Lehrstuhl für Effiziente Algorithmen
Fakultät für Informatik
Technische Universität München

# Threshold Phenomena in Branching Trees and Sparse Random Graphs

**Ulrich Voll**

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

| | |
|---|---|
| Vorsitzender: | Univ.-Prof. T. Nipkow, Ph.D. |
| Prüfer der Dissertation: | |
| 1. | Univ.-Prof. Dr. A. Steger |
| 2. | Univ.-Prof. Dr. A. Goerdt, Technische Universität Chemnitz |

Die Dissertation wurde am 26.9.2001 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 21.1.2002 angenommen.

# Abstract

We consider phase transitions in random graphs ($\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$) with constant average degree, like the sudden appearance of the giant component or the giant $k$-core ([PSW96, GM00]) at some critical average degree. Small neighbourhoods in such random graphs closely resemble branching trees. Frequently, the exact numerical values of the critical average degrees and the expected sizes of the aforementioned giant subgraphs can be 'predicted' in a semi-heuristic manner from studying 'corresponding' branching trees instead, which is far simpler than the rigorous analysis of the phase transition in the random graphs. It is a major goal to turn this observation, which - following [PSW96] - we shall call the 'Branching Tree Connection', into a rigorous proof technique. Goerdt and Molloy have achieved this for the $k$-core in the model of random faulty configurations in [GM00].

We prove – in the $\mathcal{G}_{n,p}$ model – the sudden appearance of a giant subgraph that is 'almost the $k$-core', its size being sharply concentrated around what is predicted by the Branching Tree Connection. Our proof, based on the Principle of Deferred Decisions and a new application of the Simple Concentration Bound, essentially employs the same recursive equations as used for analysing a 'corresponding' phase transition in branching trees, thus providing structural insight into why the Branching Tree Connection predicts the correct values. Adapting ideas from [GM00] to the model $\mathcal{G}_{n,m}$, we show that the aforementioned subgraph which is 'almost the $k$-core' can be 'purged', yielding a giant $k$-core upon deletion of only $o(n)$ nodes.

Motivated by a new phase transition phenomenon in branching trees we have found a novel subgraph of $k$-partite graphs, the *magic subgraph*. We can empirically demonstrate its sudden appearance. Both critical average degree and size are in good accordance with the Branching Tree Connection. The fact that it appears at a critical average degree of $4.91(\ldots)$ suggests a relation with the Antivoter Phenomenon ([PW89]). Moreover, when it appears, the magic subgraph seems to be 'almost uniquely $k$-colourable', which may turn out to be relevant for understanding the threshold for $k$-colourability in non-$k$-partite random graphs with constant average degree.

# Acknowledgements

First of all I would like to express my gratitude to Prof. Dr. A. Steger, my Ph.D. supervisor, who introduced me to the subject of random graphs and combinatorial phase transitions, constantly supporting and guiding my work. Many thanks are also due to Prof. Dr. E.W. Mayr. I was provided with first class resources.

I am grateful to all present and former members of our group LEA (Lehrstuhl für Effiziente Algorithmen, Fakultät für Informatik der Technischen Universität München), who have made for such a pleasant working environment. Dr. A. Bernasconi, Dr. S. Bischof, Prof. Dr. T. Erlebach, J. Ernst, Dr. T. Friedetzky, Dr. S. Gerke, A. Hall, Dr. V. Heun, K. Holzapfel, I. Rohloff, Dr. M. Scharbrodt, H. Täubig and Dr. J. Träff provided support and advice during many fruitful discussions. I am particularly indebted to T. Schickinger for his extremely helpful comments.

Prof. Dr. N. Wormald introduced me to the Branching Tree Connection. The central observation that there is a phase transition at $c = 4.91\ldots$ in coloured branching trees, eventually leading to the definition and 'discovery' of the magic subgraph, emerged from a very fruitful discussion with him in 1998. The present results could not have been achieved without him.

My dear brother C. Voll, Dr. D. Günther, J. Pratschke and many other friends kept me on course and encouraged me to continue with my research when it was most needed, they were grand coaches.

# Contents

# Chapter 1

# Introduction

The main focus our work is understanding combinatorial phase transitions in sparse random graphs and their algorithmic implications. We regard our results as contributions towards understanding the average case behaviour of $k$-graph-colourability for $k \geq 3^*$, which has attracted a lot of attention in the theory of random graphs. Determining, and understanding the mechanisms behind, the so called *$k$-colourability threshold* is a central open research problem posed by P. Erdős, according to M. Molloy in [Mol01]. The aforementioned reference is a recent survey on the subject including a thorough exposition of the closely related threshold phenomenon in random $k$-SAT (satisfiability). We encourage the reader to consult [Mol01] as an excellent supplement to our introductory sections.

Before seriously getting started we will outline important aspects of the subject of 3-colouring random graphs in Section 1.1, and point out some connections to statistical physics, in Section 1.2. In Section 1.3 we will introduce the Branching Tree Connection, and Section 1.4 contains a synopsis of the thesis.

## 1.1 3-Colouring and Phase Transitions

It is well known that the decision problem ($\chi$ denotes the chromatic number)

$$\chi(G) \stackrel{?}{\leq} 3$$

is NP-complete for graphs $G$ on $n$ nodes. Not even an $n^{1-\varepsilon}$-approximation algorithm for $\chi$ exists unless $co - \mathsf{RP} = \mathsf{NP}$, according to [FJ96]. For tripartite graphs the decision problem is easy, yet finding 3-colourings of such

*Generalisations to $k > 3$ frequently are analogous if not straightforward and we shall therefore often assume $k = 3$.

graphs is still NP-hard, in the (loose) sense that a polynomial time algorithm able to explicitely find 3-colourings could be employed in a rather straightforward manner to solve the above decision problem. Note that the existence of polynomial randomised algorithms for the decision problem in the general case, and for the constructive problem in the tripartite case, would imply RP = NP. Thus, considering the standard assumptions of complexity theory, there is little hope for algorithms, both randomised and deterministic, solving those problems efficiently, i.e. in polynomial time (w.r.t. the number of nodes $n$), in the *worst case*.

The picture looks rather different in the *average case*. We consider random graphs parametrised with their average degrees $c = c(n)$. We shall see below in Section 2.2 that the decision problem is indeed trivial for 'most' values of $c$ in the following sense. For some $c_{\mathrm{crit}}(n)$ answering 'yes' when $c < c_{\mathrm{crit}}$ and 'no' for $c > c_{\mathrm{crit}}$ is correct with probability approaching one, asymptotically in $n$. The rate of convergence may depend on $c$ and one expects it to deteriorate the closer $c$ gets to $c_{\mathrm{crit}}$. Unfortunately the exact asymptotics of $c_{\mathrm{crit}}(n)$ are unknown. Yet $\liminf c_{\mathrm{crit}}(n)$ and $\limsup c_{\mathrm{crit}}(n)$ provably lie between 3.85(...) and 4.99(...), and $c_{\mathrm{crit}}(n)$ is conjectured to converge to an absolute constant, c.f. Section 2.5.

There is a (randomised) heuristic for 3-colouring random tripartite graphs that seems to be inefficient only for some critical average degrees of approximately 5 ([PW89], our Section 2.6). There also are algorithms that provably find colourings for all average degrees greater than some (rather large) constant (Section 2.6, [BS95], also [AK94], [Alo98]). Deciding 3-colourability is trivial with high probability, when the average degree is less than some other constant 3.35(...), c.f. Sections 2.3 and 2.4.

Note that for graphs the property of being 3-colourable is *non-local* in the sense that the presence of certain small subgraphs (such as $K_4$) may imply non-3-colourability but the converse is not true, c.f. the existence of uniquely 3-colourable graphs with large girth (Section 2.6).

