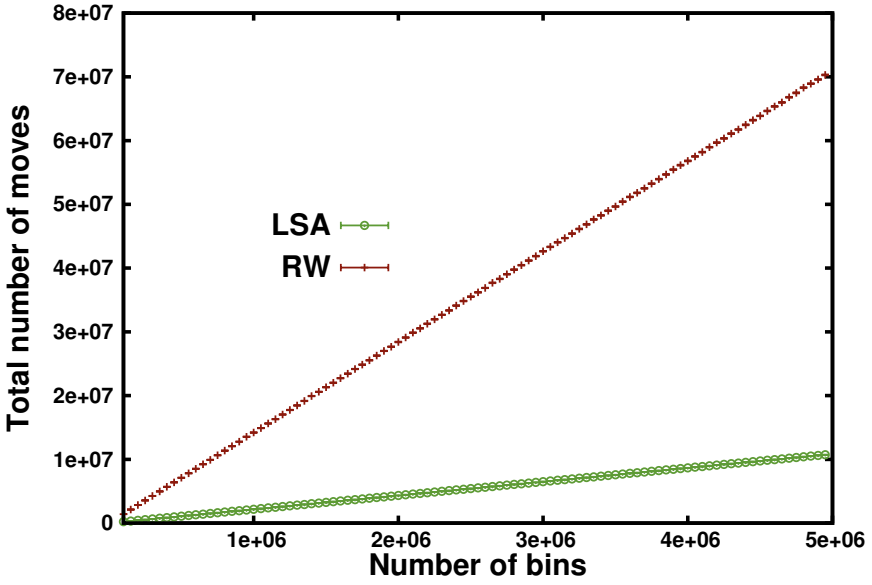
Recall that a move is either placing an item at a free location or replacing it with other item. In Figure 3.1 we give a comparison of the total number of moves (averaged over 100 random instances) performed by local search and random walk methods for $k = 3$ and $k = 4$. Figure 3.2 compares the maximum number of moves (averaged over 100 random instances) for a single insertion performed by local search and random walk methods. Figure 3.3 shows a comparison when the number of balls are fixed and density (ratio of number of balls to that of bins) approaches the threshold density. Note that the time required to obtain an allocation by random walk or local search methods is directly proportional to the number of moves performed.

We remark that local search allocation has some additional cost, i.e., the extra space required to store the labels. Though this space is $O(n)$, local search allocation is still useful for the applications where the size of objects (representing the balls) to be allocated is much larger than the labels which are integers. Moreover, with high probability, the maximum label of any vertex is $O(\log n)$. Many integer compression methods [32] have been proposed for compressing small integers and can be potentially useful in our setting for further optimizations. Also in most of the load balancing problems, the speed of finding an assignment is a much desired and the most important requirement.

We also consider the case when each bin can hold more than one ball. To adapt LSA for this setting we make a small change, i.e., the label of a vertex (bin) stays 0 until it is fully filled. Algorithm 2 gives the modified procedure for the general bin capacities. Here $\textsc{Balls}(v)$ gives the number of balls already placed in v. Let the bin capacity or maximum load allowed be s. Figure 3.4 suggests that the total number of moves are linear in the number of bins for the cases $k = 3, 4$ where the maximum bin capacity is greater than 1.

