Pontifícia Universidade Católica
do Rio de Janeiro

Igor Albuquerque Araujo

The differential equations method and independent sets in hypergraphs

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Matemática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Matemática .

Advisor: Prof. Simon Griffiths

Rio de Janeiro
June 2019

Pontifícia Universidade Católica do Rio de Janeiro

Igor Albuquerque Araujo

The differential equations method and independent sets in hypergraphs

Dissertation presented to the Programa de Pós–graduação em Matemática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Matemática . Approved by the undersigned Examination Committee.

Prof. Simon Griffiths
Advisor
Departamento de Matemática – PUC-Rio

Prof. Robert Morris
Instituto de Matemática Pura e Aplicada – IMPA

Prof. Roberto Imbuzeiro Oliveira
Instituto de Matemática Pura e Aplicada – IMPA

Prof. Nicolau Corção Saldanha
Departamento de Matemática – PUC-Rio

Rio de Janeiro, June the 24th, 2019

Igor Albuquerque Araujo

Majored in mathematics by Pontíficia Universidade Católica do Rio de Janeiro (Rio de Janeiro, Brazil)

Bibliographic data

Albuquerque Araujo, Igor

The differential equations method and independent sets in hypergraphs / Igor Albuquerque Araujo; advisor: Simon Griffiths. – Rio de janeiro: PUC-Rio , Departamento de Matemática, 2019.

v., 89 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática.

Inclui bibliografia

1. Matemática – Teses. 2. Combinatória – Teses. 3. Probabilidade – Teses. 4. Método de equções diferenciais;. 5. Grafos aleatórios;. 6. Martingales;. 7. Processos aleatórios;. 8. Hipergrafos;. I. Griffiths, Simon. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. III. Título.

CDD: 510

To my parents and siblings, for their
support and encouragement.

Acknowledgments

I would like to offer my special thanks to my advisor Simon Richard Griffiths. Without him this work would not be possible.

I also thank CNPq for their financial support.

I wish to acknowledge the Math department for the help provided and all the professors, employees and colleagues from PUC.

Finally, I wish to thank my family and friends for their support and encouragement throughout my studies.

Abstract

Albuquerque Araujo, Igor; Griffiths, Simon (Advisor). **The differential equations method and independent sets in hypergraphs**. Rio de Janeiro, 2019. 89p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

In this dissertation, we will discuss Wormald's differential equations method, which has recently had many intriguing applications in Combinatorics. This method explores the interplay between discrete and continuous mathematics and it can be used to prove concentration in a number of discrete random processes. In particular, we will discuss the H-free process and the random greedy algorithm to obtain independent sets in hypergraphs. These processes had been extensively studied through the past few years, culminating in the recent breakthrough of Tom Bohman and Patrick Bennett in 2016, who obtained a lower bound for hypergraphs with certain density conditions. We not only reproduce the proof given by them but also obtain a stronger result (expanding their result to sparser hypergraphs) and we analyze the case of linear hypergraphs, in order to make progress towards a conjecture by Johnson and Pinto concerning the Q_2-free process in the hypercube Q_d.

Keywords

Differential Equations Method; Random Graphs; Martingales; Random Processes; Hypergraphs;

Resumo

Albuquerque Araujo, Igor; Griffiths, Simon. **O método de equações diferenciais e conjuntos independentes em hipergrafos**. Rio de Janeiro, 2019. 89p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

Nesta dissertação, discutiremos o método de equações diferenciais de Wormald, que possui muitas aplicações recentes em Combinatória. Esse método explora a interação entre a matemática discreta e contínua e pode ser usado para provar concentração em uma grande quantidade de processos aleatórios discretos. Em particular, estudaremos o processo livre de H e o algoritmo guloso aleatório para gerar conjuntos independentes em hipergrafos. Esses processos tem sido amplamente estudados nos últimos anos, culminando com o recente grande avanço de Tom Bohman e Patrick Bennett em 2016, que obtiveram uma cota inferior para hipergrafos com certas condições de densidade. Nós não só reproduzimos sua demonstração mas também obtemos um resultado mais forte (expandindo seu resultado para hipergrafos mais esparsos) e analisamos o caso de hipergrafos lineares, com o intuito de progredir rumo a uma conjectura de Johnson e Pinto sobre o processo livre de Q_2 no hipercubo Q_d.

Palavras-chave

Método de equções diferenciais; Grafos aleatórios; Martingales; Processos aleatórios; Hipergrafos;

Table of contents

We'll continue tomorrow - if I live.

Paul Erdős, *The Magician of Budapest.*

1
Introduction

1.1
Our goals

In this dissertation, we will discuss the differential equations method, which has recently had many intriguing applications in Combinatorics. This method explores the interplay between discrete and continuous mathematics and it can be used to prove concentration in a number of discrete random processes. The origins of the method can be found in the work of Kurtz ([1]). The first application to random graphs is due to Karp and Sipser in 1981 ([2]), but the general method was set by Nick Wormald in the '90s ([3],[4],[5]).

In particular, a large variety of extremal combinatorics problems can be stated in terms of maximal independent sets (with some properties) in hypergraphs. So we are interested in the application of the method to the random greedy independent set algorithm for hypergraphs, as studied by Patrick Bennett and Tom Bohman ([6]).

1.2
The probabilistic method

Probabilistic techniques have been widely used in Combinatorics in the past years. This is mainly due to the difficulty in giving explicit constructions. One way to avoid this problem, by showing that such a structure exists instead of defining it, is the so-called probabilistic method. This method consists in showing that some randomly chosen object has a positive probability of having the desired properties. While applications of the probabilistic method (due to Szele and Shannon) appeared in the early 40's, it was Paul Erdős who developed the method and showed its true power over the last century.

Building on the ideas introduced by Erdős, the probabilistic method has become a vital technique for anyone studying discrete mathematics. A good reference for it is The Probabilistic Method by N. Alon and J. Spencer ([7]), which gives wide ranging applications and an overview of the method. As we can see there, a motivation for the method comes from the Ramsey Theory. In fact, an easy example of the method appears in a paper of Erdős in 1947.

1.2.1 Ramsey numbers

The Ramsey number $R(k,\ell)$ is the smallest integer n such that in any (two-)coloring of the edges of a complete graph with n vertices by red and blue, either there exists a monochromatic red K_k or a monochromatic blue K_ℓ.

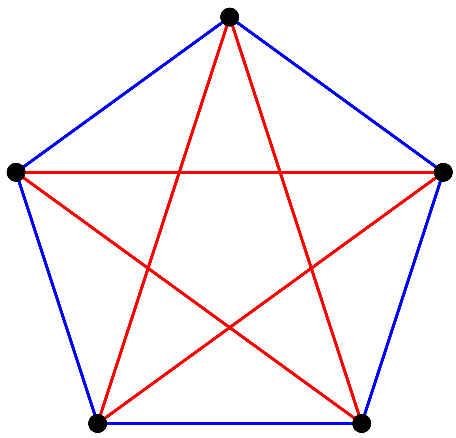


Figure 1.1: A 2-coloring of the edges of K_5 without monochromatic K_3, showing that $R(3,3) \geq 6$.

The study of these numbers started in 1929 when Ramsey showed that $R(k,\ell)$ is finite for any integers k and ℓ. Some years later, Erdős obtained a lower bound for the diagonal Ramsey numbers $R(k,k)$ by showing that the desired coloring exists in a nonconstructive way, using probabilistic arguments.

Proposition 1.1 (Erdős, 1947)**.** *If* $\binom{n}{k}2^{1-\binom{k}{2}} < 1$*, then* $R(k,k) > n$*. Thus* $R(k,k) > \lfloor 2^{k/2} \rfloor$ *for all* $k \geq 3$*.*

Proof. Consider a random coloring of the complete graph K_n, in which each edge has probability $1/2$ to be either colored red or blue, independent of the other edges. For any set S of k vertices from K_n, let A_S be the event that the subgraph induced by S is monochromatic in our coloring. Then $\mathbb{P}(A_S) = 2^{1-\binom{k}{2}}$. Since there are $\binom{n}{k}$ sets of k vertices in K_n, the probability that one of the events occur is at most $\binom{n}{k}2^{1-\binom{k}{2}} < 1$. Then, there is a positive probability that none of the events A_S occur. This means that there exists a coloring of K_n without monochromatic K_k and thus $R(k,k) > n$. The last part of the proposition follows noticing that if $n = \lfloor 2^{k/2} \rfloor$, then

$$\binom{n}{k}2^{1-\binom{k}{2}} < \frac{2^{1+\frac{k}{2}}}{k!} \cdot \frac{n^k}{2^{k^2/2}} < 1.$$

□

Another way to look at the Ramsey numbers is by considering a graph G on n vertices rather than a coloring of K_n (the red edges corresponds to the edges of G and the blue ones to the edges not present in G). In this context, the proof of Erdős is showing us that the random graph $G(n, 1/2)$ (in which each edge has probability $1/2$ of being present) has a positive (high) probability of both not containing a large clique and having small independence number. Then the random Erdős-Rényi graph gives the lower bound for the diagonal Ramsey problem.

Alternatively to the diagonal case $R(k, k)$, the next most studied case is the so-called off-diagonal problem $R(3, k)$. In 1961 Erdős proved that

$$R(3,k) = \Omega\left(\frac{k^2}{(\log k)^2}\right)$$

by applying a deterministic algorithm to $G(n, p)$. But based on the proof above, to obtain a lower bound on $R(3, k)$ we could consider a "random" triangle-free graph G (i.e., a graph that does not have K_3 as subgraph) and hope that it has small independence number with positive probability. In fact, since a random graph should have smaller independence number than the explicit known triangle-free graphs, such a graph should indeed represent an extremal coloring. However, it was only in 1990 (in the "Quo Vadis, Graph Theory?" conference) when Erdős and Bollobás first suggested the triangle-free process defined as follows.

1.3
The triangle-free process

Consider the following process: We begin with the empty graph G_0 on n vertices. At each step i we obtain G_{i+1} by adding to G_i one edge chosen uniformly at random from the collection of edges that neither are edges of G_i nor create a copy of K_3 with the edges of G_i. The process ends with a maximal triangle-free graph on n vertices, denoted by G_M (M is the random variable that counts the total number of steps of the process).

In 1995, Kim ([8]) used a similar semi-random process to prove that

$$R(3,k) = \Omega\left(\frac{k^2}{\log k}\right),$$

matching the order of magnitude of a previous upper bound from Ajtai, Komlós and Szemerédi (from 1980, [9], [10]), refined by Shearer in [11], who obtained

$$R(3,k) \leq (1 + o(1))\frac{k^2}{\log k}.$$

In 2008, Tom Bohman used the differential equations method to show that G_M has asymptotically almost surely $M = \Theta\left(n^{3/2}\sqrt{\log n}\right)$ edges and independence number $\Theta\left(\sqrt{n \log n}\right)$. This result ([12]) not only provides another proof of the lower bound $R(3,k) = \Omega(k^2/\log k)$ but also showed that the triangle-free process is very likely to generate a good Ramsey $R(3,k)$ graph for larger values of k.

Improving on the results of Bohman for the triangle-free process, Bohman and Keevash ([13]) and, independently, Fiz Pontiveros, Griffiths and Morris ([14]) proved that G_M has

$$M = \left(\frac{1}{2\sqrt{2}} + o(1)\right) n^{3/2}\sqrt{\log n}$$

edges and independence number at most $(\sqrt{2} + o(1))\sqrt{n \log n}$ with high probability, which implies that

$$R(3,k) \geq \left(\frac{1}{4} - o(1)\right) \frac{k^2}{\log k}$$

(only a factor of $4 + o(1)$ far from the best known upper bound from Shearer).

Actually, shortly afterwards the result from Bohman that the triangle-free process is very likely to generate a good Ramsey $R(3,k)$ graph, Bohman and Keevash ([15]) extended this result, by showing that for any strictly balanced graph H the number of edges in the final graph obtained in the H-free process is

$$\Omega\left(n^{2-\frac{v(H)-2}{e(H)-1}} (\log n)^{\frac{1}{e(H)-1}}\right).$$

Then, in 2015, Bennett and Bohman ([4]) extended this result even further by considering a random greeedy independent set algorithm on hypergraphs.

Our main goal in the present dissertation is to study this algorithm. By using the differential equations method we will be able to obtain a lower bound for the number of steps in the process, which will have many applications (for instance, not only in Ramsey Theory but in additive combinatorics as well).

1.4
Overview of the dissertation

The layout of the dissertation is as follows. In Chapter 2, we introduce some standard probability theory. In particular, we discuss martingales, which play an important role in all applications of the differential equations method. Also, we name some useful probability inequalities that will be extensively used throughout the dissertation.

In Chapter 3, we introduce the differential equations method and a very general theorem of Wormald (Theorem 3.1, for a recent slightly improved result see [16]), which applies in a wide variety of contexts. We explore a direct application of the theorem and conclude by discussing another way of applying the method (the wholistic approach) which is more versatile in many settings, including most recent applications of the differential equations method.

In Chapter 4, we define the H-free process and its generalization (the random greedy independent set algorithm on hypergraphs). We state the main theorem proved by Patrick Bennett and Tom Bohman (Theorem 4.1) about this algorithm, giving a lower bound on the size of the independent set obtained by the process in the case when the original hypergraph satisfies some density conditions. Furthermore, we show that their result implies a lower bound on the number of steps of the H-free process when H is a strictly balanced graph (that is, their result is indeed an extension of the previous result of Bohman and Keevash).

In Chapter 5, we discuss a variant of the H-free process in which the original graph of allowed edges is not the complete graph K_n. In particular, we discuss the results of J. Robert Johnson and Trevor Pinto ([17],[18]) on the C_4-free process in the hypercube Q_d. Furthermore, we obtain a generalization of one of their results.

In Chapter 6, we reproduce the proof of Theorem 4.1 given by Bennett and Bohman with small modifications, obtaining a stronger result than the stated in [6] (see Theorem 6.2). In Chapter 7, we show that the proof may be adjusted for linear hypergraphs (see Theorem 7.1), in which case we need weaker density conditions. Finally, in the appendix A, we address some technical standard results (that we use in Chapter 5) about the well-known gamma function.

2
Probabilistic tools

Throughout the dissertation we will use standard Probability Theory definitions. In particular, we assume the reader is familiar with standard definitions such as σ-fields, probability spaces, random variables and (conditional) expectation. In this chapter, we recall basic definitions and notation related to martingales (in Section 2.1), and give a number of useful inequalities related to deviation probabilities of martingales (in Section 2.2). These inequalities are essential tools used throughout the dissertation.

2.1
Probability background

Random graph processes

A random process is simply a sequence of random variables $X(t)$ indexed by t (we think of t as being time). A random graph process is a sequence of graphs $(G_0, G_1, \dots)$ in which every graph G_m is chosen randomly by some distribution that depends on the previous graphs $G_0, G_1 \dots, G_{m-1}$. Given a random graph process we can associate a random variable X_t (which depends only on G_t) and the natural σ-field $\mathcal{F}_t$ generated by the process.

Martingales

Let $\mathcal{F}_n$ be a filtration (an increasing sequence of σ-fields). A sequence of random variables X_n (with X_n being $\mathcal{F}_n$-measurable for each n) is said to be a *martingale* if, for all n,

- $\mathbb{E}|X_n| < \infty$ and
- $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$.

If the last condition is replaced by $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \leq X_n$, then X_n is called a *supermartingale*. Moreover, if $\mathbb{E}[X_{n+1}|\mathcal{F}_n] \geq X_n$, X_n is called a *submartingale*.

Remark 2.1. *Given a random graph process and a sequence of random variables associated with the process, we call this sequence a martingale when it is a martingale with respect to the natural filtration generated by the process.*

Stopping time

An important tool that we will use in the next chapters is the notion of a stopping time for a supermartingale. Given a filtration $\mathcal{F}_n$, a *stopping time*

T is any random variable with values in $\{0, 1, 2, \dots\} \cup \{\infty\}$ such that $\{T = n\}$ is $\mathcal{F}_n$-measurable for every n. Denoting $\min\{i, T\}$ by $i \wedge T$, the crucial result about stopping times is the following.

Lemma 2.1. *If X_i is a supermartingale and T is a stopping time, then $X_{i\wedge T}$ is also a supermartingale.*

2.2 Useful probability inequalities

If X is a positive valued random variable with finite expectation and $k > 0$, then as $X \geq k \cdot 1_{\{X \geq k\}}$ (here 1_A denotes the indicator variable of the event A), by monotonicity of expectation we obtain Markov's inequality:

$$\mathbb{E}[X] \geq k \cdot \mathbb{P}(X \geq k).$$

Applying it to the random variable $(X - \mathbb{E}[X])^2$ we obtain that

$$\mathrm{Var}(X) = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] \leq k^2 \cdot \mathbb{P}\left(|X - \mathbb{E}[X]|^2 \geq k^2\right).$$

In other words, we obtain Chebyshev's inequality:

Lemma 2.2 (Chebyshev's inequality)**.** *Let X be a random variable with expectation $\mathbb{E}[X]$ and variance $\mathrm{Var}(X)$. Then for any $k > 0$,*

$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq k\sqrt{\mathrm{Var}(X)}\right) \leq \frac{1}{k^2}.$$

Martingale inequalities

Now, suppose we have a supermartingale $X_0, X_1, \dots$, i.e., $\mathbb{E}[X_{i+1}|\mathcal{F}_i] \leq X_i$ for every $i \geq 0$. Then the increments $\Delta X_i = X_i - X_{i-1}$ satisfy $\mathbb{E}[\Delta X_i|\mathcal{F}_{i-1}] \leq 0$. If we knew that $|\Delta X| \leq c$ for some constant c then by the convexity of the exponential we would have that (for $y \in [-c, c]$ and $\lambda > 0$)

$$\begin{aligned}\exp(\lambda y) &= \exp\left(\lambda c\left(\frac{1}{2} + \frac{y}{2c}\right) + (-\lambda c)\left(\frac{1}{2} - \frac{y}{2c}\right)\right)\\ &\leq \frac{e^{\lambda c} + e^{-\lambda c}}{2} + \frac{y}{c} \cdot \frac{e^{\lambda c} - e^{-\lambda c}}{2}\end{aligned}$$

and then

$$\begin{aligned}\mathbb{E}[e^{\lambda \Delta X}] &\leq \mathbb{E}\left[\frac{e^{\lambda c} + e^{-\lambda c}}{2} + \frac{\Delta X}{c} \cdot \frac{e^{\lambda c} - e^{-\lambda c}}{2}\right]\\ &= \frac{e^{\lambda c} + e^{-\lambda c}}{2} + \frac{\mathbb{E}[\Delta X]}{c} \cdot \frac{e^{\lambda c} - e^{-\lambda c}}{2}\end{aligned}$$

$$\leq \frac{e^{\lambda c} + e^{-\lambda c}}{2} = \cosh(\lambda c) \leq e^{\frac{\lambda^2 c^2}{2}},$$

where in the last inequality we used that, for every $a > 0$,

$$\frac{e^a + e^{-a}}{2} = \cosh(a) \leq e^{a^2/2}.$$

Now, if $|\Delta X_i| \leq c_i$ a.s. for each i, then $\mathbb{E}\left[e^{\lambda \Delta X_m} | \mathcal{F}_{m-1}\right] \leq e^{\lambda^2 c_m^2/2}$ and we can compute, by induction on m,

$$\begin{aligned}
\mathbb{E}\left[e^{\lambda(X_m - X_0)}\right] &= \mathbb{E}\left[e^{\lambda\left(\sum_{i=1}^{m} \Delta X_i\right)}\right] \\
&= \mathbb{E}\left[e^{\lambda \Delta X_m} e^{\lambda\left(\sum_{i=1}^{m-1} \Delta X_i\right)}\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[e^{\lambda \Delta X_m} e^{\lambda\left(\sum_{i=1}^{m-1} \Delta X_i\right)} | \mathcal{F}_{m-1}\right]\right] \\
&= \mathbb{E}\left[e^{\lambda\left(\sum_{i=1}^{m-1} \Delta X_i\right)} \mathbb{E}\left[e^{\lambda \Delta X_m} | \mathcal{F}_{m-1}\right]\right] \\
&\leq \mathbb{E}\left[e^{\lambda\left(\sum_{i=1}^{m-1} \Delta X_i\right)} e^{\lambda^2 c_m^2/2}\right] \\
&= e^{\lambda^2 c_m^2/2} \mathbb{E}\left[e^{\lambda\left(\sum_{i=1}^{m-1} \Delta X_i\right)}\right] = \exp\left(\frac{\lambda^2}{2} \sum_{i=1}^{m} c_i^2\right).
\end{aligned}$$

So, by Markov's inequality, for all $a, \lambda > 0$,

$$\mathbb{P}(X_m - X_0 > a) = \mathbb{P}\left(e^{\lambda(X_m - X_0)} > e^{\lambda a}\right)$$

$$\leq \frac{\mathbb{E}\left[e^{\lambda(X_m - X_0)}\right]}{e^{\lambda a}} \leq \exp\left(\frac{\lambda^2}{2} \sum_{i=1}^{m} c_i^2 - \lambda a\right).$$

Hence, choosing $\lambda = \frac{a}{\sum\limits_{i=1}^{m} c_i^2}$, we obtain (Hoeffding-Azuma inequality) that

$$\mathbb{P}(X_m - X_0 > a) \leq \exp\left(\frac{\lambda^2}{2} \sum_{i=1}^{m} c_i^2 - \lambda a\right) = \exp\left(\frac{-a^2}{2\sum_{i=1}^{m} c_i^2}\right).$$

Lemma 2.3 (Hoeffding-Azuma inequality)**.** *Let X_i be a supermartingale such that $|\Delta X_i| \leq c_i$ a.s. for all i. Then*

$$\mathbb{P}(X_m - X_0 > d) \leq \exp\left(-\frac{d^2}{2\sum\limits_{i \leq m} c_i^2}\right).$$

With a slight modification in the proof above we can get the similar bound:

Lemma 2.4 (Asymmetric Hoeffding-Azuma inequality). *Let X_i be a supermartingale such that $-M \leq \Delta X_i \leq \eta$ a.s. for all i, for some $\eta < \frac{M}{10}$. Then for any $d < \eta m$ we have*

$$\mathbb{P}(X_m - X_0 > d) \leq \exp\left(-\frac{d^2}{3m\eta M}\right).$$

Sometimes the random variables $|\Delta X|$ have small probability of being close to the extremal value c_i, in which cases the above bounds are not good estimates. In these cases we will then use the following result due to Freedman ([19]).

Lemma 2.5 (Freedman inequality). *Let X_i be a supermartingale, with $|\Delta X_i| \leq C$ a.s. for all i, and $V(i) := \sum_{k\leq i} \mathrm{Var}[\Delta X_k|\mathcal{F}_{k-1}]$. Then*

$$\mathbb{P}[\exists i : V(i) \leq v, X_i - X_0 \geq d] \leq \exp\left(-\frac{d^2}{2(v + Cd)}\right).$$

We can also note that when we have asymmetric bounds $-M \leq \Delta X \leq \eta$, but $\eta > M$ we can use Freedman inequality, together to the following easy fact

Claim 2.6. *If X is a random variable such that $\mathbb{E}[X] \leq 0$ and $X \in [-a, b]$ almost surely (with $b > a > 0$) then $\mathrm{Var}[X] \leq ab$.*

Proof. Since $X \in [-a, b]$, we have $(X + a)(b - X) \geq 0$ which implies $X^2 \leq (b - a)X + ab$ and $\mathrm{Var}[X] \leq \mathbb{E}[X^2] \leq (b - a)\mathbb{E}[X] + ab \leq ab$. □

And we obtain the similar estimate

$$\mathbb{P}[X_m - X_0 \geq d] \leq \exp\left(-\frac{d^2}{2\eta(mM + d)}\right).$$

Lemma 2.7. *Let X_i be a supermartingale such that $-M \leq \Delta X \leq \eta$ a.s. for all i, for some $\eta > M$. Then for any $d < mM$ we have*

$$\mathbb{P}(X_m - X_0 > d) \leq \exp\left(-\frac{d^2}{4m\eta M}\right).$$

If the sequence X_i is not a supermartingale, but we can bound its expectation and the absolute value of its increments, we can create a auxiliar supermartingale and apply the above estimates as follows. Suppose X_i is a sequence of random variables such that $\mathbb{E}[\Delta X_i|\mathcal{F}_{i-1}] \leq b$ and $|\Delta X_i - b| \leq c_i$ for every i. Then, $Y_i = X_i - X_0 - ib$ is a supermartingale and we can define the stopping time T as being the first i for which $Y_i \geq \alpha$. Then, by Hoeffding-Azuma inequality applied to $Y_{i\wedge T}$, we obtain that

Lemma 2.8. *Let $\mathcal{F}_0, \mathcal{F}_1, \ldots$ be a filtration and $X_0, X_1, \ldots$ random variables with X_i measurable with respect to $\mathcal{F}_i$, $0 \leq i \leq t$. Suppose that for some real b and constants $c_i > 0$,*

$$\mathbb{E}(X_i - X_{i-1}|\mathcal{F}_{i-1}) < b \text{ and}$$
$$|X_i - X_{i-1} - b| \leq c_i \text{ a.s. for all } 1 \leq i \leq t.$$

Then for all $\alpha > 0$,

$$\mathbb{P}(\exists i\,(0 \leq i \leq t) : X_i - X_0 \geq ib + \alpha) \leq \exp\left(-\frac{\alpha^2}{2\sum c_j^2}\right).$$

3
The differential equations method

In this chapter, we will introduce the differential equations method by first explaining the purpose of it and introducing a general setting where we can apply it. Next, we give a number of applications of the method to various random processes. The main goal here is to make the reader familiar with the method that will be used in a more intricate setting in the next chapters. Our main reference is Wormald's lecture notes ([4]).