Very loosely speaking, a combinatorial phase transition occurs if the typical behaviour of certain (sets of) combinatorial objects associated with the random graph change discontinuously (or at least non-analytically), as a function of the average degree $c$. In view of the sudden jump in chromatic number described above, the set of proper 3-colourings is such a combinatorial object. The well known giant component phenomenon ([ER60]) and the appearance of a giant subgraph called 3-core ([PSW96], explained below in Sections 2.3 and 2.4) are fairly well understood examples for phase transitions. In both those latter phenomena certain *subgraphs* of 'giant' size $\theta(n)$ suddenly appear.

When the chromatic number jumps to $\chi = 4$ a non-empty (edge-)minimally 4-chromatic subgraph, possibly giant, must have appeared. It was conjectured by Bollobás that the jump in chromatic number should coincide with the appearance of the 3-core, because the 3-core might contain a minimally 4-chromatic subgraph. This turned out to be wrong. Still the search for the appearance of some *subgraph* that is, or implies the existence of, a minimally 4-chromatic subgraph continues (c.f. [MR99, Mol01]).

Some may argue that average case analysis has some *ad hoc* flavour, since one distribution on the set of inputs is as good as any other, the uniform distribution being the only canonical choice. However, having shown that a problem is NP-hard we know it will be difficult to solve under the standard assumptions of complexity theory, but we do not really *understand* why. Is it just because some 'artificially constructed' inputs seem to 'confuse' all algorithms feasible on Turing machines? Are all inputs equally hard? Or is it because some subset of inputs is 'difficult', and if this is true, what makes those inputs special?

Average case analysis employing distributions that are concentrated on certain subsets of the input set may help in answering those questions, by 'focusing' on subsets of the inputs [†]. It is generally believed that combinatorial phase transitions are somehow connected to those questions, in that random inputs generated according to a distribution with parameters (like the average degree $c$) tend to be 'easy', when the parameters are bounded away from phase transitions and 'difficult' otherwise. This suggests that generating inputs with the parameters set upon the critical values should be an effective way of generating 'typical' hard inputs. However, there is no general rigorous theory justifying this belief. Most certainly such a theory presupposes a rigorous definition and classification of phase transitions. We believe that a lot could be learned from statistical physics, where there is a theory of phase transitions developed by mathematical physicists. We will therefore give a short and very incomplete overview in the next subsection.

## 1.2 Colouring and Statistical Physics

We include this little excursion to demonstrate that using the term 'phase transition' in the context of random graphs is more than a superficial analogy motivated by the fact that 'something suddenly jumps'.

[†] One should keep in mind, though, that there may be other large subsets of the set of feasible inputs that are also hard, but never in the 'focus' of the class of probability distributions considered.

There are many connections between algorithmic complexity issues, as studied by computer scientists and discrete mathematicians, and phase transitions, as studied by physicists. There may be hope for a 'unifying' theory, that may yield certain inapproximability results in the average case analysis of NP-hard problems by studying the singularities of Gibbs measures as explained below. As far as we are aware there is no such theory. However note the recent results by Brightwell and Winkler in [BW99b, BW99a, BW00].

In order to draw (further) attention to this situation, we find it worthwhile to explain our limited understanding of the analogies, if nothing else, offering a different perspective on the problem and explaining some terminology as we go along.

Studying the 3-colouring problem of a graph $G$ is equivalent to studying the *anti-ferromagnetic 3-spin Potts model* associated with $G$. The *states*, i.e. families of three different 'spins' labelled by the nodes of $G$, are nothing but colourings of $G$, not necessarily proper. The *Hamiltonian* $H$ is a mapping from colourings to the number of monochromatic edges therein. To a physicist the Hamiltonian is the energy of the states. The *Gibbs measure* is a measure proportional to $\exp(-1/T \cdot H)$ on the set of all colourings, where the parameter $T$ may be interpreted as a *temperature*. The specific structure of the system is encoded in (the functional form of) the Hamiltonian, and the Gibbs measure describes the 'occupancy' of states in thermodynamic equilibrium at temperature $T$. When the temperature $T$ goes to zero, the Gibbs measure is the uniform distribution on the colourings with the minimum number of monochromatic edges, also called the *ground states of* $H$.

There is plenty of evidence in the theoretical computer science literature for the fruitful interplay between statistical physics and theoretical computer science. Theoretical computer science has benefited from ideas brought up by physicists, frequently cast in non-rigorous form. We agree with Talagrand ([Tal98]) that 'the physicists' tend to produce non-rigorous speculative results, settling for 'proofs' relying on assumptions that are implicit, overly optimistic and *ad hoc*. At the same time, ideas both for problems and even for proof techniques (e.g. hard-core lattice gas model, Dobrushin uniqueness, c.f. [BD97]) have had a strong impact on research in discrete probability and theoretical computer science.

In Physics there is a 'theory of critical phenomena' or 'scaling', which seems to be not consistently mathematically rigorous (however, see [Bax89] for mathematically rigorous results on statistical physics). Our understanding is that quantities like 'correlation-lengths' derived from the Gibbs measure, diverge as a function of the parameters according to some 'universal scaling laws', like for example $(T - T_{\text{crit}})^{\alpha}$, believed to be fairly independent of many specific details of the system (i.e. Hamiltonian) under consideration.

We remark that Bollobás et.al. in [BBC$^+$99] relate their results to 'scaling'.

Based on our acquaintance with D. Ruelle's seminal *Thermodynamic Formalism* ([Rue78]), we will therefore sketch his theory of phase transitions, and state some observations, in a rather non-rigorous fashion. Carrying out the thermodynamic limit, to be explained presently, is a rather subtle and difficult issue, our point is that there exists a mathematically rigorous theory.

The Potts-model Gibbs measure has been studied on regular lattices such as tori with $n = l^d$ nodes, where $d = 1, 2, 3, (\ldots)$. In the limit $n \to \infty$ the corresponding sequence of Gibbs measures induces a translation invariant *infinite volume Gibbs measure* on colourings of $\mathbf{Z}^d$. This process is referred to as the *thermodynamic limit*. This infinite volume Gibbs measure depends on $T, d$ and all 'macroscopic' parameters the Hamiltonian $H$ might additionally depend on. A *phase transition* is a point of non-analyticity (in some appropriate weak sense) of the infinite volume Gibbs measure.

Note that the functional form of the Gibbs measure, $\exp(-1/T \cdot H)$, suggests an essential singularity at $T = 0$. Any deterministic algorithm finding a ground state efficiently could be employed to decide 3-colourability in polynomial time, which would imply $\mathsf{P} = \mathsf{NP}$. Even a randomised algorithm efficiently sampling from the Gibbs measure at $T = 0$ would imply $\mathsf{RP} = \mathsf{NP}$. Both $\mathsf{P} = \mathsf{NP}$ and $\mathsf{RP} = \mathsf{NP}$ are generally believed to be wrong. Yet there are at first sight efficient Markov Chain Monte Carlo (MCMC) algorithms sampling from the Gibbs measure at finite temperatures. Not surprisingly in the light of the above complexity-theoretic considerations, the performance of those samplers breaks down when $T \to 0$. It is well known, at least empirically, that setting the parameters close to points of phase transitions may spoil the performance of otherwise efficient MCMC samplers. This phenomenon is referred to as *critical slowing down* (c.f. [GJ97]).