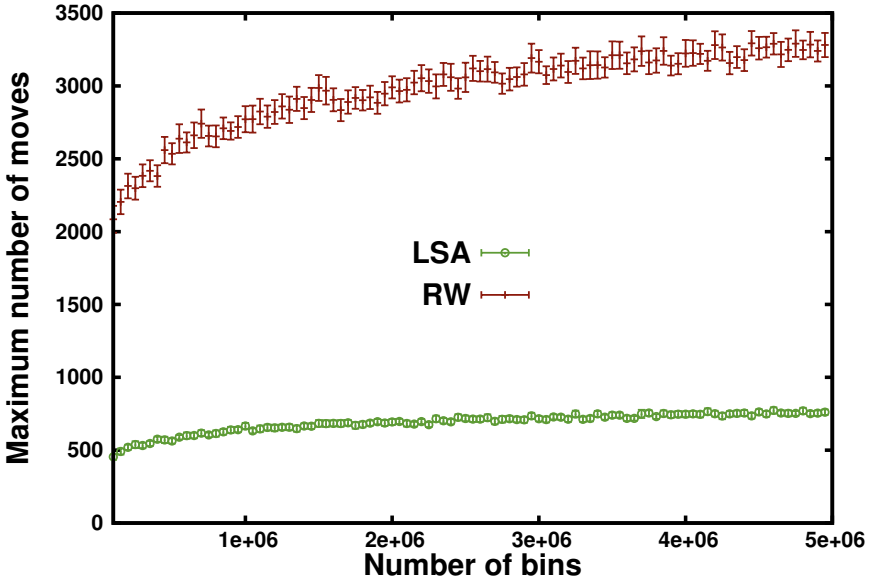


(a) $k = 3$, $c = 0.90$ ($c_3^* \approx 0.917$)

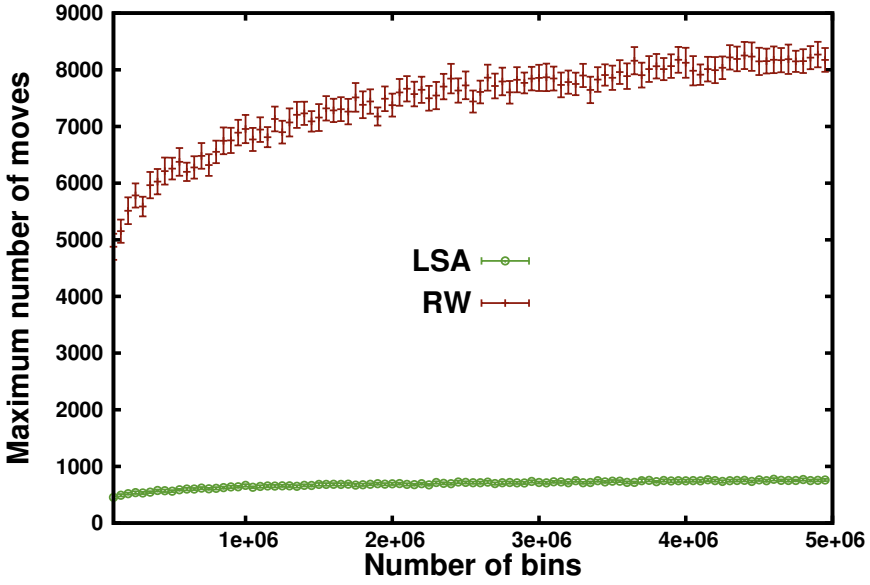


(b) $k = 4$, $c = 0.97$ ($c_4^* \approx 0.976$)

FIGURE 3.1: Comparison of total number of moves performed by local search and random walk methods.



(a) $k = 3$, $c = 0.90$ ($c_3^* \approx 0.917$).