The differential equations method is a method to prove that some random variables associated with a certain random process stay close to a solution of a system of differential equations calculated based on the expected changes in the steps of the process. This theory applies to a wide variety of random processes, in particular in the study of randomized algorithms and random graphs. The idea is that the solutions of the associated differential equations offer a deterministic good approximation to the trajectories of the random variables.

Formally, we consider a sequence of random variables $Y(t)$ indexed by time. In all applications, the proof is separated into two parts: the first in which we consider the expected changes between $Y(t)$ and $Y(t+1)$ to obtain a differential equation whose solution will be the expected path followed by Y; in the second part, we make our guess precise by using concentration inequalities and showing the desired convergence of the variable.

3.1
The general setting

Consider a sequence of discrete time random processes indexed by n. For simplicity, we omit the dependency on n from the notation. For each n, we study a sequence of discrete time random variables $Y_1(t), \dots, Y_a(t)$ associated with a random process. We write $\mathcal{F}_t$ for the corresponding natural filtration. That is, $\mathcal{F}_t = \sigma(\{Y_i(s) : 1 \leq i \leq a, 0 \leq s \leq t\})$ for all $t \geq 0$.

For many applications, it is better to consider scaled variables and time, as it will give a single differential equation rather than different equations for each n. In the main theorem below we scale by a factor of n, which will be sufficient for most applications. In some cases, when necessary, we may prescale variables before applying the theorem.

Associated with the process we have random variables $Y_1, \ldots, Y_a$ and we consider the stopping time T_D to be the minimum t such that $(t/n, Y_1(t)/n, \ldots, Y_a(t)/n) \notin D$, where $D \subset \mathbb{R}^{a+1}$ is an appropriate bounded connected open set. We think of D as being the neighborhood of the predicted trajectories. So that T_D is the first time we leave the neighborhood of the trajectories.

A theorem using the differential equations method can be stated with great generality but for our purposes, we will prove a simplified version and discuss how it can be extended to other cases later in this section.

3.1.1
The main result: Wormald's theorem

Theorem 3.1 (Wormald)**.** *Consider a sequence of discrete time random processes indexed by n and variables Y_ℓ associated with the processes (a sequence of variables indexed by n for each ℓ), for $1 \leq \ell \leq a$, where a is fixed, and functions $f_\ell : \mathbb{R}^{a+1} \to \mathbb{R}$, such that for some constant C_0 and all ℓ, we have*

$$|Y_\ell| < C_0 n \text{ for all } \ell \text{ and all } n \text{ whenever during the process.}$$

Assume the following three conditions hold, where, in (ii) and (iii), D is some bounded connected open set containing the closure of

$$\{(0, z_1, \ldots, z_a) : \mathbb{P}(Y_\ell(0) = z_\ell n, 1 \leq \ell \leq a) \neq 0 \text{ for some } n\}.$$

(i) (Boundedness hypothesis.) For some functions $\beta = \beta(n) \geq 1$ and $\gamma = \gamma(n)$, the probability that

$$\max_{1 \leq \ell \leq a} |Y_\ell(t+1) - Y_\ell(t)| \leq \beta,$$

conditional upon $\mathcal{F}_t$, is at least $1 - \gamma$ for $t < T_D$.

(ii) (Trend hypothesis.) For some function $\lambda_1 = \lambda_1(n) = o(1)$, for all $\ell \leq a$

$$\left| \mathbb{E}(Y_\ell(t+1) - Y_\ell(t) | \mathcal{F}_t) - f_\ell\left(\frac{t}{n}, \frac{Y_1(t)}{n}, \ldots, \frac{Y_a(t)}{n}\right) \right| \leq \lambda_1$$

for $t < T_D$.

(iii) (Lipschitz hypothesis.) Each function f_ℓ is continuous, and satisfies a

Lipschitz condition, on

$$D \cap \{(t, z_1, \ldots, z_a) : t \geq 0\},$$

with the same Lipschitz constant for each ℓ*.*

Then the following are true.

(a) For $(0, \hat{z}_1, \ldots, \hat{z}_a) \in D$ *the system of differential equations*

$$\frac{dz_\ell}{dx} = f_\ell(x, z_1, \ldots, z_a), \qquad \ell = 1, \ldots, a$$

has a unique solution in D *for* $z_\ell : \mathbb{R} \to \mathbb{R}$ *passing through*

$$z_\ell(0) = \hat{z}_\ell,$$

$1 \leq \ell \leq a$*, and which extends to points arbitrarily close to the boundary of D;*

(b) Let $\lambda > \lambda_1 + C_0 n\gamma$ *with* $\lambda = o(1)$*. For a sufficiently large constant* C*, with probability* $1 - O\left(n\gamma + \frac{\beta}{\lambda}\exp\left(-\frac{n\lambda^3}{\beta^3}\right)\right)$*,*

$$Y_\ell(t) = n \cdot z_\ell\left(\frac{t}{n}\right) + O(\lambda n)$$

uniformly for $0 \leq t \leq \sigma n$ *and for each* ℓ*, where* $z_\ell(x)$ *is the solution in (a) with* $\hat{z}_\ell = \frac{1}{n}Y_\ell(0)$*, and* $\sigma = \sigma(n)$ *is the supremum of those* x *to which the solution can be extended before reaching within* ℓ^∞*−distance* $C\lambda$ *of the boundary of* D*.*

Remark 3.1. *"Uniformly" in the statement of theorem refers to the fact that the implicit constant in the* $O(\lambda n)$ *term does not depend on* t*.*

Remark 3.2. *The theorem also holds when* D *depends on* n *but all Lipschitz constants are uniformly bounded.*

Remark 3.3. *The theorem also holds when "a" is a function of* n *changing the probability in (b) by*

$$1 - O\left(an\gamma + \frac{a\beta}{\lambda}\exp\left(-\frac{n\lambda^3}{\beta^3}\right)\right),$$

and if each function f_ℓ *depends only on* x *and* $z_1, \ldots, z_\ell$*.*

3.1.2
An application: The coupon collector

As an example of a simple application of the theorem, we will consider the following process. At each step one of n different labeled coupons is selected uniformly at random with repetition. It is well-known that the coupon collector needs, in average, to wait time $n \log n$ in order to complete the entire collection. Here we want to track the number of coupons left to be obtained throughout the process. Formally, this means that $Y(0) = n$ and, for each i,

$$Y(i+1) = \begin{cases} Y(i) - 1 & \text{, with probability } \frac{Y(i)}{n} \\ Y(i) & \text{, with probability } 1 - \frac{Y(i)}{n} \end{cases}$$

Then we have, for every t,

$$|Y(t)| \leq n,$$
$$|Y(t+1) - Y(t)| \leq 1 \text{ and}$$
$$\mathbb{E}[Y(t+1) - Y(t) | \mathcal{F}_t] = -\frac{Y(t)}{n}.$$

Thus, applying Theorem 3.1 with $C_0 = 1$, $\beta = 1$, $\gamma = 0$, $\lambda_1 = 0$ and the Lipschitz function $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = -y$, we have that the unique solution of

$$\begin{cases} \frac{dz}{dx} = f(x, z) = -z \\ z(0) = 1 \end{cases}$$

is given by $z(x) = \exp(-x)$ and then for every $\lambda = o(1) > 0$, with probability $1 - O\left(\frac{1}{\lambda} \exp\left(-n\lambda^3\right)\right)$,

$$Y(t) = n \cdot \exp\left(-\frac{t}{n}\right) + O(\lambda n).$$

The choice of λ can be made either to make the probability smaller or to have a smaller error term. For example, setting $\lambda = n^{-\frac{1}{3}+\varepsilon}$ we get

$$Y(t) = n \cdot \exp\left(-\frac{t}{n}\right) + O\left(n^{2/3+\varepsilon}\right)$$

$$\text{with probability } 1 - O\left(n^{1/3-\varepsilon} \cdot \exp(-n^{3\varepsilon})\right),$$

and letting $\lambda = \left(\frac{\log n}{n}\right)^{1/3}$ we have

$$Y(t) = n \cdot \exp\left(-\frac{t}{n}\right) + O\left(n^{2/3} \cdot (\log n)^{1/3}\right)$$

$$\text{with probability } 1 - O\left(n^{-2/3} \cdot (\log n)^{-1/3}\right).$$

Remark 3.4. *The values of t for which our approximations are still valid depends on the set D. In this case the function f is everywhere Lipschitz, we can choose $D \supset [0, A] \times [0, 1]$ and then the approximation holds for $t < An$ and it is important to note that the implicit constants on the $O(\cdot)$ terms depend on A.*

3.1.3
Proof of Wormald's theorem: a piecewise approach

Proof in special case $a = 1$ and $\gamma = 0$.

The first part of the theorem is well-known from the theory of differential equations (see [20]). For the second part we will write Y, z and f for Y_1, z_1 and f_1, respectively. Define

$$\omega = \left\lceil \frac{n\lambda}{\beta} \right\rceil$$

and note we can assume $\beta/\lambda < n^{1/3}$ and $\lambda < 1$ (otherwise our conclusion is trivially obtained). From now on we assume $\omega > n^{2/3}$. In order to show that the trajectory is followed almost surely, we will prove concentration of $Y(t+\omega) - Y(t)$.

For now consider that $\left(\frac{t}{n}, \frac{Y(t)}{n}\right)$ is ℓ^∞-distance at least $C\lambda$ from the boundary of D for some large constant C that will depend only on the Lipschitz constant of f. [1]

By the trend hypothesis we have

$$\mathbb{E}[Y(t+k+1) - Y(t+k)|\mathcal{F}_{t+k}] = f\left(\frac{t+k}{n}, \frac{Y(t+k)}{n}\right) + O(\lambda_1).$$

Then, as for $0 \le k < \omega$ we have $|Y(t+k) - Y(t)| \le k\beta$ and $\frac{k\beta}{n} = O(\lambda)$, by the Lipschitz hypothesis

$$\mathbb{E}[Y(t+k+1) - Y(t+k)|\mathcal{F}_{t+k}] = f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) + O(\lambda)$$

and there is a function $g(n) = O(\lambda)$ such that (conditioned on $\mathcal{F}_t$)

$$Y(t+k) - Y(t) - k \cdot f\left(\frac{t}{n}, \frac{Y(t)}{n}\right) - k \cdot g(n) \text{ is a supermartingale in } k.$$ [2]

[1] Observe then that throughout the proof we will use assumptions (i) and (ii) for $t + \omega < T_D$, since $\omega\beta/n = O(\lambda)$.

[2] Notice we can choose g such that $-Y(t+k) + Y(t) + kf\left(\frac{t}{n}, \frac{Y(t)}{n}\right) - kg(n)$ is also a supermartingale.

The step changes in these supermartingales are at most

$$|Y(t+k+1)-Y(t+k)|+\left|f\left(\frac{t}{n},\frac{Y(t)}{n}\right)\right|+|g(n)|\leq\beta+(\beta+\lambda_1)+O(\lambda)\leq\kappa\beta$$

for some $\kappa>0$ as $\beta\geq 1$.

Applying the Hoeffding-Azuma inequality,[3] we obtain for $\alpha=\frac{n\lambda^3}{\beta^3}$

$$\mathbb{P}\left(\left|Y(t+\omega)-Y(t)-\omega f\left(\frac{t}{n},\frac{Y(t)}{n}\right)\right|\geq\omega g(n)+\kappa\beta\sqrt{2\omega\alpha}\,\middle|\,\mathcal{F}_t\right)\leq 2\exp(-\alpha). \quad (3.1)$$

Now we want to prove by induction the following claim.

Claim 3.2. *Defining $k_i=i\omega$ for $i=0,1,\ldots,i_0$, with $i_0=\left\lfloor\frac{\sigma n}{\omega}\right\rfloor$. Then for each i,*

$$\mathbb{P}\left(\left|Y(k_j)-z\left(\frac{k_j}{n}\right)\cdot n\right|\geq B_j \text{ for some } j\leq i\right)=O(ie^{-\alpha})$$

where $B_j=(\lambda n+\omega)\left[\left(1+\frac{B\omega}{n}\right)^j-1\right]$.

The base case follows from the fact that $z(0)=\frac{Y(0)}{n}$. For the induction step, note that

$$\left|Y(k_{i+1})-z\left(\frac{k_{i+1}}{n}\right)\cdot n\right|=|A_1+A_2+A_3+A_4|$$

where

$$A_1=Y(k_i)-z\left(\frac{k_i}{n}\right)\cdot n,$$
$$A_2=Y(k_{i+1})-Y(k_i)-\omega\cdot f\left(\frac{k_i}{n},\frac{Y(k_i)}{n}\right),$$
$$A_3=\omega z'\left(\frac{k_i}{n}\right)+z\left(\frac{k_i}{n}\right)\cdot n-z\left(\frac{k_{i+1}}{n}\right)\cdot n,$$
$$A_4=\omega\cdot f\left(\frac{k_i}{n},\frac{Y(k_i)}{n}\right)-\omega\cdot z'\left(\frac{k_i}{n}\right).$$

Part I: By the induction hypothesis, we have that $\left|Y(k_j)-z\left(\frac{k_j}{n}\right)\cdot n\right|<B_j$ for all $j<i$ and

$$|A_1|<B_i \text{ with probability } 1-O(ie^{-\alpha}). \quad (3.2)$$

[3]For this supermartingale and the supermartingale in the previous footnote.

From now on we will suppose that these events happen.

Part II: By (3.1) with $t = k_i$, since $g(n) = O(\lambda)$ and $\beta\sqrt{\omega\alpha} = \omega\lambda$[4] we have, for some constant B' that does not depend on i,

$$|A_2| < B'\omega\lambda \text{ with probability } 1 - O(\exp(-\alpha)). \tag{3.3}$$

Part III: As z is differentiable, by the Taylor formula we have

$$z\left(\frac{k_{i+1}}{n}\right) = z\left(\frac{k_i}{n}\right) + \frac{\omega}{n}z'\left(\frac{k_i}{n}\right) + O\left(\frac{\omega^2}{n^2}\right)$$

and then, for some constant B'' that does not depend on i,

$$|A_3| \leq \frac{B''\omega^2}{n}. \tag{3.4}$$

Part IV: As z is the solution given in (a), we have $z'\left(\frac{k_i}{n}\right) = f\left(\frac{k_i}{n}, z\left(\frac{k_i}{n}\right)\right)$. Since f is Lipschitz, conditioned to (3.2), we conclude (for some constant $B''' > 0$)

$$\begin{aligned}|A_4| = \omega\left|f\left(\frac{k_i}{n}, \frac{Y(k_i)}{n}\right) - f\left(\frac{k_i}{n}, z\left(\frac{k_i}{n}\right)\right)\right| \\ \leq B'''\omega\left|\frac{Y(k_i)}{n} - z\left(\frac{k_i}{n}\right)\right| < \frac{B'''\omega B_i}{n}.\end{aligned} \tag{3.5}$$

Set $B = \max\{B', B'', B'''\}$. Then as

$$B_i\left(1 + \frac{B\omega}{n}\right) + \frac{B\omega}{n}(\lambda n + \omega) = B_{i+1},$$

by (3.2)-(3.5), with probability $1 - O((i+1)e^{-\alpha})$, $\left|Y(k_j) - z\left(\frac{k_j}{n}\right)\cdot n\right| < B_j$ for all $j \leq i+1$ and we are done with the induction.

To complete the proof, for $t \leq \sigma n$, we put $i = \left\lfloor\frac{t}{\omega}\right\rfloor$. Note that $B_i = O(n\lambda + \omega) = O(n\lambda)$[5], the change from k_i to t in Y is at most $\beta|t - k_i| \leq \omega\beta = O(\lambda n)$ and the change in z is at most $\omega(\beta + \lambda_1) = O(\lambda n)$. Thus with probability $1 - O\left(\frac{n}{\omega}\cdot e^{-\alpha}\right)$ we have

$$\left|Y(t) - z\left(\frac{t}{n}\right)\cdot n\right| = O(\lambda n).$$

□

Now we discuss how to deduce the full version of Theorem 3.1 from the

[4] Actually we have $\sqrt{\left\lceil\frac{n\lambda}{\beta}\right\rceil\left(\frac{n\lambda}{\beta}\right)}$ and not ω, but we will only use the fact that this is $O(\omega)$.

[5] Since $\left(1 + \frac{B\omega}{n}\right)^{\sigma n/\omega} = O(1)$.

special case proved above. For an arbitrary $a > 1$ the inductive hypothesis of the Claim 3.2 is modified to

$$\mathbb{P}\left(\left|Y_\ell(k_j) - z_\ell\left(\frac{k_j}{n}\right)\cdot n\right| \geq B_j \text{ for some } j \leq i\right) = O(aie^{-\alpha})$$

for every $\ell \leq a$. For the induction step, the statement has to be verified for all a variables and then the failure probability is multiplied by a. Then the theorem follows.

For an arbitrary γ, we need to condition on the event that

$$\max_{1\leq\ell\leq a} |Y_\ell(t+1) - Y_\ell(t)| \leq \beta$$

holds at each step. This alters the expected change of Y but as $|Y| \leq C_0 n$ it will be changed by at most $C_0 n\gamma$. Then replacing λ_1 by $\lambda_1 + C_0 n\gamma$ it suffices to note that the probability of failure throughout the process is $O(n\gamma)$ and the theorem follows for arbitrary γ.

3.2 The bounded 2-degree process

Not all applications of the differential equations method are direct applications of Theorem 3.1. Remarkably, it is not always necessary to formally define (and solve) the underlying system of differential equations. In this section, we show that it is sometimes possible to *guess* the trajectories corresponding to the random variables $Y_i(t)$, which turns out to be equivalent to solving the system of differential equations. This "guess" is made using heuristic probabilistic arguments. We then prove concentration of the number of isolated vertices in the bounded 2-degree process (Theorem 3.3).

The bounded 2-degree process is the process to generate a n-vertex graph in which at each step an edge is chosen uniformly at random conditioned in the event that neither vertex has degree more than 2. In other words, you choose one edge among all (not already added) edges between vertices of degrees 0 or 1. Formally, we set G_0 to be the empty graph on the n vertices and G_{i+1} is the graph obtained from G_i by adding the edge e_{i+1}, taken uniformly at random from the set $\{uv \notin E(G_i) : d_{G_i}(u) < 2, d_{G_i}(v) < 2\}$, which we will call the "available edges".

Let $Y_j(t)$ be the number of vertices of degree j in G_t for $j = 0, 1$ or 2. Note that the number of vertices of the graph is $n = Y_0(t) + Y_1(t) + Y_2(t)$ and by double counting the number of edges in G_t we have $t = \frac{1}{2}(Y_1(t) + 2Y_2(t))$.

It follows that $n - t = Y_0(t) + \frac{1}{2}Y_1(t)$, so

$$Y_1(t) = 2n - 2t - 2Y_0(t) \tag{3.6}$$
$$\text{and } Y_2(t) = -n + 2t + Y_0(t). \tag{3.7}$$

Equations (3.6) and (3.7) tell us that in order to investigate the trajectory of the number of vertices of every given degree during the process it suffices to study the number of isolated vertices.

3.2.1 Isolated vertices in the bounded 2-degree process

Let $A(t)$ the number of available edges in G_t (i.e., at time t). Note that

$$A(t) = \binom{n - Y_2(t)}{2} - F(t),$$

where $F(t)$ is the number of edges already present between vertices of degree less than 2. By the handshake lemma, the number of such edges can be at most n and, using (3.7), we can write

$$A(t) = \frac{1}{2}(n - Y_2(t))^2 + O(n) = \frac{1}{2}(2n - 2t - Y_0(t))^2 + O(n). \tag{3.8}$$

The probability of e_{t+1} to be any given available edge is $\frac{1}{A(t)}$. The number of isolated vertices $Y_0(t)$ can drop at each step by two or one, depending on whether the edge e_{t+1} added is between vertices that previously have both degree 0 or have degrees 0 and 1, respectively. It follows that the expected change in Y_0 is

$$\mathbb{E}[Y_0(t+1) - Y_0(t)|\mathcal{F}_t] = -2 \cdot \frac{\binom{Y_0(t)}{2}}{A(t)} - 1 \cdot \frac{Y_0(t)Y_1(t)}{A(t)}$$

Then, from (3.6)-(3.8), we can compute

$$\begin{aligned} E[Y_0(t+1) - Y_0(t)|\mathcal{F}_t] &= -2 \cdot \frac{\binom{Y_0(t)}{2}}{A(t)} - 1 \cdot \frac{Y_0(t)Y_1(t)}{A(t)} \\ &= \frac{-Y_0(t)(Y_0(t)-1) - Y_0(t)(2n - 2t - 2Y_0(t))}{\frac{1}{2}(2n - 2t - Y_0(t))^2 + O(n)} \\ &= \frac{-Y_0(t)(2n - 2t - Y_0(t) - 1)}{\frac{1}{2}(2n - 2t - Y_0(t))^2 + O(n)} \\ &= \frac{-2Y_0(t) + \frac{2Y_0(t)}{(2n-2t-Y_0(t))}}{(2n - 2t - Y_0(t)) + O\left(\frac{2n}{2n-2t-Y_0(t)}\right)}. \end{aligned}$$

Now, supposing $\frac{\sqrt{n}}{n-t} = o(1)$, we can use [6]

$$\frac{n}{(2n-2t-Y_0(t))^2} \leq \frac{n}{(n-t)^2} = o(1)$$

to obtain that

$$\frac{1}{(2n-2t-Y_0(t)) + O\left(\frac{2n}{2n-2t-Y_0(t)}\right)} = \frac{1+O\left(\frac{n}{(2n-2t-Y_0)^2}\right)}{2n-2t-Y_0(t)} \leq \frac{1+O\left(\frac{n}{(n-t)^2}\right)}{2n-2t-Y_0(t)}$$

and noticing that

$$0 \leq \frac{Y_0(t)}{2n-2t-Y_0(t)} \leq \frac{n-t}{2n-2t-(n-t)} = 1$$

$$\text{and } \frac{2Y_0(t)}{(2n-2t-Y_0(t))^2} \leq \frac{2(n-t)}{(2n-2t-(n-t))^2} = O\left(\frac{n}{(n-t)^2}\right),$$

we finally conclude

$$\begin{aligned}
E[Y_0(t+1)-Y_0(t)|\mathcal{F}_t] =& \frac{-2Y_0(t) + \frac{2Y_0(t)}{(2n-2t-Y_0(t))}}{(2n-2t-Y_0(t)) + O\left(\frac{2n}{2n-2t-Y_0(t)}\right)} \\
=& \frac{-2Y_0(t) + \frac{2Y_0(t)}{(2n-2t-Y_0(t))}}{2n-2t-Y_0(t)} \cdot \left(1 + O\left(\frac{n}{(n-t)^2}\right)\right) \\
=& \frac{-2Y_0}{2n-2t-Y_0} + O\left(\frac{n}{(n-t)^2}\right) \\
=& \frac{-2Y_0(t)}{2n-2t-Y_0(t)} + o(1).
\end{aligned}$$

Which suggests that $y_0(x) \approx \frac{1}{n}Y_0(xn)$ satisfies the differential equation

$$y_0'(x) = \frac{-2y_0(x)}{2-2x-y_0(x)}, \tag{3.9}$$

with general solution

$$y_0(x)(C - \log y_0(x)) = 2 - 2x.$$

As $y_0(0) = \frac{Y_0(0)}{n} = 1$ we have $C = 2$ and conclude

$$y_0(x)(2 - \log y_0(x)) = 2 - 2x. \tag{3.10}$$

[6]Here we are using the bound $Y_0(t) \leq n - t$, which can be inferred from (3.6).

3.2.2
Probabilistic intuition - How to guess the trajectories

Instead of finding the differential equation (3.9), we could derive (3.10) directly by estimating the probabilities that a vertex has degree 0, 1 or 2 during the process. Although by proceeding this way we don't arrive at a differential equation we will still call it the differential equations method since it is only an alternative to find the guessed trajectory of our random variable when we do not want to (or cannot) compute the expected changes during the process.

We write y_i for the continuous approximation of Y_i. From (3.8), we know that

$$A(t) \sim \frac{1}{2}n^2\alpha(t)^2, \text{ where } \alpha(t) = 1 - y_2(t).$$

The probability that a given vertex u (with degree < 2 at step t) is incident to e_{t+1} is the number of available edges incident to u over the total number of available edges $A(t)$ and the number of available edges incident to u is expected to be close to $n - 1 - Y_2(t) \sim n\alpha(t)$ (neglecting edges already present between vertices of degree less than 2). So

$$\mathbb{P}(u \text{ incident to } e_{t+1}) \sim \frac{n\alpha(t)}{A(t)} \sim \frac{2}{n\alpha(t)} \text{ and}$$
$$\mathbb{P}(u \text{ not incident to } e_{t+1}) \sim 1 - \frac{2}{n\alpha(t)} \sim \exp\left(-\frac{2}{n\alpha(t)}\right).$$

Letting $\lambda(s) \sim \sum_{t=0}^{s-1} \frac{2}{n\alpha(t)}$. Then the probability that the vertex u remains isolated after s steps should be close to

$$\prod_{t=0}^{s-1} \exp\left(-\frac{2}{n\alpha(t)}\right) \sim \exp\left(-\sum_{t=0}^{s-1}\frac{2}{n\alpha(t)}\right) \sim \exp(-\lambda(s)).$$

And the probability that it has degree 1 should be

$$\left(\prod_{t=0}^{s-1} \exp\left(-\frac{2}{n\alpha(t)}\right)\right) \sum_{t=0}^{s-1} \frac{\frac{2}{n\alpha(t)}}{\exp\left(-\frac{2}{n\alpha(t)}\right)} \sim \lambda(s)\exp(-\lambda(s)).$$

We deduce that the functions y_0 and y_1 should be

$$y_0(x) = e^{-\lambda(x)}$$
$$y_1(x) = \lambda(x)e^{-\lambda(x)},$$

where from (3.6) we obtain

$$(2 + \lambda(x))e^{-\lambda(x)} = 2 - 2x,$$

as in (3.10), which implicitly determine y_0.