The singularities appear only in the thermodynamic limit $n \to \infty$, but for systems with large constant $n$ the finite volume Gibbs measures vary rapidly when the parameters are close to critical points. Discontinuities appear slightly 'rounded off'. The physicists' intuition is that a system with many, say $10^{23}$, atoms resembles very closely the idealised 'limiting system' with an infinite number of atoms.

Note that the sequence of finite lattices 'everywhere look the same, locally and for all $n$'. More precisely they have a translation invariance, or isotropy, which seems to play an important role for the advent of phase transitions.

Now let us compare this to the situation in random graphs with average degree $c(n)$. They also 'everywhere look the same, locally', but in a stochastic sense. When $c(n)$ is *not constant*, this will depend on $n$, but otherwise the random graphs, too, 'everywhere look the same, locally *and for all* $n$'. To a physicist, average degree means average coordination number and should

be an extensive variable in the sense of thermodynamics, that is, it should be independent of system size $n$. This is clearly satisfied by the tori ('lattice graphs') above. Random graphs with average degree explicitly depending on $n$ do not comply with this. Thus graphs with *constant* average degrees seem best suited for a physical interpretation of a 'thermodynamic limit'. We find it remarkable that (c.f. above) the jump in chromatic number happens in exactly this regime of constant average degree.

We conclude by observing that, when the tori resemble regular crystal lattices, our random graphs are resembling 'glass'. Indeed the area of *spin glasses* seems to be very closely related to our setting, c.f. Talagrand ([Tal98]).

## 1.3 Branching Trees and Random Graphs

We have studied random graphs with *constant average degree* $c$, and 'corresponding' branching trees with expected growth rate $c$. Small neighbourhoods in random graphs with constant average degree $c$ closely resemble branching trees with expected growth rate $c$.

Branching trees are far simpler to study than the 'corresponding' random graphs, mainly because disjoint sub-trees are independently and identically distributed as the branching tree itself. Moreover, branching trees also exhibit phase transitions intuitively corresponding to phase transitions in the random graphs. This correspondence is called[‡] the *Branching Tree Connection* in [PSW96], and has been observed by many researchers in one way or another. For example, the famous 'appearance of the giant component' ([ER60]) corresponds to the 'appearance of an infinite subtree' in the branching tree, and the 'appearance of a giant 3-core' ([PSW96]) corresponds to the 'appearance of an infinite (complete) binary subtree'. This correspondence predicts the discontinuous jump not only qualitatively. The critical values $c_{\text{crit}}$ coincide and the expected relative sizes at which the giant subgraphs appear are exactly equal to the probabilities that the branching trees contain the corresponding infinite subtrees. The Branching Tree Connection even holds when we consider random graph models that differ from the classical binomial or uniform ones, only the progeny distributions in the corresponding branching trees needs to be adapted accordingly ([MR95a, MR98, GM00]).

The Branching Tree Connection suggests a *non-rigorous* 'explanation' of the basic mechanism of the appearance of giant subgraphs, that can be put as follows. The random graph 'behaves similarly' to a disjoint union of independent branching trees, the roots of the ones containing infinite subtrees forming something analogous to the giant subgraph. Obviously, this is *not*

[‡]Actually, in [PSW96] the term 'branching process connection' is used.

what is going on in random graphs. Neighbourhoods are never infinite, they resemble branching trees only for radii much smaller than the diameter of the graph and, finally, the neighbourhoods of different nodes do overlap.

Our overall goal is turning this semi-heuristic 'explanation' into rigorous proofs for phase transitions. Our contribution can be summarised as follows. We provide a little theory for treating phase transitions in branching trees. Then we provide a novel technique, in close accordance to the treatment of the corresponding branching trees, for proving the 'appearance of giant *marked sets*' which are 'almost the same' as the giant $k$-core or the giant magic subgraph. For the appearance of the $k$-core we can actually show that the 'giant marked set' differs from a giant $k$-core only by $o(n)$ nodes, adapting the proof strategy of Goerdt and Molloy ([GM00]) to the $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model of random graphs. Inspired by the theory of branching trees we have discovered a new 'sudden appearance of a giant subgraph' phenomenon in random tripartite graphs. We called this subgraph the *magic subgraph*. Empirically, magic subgraphs as they appear in random graphs are 'almost uniquely' colourable, which makes them interesting candidates for explaining the jump in chromatic number (see our discussion on p. 3).

As explained above it may appear natural for a physicist to study phase transitions in random graphs with constant average degree $c(n) = c$. When analysing the average case behaviour of the chromatic number, such random graphs are the most interesting from a purely mathematical point of view, too. But note that there are other important combinatorial phase transitions where $c$ explicitely depends on $n$, like the appearance of a Hamilton cycle. Actually, this problem helps to illustrate why it is promising to study small neighbourhoods in random graphs. Deciding Hamiltonicity is well known to be NP-complete. Yet it turns out that a random graph almost surely contains a Hamilton cycle, as soon as the average degree is such that every node has degree at least two, almost surely. For deciding the latter it clearly suffices to merely inspect all 1-neighbourhoods of the nodes, which can be performed in polynomial time. Note that the degrees of the nodes are distributed approximately like 'fairly independent' $\mathrm{Po}_c$ random variables, which can be interpreted as a 'Branching Tree Connection for 1-neighbourhoods'.

Remember that it would be appealing to explain the jump in chromatic number by the appearance of some appropriate giant subgraph which might be much easier to decide than the NP-complete problem whether the graph is 3-colourable. Our work shows that the appearance of certain giant subgraphs may be ('approximately') decided upon inspecting the $r$-neighbourhoods of the nodes, where $r = r(n)$ grows slowly with $n$, not unlike the aforementioned 'approximation' to Hamiltonicity based upon inspection all 1-neighbourhoods.

## 1.4 Synopsis

In **Chapter 2** we define basic terms, review some probabilistic concepts, and report on a number of results that are relevant in the context of our work. Only Section 2.3 contains some original definitions and insights that are so central to our work that we wanted them to be available as soon as possible.

**Chapter 3** contains a rigorous treatment of two important phase transitions in branching trees, using certain recurrence equations. These two phase transitions in branching trees can be regarded as the 'analogues' of the appearance of the $k$-core and the magic subgraph (see below) in random graphs. We include some additional material concerning branching trees.

In **Chapter 4** we formally define the *magic subgraph*, a new type of subgraph in $k$-partite graphs that is somewhat similar to the $k$-core. We discuss certain deterministic properties of magic subgraphs. In Section 4.3 we give striking empirical evidence for the sudden appearance of a giant magic subgraph at an average degree $c = 4.91(\ldots)$. The empirically observed numerical values of both the threshold and the expected size closely coincide with what would be expected from the 'corresponding branching tree theory' as laid out in Section 3.2. Moreover, this coincidence manifests itself in a way completely analogous to how empirically observed $k$-cores coincide with what would be expected from the 'corresponding branching tree theory' as laid out in Section 3.1. We present evidence that empirically observed magic subgraphs appear to be 'almost uniquely colourable'.

In **Chapter 5** we show how to prove the 'sudden appearance of a giant subgraph that is almost the $k$-core' in the $\mathcal{G}_{n,p}$-model with average degree $c$, employing *the same recurrence equations* as used for the analysis of the 'corresponding' phase transition discussed in Section 3.1. This gives a good deal of structural insight into why the Branching Tree Connection works so well. We finally 'translate' our results to the closely related $\mathcal{G}_{n,m}$-model with average degree $c$.

The aforementioned 'giant subgraph that is almost the $k$-core' is indeed closely related to the giant $k$-core. In **Chapter 6** we extend our results from Chapter 5 to a new proof for the sudden appearance of a giant $k$-core in the $\mathcal{G}_{n,m}$-model with average degree $c$, adapting the proof strategy of Goerdt and Molloy ([GM00], appearance of a giant $k$-core in random faulty configurations).