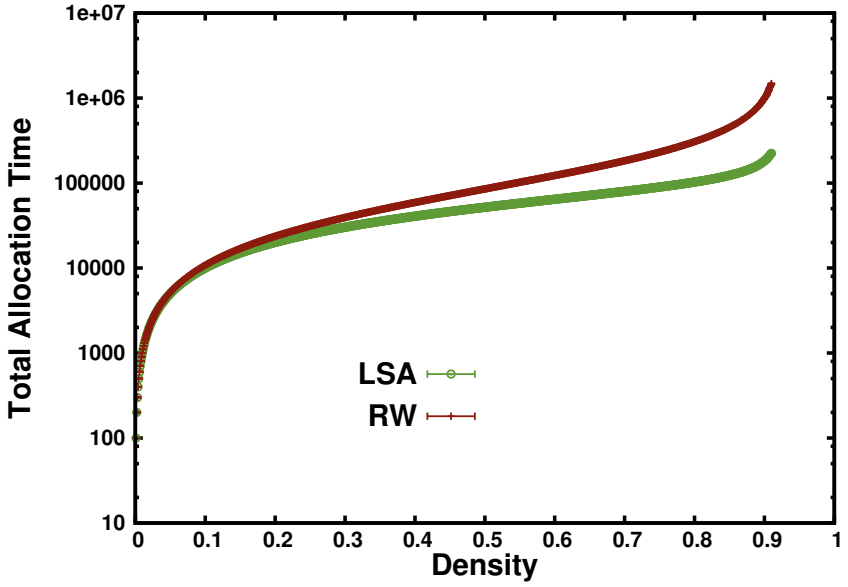


(b) $k = 4$, $c = 0.97$ ($c_4^* \approx 0.976$).

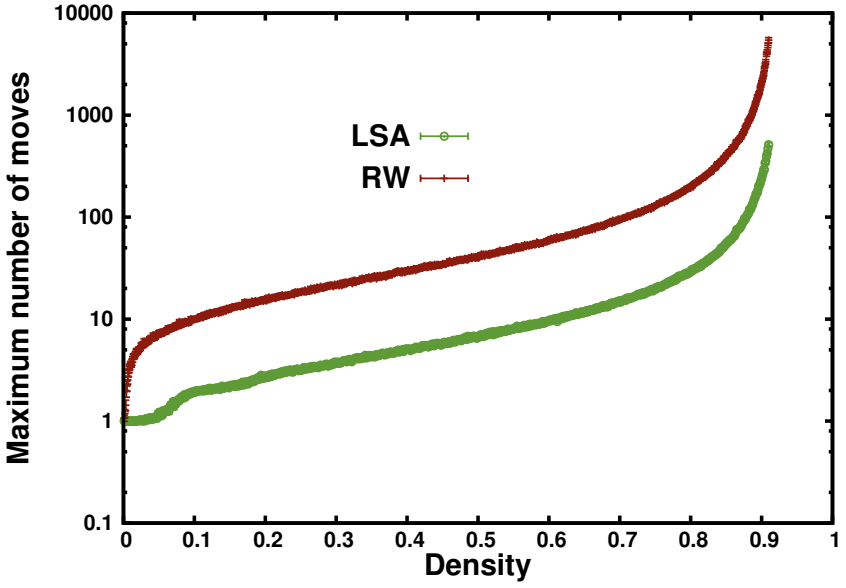
FIGURE 3.2: Comparison of maximum number of moves performed by local search and random walk methods.

3.4 Conclusion and Future Directions

We have developed a very simple and efficient method which we call local search allocation (LSA) to find an optimal allocation in a special kind of balls-into-bins process which



(a) $k = 3$, $c \leq 0.915$ $(c_3^* \approx 0.917)$



(b) $k = 3$, $c \leq 0.915$ $(c_3^* \approx 0.917)$

FIGURE 3.3: Comparison of total number of moves and maximum number of moves (for fixed number of locations, $n = 10^5$) performed by local search and random walk methods when density c approaches c_k^*.