3.2.3 Wholistic approach

In the proof of the main theorem, we introduced a large enough variable ω and used concentration of $Y(t+\omega) - Y(t)$ to show that the variable Y remains close to its predicted trajectory for multiples of ω. However, we are not required to use this "piecewise" method. As we know that $Y(t)$ should be close to $n \cdot y\left(\frac{t}{n}\right)$, we expect that $Y(t) - n \cdot y\left(\frac{t}{n}\right)$ should be close to zero and summing or subtracting a function $f(t)$ we would obtain a supermartingale or submartingale.

This "wholistic" approach can be used when the Lipschitz condition fails or when we know something more about the solution of the differential equation. In some cases, we may obtain better approximations for larger values of t. Formally, the idea of summing or subtracting a function will be made essentially by applying Lemma 2.8 (see Chapter 2).

Now we are able to state and prove the desired conclusion about the number of isolated vertices on the bounded 2-degree process.

Theorem 3.3. *Take $0 < \delta < \min\left\{3\varepsilon, \frac{1}{6} + \frac{\varepsilon}{2}\right\}$. With probability at least $1 - \exp\left(-\Omega\left(n^{1/3+\varepsilon-2\delta}\right)\right)$,*

$$n - t = Y_0(t)\left(1 + O\left(n^{-\delta}\right) + \frac{1}{2}\log\frac{n}{Y_0(t)}\right)$$

for all $0 \leq t < \lfloor n - n^{2/3+\varepsilon} \rfloor$.

Proof. As we know that $Y_0(t)$ should follow $n \cdot y_0\left(\frac{t}{n}\right)$, then remembering (3.10) we have

$$y_0\left(\frac{t}{n}\right) \cdot \left(2 - \log\left(y_0\left(\frac{t}{n}\right)\right)\right) = 2 - \frac{2t}{n},$$

and the solution of the differential equation for Y_0 is

$$\frac{Y_0(t)}{n} \cdot \left(2 - \log\left(\frac{Y_0(t)}{n}\right)\right) = \frac{2(n-t)}{n},$$

or, rearranging the terms,

$$2 = \frac{2(n-t)}{Y_0(t)} + \log\left(\frac{Y_0(t)}{n}\right).$$

Letting $F(x,y) = \frac{2(n-x)}{y} + \log\left(\frac{y}{n}\right)$ and $v_t = (t, Y_0(t))$. Then the solution above is $F(v_t) = 2$. The idea of the proof will be to show that $F(v_t)$ remains close to 2 with high probability. But since $F(v_t)$ is not defined if $Y_0(t) = 0$, we need to define a stopping time when Y_0 comes near zero. This can be done by setting

$$T = \min\{t : |F(v_t) - 2| \geq n^{-\delta}\}.$$

If $\frac{n-t}{Y_0(t)} \geq \log n$ then

$$\frac{2(n-t)}{Y_0} + \log\frac{Y_0}{n} - 2 \geq \log n + \log Y_0 - 2 \geq n^{-\delta}$$

and $|F(v_t) - 2| \geq n^{-\delta}$. This means we can assume $Y_0(t) > \frac{n-t}{\log n}$ for $t < T$.

As grad $F = \left(-\frac{2}{y}, -\frac{2(n-x)}{y^2} + \frac{1}{y}\right)$ and the second derivatives are all $O\left(\frac{1}{y^2}\right)$, we obtain

$$F(v_{t+1}) - F(v_t) = (v_{t+1} - v_t) \cdot \text{grad } F(v_t) + O\left(\frac{1}{Y_0^2}\right) \tag{3.11}$$

Now we just need to note, by our previous calculations,

$$\begin{aligned}\mathbb{E}[v_{t+1} - v_t|\mathcal{F}_t] &= (1, \mathbb{E}[Y_0(t+1) - Y_0(t)|\mathcal{F}_t]) \\ &= \left(1, \frac{-2Y_0}{2n-2t-Y_0} + O\left(\frac{n}{(n-t)^2}\right)\right)\end{aligned}$$

and $\left(1, \frac{-2Y_0}{2n-2t-Y_0}\right) \cdot \text{grad } F(v_t) = 0$[7]. Then

$$\begin{aligned}\mathbb{E}[F(v_{t+1}) - F(v_t)|\mathcal{F}_t] &= O\left(\frac{(2n-2t-Y_0)n}{Y_0^2(n-t)^2}\right) + O\left(\frac{1}{Y_0^2}\right) \\ &= O\left(\frac{n}{Y_0^2(n-t)}\right).\end{aligned}$$

Now, using that $Y_0 > \frac{n-t}{\log n}$,

$$\begin{aligned}\mathbb{E}[F(v_{t+1}) - F(v_t)|\mathcal{F}_t] &= O\left(\frac{n}{Y_0^2(n-t)}\right) \\ &= O\left(\frac{n(\log n)^2}{(n-t)^3}\right) \quad \text{for } t \leq n - n^{1/2+\varepsilon}.\end{aligned} \tag{3.12}$$

As $|Y_0(t+1) - Y_0(t)| \leq 2$ and by (3.11) we also have

$$\begin{aligned}|F(v_{t+1}) - F(v_t)| &\leq \frac{2}{Y_0} + 2 \cdot \frac{2n-2t-Y_0}{Y_0^2} + O\left(\frac{1}{Y_0^2}\right) \\ &= O\left(\frac{n-t}{Y_0^2}\right) = O\left(\frac{(\log n)^2}{n-t}\right)\end{aligned} \tag{3.13}$$

[7]Since grad $F(v_t) = \left(-\frac{2}{Y_0(t)}, -\frac{2n-2t-Y_0(t)}{Y_0(t)^2}\right)$.

Then, by (3.12) and (3.13) , applying Lemma 2.8 with

$$b = O\left(\frac{(\log n)^2}{n^{1+3\varepsilon}}\right),$$
$$c_t = O\left(\frac{(\log n)^2}{n-t}\right),$$
$$t \le t_0 = n - n^{2/3+\varepsilon}$$
$$\text{and } \alpha = \frac{1}{2}n^{-\delta},$$

noticing

$$\frac{n(\log n)^2}{(n-t)^3} \le \frac{(\log n)^2}{n^{1+3\varepsilon}} \text{ for } t < n - n^{2/3+\varepsilon},$$
$$ib + \alpha \le (n - n^{2/3+\varepsilon})O\left(\frac{(\log n)^2}{n^{1+3\varepsilon}}\right) + \frac{1}{2}n^{-\delta}$$
$$= O\left(\frac{(\log n)^2}{n^{3\varepsilon}}\right) + \frac{1}{2}n^{-\delta} < \frac{2}{3}n^{-\delta} \text{ (since } \delta < 3\varepsilon)$$
$$\text{and } \sum c_j^2 \le O((\log n)^4)(n - n^{2/3+\varepsilon})\frac{1}{n^{4/3+2\varepsilon}}$$
$$\le O((\log n)^4 n^{-1/3-2\varepsilon}) = O\left(n^{-1/3-\varepsilon}\right),$$

we obtain

$$\mathbb{P}\left(\exists i\,(0 \le i \le t_0) : |F(v_i) - F(v_0)| \ge \frac{2}{3}n^{-\delta}\right) \le \exp\left(-\Omega\left(n^{1/3+\varepsilon-2\delta}\right)\right).$$

As $F(v_0) = 2$, this implies that $T < t_0$ with the same low probability and $T \ge t_0$ implies $|F(v_i) - F(v_0)| \le n^{-\delta}$ for $i < t_0$ as well, the result follows (changing the implicit constant in the $\Omega(\cdot)$ term if needed). □

4
The random greedy independent set algorithm in dense hypergraphs and the H-free process for strictly balanced graphs

In this chapter, we will study the H-free process for strictly balanced graphs H. First, we introduce the H-free process for general H and describe an auxiliary hypergraph $\mathcal{H}_H$. Next, we see that the H-free process is analogous to the random greedy independent set algorithm on $\mathcal{H}_H$. We then state the main theorem, obtained by Bennett and Bohman in [6], which asserts that with high probability this algorithm produces a large independent set provided that the hypergraph satisfies certain degree conditions. Finally, we conclude by showing that this result, when applied to $\mathcal{H}_H$, generalizes a lower bound, obtained by Bohman and Keevash in [15], on the number of steps in the H-free process for strictly balanced graphs. The proof of the main theorem, which uses the differential equations method to track several variables throughout the process, is presented in Chapter 6.

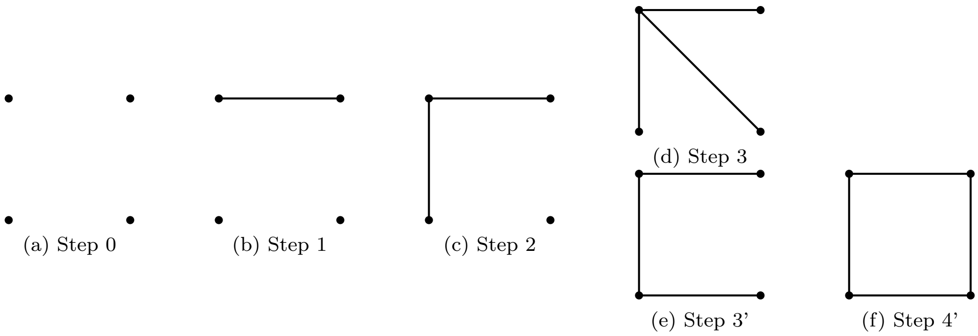
4.1
The H-free process

4.1.1
The process

Let H be a fixed graph. We call a graph G with n vertices a maximal H-free graph (also called H*-saturated* graph) if it has no copy of H as a subgraph and the addition of any new edge to G would create such a copy. The H-free process is a random process to obtain a maximal H-free graph as follows.

The process generates a nested sequence $G_0, \dots, G_M$ of graphs with the same vertex set. We begin with the empty graph G_0 with n vertices. At each step, we add uniformly at random a new edge (among all edges that would not create a copy of H) to the graph G_i to obtain G_{i+1}. The process stops when we arrive at a maximal H-free graph G_M.

As a random H-saturated graph, G_M has some interesting properties and we are particularly interested in the behavior of the random variable M.



(a) Step 0 (b) Step 1 (c) Step 2 (d) Step 3 (e) Step 3' (f) Step 4'

Figure 4.1: The K_3-free process with 4 vertices. Notice that the process can stop with a different number of edges depending on the chosen edges.

4.1.2
The underlying hypergraph

Associated with the H-free process, we can consider a hypergraph $\mathcal{H}_H$ with vertex-set the edges of the complete graph K_n, i.e. $V = \binom{[n]}{2}$ and the edge set being all copies of H in K_n.

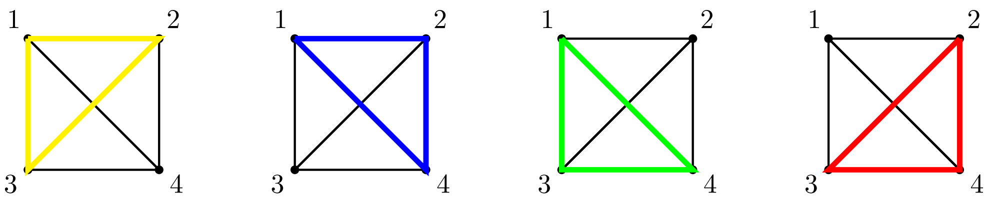


Figure 4.2: The copies of K_3 in K_4.

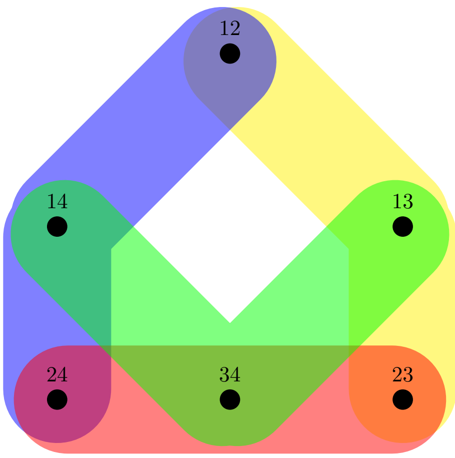


Figure 4.3: $\mathcal{H}_{K_3}$ for $n = 4$.
The vertex ij of $\mathcal{H}_{K_3}$ represents the edge connecting i and j on K_4.
The edges of $\mathcal{H}_{K_3}$ correspond to the triangles of K_4.

In this context, choosing a new edge at random that does not create a copy of H is analogous to choosing a new vertex at random that does not form an edge in the hypergraph (i.e., a vertex that forms an independent set with the previously chosen vertices). Then, the H-free process can be generalized in the hypergraph setting by the following algorithm.

4.2 The random greedy independent set algorithm

4.2.1 The algorithm

The random greedy independent set algorithm is the algorithm that forms an independent set by choosing one vertex at a time such that no edge is entirely chosen. In other words, at each step, we take uniformly at random a vertex that does not create an edge together with the already chosen vertices.

Formally we have at the beginning $\mathcal{H}(0) = \mathcal{H}$, $V(0) = V$ and $I(0) = \emptyset$. Given $i \geq 0$, an independent set $I(i)$ and a hypergraph $\mathcal{H}(i)$ on the vertex-set $V(i)$, we choose uniformly at random a vertex $v_{i+1} \in V(i)$. Then we take

- $I(i+1) := I(i) \cup \{v_{i+1}\}$, which will be an independent set of $\mathcal{H}$ with $i+1$ vertices;
- $V(i+1)$, the new vertex-set, is $V(i)$ without v_{i+1} and every other vertex u such that $\{u, v_{i+1}\}$ is an edge of $\mathcal{H}(i)$;
- $\mathcal{H}(i+1)$, the new hypergraph, is $\mathcal{H}(i)$ removing v_{i+1} from every edge that contains v_{i+1} and at least 2 other vertices and removing every edge incident to vertex (different from v_{i+1}) removed from $V(i)$.

Note that $\mathcal{H}(i)$ is no longer an uniform hypergraph.

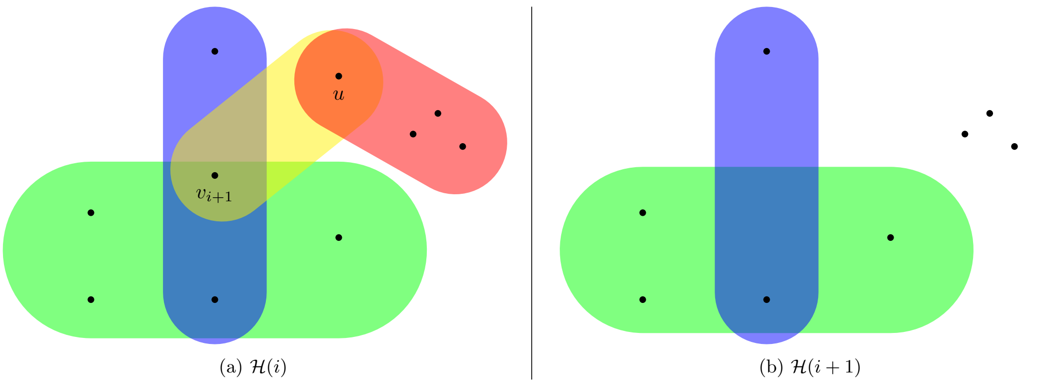


Figure 4.4: When the vertex v_{i+1} is chosen, notice that the vertex u is deleted and then the red edge that contained u is also deleted.

Remark 4.1. If we arrive at a hypergraph $\mathcal{H}(i)$ that has edges e, e' such that $e \subset e'$ we can remove e' from $\mathcal{H}$ without causing any differences in the process, as the presence of e ensures that we will never have $e' \subset I(j)$. As a convention, we will then remove all such e' edges in the study of the process.

4.2.2
The main result: Bennett-Bohman's theorem

For a subset $A \subset V$ we define the degree of A in $\mathcal{H}$, denoted $\deg(A)$, as the number of edges in $\mathcal{H}$ that contain A. We highlight that by $\deg(A)$ we denote the degree of A in the initial hypergraph $\mathcal{H}$.

Definition 4.1 ($\Delta_\ell(\mathcal{H})$)**.** The ℓ-maximum degree $\Delta_\ell(\mathcal{H})$ of a hypergraph $\mathcal{H}$ with vertex-set V is the maximum degree of A over all $A \in \binom{V}{\ell}$.

Definition 4.2 ($\Gamma(\mathcal{H})$)**.** The $(r-1)$-codegree of v and v' (where $v, v' \in V$ and $v \neq v'$) is the number of pairs of edges (e, e') such that $v \in e \backslash e'$, $v' \in e' \backslash e$ and $|e \cap e'| = r - 1$.

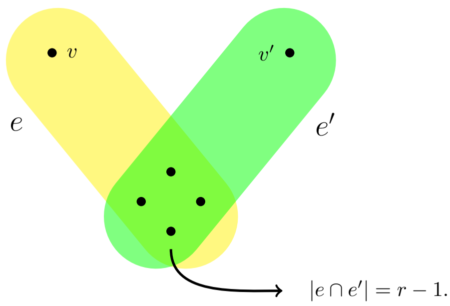


Figure 4.5: The $(r-1)$-codegree of v and v' is the number of pairs (e, e') as above.

Then $\Gamma(\mathcal{H})$ is the maximum $(r-1)$-codegree over all choices of $v, v' \in V$.

Theorem 4.1 (Bennett-Bohman [6]; proof in Chapter 6)**.** *Let r and $\varepsilon > 0$ be fixed. Let $\mathcal{H}$ be a r-uniform, D-regular hypergraph on N vertices such that $D > N^\varepsilon$. If*

$$\Delta_\ell(\mathcal{H}) < D^{\frac{r-\ell}{r-1}-\varepsilon} \qquad \textit{for } \ell = 2, \ldots, r-1 \tag{4.1}$$

and $\Gamma(\mathcal{H}) < D^{1-\varepsilon}$ then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ with

$$|I| = \Omega\left(N \cdot \left(\frac{\log N}{D}\right)^{\frac{1}{r-1}}\right) \tag{4.2}$$

with probability $1 - \exp\{-N^{\Omega(1)}\}$.

4.3
Strictly balanced graphs

Definition 4.3 (Strictly balanced graph)**.** We call a graph $H = (V_H, E_H)$ *strictly balanced* if it has at least 3 vertices, 3 edges and

$$\frac{e_{H[W]} - 1}{|W| - 2} < \frac{e_H - 1}{v_H - 2} \quad \text{for all } W \subsetneq V_H \text{ such that } |W| \geq 3, \tag{4.3}$$

where $v_H = |V_H|$ is the number of vertices in H, $e_H = |E_H|$ is the number of edges in H and $e_{H[W]}$ is the number of edges in the subgraph $H[W]$ of H induced by the subset W.

We will see that the hypothesis (4.1) is valid for $\mathcal{H}_H$ if and only if H is strictly balanced. Hence, Theorem 4.1 is a generalization of the following result due to Bohman and Keevash.

Theorem 4.2 (Bohman-Keevash [13]). *Let H be a strictly balanced graph and G be the maximal H-free graph on n vertices obtained by the H-free process. Then, with high probability, the number of edges of G is at least*

$$\Omega\left(n^{2-\frac{v_H-2}{e_H-1}}(\log n)^{\frac{1}{e_H-1}}\right).$$

First, let's see that eq.(4.1) holds for $\mathcal{H}_H$ if and only if H is strictly balanced. For $a \geq 2$, let v_a be the minimum number of vertices spanned by a edges of H. The degree of a set of a vertices in $\mathcal{H}_H$ is $\Theta(n^v)$ where v is the number of vertices not present in the correspondent a edges of H. So

$$\Delta_a(\mathcal{H}_H) = \Theta\left(n^{v_H - v_a}\right) = \Theta\left(n^{(v_H-2)\left[1-\frac{v_a-2}{v_H-2}\right]}\right).$$

Also, $\mathcal{H}_H$ is D-regular with $D = \Theta\left(n^{v_H-2}\right)$ and

$$D^{\frac{e_H-a}{e_H-1}} = \Theta\left(n^{\frac{(v_H-2)(e_H-a)}{e_H-1}}\right) = \Theta\left(n^{(v_H-2)\left[1-\frac{a-1}{e_H-1}\right]}\right).$$

Comparing above equations, we have that (4.1) is valid for $\mathcal{H}_H$ (for some $\varepsilon > 0$) if and only if $\frac{v_a-2}{v_H-2} > \frac{a-1}{e_H-1}$ for all $a \geq 2$, which holds if and only if (4.3) holds.

Now, to prove Theorem 4.2, let H be strictly balanced and take

$$\varepsilon < \min\left\{\frac{v_H-2}{2}, \frac{1}{v_H-2}, \frac{v_a-2}{v_H-2} - \frac{a-1}{e_H-1} \,:\, \text{for all } 2 \leq a < e_H\right\}.$$

The result will follow by applying Theorem 4.1 to $\mathcal{H}_H$. Notice that $\mathcal{H}_H$ has $N = \binom{n}{2}$ vertices, is e_H-uniform and D-regular with $D = \Theta\left(n^{v_H-2}\right)$. Since $\varepsilon < \frac{v_H-2}{2}$, we have $D > N^{\varepsilon}$.

As we discussed above, $\mathcal{H}_H$ satisfies (4.1) because H is strictly balanced. To be precise, since $\varepsilon < \frac{v_a-2}{v_H-2} - \frac{a-1}{e_H-1}$ for all $2 \leq a < e_H$,

$$\Delta_a(\mathcal{H}_H) = \Theta\left(n^{(v_H-2)\left[1-\frac{v_a-2}{v_H-2}\right]}\right) < \Theta\left(n^{(v_H-2)\left[1-\frac{a-1}{e_H-1}-\varepsilon\right]}\right) = D^{\frac{e_H-a}{e_H-1}-\varepsilon}.$$

As H is strictly balanced, it has no isolated vertices nor vertices with degree 1. Then two distinct copies of H with $e_H - 1$ edges in common can together span at most v_H vertices. This means that $\Gamma(\mathcal{H}_H) = O(n^{v_H - 3})$. And, since $\varepsilon < \frac{1}{v_H - 2}$ implies that $\frac{v_H - 3}{v_H - 2} = 1 - \frac{1}{v_H - 2} < 1 - \varepsilon$, we conclude that $\Gamma(\mathcal{H}_H) = O(n^{v_H - 3}) < \Theta\left(n^{(v_H - 2)(1-\varepsilon)}\right) = D^{1-\varepsilon}$.

As all hypotheses of Theorem 4.1 are satisfied, we have that the number of edges of the maximal H-free graph obtained in the H-free process is at least

$$\Omega\left(N \cdot \left(\frac{\log N}{D}\right)^{\frac{1}{e_H - 1}}\right) = \Omega\left(n^{2 - \frac{v_H - 2}{e_H - 1}} (\log n)^{\frac{1}{e_H - 1}}\right).$$

5
The random greedy independent set algorithm in linear hypergraphs and the H-free process on a host graph F

In this chapter, we will give a lower bound on the number of vertices obtained in the random greedy independent set algorithm by computing just some of the vertices added. This method covers the case of linear hypergraphs (i.e., hypergraphs $\mathcal{H}$ such that $\Delta_2(\mathcal{H}) = 1$). However, the bound is a logarithmic factor far from the right order of magnitude guessed by the differential equations heuristic. Then we introduce the H-free process where we have a host graph F instead of the complete graph K_n.

Our first theorem is a generalization of the lower bound on the number of edges in the the Q_2-free process on the hypercube Q_d given by J. Robert Johnson and Trevor Pinto in [17], which we obtain in Section 5.2.1 as a corollary. The proof of the general theorem (Theorem 5.1) is obtained by a direct adaptation of the proof by Johnson and Pinto.

5.1
A different approach - Counting good vertices

We will produce a random permutation of the vertices of the hypergraph $\mathcal{H}$ by considering independent random variables T_v, for each vertex v, distributed uniformly in $[0, 1]$. We think of the random algorithm as adding the vertices in order from the smaller variable to the greater. We call a vertex *good* if it is not the last vertex of any of the edges of $\mathcal{H}$. With this definition, a good vertex will always appear in the final independent set. To bound the number of edges obtained in the random greedy algorithm we will bound the number of good edges as below.