In **Chapter 7** we explain to what extent our new proof techniques that enabled us to give a new proof for the appearance of a giant $k$-core, can be employed for proving the empirically observed sudden appearance of a giant magic subgraph. In particular, we report on where we failed in generalising the material from Chapter 6 to the magic subgraph and discuss alternative

strategies. The appearance of the (empirically, 'almost') uniquely colourable magic subgraph could well be regarded as candidate mechanism for the aforementioned 'jump' of the chromatic number of random graphs from $k$ to $k+1$, if it were not only well defined for $k$-partite graphs. We lay out what we can say about comparing the $k$-partite and the non-$k$-partite model in Section 7.3. By 'discovering' the magic subgraph we have demonstrated that phase transitions in branching trees described by certain recursive equations sometimes 'indicate' the appearance of certain 'corresponding' giant subgraphs. We shall discuss a few recursive equations that *may* lead to the discovery of new giant subgraphs, including one recursive equation related to random 3-SAT.

In **Chapter 8** we summarise our results, state open questions and make suggestions for future research.

# Chapter 2

# Context and Related Work

In this chapter we will attempt to set the scene for later discussions by defining the terms and concepts that will be used and by reviewing the work of other authors. Once again, we refer those seeking additional information to the excellent overview provided by M. Molloy [Mol01].

In Section 2.1 we state various models of random graphs, and review a few probabilistic tools. Section 2.2 is devoted to the average case analysis of graph properties in random graphs with constant average degree, using Fourier Analysis. The probabilities of certain 'non-local' graph properties discontinuously jump as a function of the average degree. In Section 2.3 we introduce some specific concepts, such as the $k$-core and branching trees. Section 2.4 contains rigorous results for the 'sudden appearance of the 3-core' for two different models of random graphs. Section 2.5 deals with what is known about the average case analysis of 3-colourability as a function of the average degree. Section 2.6 is devoted to 3-colourability in the 3-partite case.

This chapter contains nothing that is new, apart from the definition and discussion of *ownership* and the *extended $k$-core* in Section 2.3. We have included this material at an early stage, such that we can describe the results of others in the language developed for our work.

## 2.1 Basic Probability Concepts

Here we will briefly review some important mathematical basics mainly for the sake of completeness and later reference. A reader familiar with the theory of random graphs, as laid out in the books by Bollobás [Bol85] or Janson et. al. [JŁR00], may safely skip most of this section.

We discuss various models of random graphs and how they are related, the Principle of Deferred Decisions and important concentration 'tools'.

### 2.1.1 Models of Random Graphs

We mainly consider simple undirected graphs $G = (V, E)$ on $|V| = n$ vertices. The number of edges will be (close to) $|E| = m = cn/2$, that is the average degree will be (close to) $c$. By $\mathcal{G} = \mathcal{G}_n$ we denote the set of all $2^{\binom{n}{2}}$ graphs on $n$ nodes. This will be the underlying probability space, more formally a sequence of probability spaces, in most future considerations. Apart from this we shall also consider so called *configurations*, closely related to multigraphs, to be explained below.

#### Binomial Random Graphs

Our favourite model is $\mathcal{G}_{n,p}$. This is just a shorthand notation indicating that we shall deal with a certain (sequence of) probability distribution on the set of all graphs $\mathcal{G}_n$. Random graphs 'in $\mathcal{G}_{n,p}$' are generated by retaining the edges of the complete graph $K_n$ with probability $p$, independently. Thus the aforementioned probability distribution is the product of $\binom{n}{2}$ Bernoulli distributions, each 'random bit' corresponding to a potential edge switched on or off.

Usually we will choose $p = p(n) = c/(n-1)$ in order to attain an *average* degree of $c$.

We shall later need random tripartite graphs with average degree $c$, denoted by $\mathcal{G}_{n,3,\frac{3c}{2n}}$. Each edge in the complete tripartite graph $K_{n/3,n/3,n/3}$ is retained with probability $\frac{3c}{2n}$, independently. Here $n$ will be assumed to be an integer multiple of 3.

#### Uniform Random Graphs

In the model $\mathcal{G}_{n,m}$ graphs are drawn uniformly from all graphs in $\mathcal{G}_n$ with exactly $m$ edges. Choosing $m = cn/2$ yields average degree $c$. Algorithmically, such graphs may be generated by *random graph processes* as follows.

- Start with the empty graph.
- As long as there are less than $m$ edges in the graph so far, select a consistent potential edge uniformly at random and add it.

A potential edge is *consistent* with the graph generated so far, if it does not introduce multiple edges when added. Clearly, this can be adapted to the tripartite case.

### Other Random Graphs

In the uniform model the only constraint is the number of edges. We also may prescribe a *degree sequence*, that is a sequence $(d_i)_i$ of integers counting the numbers of nodes with degree $i$. We deviate slightly from the standard definition, following [MR95a]. Instead of writing down the degrees of all nodes in increasing order, forming a string of length $n$, we count how often a certain number (i.e. degree) appears in this string, which is clearly equivalent. Note that it is also equivalent to consider the *integrated degree sequence* formed by the $D_i := \sum_{j \geq i} d_j$. We may also consider random graphs that are chosen uniformly amongst all graphs with a given degree sequence.

The aforementioned reference [MR95a] contains an interesting generalisation of the well known 'appearance of a giant component' for random graphs with a given degree sequence.

It is not easy to algorithmically construct such graphs with a given degree sequence uniformly at random. Therefore it has become standard practice to use the model of *random configurations*, and translate results back to the corresponding random graphs by 'translation tools' (see below).

- For each $i$ create $d_i$ 'meta-nodes', consisting of $i$ vertices each.
- Generate a random perfect matching on the vertices. This can be done by a random process similar to the one described above. Note that $\sum_i id_i = \sum_{v \in V} \deg(v) = 2m$ is even, whenever $(d_i)_i$ is the degree sequence of some graph.

We finally remark that Janson et. al. in [JKŁP93] have proposed a model of generating directed random graphs, where each pair of vertices is incorporated independently with probability $c$. This model seems to be particularly well adapted to analysis using generating functions. We shall make no further reference to this model.

### Relating the Models

The models $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$, when $m$ is (close to) the expected number of edges in $\mathcal{G}_{n,p}$, are not fully equivalent, yet they are very similar in spirit. Indeed there are 'translation tools' enabling us to carry over results from one model to the other, an in-depth-treatment may be found in the aforementioned references [Bol85] and [JŁR00]. We merely quote two facts from [Bol85] for later reference. Denote by $\mathcal{G}_n$ the set of all graphs on $n$ nodes and by $\mathcal{Q}^{(n)}$ a subset of $\mathcal{G}_n$. Strictly speaking $\{\mathcal{Q}^{(n)}\}_n$ is a family of respective subsets of $\mathcal{G}_n$, but we shall say that $\mathcal{Q}^{(n)}$ is an event. An event is *convex*, if for all $F \subseteq G \subseteq H$, $F \in \mathcal{Q}^{(n)}$ and $H \in \mathcal{Q}^{(n)}$ implies $G \in \mathcal{Q}^{(n)}$.