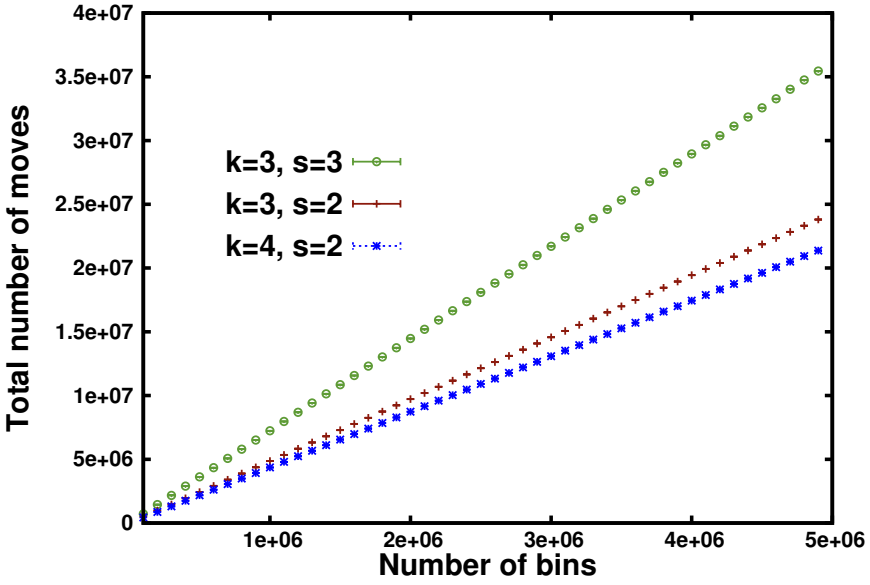


FIGURE 3.4: Total number of moves for the cases where bin capacities (maximum load, s) is greater than 1. The number of balls for all the shown cases is greater than $(c^*_{k,\ell} - 0.01)n$.

Algorithm 2 AssignBall $(x, \mathbf{L}, \mathbf{T})$

```
1: Choose a bin v among the k choices of x with minimum label L(v).
2: if (L(v) >= n − 1) then
3:    EXIT                          ▷Allocation does not exist
4: else
5:    if (BALLS(v) > s − 1) then
6:       L(v) ← 1 + min (L(u)|u ≠ v and u ∈ x)
7:    if (BALLS(v) == s) then
8:       Choose a ball (call it b) randomly from the s balls in v
9:       y ← b                      ▷Move that replaces a ball
10:      Place x in v
11:      CALL AssignBall(y, L, T)
12:   else
13:      Place x in v               ▷Move that places a ball
```

has applications in various other problems like load balancing, hashing and maximum matchings in bipartite graphs. Our algorithm runs in linear time with high probability. We performed simulations to compare our method with the state of the art method and found an order of magnitude improvement using LSA.

The most interesting aspect for continuing work is to bound the maximum allocation time, i.e., the maximum time it requires to place any ball. Our simulations show that LSA performs about 10 times better than the random walk method. A second open

question is with respect to the bin capacities. The thresholds for the existence of a proper allocation in case of arbitrary bin capacities is known (see Chapter 2). We believe that our algorithm requires linear time for finding optimal allocations even for this case. We have presented some simulations to support the same. It would therefore be interesting to provide theoretical guarantees for this case. The main obstacle (in my view) is the technical difficulty associated with proving a lemma equivalent to Lemma 3.7 in that case.

Our algorithm finds maximum matchings in large sparse k-regular random bipartite graphs in linear time with high probability. It would be an interesting direction to extend LSA for finding maximum cardinality matchings in non bipartite graphs. The main idea would be to find a representation of the matching (for example the allocation graph in the present case) with an appropriate way of labeling.

Bibliography

[1] Y. Azar, A. Z. Broder, A. R. Karlin, and E. Upfal. Balanced allocations. *SIAM Journal of Computing*, 29(1):180–200, 1999.

[2] R. Pagh and F. F. Rodler. Cuckoo hashing. *Journal of Algorithms*, 51(2):122–144, 2004.

[3] D. Fotakis, R. Pagh, P. Sanders, and P. Spirakis. Space efficient hash tables with worst case constant access time. In *Proceedings of the 20th Annual Symposium on Theoretical Aspects of Computer Science (STACS 2003)*, volume 2607 of *Lecture Notes in Computer Science*, pages 271–282. 2003.

[4] P. Sanders, S. Egner, and J. Korst. Fast concurrent access to parallel disks. In *Proceedings of the 11th annual ACM-SIAM Symposium on Discrete Algorithms (SODA 1999)*, pages 849–858, 1999.

[5] D. Fernholz and V. Ramachandran. The k-orientability thresholds for $G_{n,p}$. In *Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms (SODA 2007)*, pages 459–468, 2007.

[6] 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 *Proceedings of the 18th annual ACM-SIAM symposium on Discrete algorithms (SODA 2007)*, pages 469–476, 2007.

[7] N. Fountoulakis and K. Panagiotou. Orientability of random hypergraphs and the power of multiple choices. In *Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP 2010)*, volume 6198 of *Lecture Notes in Computer Science*, pages 348–359. 2010.