Theorem 5.1. *Let $r \geq 2$ be a fixed integer. Let $\mathcal{H} = (V, \mathcal{E})$ be a r-uniform hypergraph with N vertices such that $\Delta_2(\mathcal{H}) = 1$. Suppose that the r-degree of each vertex $v \in V$ is denoted by $d(v)$ and the minimum degree δ satisfies $\delta \to \infty$ as $N \to \infty$. Let Δ be the maximum of such degrees. Then, with probability $1 - O\left(\frac{\Delta \cdot \left(\sum_{v\in V} d(v)\right)}{\left(\sum_{v\in V} d(v)^{-\frac{1}{r-1}}\right)^2}\right)$, the number of good vertices in a uniformly* *distributed random permutation of the vertices of $\mathcal{H}$ is* $\Theta\left(\sum_{v\in V} d(v)^{-\frac{1}{r-1}}\right)$.

Before we prove the theorem above, we will state some lemmas about the Gamma function that we will need to use. The proofs of these lemmas can be found in Appendix A.

Definition 5.1 (The Gamma function)**.** *The gamma function is defined for complex numbers with positive real part by*

$$\Gamma(z) = \int_0^\infty u^{z-1}e^{-u}\, du.$$

Lemma 5.2. *For all complex numbers a, b with positive real part the following equation holds*

$$\int_0^1 (1-x^a)^b\, dx = \frac{\Gamma\left(1+\frac{1}{a}\right)\Gamma(b+1)}{\Gamma\left(b+1+\frac{1}{a}\right)}.$$

Lemma 5.3. *Let α be a positive real number. Then*

$$\lim_{n\to\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)\cdot n^\alpha} = 1.$$

Proof of Theorem 5.1. Let A_v be the indicator variable of v being good and $A = \sum_{v\in V} A_v$. Hence, conditioning on the event that $T_v = x$, the probability that v is the last vertex added to an edge is x^{r-1}. Also, as $\Delta_2(\mathcal{H}_H) = 1$, the events of v being the last vertex of different edges are independent since these events depend on variables T_u for disjoint sets of vertices u. Then

$$\begin{aligned}\mathbb{P}(A_v = 1) &= \int_0^1 (1-x^{r-1})^{d(v)}\, dx \\ &= \frac{\Gamma\left(1+\frac{1}{r-1}\right)\Gamma(d(v)+1)}{\Gamma\left(d(v)+1+\frac{1}{r-1}\right)} \\ &\sim \Gamma\left(1+\frac{1}{r-1}\right) d(v)^{-\frac{1}{r-1}},\end{aligned}$$

where we used Lemmas 5.2 and 5.3 as well as the fact that $d(v)\to\infty$.

By linearity of expectation,

$$\mathbb{E}[A] = \Theta\left(\sum_{v\in V} d(v)^{-\frac{1}{r-1}}\right).$$

As $\Delta_2(\mathcal{H}) = 1$, the event A_v depends only on variables T_u for at most $(r-1)d(v)$ values of u different from v. Then A_v is independent of all but at most $(r-1)\Delta \cdot (r-1)d(v)$ other A_f. Therefore

$$\mathrm{Var}(A) = \sum_{v\in V} \mathrm{Var}(A_v) + \sum_{u\neq v} \mathrm{Cov}(A_v, A_u) \leq \mathbb{E}[A] + (r-1)^2\Delta \cdot \sum_{v\in V} d(v).$$

By Chebyshev's inequality we have

$$\mathbb{P}\left(|A - \mathbb{E}[A]| \geq k\sqrt{\mathrm{Var}(A)}\right) \leq \frac{1}{k^2}$$

and $\left(\text{choosing } k\sqrt{\mathrm{Var}(A)} = \frac{\mathbb{E}[A]}{10} \text{ for example we have } \frac{1}{k^2} = \frac{100\,\mathrm{Var}(A)}{\mathbb{E}[A]^2}\right)$ we obtain that $A = \Theta\left(\sum_{v\in V} d(v)^{-\frac{1}{r-1}}\right)$ with probability $1 - O\left(\frac{\Delta\cdot\left(\sum_{v\in V} d(v)\right)}{\left(\sum_{v\in V} d(v)^{-\frac{1}{r-1}}\right)^2}\right)$. □

The previous theorem is very general but in most applications we work with regular hypergraphs. So we will state the very same result for D-regular hypergraphs, in which case the statement is cleaner. Notice that in this case we have $\Delta = D$ and $d(v) = D$ for all $v \in V$.

Corollary 5.4. *Let $r \geq 2$ be a fixed integer and D be a function of N. Let $\mathcal{H} = (V, \mathcal{E})$ be a D-regular r-uniform hypergraph with N vertices such that $\Delta_2(\mathcal{H}) = 1$. Then, with probability $1 - O\left(\frac{D^{2+\frac{2}{r-1}}}{N}\right)$, the number of good vertices in a uniformly distributed random permutation of the vertices of $\mathcal{H}$ is $\Theta\left(ND^{-\frac{1}{r-1}}\right)$.*

Remembering that a good vertex will always appear in the final independent set obtained by the random greedy algorithm, we have that the number of good vertices is always a lower bound on the number of steps of the algorithm and then we obtain the following:

Corollary 5.5. *Let $\mathcal{H}$ be a D-regular, r-uniform hypergraph with N vertices and such that $\Delta_2(\mathcal{H}) = 1$. Then, with probability $1 - O\left(N^{-1}D^{\frac{2r}{r-1}}\right)$, the random greedy independent set algorithm creates an independent set of size $\Omega\left(ND^{-\frac{1}{r-1}}\right)$.*

It is important to note that as we obtain not only lower bounds but also upper bounds in Theorem 5.1 and Corollary 5.4 it shows us that we cannot expect to obtain better results using this approach. In other words, by counting good vertices we can only obtain the lower bound $\Omega\left(ND^{-\frac{1}{r-1}}\right)$. As we saw in the statement of Bennett-Bohman's theorem and we will see later when we compute the differential equations heuristic, this is a logarithmic factor far from the estimated correct value.

5.2 The H-free process on a host graph F

In the H-free process, we consider a fixed graph H and we create a random H-saturated graph. Instead of creating a general graph with no copy of H, we can try to find such a graph as a subgraph of other given graph F. The classical H-free process is viewed now as the H-free process on the complete graph K_n and for a general graph F on n vertices we have the H-free process on the host graph F as follows.

The process generates a nested sequence $G_0, \ldots, G_M$ of graphs with the same vertex set of F. We begin with the empty graph G_0 with n vertices. At each step, we add uniformly at random a new edge (among all edges of F that would not create a copy of H) to the graph G_i to obtain G_{i+1}. The process stops when we arrive at a maximal H-free subgraph of F.

As before, we can associate a hypergraph to the H-free process on F. $\mathcal{H}_{H,F}$ will denote the hypergraph with vertex-set the edges of F, i.e. $V = E(F)$ and the edge set being all copies of H in F. Now, the H-free process on F is analogous to the random greedy independent set algorithm on $\mathcal{H}_{H,F}$.

5.2.1 The Q_2-free process on the hypercube Q_d

Theorem 5.6 (J. Robert Johnson, T. Pinto [17])**.** *The Q_2-free process on the hypercube Q_d with high probability generates a Q_2-free subgraph with at least $\Omega(d^{2/3}2^d)$ edges.*

Proof. Consider the hypergraph $\mathcal{H}_{Q_2,Q_d}$. As Q_d has $d2^{d-1}$ edges (and Q_2 has 4 edges), this hypergraph is 4-uniform and has $d2^{d-1}$ vertices. Also, every edge of Q_d is contained in $d-1$ Q_2's and those Q_2 share only one edge. Thus $\mathcal{H}_{Q_2,Q_d}$ is $(d-1)$-regular and $\Delta_2\left(\mathcal{H}_{Q_2,Q_d}\right) = 1$.

Applying Corollary 5.5 to $\mathcal{H}_{Q_2,Q_d}$ (with $N = d2^{d-1}$, $D = d-1$ and $r = 4$) we then obtain that with probability $1 - O\left(\frac{d^{8/3}}{d2^{d-1}}\right)$ (i.e., with high probability) the Q_2-free process on Q_d generates a graph whose number of edges is at least

$$\Omega\left(\frac{N}{D^{\frac{1}{r-1}}}\right) = \Omega\left(\frac{d2^{d-1}}{(d-1)^{\frac{1}{3}}}\right) = \Omega\left(d^{2/3}2^d\right).$$

□

6
Random greedy independent set algorithm

In this chapter, we present the proof of Theorem 4.1 by Bennett and Bohman as in [6]. We recall that the theorem asserts that with high probability this algorithm produces a large independent set provided that the hypergraph satisfies certain density conditions. As we saw in Chapter 4, this result generalizes a lower bound on the number of steps in the H-free process obtained by Bohman and Keevash in [15]. In fact, we will see that the proof actually gives a stronger result (see Theorem 6.2).

The method used here is a more sophisticated version of the "wholistic approach" presented in Chapter 3. In fact, we want to track the number of vertices $V(i)$ of the hypergraph during the algorithm. For this, we will find variables that together, when controlled, allow us to study the corresponding martingale. Besides that, those variables must form a "closed system" (in the sense that we can be able to write the expected values and variances of changes in those variables only in function of them) and then we will calculate the right error functions to obtain auxiliary supermartingales.

The chapter is presented as follows. First, we define the variables we will track and compute their expected trajectories by assuming that our independent set should look like a random set. We then will be able to state claims that together clearly imply the theorem. Last, we will finally be able to prove these claims, always following the ideas presented in Chapters 2 and 3.

6.1
Bennett-Bohman's theorem

We restate Theorem 4.1 for convenience.

Theorem 6.1 (Bennett-Bohman)**.** *Let r and $\varepsilon > 0$ be fixed. Let $\mathcal{H}$ be a r-uniform, D-regular hypergraph on N vertices such that $D > N^{\varepsilon}$. If*

$$\Delta_\ell(\mathcal{H}) < D^{\frac{r-\ell}{r-1}-\varepsilon} \quad \text{for } \ell = 2, \dots, r-1 \tag{6.1}$$

and $\Gamma(\mathcal{H}) < D^{1-\varepsilon}$ then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ with

$$|I| = \Omega\left(N \cdot \left(\frac{\log N}{D}\right)^{\frac{1}{r-1}}\right) \tag{6.2}$$

with probability $1 - \exp\{-N^{\Omega(1)}\}$.

First, we want to remark that the hypothesis of Theorem 6.1 implies that $\mathcal{H}$ is not too dense. To see this we can double count the number of edges $|\mathcal{H}|$: The number of pairs (v, e) where v is a vertex and e an edge such that $v \in e$ is $ND = r|\mathcal{H}|$. The number of triples (v, v', e) where v and v' are vertices incident to the edge e is $|\mathcal{H}|\binom{r}{2} \leq \binom{N}{2}\Delta_2(\mathcal{H})$. Then we have, using $\Delta_2(\mathcal{H}) < D^{\frac{r-2}{r-1}-\varepsilon}$:

$$\frac{ND}{r} = |\mathcal{H}| \leq \frac{1}{\binom{r}{2}}\binom{N}{2} D^{\frac{r-2}{r-1}-\varepsilon}.$$

And it follows that

$$N \geq 1 + (r-1)D^{\frac{1}{r-1}+\varepsilon} = \Omega\left(D^{\frac{1}{r-1}+\varepsilon}\right). \tag{6.3}$$

This estimate, which we use throughout this chapter, corresponds to $D = O\left(N^{\frac{r-1}{1+\varepsilon(r-1)}}\right)$. Since this is $o(N^{r-1})$, this means that $\mathcal{H}$ is not too dense, as stated before.

We also want to remark that we will prove the theorem with (6.2) replaced by

$$|I| = \Omega\left(N \cdot \left(\frac{\log D}{D}\right)^{\frac{1}{r-1}}\right),$$

this means, with $\log D$ instead of $\log N$. Although Bennett and Bohman proved their theorem with $\log N$, these results are equivalent since we are assuming $D > N^{\varepsilon}$ and, as we will see later on the heuristics (Section 6.4), we hope that the result with $\log D$ can be extended for smaller D, while the result with $\log N$ can only hold when D is at least polynomial in N (otherwise, $|I|$ would be even greater than N). Furthermore, we will be able to obtain our estimates without using that $D > N^{\varepsilon}$, obtaining the stronger version:

Theorem 6.2. *Let r and $\varepsilon > 0$ be fixed. Let $\mathcal{H}$ be a r-uniform, D-regular hypergraph on N vertices. If*

$$\Delta_\ell(\mathcal{H}) < D^{\frac{r-\ell}{r-1}-\varepsilon} \quad \textit{for } \ell = 2, \ldots, r-1$$

and $\Gamma(\mathcal{H}) < D^{1-\varepsilon}$ *then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ with*

$$|I| = \Omega\left(N \cdot \left(\frac{\log D}{D}\right)^{\frac{1}{r-1}}\right)$$

with probability $1 - N^{O(1)} \cdot \exp\{-D^{\Omega(1)}\}$.

Remark 6.1. We highlight that the probability in Theorem 6.2 is trivial unless D is at least a power of $\log N$. In fact, a careful examination of our proof can lead to the result that, if $D > (\log N)^{\frac{9r}{\varepsilon}}$ (here we made no attempt in finding the best constant in the exponent), then the probability of failure is at most $\exp(-(\log N)^{1+\Omega(1)})$. On the other hand, under the stronger condition that the hypergraph is linear (i.e., $\Delta_2(\mathcal{H}) = 1$) we may obtain the constant $2(r-1)+\sigma$, for any $\sigma > 0$, see Theorem 7.1.

6.2 Scope of proof

The idea of the proof is as follows: First, we will use the differential equations method to define the trajectories that several variables ought to follow during the process. Then we will bound one step changes of each of these variables so that we can use martingale deviation inequalities as lemmas.

With these lemmas in hand (see Chapter 2) we will be able to prove concentration of $|V(i)|$ (i.e., prove that with high probability $|V(i)|$ remains close to the expected trajectory guessed by the differential equations method), which is our main goal since this is the number of vertices that remain in the hypergraph. For this, we will have to track some other variables as well.

Definition 6.1. *For every vertex $v \in V(i)$ and $\ell \in \{2, \ldots, r\}$, let $d_\ell(i, v) = d_\ell(v)$ be the number of edges of cardinality ℓ in $\mathcal{H}(i)$ that contain v.*

Definition 6.2 (Degrees of Sets)**.** *For a set A of at least 2 vertices, let $d_{A\uparrow b}(i)$ be the number of edges of size b containing A in $\mathcal{H}(i)$.*

Definition 6.3 (Co-degrees)**.** *For a pair of vertices v, v', let $c_{a,a'\to k}(v, v', i)$ be the number of pairs of edges e, e', such that $v \in e\backslash e'$, $v' \in e'\backslash e$, $|e| = a$, $|e'| = a'$ and $|e \cap e'| = k$ and $e, e' \in \mathcal{H}(i)$.*

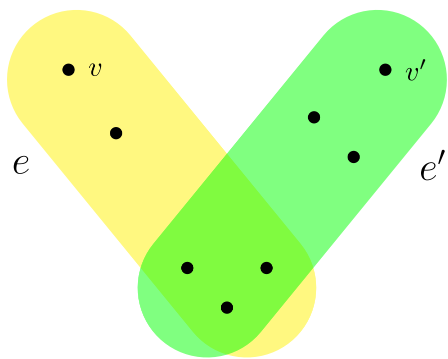


Figure 6.1: For example, $c_{5,6\to 3}(v, v')$ counts the number of pairs (e, e') as above.

6.3 Notation and conventions

Before we start, here are some notation and conventions we will use throughout the proof: ΔX represents the one-step change of a variable X, this means $\Delta X := X(i+1) - X(i)$. Since all expectations are conditional on the first steps we omit this conditioning for brevity, i.e., we write $\mathbb{E}[\cdot]$ instead of $\mathbb{E}[\cdot|\mathcal{F}_i]$ (where $\mathcal{F}_0, \mathcal{F}_1, \ldots, \mathcal{F}_i, \ldots$ is the natural filtration associated with the process).

We use $a \pm b$ to denote the interval $[a-b, a+b]$. If f and g are functions of D with the property that f is bounded from above by g times some poly-logarithmic function (in D) then we write $f = \tilde{O}(g)$. Furthermore, if f is bounded from below by g times some poly-logarithmic function (in D) then we write $f = \tilde{\Omega}(g)$.

6.4 The differential equations heuristic

Vertices

To guess the trajectories of our variables we will assume that our independent set behaves like a binomial random set. So let $S(i)$ denote the random vertex set such that each vertex v is in $S(i)$ independently with probability $p = p(i) = i/N$. Then, for a fixed vertex v, the expected number of edges $e \in \mathcal{H}$ such that $v \in e$ and $e \backslash \{v\} \subset S(i)$ is

$$Dp^{r-1} = D\left(\frac{i}{N}\right)^{r-1}.$$

Then we can think the probability of a vertex to be in $V(i)$ as $q \approx 1 - D\left(\frac{i}{N}\right)^{r-1}$ and we must have $|V(i)| \approx qN$. It turns out that the natural parametrization of time is

$$t := \frac{D^{\frac{1}{r-1}}}{N} \cdot i.$$

So we want to prove a bound of the type

$$|V(i)| \in Nq \pm ND^{-\delta} f_v, \tag{6.4}$$

where $\delta > 0$ is a constant and f_v is a function of t that is small enough so that the error term is little-o of the main term. Actually, we set $q = q(i) := \exp\left(-D\left(\frac{i}{N}\right)^{r-1}\right)$ or, in other words, $q = q(t) := \exp\left(-t^{r-1}\right)$.

Degrees of vertices

Because the expected number of edges of size ℓ is the number of possible ℓ-edges, which is $D\binom{r-1}{\ell-1}$, times the probability that a given ℓ-set forms an edge, which is $q^{\ell-1}p^{r-\ell}$, $d_\ell(v)$ should follow

$$s_\ell := D\binom{r-1}{\ell-1}q^{\ell-1}p^{r-\ell}.$$

However, it will be easier to separate the positive and negative contributions to the variable. Formally, we will write $d_\ell(v) = d_\ell^+(v) - d_\ell^-(v)$, where $d_\ell^\pm$ means the number of edges of size ℓ containing v that are created and destroyed (respectively) during the process. And we have that

$$s_\ell(t) = D\binom{r-1}{\ell-1}q^{\ell-1}\left(\frac{t}{D^{\frac{1}{r-1}}}\right)^{r-\ell} = \binom{r-1}{\ell-1}D^{\frac{\ell-1}{r-1}}t^{r-\ell}q^{\ell-1}$$

satisfies the differential equation

$$s_\ell' = \frac{\ell s_{\ell+1} - (\ell-1)s_\ell s_2}{D^{\frac{1}{r-1}}q}.$$

Indeed, remembering $q = \exp(-t^{r-1})$, we have that $q' = -(r-1)t^{r-2}q$,

$$\begin{aligned}\frac{d}{dt}\left(t^{r-\ell}q^{\ell-1}\right) &= (r-\ell)t^{r-\ell-1}q^{\ell-1} + t^{r-\ell}(\ell-1)q^{\ell-2}\Big(-(r-1)t^{r-2}q\Big)\\ &= t^{r-\ell-1}q^{\ell-1}\Big((r-\ell) - (\ell-1)(r-1)t^{r-1}\Big)\end{aligned}$$

and also

$$\begin{aligned}\ell s_{\ell+1} &= \ell\binom{r-1}{\ell}D^{\frac{\ell}{r-1}}t^{r-\ell-1}q^{\ell}\\ &= (r-\ell)\binom{r-1}{\ell-1}D^{\frac{\ell}{r-1}}t^{r-\ell-1}q^{\ell} \text{ and}\\ (\ell-1)s_\ell s_2 &= (\ell-1)\binom{r-1}{\ell-1}\binom{r-1}{1}D^{\frac{\ell-1}{r-1}}D^{\frac{1}{r-1}}t^{r-\ell+r-2}q^{\ell}.\end{aligned}$$

Therefore

$$\begin{aligned}&\ell s_{\ell+1} - (\ell-1)s_\ell s_2 =\\ &\qquad\binom{r-1}{\ell-1}D^{\frac{\ell}{r-1}}t^{r-\ell-1}q^{\ell}\Big((r-\ell) - (\ell-1)(r-1)t^{r-1}\Big) = D^{\frac{1}{r-1}}q(s_\ell)'.\end{aligned}$$

We then define

$$s_\ell^+(t) := D^{-\frac{1}{r-1}}\int_0^t \frac{\ell s_{\ell+1}(\tau)}{q(\tau)}d\tau \qquad s_\ell^-(t) := D^{-\frac{1}{r-1}}\int_0^t \frac{(\ell-1)s_\ell(\tau)s_2(\tau)}{q(\tau)}d\tau,$$

so that

$$s'_\ell = (s_\ell^+)' - (s_\ell^-)',$$

we have $s_\ell = s_\ell^+ - s_\ell^-$ and claim that $d_\ell^\pm \approx s_\ell^\pm$. We then want to prove bounds of the type

$$d_\ell^\pm(v) \in s_\ell^\pm \pm D^{\frac{\ell-1}{r-1}-\delta} f_\ell \quad \text{for } \ell = 2, \ldots, r \text{ and all } v \in V(i), \tag{6.5}$$

where $\delta > 0$ is a constant (the same δ from (6.4)) and $f_2, \ldots, f_r$ are functions of t that are small enough so that the error terms are little-o of the main terms.

Degrees of sets and co-degrees

For $d_{A\uparrow b}$ we can give an upper bound: it should follow (letting $a = |A|$)

$$\deg(A)\binom{r-a}{b-a} q^{b-a} p^{r-b} \leq \Delta_a(\mathcal{H})\binom{r-a}{b-a} q^{b-a} p^{r-b}.$$

And, since we have the condition $\Delta_\ell(\mathcal{H}) < D^{\frac{r-\ell}{r-1}-\varepsilon}$ and $p = tD^{-\frac{1}{r-1}}$, the function $d_{A\uparrow b}$ should satisfies

$$\begin{aligned} d_{A\uparrow b} &\leq \Delta_a(\mathcal{H})\binom{r-a}{b-a} q^{b-a} p^{r-b} \\ &\leq D^{\frac{r-a}{r-1}-\varepsilon}\binom{r-a}{b-a} q^{b-a} t^{r-b} D^{-\frac{r-b}{r-1}} \\ &\leq D^{\frac{b-a}{r-1}-\varepsilon}\binom{r-a}{b-a} q^{b-a} t^{r-b}, \end{aligned}$$

which gives a hint of how the upper bound on $d_{A\uparrow b}$ must look like.

Similarly, we can bound $c_{a,a'\to k}(v, v', i)$ using the co-degree of the pair of vertices v, v' in the original hypergraph $\mathcal{H}$ and derive how the upper bound on $c_{a,a'\to k}$ should look like but, for briefness, we will omit these calculations.

For those variables, we want to prove only upper bounds of the type

$$d_{A\uparrow b} \leq D_{a\uparrow b} \quad \text{for } 2 \leq a < b \leq r \text{ and all } A \in \binom{V(i)}{a} \tag{6.6}$$

$$c_{a,a'\to k}(v, v') \leq C_{a,a'\to k} \quad \text{for all } v, v' \in V(i) \tag{6.7}$$

where $D_{a\uparrow b}$ and $C_{a,a'\to k}$ are functions of D (and not of t).

Stopping time

Finally we introduce a constant $\zeta > 0$ that will be chosen so that $\zeta \ll \delta \ll \varepsilon$ (meaning that given $\varepsilon > 0$ we choose $\delta > 0$ sufficiently small and then $\zeta > 0$ sufficiently small with respect to δ) and we claim that the variables with high probability follow their trajectories until time $t_{\max} := \zeta \log^{\frac{1}{r-1}} D$ (i.e., $i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$). Hence, let the stopping time T be the minimum between $i_{\max}$ and the first step when any of (6.4), (6.5), (6.6) or (6.7) fails to hold.

We know that, for all $i > 0$ (as long as our heuristic assumptions are good approximations), i and $i+|V(i)|$ are lower and upper bounds to the number of steps of the algorithm. We then highlight that $i_{\max}$ is the natural choice of a end time because it is the moment when we expect $|V(i)|$ and i to have the same order of magnitude. This is the main difference between the proof from [6] and the present one. Bennett and Bohman considered $i_{\max} = \zeta N D^{-\frac{1}{r-1}} (\log N)^{\frac{1}{r-1}}$ because this is when $|V(i)| \approx$ constant. The issue with this choice is that we need to believe our heuristic assumptions are good approximations for a much longer period, which may not be the case for smaller D. By setting $i_{\max} = \zeta N D^{-\frac{1}{r-1}} (\log D)^{\frac{1}{r-1}}$ we are able to obtain a more general result (for instance, D can be poly-logarithmic in N and does not need to be at least polynomial in N).