**Fact 1** *If the probability of a convex event $\mathcal{Q}^{(n)}$ converges to one in $\mathcal{G}_{n,p}$ then it also converges to one in $\mathcal{G}_{n,m}$, when $p = p(n)$ is such that the expected number of edges is $m = m(n)$.*

**Fact 2** *The probability of any event $\mathcal{A}^{(n)}$ in $\mathcal{G}_{n,m}$ is bounded by $O(m(n)^{1/2})$ times its probability in $\mathcal{G}_{n,p}$, when $p = p(n)$ is such that the expected number of edges is $m = m(n)$.*

Also, random configurations with degree sequence $(d_i)_i$ are closely related to random graphs with degree sequence $(d_i)_i$. The essential idea for a 'translation tool' is to contract 'meta-nodes' to single vertices. This will however yield multigraphs, with multiple edges and self-loops. First note that a random configuration *conditional* on no self-loops or multiple edges yields a random graph with degree sequence $(d_i)_i$, uniformly amongst all such graphs, because every graph is the 'contraction' of $\prod_i (d_i)!$ configurations. Moreover, it can be shown that the event 'no self-loops or multiple edges' has a constant probability. Thus any event that has probability $o(1)$ in the configuration model has probability $o(1)$ in the corresponding random graph model, too. See, for example, [MR95a] for a more thorough discussion.

### 2.1.2 Principle of Deferred Decisions

When working in the $\mathcal{G}_{n,p}$ and related models, we may encode the graphs as strings in $\{0,1\}^M$, $M = \binom{n}{2}$, each *independent* bit standing for a potential edge in $K_n$ being switched on or off. When analysing algorithms on the graphs in the course of 'time' it is sometimes convenient to think of the random choices being made at the very time when we 'ask' for their outcome. As long as we have not 'looked' at the corresponding potential edges, their outcomes are independent of all other bits. Instead of stating this more formally we will give a couple of examples.

The $n$ vertices of a graph are (labelled by) the integers in $[n]$. At each 'time' $t$ we consider the (potential) edges connecting node $t$ with $[t-1]$. This is tantamount to building up the graph by 'glueing' node $t$ to the 'graph so far'. This specific way of 'grouping' the edges is referred to as *vertex exposure*.

Consider a breadth first search (BFS) or depth first search (DFS) in the random graph. Instead of generating the full random graph and then running a BFS or DFS we can do the following. At 'time' $t$ we know that some nodes are fully explored, some are on the boundary and the rest are unexplored. According to the protocol (DFS or BFS with appropriate tie-breaking) we choose some node $v_t$ from the boundary. Then we 'ask' for the outcome of

the random bits corresponding to (potential) edges connecting $v_t$ with other boundary and unexplored nodes, moving $v_t$ from being 'boundary' to being 'explored' and the formerly 'unexplored' newly discovered neighbours to the boundary. Whatever the outcome of the random bits so far, the bits we 'ask' for at 'time' $t$ are completely 'fresh'.

Another 'BFS-like' *fanning-out process* in the above spirit is obtained when the boundary nodes to be expanded are chosen uniformly at random. We shall later refer to it as *Karp's fanning-out process*, see [Kar90].

### 2.1.3 Concentration Tools

In the asymptotic theory of random graphs we frequently encounter discrete random variables $X^{(n)} = \sum_{i \in I^{(n)}} Y_i^{(n)}$, counting certain combinatorial objects associated with the graphs ($n$ is the number of nodes). Think of the number of (proper!) colourings, or the number of nodes participating in a triangle, for example. Thus $X^{(n)}$ takes values in the non-negative integers. The indicators $Y_i^{(n)}$ in the sum constituting $X^{(n)}$ are usually *not independent*, which renders analysis more difficult.

Note that in the independent case due to the law of large numbers $X^{(n)}$ will be sharply concentrated around $\mathbb{E}[X^{(n)}]$, which itself is simply the sum of the $\mathbb{E}[Y_i^{(n)}]$. For sufficiently large $n$ the behaviour of $X^{(n)}$ would be basically determined by $\mathbb{E}[X^{(n)}]$.

Fortunately, linearity of expectation also holds in the dependent case, and we may still be able to calculate $\mathbb{E}[X^{(n)}]$ using the $\mathbb{E}[Y_i^{(n)}]$. Sometimes it is even possible to calculate the covariance matrix and we may get hold of $\mathrm{Var}[X^{(n)}]$ from the $\mathbb{E}[Y_i^{(n)} Y_j^{(n)}]$. If we can not calculate those quantities we might still be able to bound them. We will frequently write $X$ instead of $X^{(n)}$, etc., the dependence on $n$ being understood.

Knowledge of moments, even of the first moment alone, may yield more insight into the behaviour of $X$ than one would naively expect. The First and Second Moment Method to be explained presently are exploiting little more than the elementary Markov Inequality for analysing $\mathbb{P}[X = 0]$.

The aforementioned concentration of $X$ in the independent case is best captured by Chernoff Bounds when the $Y_i$ are independent indicators. As a rule of thumb, when correlations are absent, with high probability the deviation from $\mathbb{E}[X^{(n)}]$ is $\theta(\sqrt{n})$, when $\mathbb{E}[X^{(n)}] = \theta(n)$. When the $Y_i$ are in some sense 'weakly' dependent indicators, or even when $X$ is no longer a sum of indicators but a more general function of random variables that 'does not vary too much' (satisfying a Lipschitz condition, see below) concentration results are available. There are various techniques yielding 'Chernoff-like' concentration inequalities some of which we will briefly describe below.

Even when correlations are not small, but 'the right way round', there are interesting results, a formal discussion of which we consider out of scope. Note that there is not even a unique canonical definition of what should be understood by 'negatively or positively correlated'. We will limit ourselves to reproducing an informal but illustrative picture of the effect of correlations on sums of random variables due to B. Reed. Consider $n$ random indicator variables, say, each being one with probability $1/2$ but not independently.

The extreme case of positive correlation is that the first variable 'decides' and all the others go along. In that case the sum is concentrated on 0 and $n$, in particular, it is not concentrated around its mean.

For negative correlation each variable does the opposite of what the majority of the others have done. Up to rounding errors the sum is exactly concentrated on $n/2$, which is the mean.

Thus, intuitively, negative correlation improves concentration whereas positive correlation spoils it.

### The First and Second Moment Method

Let us now suppose that we were somehow able to find out, or at least bound, the asymptotic behaviour of $\mathbb{E}[X]$, and possibly also of $\mathrm{Var}[X]$. As mentioned above, this is usually achieved by considering cleverly-chosen sums of (possibly dependent) indicator random variables, using linearity of expectation. When $X$ is counting certain objects, we are particularly interested in $\mathbb{P}[X = 0]$, i.e. the probability that no such object exists. Applying the well known and elementary inequalities of Markov and Chebychev we can learn about $\mathbb{P}[X = 0]$, purely from our knowledge of $\mathbb{E}[X]$ and $\mathrm{Var}[X]$.

I) When $\mathbb{E}[X] \to 0$ then $\mathbb{P}[X = 0] \to 1$, by Markov's inequality.

   Proving that $\mathbb{P}[X = 0] \to 1$ using this implication is called the *First Moment Method*.

II) What if $\mathbb{E}[X]$ diverges, does this imply $\mathbb{P}[X = 0] \to 0$? As long as $X$ is reasonably concentrated around $\mathbb{E}[X]$ one should expect so. And indeed, employing Chebychev's inequality yields that $\mathbb{P}[X = 0] \to 0$ as long as $\mathrm{Var}[X]$ does not diverge too badly.

   Proving $\mathbb{P}[X = 0] \to 0$ by this implication is called the *Second Moment Method*.

Note that $\mathbb{E}[X] \to \infty$ in general does *not* imply $\mathbb{P}[X = 0] \to 0$, as can be seen from the following counterexample. Let $\mathbb{P}[X = i] \equiv 0$ except $\mathbb{P}[X = 0] = 1 - 1/n$ and $\mathbb{P}[X = 2^n] = 1/n$. Clearly, $\mathbb{E}[X]$ diverges but $\mathbb{P}[X = 0] \to 1$.