[8] N. Fountoulakis and K. Panagiotou. Sharp load thresholds for cuckoo hashing. *Random Structures & Algorithms*, 41(3):306–333, 2012.

[9] A. M. Frieze, P. Melsted, and M. Mitzenmacher. An analysis of random-walk cuckoo hashing. In *APPROX-RANDOM*, pages 490–503, 2009.

[10] P. Gao and N. C. Wormald. Load balancing and orientability thresholds for random hypergraphs. In *Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC 2010)*, pages 97–104, 2010.

[11] M. Lelarge. A new approach to the orientation of random hypergraphs. In *Proceedings of the 23th ACM-SIAM Symposium on Discrete Algorithms (SODA 2012)*, pages 251–264, 2012.

[12] M. Leconte, M. Lelarge, and L. Massoulié. Convergence of multivariate belief propagation, with applications to cuckoo hashing and load balancing. In *Proceedings of the 24th ACM-SIAM Symposium on Discrete Algorithms (SODA 2013)*, pages 35–46, 2013.

[13] A. Czumaj and V. Stemann. Randomized allocation processes. *Random Structures & Algorithms*, 18(4):297–331, 2001.

[14] L. Devroye and P. Morin. Cuckoo hashing: Further analysis. *Information Processing Letters*, 86(4):215 – 219, 2003.

[15] N. Fountoulakis, K. Panagiotou, and A. Steger. On the insertion time of cuckoo hashing. *CoRR*, abs/1006.1231, 2010.

[16] H. Bast, K. Mehlhorn, G. Schäfer, and H. Tamaki. Matching algorithms are fast in sparse random graphs. *Theory of Computing Systems*, 39(1):3–14, 2006.

[17] R. Motwani. Average-case analysis of algorithms for matchings and related problems. *Journal of the ACM*, 41(6):1329–1356, 1994.

[18] A. Goel, M. Kapralov, and S. Khanna. Perfect matchings via uniform sampling in regular bipartite graphs. *ACM Transactions. Algorithms*, 6(2):27:1–27:13, 2010.

[19] M. Dietzfelbinger, H. Peilke, and M. Rink. A more reliable greedy heuristic for maximum matchings in sparse random graphs. In *Experimental Algorithms*, volume 7276 of *Lecture Notes in Computer Science*, pages 148–159. 2012.

[20] M. Dietzfelbinger, A. Goerdt, M. Mitzenmacher, A. Montanari, R. Pagh, and M. Rink. Tight thresholds for cuckoo hashing via XORSAT. In *Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP 2010)*, volume 6198 of *Lecture Notes in Computer Science*, pages 213–225. 2010.

[21] N. Fountoulakis, M. Khosla, and K. Panagiotou. The multiple-orientability thresholds for random hypergraphs. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2011)*, pages 1222–1236, 2011.

[22] M. Khosla. Balls into bins made faster. In *Algorithms–ESA 2013*, volume 8125 of *Lecture Notes in Computer Science*, pages 601–612. 2013.

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

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

[25] J. H. Kim. Poisson cloning model for random graphs. Manuscript, 2004.

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

[27] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1:311–316, 1980.

[28] E.A. Bender and E.R. Canfield. The asymptotic number of labelled graphs with given degree sequence. *Journal of Combinatorial Theory, Series A*, 24(3):296 – 307, 1978.

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

[30] M. Dietzfelbinger, M. Mitzenmacher, and M. Rink. Cuckoo hashing with pages. In *Algorithms-ESA 2011*, volume 6942 of *Lecture Notes in Computer Science*, pages 615–627. 2011.

[31] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, M. Booth, and F. Rossi. Gnu scientific library reference manual. *URL:http://www. gnu. org/software/gsl*, 2003.

[32] B. Schlegel, R. Gemulla, and W. Lehner. Fast integer compression using simd instructions. In *Workshop on Data Management on New Hardware (DaMoN 2010)*, pages 34–40, 2010.