Summarizing

Formally, we prove Theorem 6.2 by proving that $\mathbb{P}(T < i_{\max}) < N^{O(1)} \exp\{-D^{\Omega(1)}\}$. For this, it suffices to bound the probability of any of inequalities to fail for some $i+1$ *given* that all of them hold for smaller values of i. In the following claims, we then understand the bounds on $\mathbb{P}(\exists i A_i)$ as being bounds on the conditional probability of A_{i+1} to occur given all (6.4), (6.5), (6.6) and (6.7) hold. So Theorem 6.2 is a consequence of the following claims:

Claim 6.3. *Let* $2 \le a < b \le r$, $\lambda = \varepsilon/8r$ *and*

$$D_{a\uparrow b} := D^{\frac{b-a}{r-1} - \varepsilon + 2(r-b)\lambda}.$$

Then

$$\mathbb{P}\left(\exists i \le i_{max} \text{ and } A \in \binom{V(i)}{a} \text{ such that } d_{A\uparrow b} \ge D_{A\uparrow b}\right) \le N^{O(1)} \cdot \exp\left\{-D^{\Omega(1)}\right\}.$$

Claim 6.4. *Let* $2 \le a, a' \le r$ *and* $1 \le k < a, a'$ *be fixed. Let* $\lambda = \varepsilon/8r$ *and*

$$C_{a,a'\to k} := 2^r D^{\frac{a+a'-k-2}{r-1} - \varepsilon + (4r-2a-2a')\lambda}.$$

Then

$$\mathbb{P}\left(\exists i \leq i_{max} \text{ and } v, v' \in V(i) \text{ such that } c_{a,a'\to k}(v,v',i) \geq C_{a,a'\to k}\right) \\ \leq N^{O(1)} \cdot \exp\left\{-D^{\Omega(1)}\right\}.$$

Claim 6.5. *Setting*

$$f_v := \left(1+t^2\right) \cdot \exp(\alpha t + \beta t^{r-1}) \cdot q^2,$$

where α, β are constants that depends only on r, we have

$$\mathbb{P}\left(\exists i \leq i_{max} \text{ such that } |V(i)| \notin Nq \pm ND^{-\delta} f_v\right) \leq N^{O(1)} \cdot \exp\left\{-D^{\Omega(1)}\right\}.$$

Claim 6.6. *Setting*

$$f_\ell = \left(1+t^{r-\ell+2}\right) \cdot \exp(\alpha t + \beta t^{r-1}) \cdot q^\ell,$$

where α, β are constants that depends only on r (in fact that will be chosen the same constants as in Claim 6.5), we have

$$\mathbb{P}\left(\exists i \leq i_{max}, \ell \in \{2,\dots,r\} \text{ and } v \in V(i) \text{ such that } d_\ell^{\pm}(v) \notin s_\ell^{\pm} \pm D^{\frac{\ell-1}{r-1}-\delta} f_\ell\right) \\ \leq N^{O(1)} \cdot \exp\left\{-D^{\Omega(1)}\right\}.$$

Claims 6.3 to 6.6 imply that, with probability $1 - N^{O(1)} \cdot \exp\left\{-D^{\Omega(1)}\right\}$, $|V(i_{\max})| > 0$ and then the independent set obtained in the random greedy algorithm has size at least $i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$, as desired.

We break the proof into several parts. We first prove Claims 6.3 and 6.4 in Section 6.5 and then we prove Claims 6.5 and 6.6 in Section 6.6 using some auxiliary supermartingales and concentration inequalities for these supermartingales. To this, we will need to bound the martingale increments.

Remark 6.2. *The martingales Z that depends on a vertex v or a set of vertices A are frozen in the sense that $Z(i) = Z(i-1)$ if the vertex v or some of the vertices of A are not in $V(i)$. Also, when we write $Z(i)$ we actually work with $Z(i \wedge T)$.*

6.5 Proof of the theorem: Part I

Since we may choose $\zeta > 0$ sufficiently small relatively to ε, and $q(t_{\max}) = \exp\left(-(\zeta \log^{\frac{1}{r-1}} D)^{r-1}\right) = D^{-\zeta^{r-1}}$, we can let $\zeta^{r-1} < \lambda$ so that $Nq(t_{\max}) > ND^{-\lambda} + r$ and then

$$|V(i)| > ND^{-\lambda} + r \tag{6.8}$$

is valid whenever we have (6.4). In this section we will assume that this holds to prove Claims 6.3 and 6.4.

Since we are also assuming (6.5), it is important to note that it is an easy calculus exercise to show that

$$s_\ell(t) = \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}} t^{r-\ell} q^{\ell-1} \leq \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}},^{1}$$

implying we can also assume

$$d_\ell \leq \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}} \leq D^{\frac{\ell-1}{r-1}+(2r-2\ell)\lambda}. \tag{6.9}$$

6.5.1 How many ℓ-edges a fixed set of vertices belongs to

Proof of Claim 6.3: Bound on $d_{A\uparrow b}$.

The proof is by reverse induction on b. The base case is for $b = r$ where we have, by definition of $d_{A\uparrow r}$ and the hypothesis on maximum degrees of Theorem 6.1,

$$d_{A\uparrow r}(i) \leq \Delta_a(\mathcal{H}) < D^{\frac{r-a}{r-1}-\varepsilon} = D_{a\uparrow r}$$

and the desired bound is valid.

Now let $b < r$ and $A \in \binom{V}{a}$ be a fixed set. Assume that Claim 6.3 is true for $b+1$. We define $N_j(i)$, the number of vertices in $V(i)$ and not in A that appear in j edges counted by $d_{A\uparrow b+1}(i)$ (i.e., the vertices v such that there is j edges with $b+1$ vertices containing $A \cup \{v\}$). Note that $N_j(i)$ can be non-zero only for $0 \leq j \leq D_{a+1\uparrow b+1}$. Hence we have that $\sum_{j=0}^{D_{a+1\uparrow b+1}} N_j(i)$ counts the vertices in $V(i)$ not in A (we then have that $\sum_j N_j(i) = |V(i)| - a$). Moreover, we can double count the pairs of vertex and edge (v, e) where $v \in e$ and $v \notin A$ by

$$\sum_{j=0}^{D_{a+1\uparrow b+1}} jN_j(i) = (b+1-a)d_{A\uparrow b+1},$$

[1]Careful calculations show us $\max_{t\geq 0} s_\ell(t) = \binom{r-1}{\ell-1}\left(\frac{r-\ell}{e(r-1)(\ell-1)}\right)^{\frac{r-\ell}{r-1}} D^{\frac{\ell-1}{r-1}}$.

and using the induction hypothesis we have (with probability $1 - N^{O(1)}\exp\left(-D^{\Omega(1)}\right)$)

$$\sum_{j=0}^{D_{a+1\uparrow b+1}} jN_j(i) \le (b+1-a)D_{a\uparrow b+1}. \tag{6.10}$$

We have that $d_{A\uparrow b}$ can increase at most j when a vertex v counted by $N_j(i)$ is chosen at the next step of the algorithm. So define the auxiliary random variable X by $X(0) = 0$ and

$$\mathbb{P}(\Delta X(i) = j) = \frac{N_j(i)}{|V(i)| - a}.$$

As observed above we have that $\mathbb{P}(\Delta d_{A\uparrow b} \ge j) \le \mathbb{P}(\Delta X \ge j)$ (in other words, $\Delta d_{A\uparrow b}$ is stochastically dominated by ΔX) and to bound $d_{A\uparrow b}(i)$ is sufficient to bound $X(i)$.

Bounds on the one step change

Before applying Freedman's inequality (Lemma 2.5), we need to calculate

$$\mathbb{E}[\Delta X] = \frac{1}{|V(i)| - a}\sum_j jN_j(i) \le \frac{(b+1-a)D_{a\uparrow b+1}}{|V(i)| - a}$$

and, since we are assuming (6.8), we have

$$\mathbb{E}[\Delta X] \le \frac{(b+1-a)D_{a\uparrow b+1}}{ND^{-\lambda}} \le \frac{r}{N}D^{\frac{b-a+1}{r-1}-\varepsilon+(2r-2b-1)\lambda}.$$

Thus we define the supermartingale

$$Y(i) := X(i) - \frac{r}{N}D^{\frac{b-a+1}{r-1}-\varepsilon+(2r-2b-1)\lambda}\cdot i$$

and calculate (using (6.8) and (6.10))

$$\begin{aligned}\mathrm{Var}[\Delta Y] = \mathrm{Var}[\Delta X] \le \mathbb{E}\left[(\Delta X)^2\right] &= \frac{1}{|V(i)| - a}\sum_{j=0}^{D_{a+1\uparrow b+1}} j^2 N_j(i)\\ &\le \frac{D_{a+1\uparrow b+1}}{|V(i)| - a}\sum_j jN_j \le \frac{D_{a+1\uparrow b+1}}{|V(i)| - a}\cdot rD_{a\uparrow b+1}\\ &\le \frac{r}{ND^{-\lambda}}\cdot D^{\frac{b-a}{r-1}-\varepsilon+2(r-b-1)\lambda}\cdot D^{\frac{b+1-a}{r-1}-\varepsilon+2(r-b-1)\lambda}\\ &\le \frac{r}{N}D^{\frac{2b-2a+1}{r-1}-2\varepsilon+(4r-4b-3)\lambda}.\end{aligned}$$

Applying Freedman's inequality

Letting $C = D_{a+1\uparrow b+1} = D^{\frac{b-a}{r-1}-\varepsilon+(2r-2b-2)\lambda}$ we have that $\Delta Y(i) \le C$ for all i. We set

$$v = (\log D) D^{\frac{2b-2a}{r-1} - 2\varepsilon + (4r-4b-3)\lambda},$$

and then, for $i \le i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$ and sufficiently large D (when $r > 2$; for $r = 2$ we choose ζ sufficiently small),

$$\sum_{k \le i} \mathrm{Var}[\Delta Y(k) | \mathcal{F}_k] \le i_{\max} \cdot \frac{r}{N} D^{\frac{2b-2a+1}{r-1} - 2\varepsilon + (4r-4b-3)\lambda}$$

$$= \zeta r \log^{\frac{1}{r-1}} D \cdot D^{\frac{2b-2a}{r-1} - 2\varepsilon + (4r-4b-3)\lambda} < v.$$

Applying Freedman's inequality with $d = D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-1)\lambda}$ we conclude

$$\mathbb{P}\left[\exists i : Y(i) \ge D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-1)\lambda}\right] \le \exp\left(-\frac{d^2}{2(v + Cd)}\right) =$$

$$\exp\left(\frac{-D^{\frac{2b-2a}{r-1} - 2\varepsilon + (4r-4b-2)\lambda}}{2\left((\log D) D^{\frac{2b-2a}{r-1} - 2\varepsilon + (4r-4b-3)\lambda} + D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-2)\lambda} \cdot D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-1)\lambda}\right)}\right)$$

$$= \exp\left(-\frac{D^{\lambda}}{2((\log D) + 1)}\right).$$

Then we have

$$\mathbb{P}\left[\exists i : X(i) \ge \left(\frac{r}{N} D^{\frac{1}{r-1}} i + 1\right) D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-1)\lambda}\right] \le \exp\left(-\frac{D^{\lambda}}{2((\log D) + 1)}\right)$$

and we also have $\left(\frac{r}{N} D^{\frac{1}{r-1}} \cdot i + 1\right) \le \left(\zeta r \cdot \log^{\frac{1}{r-1}} D + 1\right) \le D^{\lambda}$ for sufficiently large D. Therefore, taking the union bound over all choices of set A with $|A| = a$,

$$\mathbb{P}\left[\exists i \text{ and } A \in \binom{V(i)}{a} : d_{A\uparrow b}(i) \ge D_{a\uparrow b}\right] \le$$

$$\le \binom{N}{a} \mathbb{P}\left[\exists i : X(i) \ge D_{a\uparrow b} = D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b)\lambda}\right]$$

$$\le \binom{N}{a} \mathbb{P}\left[\exists i : X(i) \ge \left(\frac{r}{N} D^{\frac{1}{r-1}} i + 1\right) D^{\frac{b-a}{r-1} - \varepsilon + (2r-2b-1)\lambda}\right]$$

$$\le \binom{N}{a} \exp\left(-\frac{D^{\lambda}}{2((\log D) + 1)}\right).$$

□

6.5.2 Controlling codegrees

Proof of Claim 6.4: Bound on $c_{a,a' \to k}(v, v')$.

First we note that Claim 6.4 follows from Claim 6.3 except for $a = a' = k+1$. If $a' > k+1$ then we have that $c_{a,a'\to k}$ (i.e., the number of pairs (e, e') counted by this variable) is at most the number of edges e of size a that contains v times the number of ways of choosing the edge e' given the specific choice of the k vertices of e that belongs to e' too. Hence we have

$$c_{a,a'\to k}(v, v') \leq d_a(v) \cdot \binom{a-1}{k} \cdot D_{k+1\uparrow a'}.$$

As $d_a(v) \leq D^{\frac{a-1}{r-1}+(2r-2a)\lambda}$ (by (6.9)) and $\binom{a-1}{k} \leq 2^a \leq 2^r$, we conclude

$$c_{a,a'\to k}(v, v') \leq$$
$$D^{\frac{a-1}{r-1}+(2r-2a)\lambda} \cdot 2^r \cdot D^{\frac{a'-(k+1)}{r-1}-\varepsilon+2(r-a')\lambda} = 2^r D^{\frac{a+a'-k-2}{r-1}-\varepsilon+(4r-2a-2a')\lambda},$$

and, since $2^r D^{\frac{a+a'-k-2}{r-1}-\varepsilon+(4r-2a-2a')\lambda} = C_{a,a'\to k}$, we are done in this case. Note that supposing $a > k+1$ we have the similar bound $c_{a,a'\to k}(v, v') \leq d_{a'}(v) \cdot \binom{a'-1}{k} \cdot D_{k+1\uparrow a}$ and this case is completely analogous.

Now we can turn to case $a = a' = k+1$ and proceed by reserve induction on k. The base case $k = r-1$ follows from the definition of $c_{r,r\to r-1}$ and the hypothesis on $\Gamma(\mathcal{H})$ in Theorem 6.1:

$$c_{r,r\to r-1} \leq \Gamma(\mathcal{H}) < D^{1-\varepsilon} \leq 2^r D^{1-\varepsilon} = C_{r,r\to r-1}.$$

Assume $k < r-1$. Note that $c_{k+1,k+1\to k}$ can increase when the algorithm chooses one vertex of the following:

- In the intersection of a pair of edges counted by $c_{k+2,k+2\to k+1}(v, v')$;
- Not contained in the intersection of a pair of edges counted by $c_{k+2,k+1\to k}(v, v')$ or $c_{k+1,k+2\to k}(v, v')$.

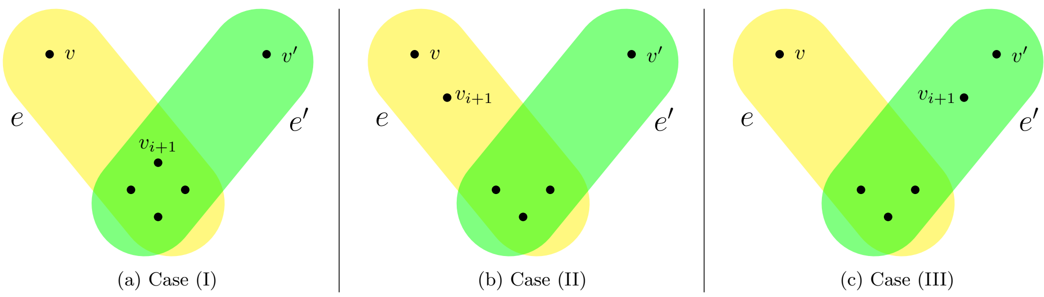


Figure 6.2: Illustration of the three cases when $k = 3$.

Then when $c_{k+1,k+1\to k}$ increases, it can increase at most by $3D_{2\uparrow k+2} = 3D^{\frac{k}{r-1}-\varepsilon+2(r-k-2)\lambda}$ as choosing a vertex u can increase $c_{k+1,k+1\to k}(v, v')$ at most by

a) The number of edges e of size $k+2$ that contains $\{u, v\}$ (and then the choice of e' with $k+2$ vertices, $k+1$ in common with e and not containing v is at most unique);

b) The number of edges e of size $k+2$ that contains $\{u, v\}$ (and then the choice of e' with $k+1$ vertices, k in common with e and not containing u and v is at most unique);

c) The number of edges e' of size $k+2$ that contains $\{u, v'\}$ (and then the choice of e with $k+1$ vertices, k in common with e' and not containing u and v' is at most unique).

For $0 \leq j \leq 3D^{\frac{k}{r-1}-\varepsilon+2(r-k-2)\lambda}$, let $N_j(i)$ denote the number of vertices from $V(i)$ that would increase $c_{k+1,k+1\to k}(v, v')$ by j if chosen in the next step of the algorithm. Then $\sum\limits_{j=0}^{3D_{2\uparrow k+2}} N_j(i) = |V(i)| - 2$ and, double-counting the number of vertex and edges (u, e, e') that would increase $c_{k+1,k+1\to k}$ (in the sense that they are as in items a), b) or c) above), we have

$$\begin{aligned}\sum jN_j(i) &\leq (k+1)C_{k+2,k+2\to k+1} + C_{k+2,k+1\to k} + C_{k+1,k+2\to k} \\ &= 2^r \cdot D^{\frac{k+1}{r-1}-\varepsilon+(4r-4k-6)\lambda} \cdot \Big((k+1)D^{-2\lambda} + 2\Big).\end{aligned}$$

We have that $(k+1)D^{-2\lambda} \leq 1$ for D sufficiently large and it follows that

$$\sum jN_j(i) \leq 3 \cdot 2^r \cdot D^{\frac{k+1}{r-1}-\varepsilon+(4r-4k-6)\lambda}.$$

Defining $X(i)$ by $X(0) = 0$ and

$$\mathbb{P}(\Delta X(i) = j) = \frac{N_j(i)}{|V(i)| - 2},$$

we have that $\Delta c_{k+1,k+1\to k}(v, v', i)$ is stochastically dominated by $\Delta X(i)$ and in order to prove our desired bound on $c_{k+1,k+1\to k}$ it's sufficient to prove it for X.

Bounds on the one step change

Hence

$$\mathbb{E}[\Delta X(i)] = \frac{1}{|V(i)| - 2} \sum_j jN_j(i) \leq \frac{3 \cdot 2^r}{ND^{-\lambda}} D^{\frac{k+1}{r-1}-\varepsilon+(4r-4k-6)\lambda},$$

we define the auxiliary supermartingale

$$Y(i) := X(i) - \frac{3 \cdot 2^r}{N} D^{\frac{k+1}{r-1}-\varepsilon+(4r-4k-5)\lambda} \cdot i$$

and calculate

$$\begin{aligned}\operatorname{Var}[\Delta Y] = \operatorname{Var}[\Delta X] \leq \mathbb{E}[(\Delta X)^2] = \\ = \frac{1}{|V(i)|-2}\sum_j j^2 N_j(i) \leq \frac{3D_{2\uparrow k+2}}{ND^{-\lambda}}\sum_j jN_j(i) \\ \leq \frac{3D^{\frac{k}{r-1}-\varepsilon+(2r-2k-4)\lambda}}{ND^{-\lambda}} \cdot 3\cdot 2^r D^{\frac{k+1}{r-1}-\varepsilon+(4r-4k-6)\lambda} \\ = \frac{9\cdot 2^r}{N} D^{\frac{2k+1}{r-1}-2\varepsilon+(6r-6k-9)\lambda}.\end{aligned}$$

Applying Freedman's inequality

Letting $C = 3D_{2\uparrow k+2} = 3D^{\frac{k}{r-1}-\varepsilon+(2r-2k-4)\lambda}$ we have that $\Delta Y(i) \leq C$ for all i. We set

$$v = (\log D)\cdot D^{\frac{2k}{r-1}-2\varepsilon+(6r-6k-9)\lambda},$$

and then, for $i \leq i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$ and sufficiently large D (when $r > 2$; for $r = 2$ we choose ζ sufficiently small),

$$\begin{aligned}\sum_{k\leq i} \operatorname{Var}[\Delta Y(k)|\mathcal{F}_k] \leq i_{\max}\cdot \frac{9\cdot 2^r}{N} D^{\frac{2k+1}{r-1}-2\varepsilon+(6r-6k-9)\lambda} \\ = 9\cdot 2^r \zeta \log^{\frac{1}{r-1}} D \cdot D^{\frac{2k}{r-1}-2\varepsilon+(6r-6k-9)\lambda} < (\log D)\cdot D^{\frac{2k}{r-1}-2\varepsilon+(6r-6k-9)\lambda} = v.\end{aligned}$$

Applying Freedman's inequality with $d = D^{\frac{k}{r-1}-\varepsilon+(4r-4k-5)\lambda}$ we conclude

$$\mathbb{P}\left[\exists i : Y(i) \geq D^{\frac{k}{r-1}-\varepsilon+(4r-4k-5)\lambda}\right] \leq \exp\left(-\frac{d^2}{2(v+Cd)}\right) =$$

$$\exp\left(\frac{-D^{\frac{2k}{r-1}-2\varepsilon+(8r-8k-10)\lambda}}{2\left((\log D)D^{\frac{2k}{r-1}-2\varepsilon+(6r-6k-9)\lambda} + 3D^{\frac{k}{r-1}-\varepsilon+(2r-2k-4)\lambda}\cdot D^{\frac{k}{r-1}-\varepsilon+(4r-4k-5)\lambda}\right)}\right)$$

$$= \exp\left(-\frac{D^{(2r-2k-1)\lambda}}{2((\log D)+3)}\right).$$

Then we have

$$\mathbb{P}\left[\exists i : X(i) \geq \left(\frac{3\cdot 2^r}{N} D^{\frac{1}{r-1}} i + 1\right) D^{\frac{k}{r-1}-\varepsilon+(4r-4k-5)\lambda}\right] \leq \exp\left(-\frac{D^{(2r-2k-1)\lambda}}{2((\log D)+3)}\right)$$

and we also have $\left(\frac{3\cdot 2^r}{N} D^{\frac{1}{r-1}} i + 1\right) \leq \left(3\cdot 2^r \cdot \zeta \cdot \log^{\frac{1}{r-1}} D + 1\right) \leq 2^r D^{\lambda}$ for sufficiently large D. Then, taking the union bound over all choices of v and v',

$$\mathbb{P}\left[\exists i \text{ and } v, v' \in V(i) : c_{k+1,k+1\to k}(v,v',i) \geq C_{k+1,k+1\to k}\right] \leq$$
$$\leq \binom{N}{2}\mathbb{P}\left[\exists i : X(i) \geq C_{k+1,k+1\to k} = 2^r D^{\frac{k}{r-1}-\varepsilon+(4r-4k-4)\lambda}\right]$$
$$\leq \binom{N}{2}\mathbb{P}\left[\exists i : X(i) \geq \left(\frac{3\cdot 2^r}{N}D^{\frac{1}{r-1}}i+1\right)D^{\frac{k}{r-1}-\varepsilon+(4r-4k-5)\lambda}\right]$$
$$\leq \binom{N}{2}\exp\left(-\frac{D^{(2r-2k-1)\lambda}}{2((\log D)+3)}\right).$$

□

6.6 Proof of the theorem: Part II

Until now we only obtained upper bounds for the variables $d_{A\uparrow b}$ and $c_{a,a'\to k}(v,v')$. From now on, we will need upper and lower bounds for the main variables $|V(i)|$ and $d_\ell(v)$, while the previous upper bounds obtained will be used as auxiliary estimates.

In this section, to prove the upper bounds in (6.4) and (6.5) (this means, upper bounds in Claims 6.5 and 6.6) we will define auxiliary supermartingales and, since they will have initial values negative and relatively large in absolute value, with the help of some martingale deviation inequalities we will obtain that is very unlikely them to ever be positive. For the correspondent lower bounds we will need analogous random variables, then similar arguments and calculations will apply the desired claims.

6.6.1 Auxiliary supermartingales

Consider

$$\begin{aligned} Z_V(i) &:= |V(i)| - Nq(t) - ND^{-\delta}f_v(t); \\ Z_\ell^+(v,i) &:= d_\ell^+(v,i) - s_\ell^+(t) - D^{\frac{\ell-1}{r-1}-\delta}f_\ell(t), \qquad \text{for } 2\leq \ell \leq r-1; \\ Z_\ell^-(v,i) &:= d_\ell^-(v,i) - s_\ell^-(t) - D^{\frac{\ell-1}{r-1}-\delta}f_\ell(t), \qquad \text{for } 2\leq \ell \leq r; \end{aligned}$$

where $f_v(t)$ and $f_\ell(t)$ are functions which will be chosen such that each of Z_V, Z_ℓ^+ and Z_ℓ^- are supermartingales.

Lemma 6.7. *If*

$$f_v(t) = \left(1+t^2\right)\cdot\exp(\alpha t + \beta t^{r-1})\cdot q^2 \text{ and}$$

$$f_\ell(t) = \left(1 + t^{r-\ell+2}\right) \cdot \exp(\alpha t + \beta t^{r-1}) \cdot q^\ell,$$

with α and β sufficiently large (depending only on r), then Z_V, Z_ℓ^+ and Z_ℓ^- are supermartingales.

Next, we will address conditions on the function f_v and f_ℓ (for $2 \leq \ell \leq r$) so that the variables above are supermartingales. Then in subsection 6.6.2 we will see that the chosen functions satisfy the desired conditions (completing the proof of Lemma 6.7) and in subsection 6.6.3 we will use auxiliary lemmas to complete the proof of the upper bounds in Claims 6.5 and 6.6. Throughout our estimates we will keep in mind that our error functions are chosen so that they all evaluate 1 at $t = 0$, are increasing in t and are well-behaved in the sense that they are smooth (then we can differentiate and use Taylor approximations).