Intuitively, the First Moment Method fails to prove $\mathbb{P}[X = 0] \to 1$, when in the (unlikely!) case that $X \neq 0$ there are 'far too many' objects counted by $X$. Following Molloy [Mol01] we shall call this a 'jackpot-phenomenon'. This suggests a refinement of the First Moment Method. Use $\tilde{X}$ instead of $X$, counting only 'some' of the objects. If those selected objects are in some sense maximal, that is $\tilde{X} = 0 \Rightarrow X = 0$, there is hope that the 'jackpot' has been sufficiently reduced such that now $\mathbb{E}[\tilde{X}] \to 0$ yields the desired result, although $\mathbb{E}[X]$ itself still diverges.

### Concentration Inequalities

We will mainly follow McDiarmid ([HMRR98]) in our exposition. A concentration inequality or 'tail bound' associated with a random variable $X$ always has the single-sided form

$$\mathbb{P}\left[X^{(n)} - \mathbb{E}\left[X^{(n)}\right] \geq \delta(n)\right] \leq \varepsilon(n),$$

or the double-sided form

$$\mathbb{P}\left[\left|X^{(n)} - \mathbb{E}\left[X^{(n)}\right]\right| \geq \delta'(n)\right] \leq \varepsilon'(n).$$

The slower $\varepsilon$ and $\delta$ grow with $n$, the better the bound.

Chernoff Bounds apply to random variables that are sums of independent real random variables, frequently but not exclusively indicator random variables

$$X^{(n)} = \sum_{i \in [n]} Y_i^{(n)}.$$

Chernoff Bounds are the paradigm for all concentration inequalities. Newcomers to the field may be confused when finding not one Chernoff Inequality but many similar-looking inequalities in the literature. This is because for many applications one may trade off some sharpness in favour of *expressiveness*, that is simpler formulae for $\varepsilon(n)$ and $\delta(n)$. Thus we are really talking about a whole class of inequalities. It takes some time to grow accustomed to this vagueness and to learn how to handle it correctly. The same applies to all other inequalities stated in this paragraph. We will use the following version of the Chernoff bounds for later reference.

**Fact 3** *(McDiarmid [HMRR98, p. 200]). Let $X = \sum_{i=1}^{n} X_i$ the sum of i.i.d. 0/1 random variables, such that $\mathbb{E}[X] = \mu$. Then for all $\delta > 0$*

$$\mathbb{P}[X - \mu > \delta] \leq e^{-\frac{\delta^2}{2(\mu(1+1/3\cdot\delta/\mu))}}.$$

As we have seen above concentration inequalities may exist even when $X$ is no longer a sum but a more general function. Consider a sequence $\mathbf{X} = (X_1, \ldots, X_n)$ of $n$ independent random variables and a function $f$ from the product $\prod_{k=1}^{n} A_k$ of the respective ranges $A_k$ of the $X_k$ to the reals. We begin with the *Simple Concentration Bound.*

**Fact 4** *(McDiarmid [HMRR98, p. 206]). Using the above notation and definitions, when $f$ satisfies a Lipschitz condition, that is if for all $k$*

$$(L) \quad |f(x_1, \ldots, x_{k-1}, x, \ldots, x_n) - f(x_1, \ldots, x_{k-1}, x', \ldots, x_n)| \leq c_k,$$

*then for any $t > 0$, $\mathbb{P}[|f(\mathbf{X}) - \mathbb{E}[f(\mathbf{X})]| \geq t] \leq 2\exp(-\frac{2t^2}{\sum c_k^2})$.*

Note that the function 'sum' $(x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} x_i$ is a special case having Lipschitz constants $c_k \equiv 1$. The Simple Concentration Bound yields similar but slightly weaker inequalities than the Chernoff Bounds, the latter being especially tailored for the case of sums of independent variables.

Note that for the last inequality to hold, the $X_i$ still need to be independent. Even this may be relaxed using martingale techniques. In combinatorial applications we frequently encounter a 'limiting' random variable $Y$ factoring over the $(x_1, \ldots, x_n)$. The quantities

$$Y_k = Y_k(x_1, \ldots, x_k) := \mathbb{E}\,[f(x_1, \ldots, x_k, X_{k+1}, \ldots, X_n)]$$

yield 'intermediate' variables on the probability space of the $(x_1, \ldots, x_k)$ 'averaged out' over the respective futures $(x_{k+1}, \ldots, x_n)$. $Y_k$ may thus be regarded as a random variable on the probability space of all possible histories up to time $k$. In particular we have $Y_n = Y$ and $Y_0 \equiv \mathbb{E}[Y]$. When the $(Y_i - Y_{i-1})_i$ form a *martingale*, Azuma's Inequality yields concentration of $Y = Y_n$. We shall not elaborate on the concept of martingales. We shall merely remark that processes $(Y_i - Y_{i-1})_i$ arising from a single random variable $Y$ on a product space by conditioning in the above way are called *Doob martingales.*

**Fact 5** *(A simple version of Azuma's Inequality, folklore.) Let $(X_i)_{i=1}^{n}$ be* not necessarily independent *random variables, and let $f : \mathbb{R}^n \to \mathbb{R}$ satisfy the following.*

*For all possible histories $(x_1, \ldots, x_{k-1}, x_k)$ the expected increase is $0$, i.e.*

$$\mathbb{E}\,[f(x_1, \ldots, x_{k-1}, X_k, \ldots, X_n)] - \mathbb{E}\,[f(x_1, \ldots, x_k, X_{k+1}, \ldots, X_n)] = 0,$$

*and $f$ satisfies a Lipschitz condition $(L)$ with constants $c_k$ as above.*

*Then*

$$\mathbb{P}\,[|f(\mathbf{X}) - \mathbb{E}\,[f(\mathbf{X})]| \geq t] \leq 2\exp(-\frac{t^2}{2\sum c_k^2}).$$

We will finally merely sketch a result of N. Wormald, which is somewhat similar in spirit to Azuma's inequality. For a proper treatment we refer the reader to [Wor95, Wor99].

Suppose we want to study a system of random variables forming a 'vector valued' discrete time stochastic process, similar to the $Y_i$ discussed above. If, after appropriate re-scaling, the expected increase yields a system of ordinary differential equations, and the deviation from the expected increase is appropriately bounded, the random variables will closely follow the solution of the (system of) differential equations with high probability.

Let $\mathbf{Y}_t$ be a vector of non-negative integers, describing the sizes of certain combinatorial quantities in the course of some process. Suppose that for $t = 0, \ldots, n-1$

$$\mathbf{Y}_{t+1}^{(n)} - \mathbf{Y}_t^{(n)} = \mathbf{F}(t, \mathbf{Y}_t^{(n)}) \Leftrightarrow \overbrace{\frac{\mathbf{Y}_{t+1}^{(n)} - \mathbf{Y}_t^{(n)}}{n}}^{\text{discrete derivative}} = \mathbf{F}(t, n \cdot \frac{\mathbf{Y}_t^{(n)}}{n})/n.$$

Then it is at least plausible to associate the ordinary differential equation

$$\frac{d}{dx}\mathbf{y}(x) = \mathbf{f}(x, \mathbf{y}(x)),$$

where

$$\mathbf{f}(x, \mathbf{y}) := \mathbf{F}(n \cdot x, n \cdot \mathbf{y})/n, \quad \mathbf{y}(0) := \mathbf{Y}_0^{(n)}/n \quad x \in [0, 1],$$

and to believe $x \mapsto \mathbf{y}(x)$ suggests an approximate solution $\hat{\mathbf{Y}}_t$

$$t \mapsto \hat{\mathbf{Y}}_t := n \cdot \mathbf{y}(t/n).$$

The true $\mathbf{Y}_t$ will follow $\hat{\mathbf{Y}}_t$ quite closely, provided that $\mathbf{f}$ again satisfies a Lipschitz condition. Note that a Lipschitz condition is also needed to ensure the existence of a unique solution of the ordinary differential equation. The method is very powerful and has important applications, two of which we will discuss below (Sections 2.5.1 and 2.4.2). Usually it is quite easy to compute the differential equation, to solve it (at least numerically) and see what is going on. But turning this into a rigorous proof, checking the conditions of the theorem and handling subtle difficulties, requires a lot of careful reasoning, which makes this method difficult to apply.