6.6.1.1 Controlling the number of vertices

Z_V is a supermartingale

Let $S_t = N/D^{\frac{1}{r-1}}$ and recall that $t = i/S_t$. We write

$$\Delta Z_V = (|V(i+1)| - |V(i)|) - N(q(t+1/S_t) - q(t)) - ND^{-\delta}(f_v(t+1/S_t) - f_v(t)), \tag{6.11}$$

and make use of the Taylor approximations

$$q(t + 1/S_t) - q(t) = \frac{q'(t)}{S_t} - O\left(\frac{q''}{S_t^2}\right) \tag{6.12}$$

$$f_v(t + 1/S_t) - f_v(t) = \frac{f_v'(t)}{S_t} - O\left(\frac{f_v''}{S_t^2}\right) \tag{6.13}$$

where q'' and f_v'' are understood to be bounds on the second derivative that hold uniformly in the interval of interest. Remembering $q(t) = \exp\left(-t^{r-1}\right)$ and $s_2(t) = (r-1)D^{\frac{1}{r-1}} t^{r-2} \cdot q$, we have that

$$q'(t) = -(r-1)t^{r-2} \exp\left(-t^{r-1}\right) = -s_2(t) \cdot D^{-\frac{1}{r-1}} \tag{6.14}$$

$$q''(t) = \left((r-1)^2 t^{2r-4} - (r-1)(r-2)t^{r-3}\right) \cdot q(t) = O(1) \tag{6.15}$$

Where the last equality we obtain because $q''(t)$ is a continuous function that satisfies $q''(0) = 0$ and $q''(t)$ tends to 0 as t goes to infinity, then q'' is bounded (with bounds that depend only on r). [2]

Then, by (6.12), (6.14), (6.15) and the definition $S_t = N/D^{\frac{1}{r-1}}$,

[2] Using that $q''(t) \leq (r-1)^2 (t^{r-1})^2 e^{-t^{r-1}}$ for $t \geq 1$ by example, it's an easy calculus exercise to show that $q'' \leq (r-1)^2$.

$$N\left(q(t+1/S_t)-q(t)\right)=N\frac{q'(t)D^{\frac{1}{r-1}}}{N}-O\left(\frac{D^{\frac{2}{r-1}}\cdot q''}{N}\right)=-s_2(t)-O\left(\frac{D^{\frac{2}{r-1}}}{N}\right). \tag{6.16}$$

Now using (6.13) and $S_t = N/D^{\frac{1}{r-1}}$ we have

$$ND^{-\delta}(f_v(t+1/S_t)-f_v(t))=D^{\frac{1}{r-1}-\delta}f_v'(t)-O\left(\frac{D^{\frac{2}{r-1}-\delta}f_v''}{N}\right) \tag{6.17}$$

To estimate $\mathbb{E}[\Delta Z_V]$ we need first to estimate $\mathbb{E}[\Delta|V(i)|]$. Notice that if we choose a vertex v in the next step of the algorithm the number of vertices of the hypergraph deleted are $d_2(v)+1$ and then

$$\mathbb{E}[\Delta|V(i)|]=\frac{-1}{|V(i)|}\sum_{v\in V(i)}(d_2(v)+1)\le -s_2(t)+2D^{\frac{1}{r-1}-\delta}f_2(t), \tag{6.18}$$

where we make use of the estimate $d_2(v)\ge s_2(t)-2D^{\frac{1}{r-1}-\delta}f_2(t)$.

Finally, remembering (6.3), we have $O\left(\frac{1}{N}\right) = O\left(D^{-\frac{1}{r-1}-\varepsilon}\right)$ and, by (6.11), (6.16), (6.17) and (6.18),

$$\begin{aligned}\mathbb{E}[\Delta Z_V]&\le\\ &\le -s_2(t)+2D^{\frac{1}{r-1}-\delta}f_2(t)+s_2(t)-D^{\frac{1}{r-1}-\delta}f_v'(t)+O\left(D^{\frac{1}{r-1}-\varepsilon}\right)+O\left(D^{\frac{1}{r-1}-\varepsilon-\delta}f_v''\right)\\ &\le D^{\frac{1}{r-1}-\delta}[2f_2-f_v']+O\left(D^{\frac{1}{r-1}-\delta-\varepsilon}f_v''+D^{\frac{1}{r-1}-\varepsilon}\right).\end{aligned} \tag{6.19}$$

From the above estimate we conclude that Z_V is a supermartingale so long as $\varepsilon>\delta$ and

$$f_v''=o(D^{\varepsilon}) \tag{6.20}$$

$$f_v'>3f_2. \tag{6.21}$$

6.6.1.2
How many ℓ-edges containing a fixed v are created

Z_ℓ^+ is a supermartingale

Now we turn to $Z_\ell^+(v)$ for $2\le\ell\le r-1$ and a fixed vertex v. We write

$$\Delta Z_\ell^+(v)=\Delta d_\ell^+(v)-(s_\ell^+(t+1/S_t)-s_\ell^+(t))-D^{\frac{\ell-1}{r-1}-\delta}(f_\ell(t+1/S_t)-f_\ell(t)) \tag{6.22}$$

and make use of the Taylor approximations

$$s_\ell^+(t+1/S_t)-s_\ell^+(t)=\frac{\left(s_\ell^+\right)'(t)}{S_t}-O\left(\frac{\left(s_\ell^+\right)''}{S_t^2}\right) \tag{6.23}$$

$$f_\ell(t+1/S_t)-f_\ell(t)=\frac{(f_\ell)'(t)}{S_t}-O\left(\frac{(f_\ell)''}{S_t^2}\right) \tag{6.24}$$

where again $\left(s_\ell^+\right)''$ and $(f_\ell)''$ are understood to be bounds on the second derivative that hold uniformly in the interval of interest. Remembering $s_\ell(t) = \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}} t^{r-\ell} q^{\ell-1}$ and $s_\ell^+(t) = D^{-\frac{1}{r-1}} \int_0^t \frac{\ell s_{\ell+1}(\tau)}{q(\tau)} d\tau$, we have that

$$\left(s_\ell^+\right)'(t) = D^{-\frac{1}{r-1}} \frac{\ell s_{\ell+1}(t)}{q(t)}$$

and then

$$\frac{\left(s_\ell^+\right)'(t)}{S_t} = \frac{D^{-\frac{1}{r-1}} \frac{\ell s_{\ell+1}(t)}{q(t)}}{N D^{-\frac{1}{r-1}}} = \frac{\ell s_{\ell+1}(t)}{N q(t)}. \tag{6.25}$$

We also have that

$$\left(s_\ell^+\right)''(t) = \ell D^{-\frac{1}{r-1}} \left[\frac{q(t)(s_{\ell+1})'(t) - s_{\ell+1}(t)}{q(t)^2}\right]$$

$$= \ell D^{-\frac{1}{r-1}} \left[\frac{q\left(D^{-\frac{1}{r-1}}\left(\frac{(\ell+1)s_{\ell+2} - \ell s_{\ell+1} s_2}{q}\right)\right) - s_{\ell+1}}{q^2}\right] \Rightarrow$$

$$\left(s_\ell^+\right)''(t) = \ell \binom{r-1}{\ell} D^{\frac{\ell-1}{r-1}} t^{r-\ell-2} q(t)^{\ell-2} \left[\left(r-\ell-1-\ell(r-1)t^{r-1}\right) q(t) - t\right].$$

As in (6.15), we conclude that $\frac{\left(s_\ell^+\right)''(t)}{D^{\frac{\ell-1}{r-1}}}$ is bounded if $\ell > 2$ (with bounds that depend only on r) and for $\ell = 2$

$$\left|\left(s_2^+\right)''(t)\right| = 2\binom{r-1}{2} D^{\frac{1}{r-1}} |t|^{r-2} \left|\frac{(r-3)q(t) - 2(r-1)t^{r-1}q(t) - t}{t^2}\right|.$$

If $r = 3$ we have

$$\left|\left(s_2^+\right)''(t)\right| = 2D^{\frac{1}{2}} \left|\frac{-4t^2 q(t) - t}{t}\right| = 2D^{\frac{1}{2}} \left|4tq(t) + 1\right| = O\left(D^{\frac{1}{2}}\right).$$

If $r = 4$ we have

$$\left|\left(s_2^+\right)''(t)\right| = 6D^{\frac{1}{3}} \left|q(t) - 6t^3 q(t) - t\right| = O\left(D^{\frac{1}{3}}|t|\right).$$

If $r \geq 4$ we have

$$\left|\left(s_2^+\right)''(t)\right| =$$

$$= 2\binom{r-1}{2} D^{\frac{1}{r-1}} |t|^{r-4} \left|(r-3)q(t) - 2(r-1)t^{r-1}q(t) - t\right|$$

$$= O\left(D^{\frac{1}{r-1}} |t|^{r-3}\right).$$

In any case, as $|t| \le t_{\max} = \zeta (\log D)^{\frac{1}{r-1}}$ we have that

$$\left(s_2^+\right)'' = O\left(D^{\frac{1}{r-1}} (t_{\max})^{r-3}\right) = O\left((\log D)^{\frac{r-3}{r-1}} D^{\frac{1}{r-1}}\right) = \tilde{O}\left(D^{\frac{1}{r-1}}\right)$$

and then

$$\frac{\left(s_\ell^+\right)''}{S_t^2} = \tilde{O}\left(\frac{D^{\frac{\ell+1}{r-1}}}{N^2}\right). \tag{6.26}$$

To compute $\mathbb{E}[\Delta d_\ell^+(v)]$ we need to notice that by choosing the vertex u in the next step of the algorithm we create $d_{\{u,v\}\uparrow \ell+1}$ new edges of size ℓ containing v. Thus

$$\mathbb{E}[\Delta d_\ell^+(v)] = \frac{1}{|V(i)|} \sum_{u \in V(i) \setminus \{v\}} d_{\{u,v\}\uparrow \ell+1}$$

and double-counting the triples (v, u, e) where $\{u, v\} \subset e$ and $|e| = \ell + 1$ we have that

$$\sum_{u \in V(i) \setminus \{v\}} d_{\{u,v\}\uparrow \ell+1} = \ell d_{\ell+1}(v),$$

therefore we conclude

$$\mathbb{E}[\Delta d_\ell^+(v)] = \frac{\ell d_{\ell+1}(v)}{|V(i)|}. \tag{6.27}$$

Gathering (6.22)-(6.27) we obtain

$$\begin{aligned}
&\mathbb{E}[\Delta Z_\ell^+(v)] = \\
&\quad \frac{\ell d_{\ell+1}(v)}{|V(i)|} - \frac{\ell s_{\ell+1}}{Nq} - \frac{D^{\frac{\ell}{r-1}-\delta}}{N} f_\ell' + \tilde{O}\left(\frac{D^{\frac{\ell+1}{r-1}}}{N^2}\right) + O\left(\frac{D^{\frac{\ell+1}{r-1}-\delta} (f_\ell)''}{N^2}\right) \\
&\qquad \le \frac{\ell\left(s_{\ell+1} + 2D^{\frac{\ell}{r-1}-\delta} f_{\ell+1}\right)}{Nq - ND^{-\delta} f_v} - \frac{\ell s_{\ell+1}}{Nq} - \frac{D^{\frac{\ell}{r-1}-\delta}}{N} f_\ell' \\
&\qquad\qquad + \tilde{O}\left(\frac{D^{\frac{\ell}{r-1}-\varepsilon}}{N}\right) + O\left(\frac{D^{\frac{\ell}{r-1}-\delta-\varepsilon}}{N} f_\ell''\right),
\end{aligned}$$

where we used the bounds $|V(i)| \ge Nq - ND^{-\delta} f_v$, $d_{\ell+1} \le s_{\ell+1} + 2D^{\frac{1}{r-1}-\delta} f_{\ell+1}$ and $N = \Omega\left(D^{\frac{1}{r-1}+\varepsilon}\right)$. As

$$\frac{\ell\left(s_{\ell+1} + 2D^{\frac{\ell}{r-1}-\delta} f_{\ell+1}\right)}{Nq - ND^{-\delta} f_v} - \frac{\ell s_{\ell+1}}{Nq} = \frac{2\ell D^{\frac{\ell}{r-1}-\delta} f_{\ell+1} Nq + \ell s_{\ell+1} ND^{-\delta} f_v}{Nq(Nq - ND^{-\delta} f_v)},$$

$$\frac{1}{Nq - ND^{-\delta} f_v} = \frac{1}{Nq(1 - D^{-\delta} q^{-1} f_v)} = \frac{1 + O\left(D^{-\delta} q^{-1} f_v\right)}{Nq} \text{ and}$$

$$s_\ell = \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}} t^{r-\ell} q^{\ell-1}$$

we conclude [3]

$$\mathbb{E}[\Delta Z_\ell^+(v)] \leq \frac{D^{\frac{\ell}{r-1}-\delta}}{N}\left[2lq^{-1}f_{\ell+1} + \ell\binom{r-1}{\ell}t^{r-\ell-1}q^{\ell-2}f_v - f_\ell' + o(1)\right] + \tilde{O}\left(\frac{D^{\frac{\ell}{r-1}-\varepsilon}}{N}\right) + O\left(\frac{D^{\frac{\ell}{r-1}-\delta-\varepsilon}}{N}f_\ell''\right). \quad (6.28)$$

Whence we have that $Z_\ell^+(v)$ is a supermartingale so long as $\delta < \varepsilon$ and

$$f_\ell'' = o(D^\varepsilon) \quad (6.29)$$

$$f_\ell' > 5lq^{-1}f_{\ell+1} \quad (6.30)$$

$$f_\ell' > 2l\binom{r-1}{\ell}t^{r-\ell-1}q^{\ell-2}f_v. \quad (6.31)$$

6.6.1.3
How many ℓ-edges containing a fixed v are destroyed

Z_ℓ^- is a supermartingale

As before, but for $Z_\ell^-(v)$, we analogously write

$$\Delta Z_\ell^-(v) = \Delta d_\ell^-(v) - (s_\ell^-(t+1/S_t) - s_\ell^-(t)) - D^{\frac{\ell-1}{r-1}-\delta}(f_\ell(t+1/S_t) - f_\ell(t)) \quad (6.32)$$

and make use of the Taylor approximations (6.24) and

$$s_\ell^-(t+1/S_t) - s_\ell^-(t) = \frac{\left(s_\ell^+\right)'(t)}{S_t} - O\left(\frac{\left(s_\ell^+\right)''}{S_t^2}\right) \quad (6.33)$$

where again $\left(s_\ell^-\right)''$ is understood to be a bound on the second derivative that hold uniformly in the interval of interest. We now have

$$\left(s_\ell^-\right)'(t) = D^{-\frac{1}{r-1}}\frac{(\ell-1)s_\ell(t)s_2(t)}{q(t)} = (\ell-1)(r-1)\binom{r-1}{\ell-1}D^{\frac{\ell-1}{r-1}}t^{2r-\ell-2}q^{\ell-1}$$

and $\left(s_\ell^-\right)''(t) =$

$$O(1)\cdot D^{\frac{\ell-1}{r-1}}\left[(2r-\ell-2)t^{2r-\ell-3}q^{\ell-1} - (\ell-1)(r-1)t^{3r-\ell-4}q^{\ell-1}\right] = O\left(D^{\frac{\ell-1}{r-1}}\right),$$

[3] Here we used that the functions f_v and f_ℓ are sufficiently small, we will see that this will follow from the choice of ζ sufficiently small with respect to α and β.

which implies

$$\frac{\left(s_\ell^-\right)'(t)}{S_t} = \frac{(\ell-1)s_\ell s_2}{Nq} \tag{6.34}$$

$$\frac{\left(s_\ell^-\right)''}{S_t^2} = O\left(\frac{D^{\frac{\ell+1}{r-1}}}{N^2}\right) \tag{6.35}$$

Gathering (6.24) and (6.32)-(6.35) we obtain

$$\Delta Z_\ell^-(v) = \Delta d_\ell^-(v) - \frac{(\ell-1)s_\ell s_2}{Nq} - \frac{D^{\frac{\ell}{r-1}-\delta} f_\ell'}{N} + O\left(\frac{D^{\frac{\ell+1}{r-1}}}{N^2}\right) + O\left(\frac{D^{\frac{\ell+1}{r-1}-\delta}(f_\ell)''}{N^2}\right) \tag{6.36}$$

Finally, we need to estimate the changes $\Delta d_\ell^-(v)$. To this note that there are 3 ways of removing an edge e from the count d_ℓ by choosing a vertex $y \in V(i)$:

1. If the vertex y is contained in e.
2. If there exists x with $\{x, y\} \in \mathcal{H}(i)$ and $\{x, v\} \subset e$.
3. If there exists $e' \in \mathcal{H}(i)$ such that $y \in e'$, $|e'| \geq 3$ and $e' \backslash \{y\} \subset e$.

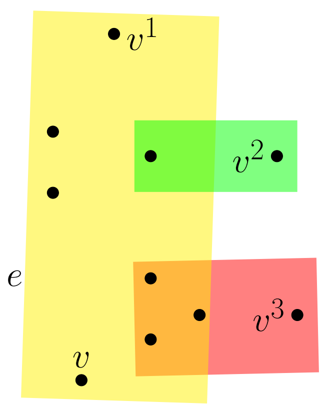


Figure 6.3: The vertices v^1, v^2 and v^3 illustrates the cases 1, 2 and 3, respectively. If one of them is chosen to be v_{i+1}, then the edge e will be removed from the count of $d_8(v)$.

The number of vertices of the first type are $\ell - 1$ and the main term in the expectation will come from the other ways of deleting an edge. First note that the sums

$$\sum_{x \in e \backslash \{v\}} d_2(x) + \sum_{A \subset e, |A| \geq 2} d_{A \uparrow |A|+1}$$

count each vertex in cases 2 and 3 at least once and at most 2^ℓ times (the number of subsets of e is 2^ℓ). Also the number of vertices that are counted more than once in the first sum is at most $\binom{\ell-1}{2}C_{2,2\to 1}$. Then the number of vertices that make e deleted is [4]

$$\sum_{x\in e\setminus\{v\}} d_2(x) + O\left(C_{2,2\to 1} + \sum_{k=2}^{\ell-1} D_{k\uparrow k+1}\right).$$

Therefore

$$\mathbb{E}[\Delta d_\ell^-(v)] = \frac{1}{|V(i)|}\left\{\sum_{e\in d_\ell(v)}\sum_{x\in e\setminus\{v\}} d_2(x) + O\left(d_\ell\cdot\left[C_{2,2\to 1} + \sum_{k=2}^{\ell-1} D_{k\uparrow k+1}\right]\right)\right\}.$$

Recording that $C_{2,2\to 1} = 2^r D^{\frac{1}{r-1}-\varepsilon+(4r-8)\lambda}$, $D_{k\uparrow k+1} = D^{\frac{1}{r-1}-\varepsilon+(2r-2k-2)\lambda}$ and $d_\ell \le D^{\frac{\ell-1}{r-1}+(2r-2\ell)\lambda}$ (the last by (6.9)) we have

$$d_\ell\cdot\left[C_{2,2\to 1} + \sum_{k=2}^{\ell-1} D_{k\uparrow k+1}\right] = O\left(D^{\frac{\ell}{r-1}-\varepsilon+(6r-2\ell-8)\lambda}\right) = O\left(D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}\right),^{5}$$

using $d_\ell \le s_\ell + 2D^{\frac{\ell-1}{r-1}-\delta} f_\ell$ and $|V(i)| \ge Nq - ND^{-\delta} f_v$ we have

$$\frac{1}{|V(i)|}\sum_{e\in d_\ell(v)}\sum_{x\in e\setminus\{v\}} d_2(x) \le \frac{(\ell-1)\cdot(s_\ell + 2D^{\frac{\ell-1}{r-1}-\delta} f_\ell)\cdot(s_2 + 2D^{\frac{1}{r-1}-\delta} f_2)}{Nq - ND^{-\delta} f_v},$$

where it is concluded that

$$\mathbb{E}[\Delta d_\ell^-(v)] \le \frac{(\ell-1)\cdot(s_\ell + 2D^{\frac{\ell-1}{r-1}-\delta} f_\ell)\cdot(s_2 + 2D^{\frac{1}{r-1}-\delta} f_2)}{Nq - ND^{-\delta} f_v} + O\left(\frac{D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}}{Nq}\right) \tag{6.37}$$

Now by (6.36) and (6.37) we get

$$\mathbb{E}[\Delta Z_\ell^-(v)] \le \frac{(\ell-1)\left(s_\ell + 2D^{\frac{\ell-1}{r-1}-\delta} f_\ell\right)\left(s_2 + 2D^{\frac{1}{r-1}-\delta} f_2\right)}{Nq - ND^{-\delta} f_v} - \frac{(\ell-1)s_\ell\cdot s_2}{Nq} - \frac{D^{\frac{\ell}{r-1}-\delta}}{N} f'_\ell + O\left(\frac{D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}}{Nq} + \frac{D^{\frac{\ell+1}{r-1}}}{N^2} + \frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2} f''_\ell\right)$$

Using again that $\frac{1}{Nq-ND^{-\delta}f_v} = \frac{1+O\left(D^{-\delta}q^{-1}f_v\right)}{Nq}$ and $s_\ell = \binom{r-1}{\ell-1} D^{\frac{\ell-1}{r-1}} t^{r-\ell} q^{\ell-1}$

[4]Here we write the number of vertices as being $a+O(b)$ meaning that it's bounded from below by $a-O(b)$ and from above by $a+O(b)$.

[5]Notice here we used our choice of $\lambda = \varepsilon/8r$.

we conclude [6]

$$\mathbb{E}[\Delta Z_\ell^-(v)] \leq \frac{D^{\frac{\ell}{r-1}-\delta}}{N} \cdot \Bigg[2(\ell-1)\binom{r-1}{\ell-1}t^{r-\ell}q^{\ell-2}f_2 + 2(\ell-1)(r-1)t^{r-2}f_\ell$$
$$+(\ell-1)(r-1)\binom{r-1}{\ell-1}t^{2r-\ell-2}q^{\ell-2}f_v - f'_\ell + o(1)\Bigg]$$
$$+ O\left(\frac{D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}}{Nq} + \frac{D^{\frac{\ell+1}{r-1}}}{N^2} + \frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2}f''_\ell\right) \quad (6.38)$$

Whence we have that $Z_\ell^-(v)$ is a supermartingale so long as $\delta < \frac{\varepsilon}{4} - \lambda$, $f''_\ell = o(D^\varepsilon)$ and

$$f'_\ell > 7(\ell-1)\binom{r-1}{\ell-1}t^{r-\ell}q^{\ell-2}f_2 \quad (6.39)$$
$$f'_\ell > 6(\ell-1)(r-1)t^{r-2}f_\ell \quad (6.40)$$
$$f'_\ell > 3(\ell-1)(r-1)\binom{r-1}{\ell-1}t^{2r-\ell-2}q^{\ell-2}f_v \quad (6.41)$$

6.6.2
Choosing the error functions

Recalling that the functions f_ℓ and f_v are chosen

$$f_\ell = \left(1+t^{r-\ell+2}\right)\cdot \exp(\alpha t + \beta t^{r-1})\cdot q^\ell,$$
$$f_v = \left(1+t^2\right)\cdot \exp(\alpha t + \beta t^{r-1})\cdot q^2.$$

Then [7]

$$f'_\ell = \left((r-\ell+2)t^{r-\ell+1}\right)\cdot \exp(\alpha t + \beta t^{r-1})\cdot q^\ell +$$
$$\left(1+t^{r-\ell+2}\right)\cdot(\alpha + (r-1)\beta t^{r-2})\exp(\alpha t + \beta t^{r-1})\cdot q^\ell +$$
$$\left(1+t^{r-\ell+2}\right)\cdot \exp(\alpha t + \beta t^{r-1})\cdot(-\ell(r-1)t^{r-2})q^\ell \geq$$
$$\left[\alpha + (\beta-\ell)(r-1)t^{2r-\ell}\right]\cdot \exp(\alpha t + \beta t^{r-1})\cdot q^\ell.$$

Analogously

$$f'_v \geq [\alpha + (\beta-2)(r-1)t^r]\cdot \exp(\alpha t + \beta t^{r-1})\cdot q^2.$$

Using the above estimates in the variation equations (6.21), (6.30), (6.31),

[6] Again we used that the functions f_v and f_ℓ are sufficiently small, which will follow from the choice of ζ.

[7] Note that here we use we will take $\beta > \ell$.

(6.39), (6.40), (6.41) we see that the exponential terms are equal in both left and right-hand size of each inequality, remaining to check the polynomial term. This means, we are only left to, respectively, check

$$\begin{aligned}
\alpha+(\beta-2)(r-1)t^{r} &\geq 3\ (1+t^{r})\\
\alpha+(\beta-\ell)(r-1)t^{2r-\ell} &\geq 5l\ (1+t^{r-\ell+1})\\
\alpha+(\beta-\ell)(r-1)t^{2r-\ell} &\geq 2l\binom{r-1}{\ell}t^{r-\ell-1}\ (1+t^{2})\\
\alpha+(\beta-\ell)(r-1)t^{2r-\ell} &\geq 7(\ell-1)\binom{r-1}{\ell-1}t^{r-\ell}\ (1+t^{r})\\
\alpha+(\beta-\ell)(r-1)t^{2r-\ell} &\geq 6(\ell-1)(r-1)t^{r-2}\ (1+t^{r-\ell+2})\\
\alpha+(\beta-\ell)(r-1)t^{2r-\ell} &\geq 3(\ell-1)(r-1)\binom{r-1}{\ell-1}t^{2r-\ell-2}\ (1+t^{2})
\end{aligned}$$

and all the variation equations hold because we choose α and β sufficiently large depending on r.

Differentiating one more time we have f''_ℓ is still of the form: a polynomial in t times $\exp(\alpha t+(\beta-\ell)t^{r-1})$[8]. As $t_{\max}=\zeta\,(\log D)^{\frac{1}{r-1}}$, we get $\exp(t_{\max}^{r-1})=D^{\zeta^{r-1}}$. Then we have that all f_ℓ, f_v, f''_ℓ and f''_v are $\tilde{O}\left(D^{\beta\zeta^{r-1}+o(1)}\right)$. Thus, choosing ζ such that

$$3\beta\zeta^{r-1}<\delta \tag{6.42}$$

we obtain the estimates $f''_\ell, f''_v = o\left(D^{\varepsilon}\right)$ and

$$\begin{aligned}
D^{-\delta}q^{-1}f_v\left[f_{\ell+1}+t^{r-\ell-1}q^{\ell-1}f_v\right] &= \tilde{O}\left(D^{2\beta\zeta^{r-1}-\delta+o(1)}\right)=o(1)\\
D^{-\delta}q^{-1}f_v\left[D^{-\delta}q^{-1}f_\ell f_2\right] &= \tilde{O}\left(D^{3\beta\zeta^{r-1}-2\delta+o(1)}\right)=o(1)\\
D^{-\delta}q^{-1}f_v\left[t^{r-\ell}q^{\ell-2}f_2+t^{r-2}f_\ell+t^{2r-\ell-2}q^{\ell-2}f_v\right] &= \tilde{O}\left(D^{2\beta\zeta^{r-1}-\delta+o(1)}\right)=o(1)\\
D^{-\delta}q^{-1}f_\ell f_2 &= \tilde{O}\left(D^{2\beta\zeta^{r-1}-\delta+o(1)}\right)=o(1)
\end{aligned}$$

implying the estimates that were remaining in (6.28) and (6.38).

6.6.3
Applying martingale variation inequalities

We recall we want to prove Claim 6.5, i.e., the bounds

$$|V(i)|\in Nq\pm ND^{-\delta}f_v.$$

[8]And the same changing $\beta-\ell$ to $\beta-2$ for f''_v.

In order to obtain the upper bound in Claim 6.5 we use the Hoeffding-Azuma inequality.

Proof of Claim 6.5: Upper bound on $|V(i)|$.

As we saw in subsection 6.6.1, choosing the vertex v at step i makes $\Delta V = -1 - d_2(v) \in -s_2 \pm D^{\frac{1}{r-1}-\delta} f_2$ and $\frac{Nq'}{S_t} = s_2$, where we derive

$$|\Delta Z_V| \leq D^{\frac{1}{r-1}-\delta} f_2 + O\left(\frac{D^{\frac{2}{r-1}} q''}{N}\right) + D^{\frac{1}{r-1}-\delta} f'_v + O\left(\frac{D^{\frac{2}{r-1}-\delta} f''_v}{N}\right)$$
$$= O\left(D^{\frac{1}{r-1}-\delta}(f_2 + f'_v)\right)$$

Then Hoeffding-Azuma inequality (Lemma 2.3) with $d = Z_V(0) = -ND^{-\delta}$ shows that for each $m \leq i_{\max} = O\left(ND^{-\frac{1}{r-1}}(\log D)^{\frac{1}{r-1}}\right)$ the probability that $Z_V(m)$ is positive is at most

$$\exp\left(-\Omega\left(\frac{(ND^{-\delta})^2}{m\left[D^{\frac{1}{r-1}-\delta}(f_2 + f'_v)\right]^2}\right)\right) \leq$$
$$\leq \exp\left(-\Omega\left(\frac{(ND^{-\delta})^2}{ND^{-\frac{1}{r-1}}(\log D)^{\frac{1}{r-1}}(D^{\frac{1}{r-1}-\delta}(f_2 + f'_v))^2}\right)\right) =$$
$$\exp\left(-\Omega\left(\frac{ND^{-\frac{1}{r-1}}}{(\log D)^{\frac{1}{r-1}}(f_2 + f'_v)^2}\right)\right) = \exp\left(-\tilde{\Omega}\left(D^{\varepsilon - 2\beta\zeta^{r-1} - o(1)}\right)\right),$$

where we used $N = \Omega\left(D^{\frac{1}{r-1}+\varepsilon}\right)$ and $f_2, f'_v = \tilde{O}\left(D^{\beta\zeta^{r-1}+o(1)}\right)$ to get the last equality. □

Remark 6.3. To show the lower bound we need to use the auxiliary variable

$$Y_V := -|V(i)| + Nq - ND^{-\delta} f_v$$

and, similar to subsection 6.6.1, we obtain the estimates

$$\mathbb{E}[\Delta Y_V] \leq$$
$$s_2(t) + 2D^{\frac{1}{r-1}-\delta} f_2(t) - s_2(t) - D^{\frac{1}{r-1}-\delta} f'_v(t) + O\left(D^{\frac{1}{r-1}-\varepsilon}\right) + O\left(D^{\frac{1}{r-1}-\varepsilon-\delta} f''_v\right)$$
$$\leq D^{\frac{1}{r-1}-\delta}[2f_2 - f'_v] + O\left(D^{\frac{1}{r-1}-\delta-\varepsilon} f''_v + D^{\frac{1}{r-1}-\varepsilon}\right),$$

where we conclude that Y_V is a supermartingale since $\delta < \varepsilon$, $f''_v = o(D^{\varepsilon})$ and $f'_v > 3f_2$. To finish the proof, using Hoeffding-Azuma inequality, we only need to see that the bound $|\Delta Y_V| = O\left(D^{\frac{\ell}{r-1}-\delta}(f_2 + f'_v)\right)$ still holds and $Y_V(0) = -ND^{-\delta}$. Then the same calculations above shows the desired result.

We recall we want to prove in Claim 6.6, the bounds

$$d_\ell^\pm(v) \notin s_\ell^\pm \pm D^{\frac{\ell-1}{r-1}-\delta} f_\ell.$$

Now to infer the upper bounds in Claim 6.6 we use Lemma 2.7.

Proof of Claim 6.6: Upper bound on $d_\ell^\pm(v)$.

As all $d_\ell^\pm$, $s_\ell^\pm$ and f_ℓ are increasing functions we have that

$$-\Delta s_\ell^\pm - D^{\frac{\ell-1}{r-1}-\delta}\Delta f_\ell \le |\Delta Z_\ell^\pm(v)| \le \Delta d_\ell^\pm(v)$$

It's sufficient now to note that choosing the vertex u at step i makes $\Delta d_\ell^+(v) = d_{\{u,v\}\uparrow \ell+1}$ and by our previous estimates in subsection 6.6.1.2 we get

$$\Delta d_\ell^+(v) \le D_{2\uparrow\ell+1} \le D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}},$$
$$\Delta d_\ell^-(v) \le O\left(D_{2\uparrow\ell} + \sum_{k=1}^{\ell-1} C_{\ell,k+1\to k}\right) = O\left(D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}\right) \text{ and}$$
$$\Delta s_\ell^\pm + D^{\frac{\ell-1}{r-1}-\delta}\Delta f_\ell = O\left(\left(s_\ell^\pm\right)' \cdot \frac{D^{\frac{1}{r-1}}}{N} + f_l' \cdot \frac{D^{\frac{l}{r-1}-\delta}}{N}\right) = O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right).$$

Since

$$Z_\ell^\pm(0) = -D^{\frac{\ell-1}{r-1}-\delta},$$
$$\frac{D^{\frac{\ell}{r-1}}}{N} = o\left(D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}\right) \text{ and}$$
$$D^{\frac{\ell-1}{r-1}-\delta} = o\left(i_{\max} \cdot \frac{D^{\frac{\ell}{r-1}}}{N}\right),$$

applying Lemma 2.7 with

$$\eta = O\left(D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}\right),$$
$$M = O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right)$$
$$\text{and } d = D^{\frac{\ell-1}{r-1}-\delta}$$

we obtain that the probability that $Z_\ell^\pm(v)$ is positive at a time $m \le i_{\max}$ is at most

$$\exp\left\{-\Omega\left(\frac{\left(D^{\frac{\ell-1}{r-1}-\delta}\right)^2}{ND^{-\frac{1}{r-1}}(\log D)^{\frac{1}{r-1}}\cdot\frac{1}{N}D^{\frac{\ell}{r-1}}\cdot D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}}\right)\right\}$$
$$\leq \exp\left\{-\Omega\left(\frac{D^{\frac{\varepsilon}{2}-2\delta}}{(\log D)^{\frac{1}{r-1}}}\right)\right\}.$$

□

Remark 6.4. To obtain the lower bounds in Claim 6.6 we use the asymmetric version of Hoeffding-Azuma inequality (Lemma 2.4).

The auxiliary variables in this case will be

$$\begin{aligned} Y_\ell^+(v) &:= -d_\ell^+(v) + s_\ell^+ - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, &&\text{for } 2\leq \ell\leq r-1;\\ Y_\ell^-(v) &:= -d_\ell^-(v) + s_\ell^- - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, &&\text{for } 2\leq \ell\leq r. \end{aligned}$$

and, similar to subsection 6.6.1.2, we obtain the estimates

$$\begin{aligned} \mathbb{E}[\Delta Y_\ell^+(v)] &\leq \frac{-\ell\left(s_{\ell+1}-2D^{\frac{\ell}{r-1}-\delta}f_{\ell+1}\right)}{Nq-ND^{-\delta}f_v} + \frac{\ell s_{\ell+1}}{Nq} - \frac{D^{\frac{\ell}{r-1}-\delta}}{N}f'_\ell \\ &\quad + \tilde{O}\left(\frac{D^{\frac{\ell}{r-1}-\varepsilon}}{N}\right) + O\left(\frac{D^{\frac{\ell}{r-1}-\delta-\varepsilon}}{N}f''_\ell\right) \\ &\leq \frac{D^{\frac{\ell}{r-1}-\delta}}{N}\left[2lq^{-1}f_{\ell+1} - \ell\binom{r-1}{\ell}t^{r-\ell-1}q^{\ell-2}f_v - f'_\ell + o(1)\right] \\ &\quad + \tilde{O}\left(\frac{D^{\frac{\ell}{r-1}-\varepsilon}}{N}\right) + O\left(\frac{D^{\frac{\ell}{r-1}-\delta-\varepsilon}}{N}f''_\ell\right). \end{aligned}$$

and

$$\begin{aligned} \mathbb{E}[\Delta Y_\ell^-(v)] &\leq \frac{-(\ell-1)\left(s_\ell-2D^{\frac{\ell-1}{r-1}-\delta}f_\ell\right)\left(s_2-2D^{\frac{1}{r-1}-\delta}f_2\right)}{Nq-ND^{-\delta}f_v} + \frac{(\ell-1)s_\ell\cdot s_2}{Nq} \\ &\quad - \frac{D^{\frac{\ell}{r-1}-\delta}}{N}f'_\ell + O\left(\frac{D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}}{Nq} + \frac{D^{\frac{\ell+1}{r-1}}}{N^2} + \frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2}f''_\ell\right) \\ &\leq \frac{D^{\frac{\ell}{r-1}-\delta}}{N}\cdot\left[2(\ell-1)\binom{r-1}{\ell-1}t^{r-\ell}q^{\ell-2}f_2 + 2(\ell-1)(r-1)t^{r-2}f_\ell\right. \\ &\quad \left.-(\ell-1)(r-1)\binom{r-1}{\ell-1}t^{2r-\ell-2}q^{\ell-2}f_v - f'_\ell + o(1)\right] \\ &\quad + O\left(\frac{D^{\frac{\ell}{r-1}-\frac{\varepsilon}{4}}}{Nq} + \frac{D^{\frac{\ell+1}{r-1}}}{N^2} + \frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2}f''_\ell\right) \end{aligned}$$

where we conclude that $Y_\ell^\pm$ are supermartingales by our established conditions on f_ℓ and f_v.

Again as all $d_\ell^\pm$, $s_\ell^\pm$ and f_ℓ are increasing functions we have that

$$-O\left(D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}\right) = -\Delta d_\ell^\pm(v) - D^{\frac{\ell-1}{r-1}-\delta}\Delta f_\ell \leq |\Delta Y_\ell^\pm(v)| \leq \Delta s_\ell^\pm = O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right).$$

Since $Y_\ell^\pm(0) = -D^{\frac{\ell-1}{r-1}-\delta}$ and the hypotheses of Lemma 2.4 are valid because $\frac{D^{\frac{\ell}{r-1}}}{N} = o\left(D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}\right)$ and $D^{\frac{\ell-1}{r-1}-\delta} = o\left(\frac{D^{\frac{\ell}{r-1}}}{N}\cdot i_{\max}\right)$, the probability that $Y_\ell^\pm(v)$ is positive at a time $m \leq i_{\max}$ is at most

$$\exp\left\{-\Omega\left(\frac{\left(D^{\frac{\ell-1}{r-1}-\delta}\right)^2}{ND^{-\frac{1}{r-1}}(\log D)^{\frac{1}{r-1}}\cdot\frac{1}{N}D^{\frac{\ell}{r-1}}\cdot D^{\frac{\ell-1}{r-1}-\frac{\varepsilon}{2}}}\right)\right\} \leq \exp\left\{-\Omega\left(\frac{D^{\frac{\varepsilon}{2}-2\delta}}{(\log D)^{\frac{1}{r-1}}}\right)\right\}.$$

7
Random greedy independent set algorithm in linear hypergraphs

In this chapter, we will discuss the issues of making the differential equations heuristics rigorous in a sparse setting. Precisely, we will present the argument given by Robert Johnson and Pinto [17] in the Q_2-free process on Q_d in Section 7.1 and extend it to the random greedy independent set algorithm setting for linear hypergraphs.

Then, we will follow the proof given in last chapter to obtain that the random variables associated with the random greedy independent set algorithm follow their expected trajectories for D-regular linear hypergraphs when $D > (\log N)^{2(r-1)+\sigma}$ for some $\sigma > 0$. In Section 7.2, we state our result and highlight some differences from the proof of Bennett and Bohman. Finally, in Sections 7.3 and 7.4 we present the estimates that are different from the previous chapter. Through this chapter, we will use the definitions and notations from the previous ones.

7.1
Q_2-free process on the hypercube Q_d revisited

Remembering that d_2 counts the number of 2-edges in the random greedy independent set algorithm, we can consider the variable $d_2(uv)$ for an edge uv in Q_d (which corresponds to a vertex in $\mathcal{H}_{Q_2,Q_d}$). This variable is zero while the vertices u and v are isolated during the Q_2-free process. Then

$$\begin{aligned}\mathbb{P}(d_2(uv,i)=0) &\geq \frac{\binom{d2^{d-1}-2d}{j}}{\binom{d2^{d-1}}{j}}\\ &= \frac{(d2^{d-1}-j)\dots(d2^{d-1}-2d+1-j)}{(d2^{d-1})\dots(d2^{d-1}-2d+1)}\\ &\geq \left(1-\frac{j}{d2^{d-1}}\right)^{2d-1}\\ &\geq \exp\left(-\frac{j}{2^{d-2}}\right),\end{aligned}$$

where $j(i)$ is the number of edges chosen in the Q_2-free process to have i edges added. If $j(i) \leq cd2^{d-1}$ then

$$\mathbb{P}(d_2(uv,i)=0) \geq \exp(-2cd)\,.$$

In expectation, we have at least $d2^{d-1}\exp(-2cd)$ edges with $d_2 = 0$. Then we have high probability that, for some edge uv, $d_2(uv)$ doesn't follow its trajectory for a constant proportion of the process.

For a general D-regular and r-uniform linear hypergraph $\mathcal{H}$, the analogue of the vertices u and v being isolated would be to consider a vertex v and two vertices from each r-edge that contain v. Then we bound the probability that $d_2 = 0$ by the probability that none of those $2D$ vertices (together with the vertex v) is chosen among the first j vertices. Therefore, we conclude

$$\begin{aligned}\mathbb{P}(d_2(v,i)=0) &\geq \frac{\binom{N-2D-1}{j}}{\binom{N}{j}} \\ &= \frac{(N-j)\dots(N-2D-j)}{N\dots(N-2D)} \\ &\geq \left(1-\frac{j}{N}\right)^{2D+1} \\ &\geq \exp\left(-\frac{j(2D+1)}{N}\right).\end{aligned}$$

So, with high probability, for some v, $d_2(v)$ doesn't follow its expected trajectory for a constant proportion of the process when $D = O(\log N)$. This means our approach can only prove that all variables follow their trajectories when $D \gg \log N$. However, we will be able to prove our heuristics when D is a power of $\log N$.

7.2 The theorem

Theorem 7.1. *Let $\sigma > 0$ and $r \geq 3$ be fixed. Let $\mathcal{H}$ be a r-uniform, D-regular hypergraph on N vertices such that $\Delta_2(\mathcal{H}) = 1$. If $D > (\log N)^{2(r-1)+\sigma}$ then the random greedy independent set algorithm produces an independent set I in $\mathcal{H}$ with*

$$|I| = \Omega\left(N \cdot \left(\frac{\log D}{D}\right)^{\frac{1}{r-1}}\right) \tag{7.1}$$

with probability at least $1 - O((\log N)^{1+\Omega(1)})$.

Remark 7.1. *We encourage the reader to compare this result with Theorem 6.2. The current result requires a weaker condition on D, provided $\mathcal{H}$ is linear* *(i.e., $\Delta_2(\mathcal{H}) = 1$). Furthermore, the linearity condition helps us simplify some aspects of the proof.*

First, we want to remark that for linear hypergraphs $N = \Omega(D)$. To see this we can double count the number of edges $|\mathcal{H}|$: The number of pairs (v, e) where v is a vertex and e an edge such that $v \in e$ is $ND = r|\mathcal{H}|$. The number of triples (v, v', e) where v and v' are vertices incident to the edge e is $|\mathcal{H}|\binom{r}{2} \leq \binom{N}{2}\Delta_2(\mathcal{H})$. Then

$$\frac{ND}{r} = |\mathcal{H}| \leq \frac{1}{\binom{r}{2}}\binom{N}{2}.$$

And it follows that

$$N \geq 1 + (r-1)D = \Omega(D). \tag{7.2}$$

We will use this estimate throughout this chapter.

Remark 7.2. *Note that* (7.2) *is one of the differences between last chapter setting and the current one. We highlight that we did not use the previous bound* $N = \Omega\left(D^{\frac{1}{r-1}+\varepsilon}\right)$ *to obtain most of the estimates.*

The heuristics from the last chapter still hold in our setting, with the difference that $d_{A\uparrow b} \leq 1$ for every set A with at least two vertices and that $c_{a,a'\to k}(v, v') = 0$ if $k \geq 2$. We then want to prove that

$$|V(i)| \in Nq \pm ND^{-\delta} f_v, \tag{7.3}$$

$$d_\ell^{\pm}(v) \in s_\ell^{\pm} \pm D^{\frac{\ell-1}{r-1}-\delta} f_\ell \quad \text{for } \ell = 2, \ldots, r \text{ and all } v \in V(i), \tag{7.4}$$

$$c_{a,a'\to 1}(v, v') \leq C_{a,a'\to 1} \quad \text{for all } v, v' \in V(i) \tag{7.5}$$

As before we consider the stopping time T as the minimum between $i_{\max} = \zeta N D^{-\frac{1}{r-1}}(\log D)^{\frac{1}{r-1}}$ and the first i such that any of the above equations fail to hold. Theorem 7.1 is a consequence of the following claims:

Claim 7.2. *Let* $2 \leq a, a' \leq r$. *Then*

$$c_{a,a'\to 1}(v, v') \leq (a-1) \cdot d_a(v)$$
$$c_{a,a'\to 1}(v, v') \leq (a'-1) \cdot d_{a'}(v').$$

Claim 7.3.

$$\mathbb{P}\left(\exists i \leq i_{max} \text{ and } v, v' \in V(i) \text{ such that } c_{2,2\to 1}(v, v', i) \geq 2 \cdot D^{\frac{1}{2(r-1)}}\right) \leq \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

Claim 7.4. *Setting*

$$f_v := \left(1 + t^2\right) \cdot \exp(\alpha t + \beta t^{r-1}) \cdot q^2,$$

where α, β are constants that depends only on r, we have

$$\mathbb{P}\left(\exists i \le i_{max} \text{ such that } |V(i)| \notin Nq \pm ND^{-\delta} f_v\right) \le \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

Claim 7.5. *Setting*

$$f_\ell = \left(1 + t^{r-\ell+2}\right) \cdot \exp(\alpha t + \beta t^{r-1}) \cdot q^\ell,$$

we have

$$\mathbb{P}\left(\exists i \le i_{max}, \ell \in \{2, \dots, r\} \text{ and } v \in V(i) \text{ such that } d_\ell^\pm(v) \notin s_\ell^\pm \pm D^{\frac{\ell-1}{r-1}-\delta} f_\ell\right)$$
$$\le \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

Claims 7.2 to 7.5 imply that, with probability $1 - \exp\left(-(\log N)^{1+\Omega(1)}\right)$, $|V(i_{\max})| > 0$ and then the independent set obtained in the random greedy algorithm has size at least $i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$, as desired.

Most of the estimates from the previous chapter are analogous for this case. For the sake of brevity, we omit them and focus only on the different computations in this new setting. We prove Claims 7.2 and 7.3 in Section 7.3 and then we prove Claims 7.4 and 7.5 in Section 7.4.

Remark 7.3. *The martingales Z that depends on a vertex v are frozen in the sense that $Z(i) = Z(i-1)$ if the vertex v is not in $V(i)$.*

7.3 Proof of the theorem: Part I

Since we may choose $\zeta > 0$ sufficiently small relatively to r, and $q(t_{\max}) = \exp\left(-(\zeta \log^{\frac{1}{r-1}} D)^{r-1}\right) = D^{-\zeta^{r-1}}$, we can let $\zeta^{r-1} < \lambda = \frac{1}{4(r-1)}$ so that $q(t_{\max}) > D^{-\lambda}$ and then

$$|V(i)| > ND^{-\lambda} + r \tag{7.6}$$

is valid whenever we have (7.3). As in the previous chapter, we also obtain that

$$d_\ell = O\left(D^{\frac{\ell-1}{r-1}}\right) \tag{7.7}$$

is valid from (7.4). In this section we will assume that these equations hold.

7.3.1 Controlling codegrees

Proof of Claims 7.2 and 7.3: Bound on $c_{a,a'\to k}(v,v')$.

First we note that Claim 7.2 follows from $\Delta_2(\mathcal{H}) = 1$. Indeed, we have that $c_{a,a'\to 1}$ (i.e., the number of pairs (e,e') counted by this variable) is at most the number of edges e of size a that contains v times the number of ways of choosing the edge e' given the specific choice of the one vertex of e that belongs to e' too. Hence we have

$$c_{a,a'\to k}(v,v') \le d_a(v)\cdot\binom{a-1}{1}.$$

Notice we have the similar bound $d_{a'}(v)\cdot\binom{a'-1}{1}$.

Now we can turn to case $k=1$ and $a=a'=2$. Note that $c_{2,2\to 1}$ can increase when the algorithm chooses one vertex of the following:

- Not contained in the intersection of a pair of edges counted by $c_{3,2\to 1}(v,v')$ or $c_{2,3\to 1}(v,v')$.

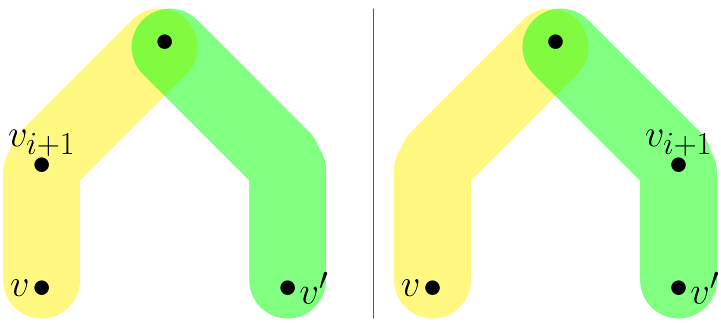


Figure 7.1: In each case, when the vertex v_{i+1} is chosen, it increases $c_{2,2\to 1}(v,v')$.

Then when $c_{2,2\to 1}$ increases, it can increase at most by 2.

For $0\le j\le 2$, let $N_j(i)$ denote the number of vertices from $V(i)$ that would increase $c_{2,2\to 1}(v,v')$ by j if chosen in the next step of the algorithm. Then $N_0(i)+N_1(i)+N_2(i) = |V(i)|-2$ and, double-counting the number of vertex and edges (u,e,e') that would increase $c_{2,2\to 1}$, we have

$$N_1(i)+2N_2(i) = c_{3,2\to 1}(v,v') + c_{2,3\to 1}(v,v') \le d_2(v)+d_2(v') \le 2\cdot D^{\frac{1}{r-1}}.$$

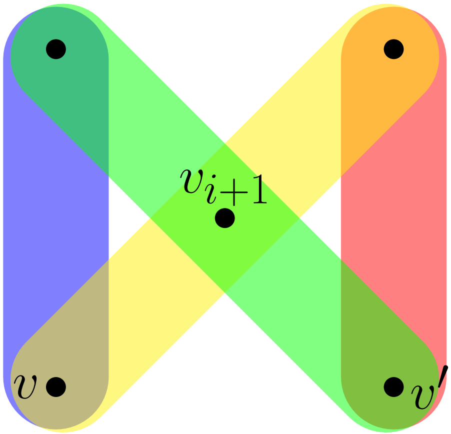


Figure 7.2: This is the only case when $c_{2,2\to 1}(v, v')$ can increase by 2.