## 2.2 Thresholds for Graph Properties

### 2.2.1 Monotone Properties of Random Graphs

A *graph property* $\mathcal{A} = (\mathcal{A}^{(n)})_n$ is a sequence of respective subsets of the sets $\mathcal{G}_n$ of all graphs on $n$ nodes. We will frequently suppress the dependence

on $n$ in our notation. We will abbreviate $\binom{n}{2}$ by $M$. It is convenient to view graphs as strings in $\{0,1\}^M$, each 'bit' standing for the presence or absence of a potential edge, after an arbitrary labelling of the edges in the complete graph $K_n$. A graph property is *symmetric* if it is invariant against re-orderings induced by permutations of the vertices and *monotone* if closed against addition of edges. The property of being non-$k$-colourable is a symmetric and monotone graph property since quite obviously the labelling of the vertices does not matter and adding edges makes a graph 'even more non-$k$-colourable'.

When we consider probability distributions $\mathbb{P}_c[\cdot]$ on $\mathcal{G}_n$, parametrised by the average degree (see Section 2.1), it is quite natural to study $\mathbb{P}_c[\mathcal{A}^{(n)}]$, where $\mathcal{A}$ denotes a non-empty, monotone and symmetric graph-property. Clearly $\mathbb{P}_0[\mathcal{A}^{(n)}] \not\equiv 0$ only if the empty graph is contained in $\mathcal{A}^{(n)}$, in which case $\mathcal{A}^{(n)} = \mathcal{G}_n$ due to monotonicity. Moreover, the function $c \mapsto \mathbb{P}_c[\mathcal{A}^{(n)}]$ is strictly increasing in $c$ for every fixed $n$, converging to 1 when $c$ goes to infinity. The point-wise limit $\lim_{n\to\infty} \mathbb{P}_c[\mathcal{A}^{(n)}]$ is still nondecreasing and converging to 1, but it may be discontinuous. For each $n$ the *threshold* $c^{(n)}_{\text{crit}}$ is the unique value of $c$ such that

$$\mathbb{P}_{c^{(n)}_{\text{crit}}}\left[\mathcal{A}^{(n)}\right] = 1/2.$$

Usually $c^{(n)}_{\text{crit}}$ will converge to a limit, although a 'weird' behaviour is feasible (however, c.f. [FK96]). Now the *threshold interval* $[c^{(n)}_-, c^{(n)}_+]$ is defined by the values $[c^{(n)}_-$ and $c^{(n)}_+]$ satisfying $\mathbb{P}_{c^{(n)}_-}[\mathcal{A}^{(n)}] = \varepsilon$ and $\mathbb{P}_{c^{(n)}_+}[\mathcal{A}^{(n)}] = 1-\varepsilon$. If

$$\frac{\left|c^{(n)}_+ - c^{(n)}_-\right|}{c^{(n)}_{\text{crit}}}$$

is bounded away from 0 (for all $\varepsilon, n$) the property $\mathcal{A}$ is said to have a *coarse threshold*, and to have a *sharp threshold* otherwise. Note that the threshold interval is called the *scaling window* in [BBC$^+$99], where a similar situation arising from random 2-SAT formulae is studied. It will turn out that non-$k$-colourability has a sharp threshold. Observe that in the case of non-$k$-colourability this yields a rather trivial algorithm for 'approximately' deciding the - in the worst case - NP-complete decision problem

$$\chi(G) \stackrel{?}{\leq} 3$$

that performs well *on average* for almost all values of $c$ we just answer 'yes' when $c < c^{(n)}_{\text{crit}}$, and 'no' otherwise. The big problem is that the limit of $c^{(n)}_{\text{crit}}$ is not known in this case, not even whether the limit exists. We will see upper bounds for $\limsup c^{(n)}_{\text{crit}}$ and lower bounds for $\liminf c^{(n)}_{\text{crit}}$ below.

### 2.2.2 Fourier-Walsh Transform of Graph Properties

Note that the set $\{0,1\}^M$ is a product group, with the product measure induced by the Bernoulli measures on the factors, and we may define a Fourier transform $\hat{f}$ of a function $f$ on $\{0,1\}^M$, called the *Fourier-Walsh transform.* The transform $\hat{f}$ is again a function on $\{0,1\}^M$ and the coefficients $\hat{f}(H)$ are labelled by what may be seen as subgraphs of the complete graph, identifying 'bit-strings' with graphs as above. The number of 'bits switched on' in $H$, i.e. the size of the labelling subgraphs may be regarded as some kind of 'frequency'. The overall idea is that 'local' properties that can be well approximated by 'low frequencies', i.e. by Fourier-coefficients labelled by small graphs, have a coarse threshold. We will occasionally identify $\mathcal{A}^{(n)}$ with the corresponding indicator, or characteristic function $f = \chi_{\mathcal{A}^{(n)}}$ on $\{0,1\}^M$.

We now set out to explain the (generalised) Fourier-Walsh-transform. Note that:

- We encode (labelled) graphs by strings in $\mathcal{G}_n := \{0,1\}^M$.
- The measure on the set of all graphs $\mathcal{G}_n$ is given by the $\mathcal{G}(n, c/n)$-model, i.e.:
$$\mathbb{P}_c[G] = (c/n)^{|E(G)|}(1-c/n)^{\binom{n}{2}-|E(G)|}.$$
- A property $\mathcal{A}^{(n)}$ is a subset of $\mathcal{G}_n$ or, equivalently, the characteristic function $f$ of that subset.
- Both the property $f$ and $\mathbb{P}_c[\cdot]$ are required to be *symmetric* with respect to the action of the permutation group $S_n$ on the complete graph $K_n$. I.e., isomorphic graphs have the same values of $f$ and the same probability $\mathbb{P}_c[\cdot]$.
- The property is *monotone* if $G \leq G' \Rightarrow f(G) \leq f(G')$, where $G \leq G'$ if $G'$ can be obtained from $G$ by adding some edges.