Defining $X(i)$ by $X(0) = 0$ and

$$\mathbb{P}(\Delta X(i) = j) = \frac{N_j(i)}{|V(i)| - 2},$$

we have that $\Delta c_{2,2\to 1}(v, v', i)$ is stochastically dominated by $\Delta X(i)$ and ir order to prove our desired bound on $c_{2,2\to 1}$ it's sufficient to prove it for X.

Bounds on the one step change

Hence

$$\mathbb{E}[\Delta X(i)] = \frac{1}{|V(i)| - 2} \sum_j jN_j(i) \leq \frac{2 \cdot D^{\frac{1}{r-1}}}{|V(i)| - 2} \leq \frac{2 \cdot D^{\frac{1}{r-1}+\lambda}}{N},$$

we define the auxiliary supermartingale

$$Y(i) := X(i) - \frac{2 \cdot D^{\frac{1}{r-1}+\lambda}}{N} \cdot i$$

and calculate

$$\mathrm{Var}[\Delta Y] = \mathrm{Var}[\Delta X] \leq \mathbb{E}[(\Delta X)^2] = \frac{1}{|V(i)| - 2} \sum_j j^2 N_j(i) \leq 2 \cdot \mathbb{E}[\Delta X] \leq \frac{4 \cdot D^{\frac{1}{r-1}+\lambda}}{N}.$$

Applying Freedman's inequality

Letting $C = 2$ we have that $\Delta Y(i) \leq C$ for all i. We set

$$v = (\log D)^{\frac{1}{r-1}} D^{\lambda},$$

and then, for $i \le i_{\max} = \zeta N D^{-\frac{1}{r-1}} \log^{\frac{1}{r-1}} D$,

$$V(i) := \sum_{k \le i} \mathrm{Var}[\Delta Y(k) | \mathcal{F}_k] \le i_{\max} \cdot \frac{4 \cdot D^{\frac{1}{r-1}+\lambda}}{N} = 4\zeta (\log D)^{\frac{1}{r-1}} D^{\lambda} \le v.$$

Applying Freedman's inequality (Lemma 2.5) with $d = D^{\frac{1}{2(r-1)}}$ we conclude

$$\begin{aligned}\mathbb{P}\left[\exists i : Y(i) \ge D^{\frac{1}{2(r-1)}}\right] &\le \exp\left(-\frac{d^2}{2(v + Cd)}\right) \\ &= \exp\left(-\frac{D^{\frac{1}{r-1}}}{2\left((\log D)^{\frac{1}{r-1}} D^{\lambda} + 2 \cdot D^{\frac{1}{2(r-1)}}\right)}\right) \\ &\le \exp\left(-\Omega\left(D^{\frac{1}{2(r-1)}}\right)\right) \le \exp\left(-(\log N)^{1+\Omega(1)}\right).\end{aligned}$$

Then we have

$$\mathbb{P}\left[\exists i : X(i) \ge \frac{2 \cdot D^{\frac{1}{r-1}+\lambda}}{N} \cdot i + D^{\frac{1}{2(r-1)}}\right] \le \exp\left(-(\log N)^{1+\Omega(1)}\right)$$

and we also have

$$\frac{2 \cdot D^{\frac{1}{r-1}+\lambda}}{N} \cdot i + D^{\frac{1}{2(r-1)}} \le 2\zeta (\log D)^{\frac{1}{r-1}} D^{\lambda} + D^{\frac{1}{2(r-1)}} \le 2 \cdot D^{\frac{1}{2(r-1)}}$$

for sufficiently large D.

Then, taking the union bound over all choices of v and v',

$$\begin{aligned}&\mathbb{P}\left[\exists i \text{ and } v, v' \in V(i) : c_{2,2\to 1}(v, v', i) \ge 2 \cdot D^{\frac{1}{2(r-1)}}\right] \\ &\qquad \le \binom{N}{2} \mathbb{P}\left[\exists i : X(i) \ge 2 \cdot D^{\frac{1}{2(r-1)}}\right] \\ &\qquad \le \binom{N}{2} \mathbb{P}\left[\exists i : X(i) \ge \frac{2 \cdot D^{\frac{1}{r-1}+\lambda}}{N} \cdot i + D^{\frac{1}{2(r-1)}}\right] \\ &\qquad \le \exp\left(-(\log N)^{1+\Omega(1)}\right).\end{aligned}$$

□

7.4
Proof of the theorem: Part II

In this section, to prove the upper bounds in (7.3) and (7.4) we will use the auxiliary supermartingales

$$\begin{aligned} Z_V :=& |V(i)| - Nq - ND^{-\delta} f_v; \\ Z_\ell^+(v) :=& d_\ell^+(v) - s_\ell^+ - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, \qquad \text{for } 2 \le \ell \le r-1; \\ Z_\ell^-(v) :=& d_\ell^-(v) - s_\ell^- - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, \qquad \text{for } 2 \le \ell \le r. \end{aligned}$$

We obtain analogous estimates for Z_V and Z_ℓ^+:

$$\begin{aligned} \mathbb{E}[\Delta Z_V] \le& -s_2(t) + 2D^{\frac{1}{r-1}-\delta} f_2(t) + s_2(t) - D^{\frac{1}{r-1}-\delta} f_v'(t) \\ &+ O\left(\frac{D^{\frac{2}{r-1}}}{N}\right) + O\left(\frac{D^{\frac{2}{r-1}-\delta} f_v''}{N}\right) \\ &\le D^{\frac{1}{r-1}-\delta}[2f_2 - f_v'] + O\left(\frac{D^{\frac{2}{r-1}}}{N} + \frac{D^{\frac{2}{r-1}-\delta} f_v''}{N}\right). \qquad (7.8) \end{aligned}$$

$$\begin{aligned} \mathbb{E}[\Delta Z_\ell^+(v)] \le& \frac{D^{\frac{\ell}{r-1}-\delta}}{N}\left[2lq^{-1} f_{\ell+1} + \ell\binom{r-1}{\ell} t^{r-\ell-1} q^{\ell-2} f_v - f_\ell' + o(1)\right] \\ &+ \tilde{O}\left(\frac{D^{\frac{\ell+1}{r-1}}}{N^2}\right) + O\left(\frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2} f_\ell''\right). \qquad (7.9) \end{aligned}$$

For Z_ℓ^-, it is a little different because we have to use our bound on $c_{2,2\to 1}$ to estimate the changes $\Delta d_\ell^-(v)$. For completeness, we repeat the arguments to bound the increments $\Delta d_\ell^-(v)$. Note that there are 2 ways of removing an edge e from the count d_ℓ by choosing a vertex $y \in V(i)$:

1. If the vertex y is contained in e.
2. If there exists x with $\{x, y\} \in \mathcal{H}(i)$ and $\{x, v\} \subset e$.

The number of vertices of the first type are $\ell - 1$ and the main term in the expectation will come from the other way of deleting an edge. First note that the sum

$$\sum_{x \in e \setminus \{v\}} d_2(x)$$

count each vertex in case 2 at least once and at most ℓ times. Also the number of vertices that are counted more than once in the sum is at most $\binom{\ell-1}{2} C_{2,2\to 1}$.

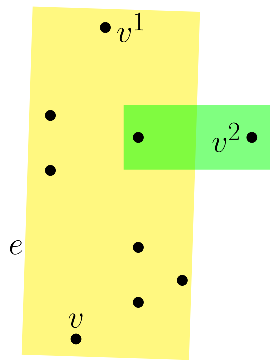


Figure 7.3: The vertices v^1 and v^2 illustrates the cases 1 and 2, respectively. If one of them is chosen to be v_{i+1}, then the edge e will be removed from the count of $d_8(v)$.

Then the number of vertices that make e deleted is [1]

$$\sum_{x \in e \setminus \{v\}} d_2(x) + O\left(C_{2,2\to 1}\right).$$

Therefore

$$\mathbb{E}[\Delta d_\ell^-(v)] = \frac{1}{|V(i)|}\left\{\sum_{e \in d_\ell(v)} \sum_{x \in e\setminus\{v\}} d_2(x) + O\left(d_\ell \cdot C_{2,2\to 1}\right)\right\}.$$

Recording that $C_{2,2\to 1} = 2 \cdot D^{\frac{1}{2(r-1)}}$ and $d_\ell = O\left(D^{\frac{\ell-1}{r-1}}\right)$ we have

$$d_\ell \cdot C_{2,2\to 1} = O\left(D^{\frac{\ell}{r-1} - \frac{1}{2(r-1)}}\right).$$

Then, as before, we conclude that

$$\mathbb{E}[\Delta Z_\ell^-(v)] \leq \frac{D^{\frac{\ell}{r-1}-\delta}}{N} \cdot \Bigg[2(\ell-1)\binom{r-1}{\ell-1} t^{r-\ell} q^{\ell-2} f_2 + 2(\ell-1)(r-1)t^{r-2} f_\ell$$
$$+(\ell-1)(r-1)\binom{r-1}{\ell-1} t^{2r-\ell-2} q^{\ell-2} f_v - f_\ell' + o(1)\Bigg]$$
$$+ O\left(\frac{D^{\frac{\ell}{r-1} - \frac{1}{2(r-1)}}}{Nq} + \frac{D^{\frac{\ell+1}{r-1}}}{N^2} + \frac{D^{\frac{\ell+1}{r-1}-\delta}}{N^2} f_\ell''\right) \quad (7.10)$$

Finally, remembering (7.2), we have $O\left(\frac{1}{N}\right) = O\left(D^{-1}\right)$, the conditions

[1]Here we write the number of vertices as being $a + O(b)$ meaning that it's bounded from below by $a - O(b)$ and from above by $a + O(b)$.

on δ for Z_V, $Z_\ell^{\pm}$ to be supermartingales are then

$$\delta < 1 - \frac{1}{r-1} \tag{7.11}$$

$$\delta < \frac{1}{2(r-1)} - \lambda = \frac{1}{4(r-1)} \tag{7.12}$$

We also need that

$$f_v'' = o\left(D^{1-\frac{1}{r-1}}\right) \tag{7.13}$$

$$f_\ell'' = o\left(D^{1-\frac{1}{r-1}}\right) \tag{7.14}$$

and the variaton equation, as before,

$$f_v' > 3f_2. \tag{7.15}$$

$$f_\ell' > 5lq^{-1}f_{\ell+1} \tag{7.16}$$

$$f_\ell' > 2l\binom{r-1}{\ell}t^{r-\ell-1}q^{\ell-2}f_v. \tag{7.17}$$

$$f_\ell' > 7(\ell-1)\binom{r-1}{\ell-1}t^{r-\ell}q^{\ell-2}f_2 \tag{7.18}$$

$$f_\ell' > 6(\ell-1)(r-1)t^{r-2}f_\ell \tag{7.19}$$

$$f_\ell' > 3(\ell-1)(r-1)\binom{r-1}{\ell-1}t^{2r-\ell-2}q^{\ell-2}f_v \tag{7.20}$$

7.4.1 Choosing the error functions

As before, all the variation equations hold because we choose α and β sufficiently large depending on r.

Since f_ℓ, f_v, f_ℓ'' and f_v'' are $\tilde{O}\left(D^{\beta\zeta^{r-1}+o(1)}\right)$, choosing ζ and δ such that $2\beta\zeta^{r-1} < \delta$ (for the estimates in (7.9) and (7.10) to hold) and

$$2\beta\zeta^{r-1} < 1 - \frac{1}{(r-1)}, \tag{7.21}$$

we obtain $f_\ell'', f_v'' = o\left(D^{1-\frac{1}{r-1}}\right)$.

7.4.2 Applying martingale variation inequalities

7.4.2.1 Controlling the number of vertices

Proof of Claim 7.4: Bound on $|V(i)|$.

Again, we will use the computations from the last chapter for brevity. Choosing ζ such that $2\beta\zeta^{r-1} < 1 - \frac{3}{2(r-1)}$, we have $1 - \frac{1}{r-1} - 2\beta\zeta^{r-1} > \frac{1}{2(r-1)}$. As before, by Hoeffding-Azuma inequality, the probability that $Z_V(m)$ is positive for some $m \leq i_{\max}$ is at most

$$\exp\left(-\tilde{\Omega}\left(\frac{ND^{-\frac{1}{r-1}}}{(f_2+f_v')^2}\right)\right) = \exp\left(-\tilde{\Omega}\left(D^{1-\frac{1}{r-1}-2\beta\zeta^{r-1}}\right)\right)$$
$$\leq \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

The same applies for the supermartingale

$$Y_V := -|V(i)| + Nq - ND^{-\delta} f_v,$$

which gives us the lower bound on $|V(i)|$. □

7.4.2.2
How many ℓ-edges a fixed vertex v belongs to

Proof of Claim 7.5: Upper bound on $d_\ell^{\pm}(v)$. Now to infer the upper bounds in Claim 7.5, we note that

$$-\Delta s_\ell^{\pm} - D^{\frac{\ell-1}{r-1}-\delta}\Delta f_\ell \leq |\Delta Z_\ell^{\pm}(v)| \leq \Delta d_\ell^{\pm}(v)$$

and

$$\Delta d_\ell^{+}(v) \leq D_{2\uparrow\ell+1} \leq 1,$$
$$\Delta d_\ell^{-}(v) \leq O\left(C_{\ell,2\to1}\right) = O\left(D^{\frac{\ell-1}{r-1}-\frac{1}{2(r-1)}}\right)$$
$$\Delta s_\ell^{\pm} + D^{\frac{\ell-1}{r-1}-\delta}\Delta f_\ell = O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right).$$

Since $Z_\ell^{\pm}(0) = -D^{\frac{\ell-1}{r-1}-\delta}$, $\frac{D^{\frac{\ell}{r-1}}}{N} = o\left(D^{\frac{\ell-1}{r-1}-\frac{1}{2(r-1)}}\right)$ and $D^{\frac{\ell-1}{r-1}-\delta} = o\left(i_{\max}\cdot\frac{D^{\frac{\ell}{r-1}}}{N}\right)$, we obtain that the probability that $Z_\ell^{\pm}(v)$ is positive at a time $m \leq i_{\max}$ is at most

$$\exp\left\{-\tilde{\Omega}\left(\frac{\left(D^{\frac{\ell-1}{r-1}-\delta}\right)^2}{ND^{-\frac{1}{r-1}}\cdot\frac{1}{N}D^{\frac{\ell}{r-1}}\cdot D^{\frac{\ell-1}{r-1}-\frac{1}{2(r-1)}}}\right)\right\}$$
$$\leq \exp\left\{-\tilde{\Omega}\left(D^{\frac{1}{2(r-1)}-2\delta}\right)\right\} \leq \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

Note that we can choose $2\delta < \frac{1}{2(r-1)} - \frac{1}{2(r-1)+\sigma}$ to obtain the last expression. $\square$

Remark 7.4. To obtain the lower bounds in Claim 7.5 we use the asymmetric version of Hoeffding-Azuma inequality to the supermartingales

$$\begin{aligned} Y_\ell^+(v) &:= - d_\ell^+(v) + s_\ell^+ - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, \qquad &&\text{for } 2 \le \ell \le r-1; \\ Y_\ell^-(v) &:= - d_\ell^-(v) + s_\ell^- - D^{\frac{\ell-1}{r-1}-\delta} f_\ell, \qquad &&\text{for } 2 \le \ell \le r. \end{aligned}$$

For $Y_\ell^-(v)$, we have

$$-O\left(D^{\frac{\ell-1}{r-1}-\frac{1}{2(r-1)}}\right) \le |\Delta Y_\ell^-(v)| \le O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right)$$

and the same calculations above apply.

For $Y_\ell^+(v)$, we have

$$-O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right) \le |\Delta Y_\ell^+(v)| \le O\left(\frac{D^{\frac{\ell}{r-1}}}{N}\right)$$

and, by the Hoeffding-Azuma inequality, the probability that Y_ℓ^+ is ever positive is at most

$$\exp\left\{-\tilde{\Omega}\left(\frac{\left(D^{\frac{\ell-1}{r-1}-\delta}\right)^2}{ND^{-\frac{1}{r-1}} \cdot \left(\frac{1}{N}D^{\frac{\ell}{r-1}}\right)^2}\right)\right\} \le \exp\left\{-\tilde{\Omega}\left(N \cdot D^{-\frac{1}{r-1}}\right)\right\} \le \exp\left(-(\log N)^{1+\Omega(1)}\right).$$

Bibliography

[1] THOMAS G. KURTZ. **Solutions of ordinary differential equations as limits of pure jump Markov processes.** Journal of Applied Probability, 7 (1):49–58, 1970.

[2] R. M. KARP; M. SIPSER. **Maximum matchings in sparse random graphs.** In: PROCEEDINGS OF THE TWENTY-SECOND ANNUAL IEEE SYMPOSIUM ON FOUNDATIONS OF COMPUTING, p. 364–375, Nashville, TN, USA, 1981.

[3] NICHOLAS C. WORMALD. **Differential equations for random processes and random graphs.** Annals of Applied Probability, 5 (4):1217–1235, 1995.

[4] NICHOLAS C. WORMALD. **The differential equation method for random graph processes and greedy algorithms.** In: LECTURES ON APPROXIMATION AND RANDOMIZED ALGORITHM, p. 73–135, PWN, Warsaw, 1999.

[5] NICHOLAS C. WORMALD. **Analysis of greedy algorithms on graphs with bounded degrees.** Discrete Mathematics, 273 (1-3):235–260, 2003.

[6] PATRICK BENNETT; TOM BOHMAN. **A note on the random greedy independent set algorithm.** Random Structures & Algorithms, 49 (3):479–502, 2016.

[7] NOGA ALON; JOEL H. SPENCER. **The Probabilistic Method.** Wiley Publishing, 4th edition, 2016.

[8] JEONG HAM KIM. **The Ramsey number R(3,t) has order of magnitude $t^2/\log t$.** Random Structures & Algorithms, 7 (3):173–207, 1995.

[9] AJTAI, M.; KOMLÓS, J.; SZEMERÉDI, E.. **A note on Ramsey numbers.** Journal of Combinatorial Theory. Series A, 29 (3):354–360, 1980.

[10] AJTAI, M.; KOMLÓS, J.; SZEMERÉDI, E.. **A dense infinite Sidon sequence.** European Journal of Combinatorics, 2 (1):1–11, 1981.

[11] JAMES B. SHEARER. **A note on the independence number of triangle-free graphs**. Discrete Mathematics, 46 (1):83–87, 1983.

[12] TOM BOHMAN. **The triangle-free process**. Advances in Mathematics, 221 (5):1653–1677, 2009.

[13] TOM BOHMAN; PETER KEEVASH. **Dynamic concentration of the triangle-free process**. In: THE SEVENTH EUROPEAN CONFERENCE ON COMBINATORICS, GRAPH THEORY AND APPLICATIONS. CRM SERIES, VOL 16., p. 489–495, Edizioni della Normale, Pisa, 2013.

[14] GONZALO FIZ PONTIVEROS; SIMON GRIFFITHS; ROBERT MORRIS. **The triangle-free process and the Ramsey number R(3,k)**. Memoirs of the American Mathematical Society, p. arXiv:1302.6279, to appear.

[15] TOM BOHMAN; PETER KEEVASH. **The early evolution of the H-free process**. Inventiones mathematicae, 181 (2):291–336, 2010.

[16] LUTZ WARNKE. **On Wormald's differential equation method**. arXiv e-prints, p. arXiv:1905.08928, May 2019.

[17] J. ROBERT JOHNSON; TREVOR PINTO. **The Q_2-free process in the hypercube**. arXiv e-prints, p. arXiv:1804.09029, 2018.

[18] TREVOR PINTO. **Extremal Problems on the Hypercube**. PhD thesis, Queen Mary University of London, London, UK, 2016.

[19] DAVID A. FREEDMAN. **On tail probabilities for martingales**. The Annals of Probability, 3 (1):100–118, 1975.

[20] MORRIS W. HIRSCH; STEPHEN SMALE; ROBERT L. DEVANEY. **Differential Equations, Dynamical Systems, and an Introduction to Chaos**. Academic Press, 3rd edition, 2012.

A
Gamma function

The goal of this appendix is to give a proof of the lemmas used in Chapter 5 about the gamma function. Just for completeness, we remember the definition of Gamma function and the statement of the lemmas.

Definition A.1 (Gamma function)**.** *The gamma function is defined for complex numbers with positive real part by*

$$\Gamma(z) = \int_0^\infty u^{z-1} e^{-u}\ du.$$

Lemma A.1. *For all complex numbers a, b with positive real part the following equation holds*

$$\int_0^1 (1 - x^a)^b\ dx = \frac{\Gamma\left(1 + \frac{1}{a}\right)\Gamma(b+1)}{\Gamma\left(b + 1 + \frac{1}{a}\right)}.$$

Lemma A.2. *Let α be a positive real number. Then*

$$\lim_{n\to\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)\cdot n^\alpha} = 1.$$

To prove the first lemma, we need to show a property of the gamma function and we will introduce the beta function. Then, the result will follow from standard calculus computations.

By integrating by parts, with $v = t^z$ and $u = -e^{-t}$, we have

$$\Gamma(z+1) = \int_0^\infty t^z e^{-t}\ dt = \left[-e^{-t}t^z\right]_{t=0}^\infty + z\int_0^\infty u^{z-1}e^{-t}\ dt$$

and, using that $\lim\limits_{t\to\infty} e^{-t}t^z = 0$, we conclude that

$$\Gamma(z+1) = z\Gamma(z). \tag{A.1}$$

Definition A.2 (Beta function)**.** *The beta function is defined for complex numbers x and y with positive real part by*

$$B(x, y) = \int_0^1 u^{x-1}(1-u)^{y-1}\ du.$$

Let x and y be complex numbers with positive real part. We will see that

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}. \tag{A.2}$$

We will prove that $\Gamma(x)\Gamma(y) = \Gamma(x+y)B(x,y)$. Begin with

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \left(\int_0^\infty e^{-u}u^{x-1}\,du\right)\cdot\left(\int_0^\infty e^{-v}v^{y-1}\,dv\right)\\ &= \int_0^\infty\int_0^\infty e^{-u-v}u^{x-1}v^{y-1}\,du\,dv,\end{aligned}$$

and changing variables by $u = f(z,t) = zt$ and $v = g(z,t) = z(1-t)$ we have that

$$\begin{aligned}\Gamma(x)\Gamma(y) &= \int_0^\infty\int_0^1 e^{-z}(zt)^{x-1}(z-zt)^{y-1}\Big|J(z,t)\Big|\,dt\,dz\\ &= \int_0^\infty\int_0^1 e^{-z}z^{x+y-1}t^{x-1}(1-t)^{y-1}\,dt\,dz,\end{aligned}$$

where $|J(z,t)| = \begin{vmatrix} t & z \\ 1-t & -z\end{vmatrix} = z$ is the Jacobian determinant of the change of variables $u = f(z,t)$ and $v = g(z,t)$. Then we conclude

$$\Gamma(x)\Gamma(y) = \left(\int_0^\infty e^{-z}z^{x+y-1}\,dz\right)\left(\int_0^1 t^{x-1}(1-t)^{y-1}\,dt\right) = \Gamma(x+y)B(x,y).$$

Proof of Lemma A.1. Changing variable by $x = u^{1/a}$ we have

$$\int_0^1 (1-x^a)^b\,dx = \int_0^1 \frac{1}{a}u^{\frac{1}{a}-1}(1-u)^b\,du,$$

and then using the definition of beta function, (A.2) and (A.1) respectively, we obtain

$$\int_0^1 (1-x^a)^b\,dx = \frac{1}{a}B\left(\frac{1}{a}, b+1\right) = \frac{\frac{1}{a}\Gamma\left(\frac{1}{a}\right)\Gamma(b+1)}{\Gamma\left(b+1+\frac{1}{a}\right)} = \frac{\Gamma\left(1+\frac{1}{a}\right)\Gamma(b+1)}{\Gamma\left(b+1+\frac{1}{a}\right)}.$$

□

To prove the second lemma we need to use Stirling's approximation formula

$$\Gamma(z) = \sqrt{\frac{2\pi}{z}}\left(\frac{z}{e}\right)^z\left(1+O\left(\frac{1}{z}\right)\right) \tag{A.3}$$

Proof of Lemma A.2. Using (A.3) we have

$$\frac{\Gamma(n+\alpha)}{\Gamma(n)} \sim \sqrt{\frac{n}{n+\alpha}}\frac{\left(\frac{n+\alpha}{e}\right)^{n+\alpha}}{\left(\frac{n}{e}\right)^n} = \sqrt{\frac{n}{n+\alpha}}\;e^{-\alpha}\;\left(\frac{n+\alpha}{n}\right)^n\;(n+\alpha)^\alpha.$$

Then, remembering that $\left(1+\frac{\alpha}{n}\right)^n \to e^\alpha$ when $n \to \infty$, we conclude

$$\lim_{n\to\infty} \frac{\Gamma(n+\alpha)}{\Gamma(n)\cdot n^\alpha} = e^{-\alpha} \cdot \left(\lim_{n\to\infty} \left(\frac{n+\alpha}{n}\right)^n\right) \cdot \left(\lim_{n\to\infty} \left(\frac{n+\alpha}{n}\right)^{\alpha-\frac{1}{2}}\right) = 1.$$

□