Associated with the set of graphs $\mathcal{G}_n$ together with the measure $\mathbb{P}_c[\cdot]$ is the Hilbert-space of square integrable functions $L^2(\mathcal{G}_n, \mathbb{P}_c[\cdot])$.

$$< f, g >_c := \sum_G f(G)g(G)\mathbb{P}_c[G].$$

We consider the orthonormal basis $\{U_H^{(c)}\}_H$ of $L^2(\mathcal{G}_n, \mathbb{P}_c[\cdot])$ which is indexed by the set of all graphs:

$$U_H^{(c)}(G) := \prod_{e \in H} U_e^{(c)}(G),$$

where

$$U_e^{(c)}(G) := \begin{cases} -\sqrt{\frac{1-c/n}{c/n}} & \text{if } e \in G \\ \sqrt{\frac{c/n}{1-c/n}} & \text{if } e \notin G \end{cases} .$$

The collection of the coefficients

$$\hat{f}_H :=< U_H^{(c)}, f >$$

is called the *Fourier-Walsh-spectrum* of $f$. It is an isometry, and Parseval's identity holds:

$$< f, f >= \sum_H \hat{f}_H^2. \tag{2.1}$$

Note that, since $f$ is a 0/1-function we have the useful identities:

$$\begin{aligned} < U_\emptyset^{(c)}, f > =< \mathbf{1}, f >= \sum_G f(G)\mathbb{P}_c[G] = \sum_G f(G)^2\mathbb{P}_c[G] =< f, f > \\ = \mathbb{P}_c[f = 1] = \mathbb{E}_c[f]. \end{aligned} \tag{2.2}$$

### 2.2.3 Threshold and Spectrum of 'G Contains a Triangle'

Obviously, the property '$G$ contains a triangle' is symmetric and monotone. Following Friedgut and Kalai in [FK96], we now turn to (approximately) computing the spectrum of this property since it serves as an illustrative example. We will actually only present the computation of two important Fourier coefficients, and even there we shall be sloppy with asymptotics. Our goal is merely to give an idea of how the argument works. Remember that we identify the property $\mathcal{A}^{(n)}$, corresponding to the set of graphs containing at least one triangle, with its characteristic function

$$f(G) = \begin{cases} 1 & G \in \mathcal{A}^{(n)} \\ 0 & G \notin \mathcal{A}^{(n)} \end{cases} .$$

We shall first (approximately) compute the distribution of the number of triangles $X$ in a sparse $\mathcal{G}(n, c/n)$ graph, where $c = c_{\text{crit}}^{(n)}$ is chosen to be the threshold value (i.e. $\mathbb{P}_c[$ $G$ contains no triangle $] = 1/2$). The number of

triangles $X(G)$ is asymptotically Poisson distributed (see e.g. [AS92]), i.e.

$$\begin{aligned} k = 0: &\quad \mathrm{Po}_\varepsilon(0) = e^{-\varepsilon} \\ k = 1: &\quad \mathrm{Po}_\varepsilon(1) = e^{-\varepsilon} \cdot \varepsilon \\ k = 2: &\quad \mathrm{Po}_\varepsilon(2) = e^{-\varepsilon} \cdot \varepsilon^2/2! \\ &\quad \vdots \\ k = l: &\quad \mathrm{Po}_\varepsilon(2) = e^{-\varepsilon} \cdot \varepsilon^l/l! \\ &\quad \vdots \end{aligned}$$

It therefore suffices to calculate the *expected* number $\varepsilon = \varepsilon(c) := \mathbb{E}_c[X]$ of triangles in $G$ to get the asymptotic distribution of $X$. The critical $c$ is implicitly defined by $\mathrm{Po}_{\varepsilon(c)}(0) = e^{-\varepsilon(c)} \stackrel{!}{=} 1/2$. Thus $c$ is the solution of

$$\log 2 \stackrel{!}{=} \varepsilon(c) = \sum_{\{v_1,v_2,v_3\} \subseteq V} \Pr[\{v_1, v_2, v_3\} \text{ form a triangle}] = \binom{n}{3}(c/n)^3. \tag{2.3}$$

Therefore the probabilities that $G$ contains $k$ triangles are:

$$\begin{aligned} k = 0: &\quad \mathrm{Po}_\varepsilon(0) = 1/2 \\ k = 1: &\quad \mathrm{Po}_\varepsilon(1) = 1/2 \cdot \varepsilon \\ k = 2: &\quad \mathrm{Po}_\varepsilon(2) = 1/2 \cdot \varepsilon^2/2! \\ &\quad \vdots \\ k = l: &\quad \mathrm{Po}_\varepsilon(2) = 1/2 \cdot \varepsilon^l/l! \\ &\quad \vdots \end{aligned}$$

In particular the probability that there is at least one triangle is

$$\mathbb{P}_c\left[\mathcal{A}^{(n)}\right] = 1 - \mathrm{Po}_{\binom{n}{3}(c/n)^3}(0).$$

The threshold $c^{(n)}_{\mathrm{crit}}$ converges to $(6 \log 2)^{1/3}$ from above. The function $c \mapsto \mathbb{P}_c[\mathcal{A}^{(n)}]$ has bounded derivatives near the threshold for all $n$ and thus the property $\mathcal{A}$ has a coarse threshold.

We want to compute the Fourier-Walsh spectrum $\hat{f}$ for $c$ set on the threshold value (c.f. Eq. 2.3) such that $\mathbb{P}_c[f = 1] = 1/2$. We will indicate how to compute all Fourier-coefficients for the case that $H$ is a union of $k \in \mathbb{N}_0$ disjoint triangles and then see indirectly that all the others have to be zero because of Eq. 2.1.

$k = 0$: The first Fourier-coefficient is simply the probability of $f = 1$:

$$\hat{f}_{\emptyset} = < U_{\emptyset}, f > = 1/2.$$

$k = 1$: We will first compute* $U_{\{e_1,e_2,e_3\}}(G)$, for a specific triangle induced by the edges $\{e_1, e_2, e_3\}$.

$$:= \begin{cases} -(\frac{1-c/n}{c/n})^{3/2} & \simeq -(c/n)^{-3/2} \text{ if } \{e_1,e_2,e_3\} \cap E(G) = \{e_1,e_2,e_3\}, \\ (\frac{1-c/n}{c/n})\sqrt{\frac{c/n}{1-c/n}} & \simeq +(c/n)^{-1/2} \text{ if } \{e_1,e_2,e_3\} \cap E(G) = \{e_1,e_2,e_3\} \setminus e_i, \\ -\sqrt{\frac{1-c/n}{c/n}}\frac{c/n}{1-c/n} & \simeq -(c/n)^{+1/2} \text{ if } \{e_1,e_2,e_3\} \cap E(G) = \{e_i\}, \\ (\frac{c/n}{1-c/n})^{3/2} & \simeq +(c/n)^{+3/2} \text{ if } \{e_1,e_2,e_3\} \cap E(G) = \emptyset. \end{cases}$$

Next we want the Fourier-coefficient corresponding to the (specific!) triangle $\{e_1, e_2, e_3\}$:

$$\hat{f}_{\{e_1,e_2,e_3\}} = < U_{\{e_1,e_2,e_3\}}, f > = \sum_{G} U_{\{e_1,e_2,e_3\}}(G) f(G) \mathbb{P}_c[G].$$

Partition the graphs according to how much of the triangle defined by $\{e_1, e_2, e_3\}$ they contain:

$$\begin{aligned} \hat{f}_{\{e_1,e_2,e_3\}} \simeq & - \sum_{\{e_1,e_2,e_3\} \cap E(G) = \{e_1,e_2,e_3\}} (c/n)^{-3/2} f(G) \mathbb{P}_c[G] \\ & +3 \sum_{\{e_1,e_2,e_3\} \cap E(G) = \{e_1,e_2\}} (c/n)^{-1/2} f(G) \mathbb{P}_c[G] \\ & -3 \sum_{\{e_1,e_2,e_3\} \cap E(G) = \{e_1\}} (c/n)^{+1/2} f(G) \mathbb{P}_c[G] \\ & + \sum_{\{e_1,e_2,e_3\} \cap E(G) = \emptyset} (c/n)^{+3/2} f(G) \mathbb{P}_c[G]. \end{aligned}$$

We interpret these terms as probabilities.

$$\begin{aligned} \hat{f}_{\{e_1,e_2,e_3\}} \simeq & -(c/n)^{-3/2} \mathbb{P}_c[G \text{ contains } \{e_1,e_2,e_3\} \text{ and } f(G) = 1] \\ & +3(c/n)^{-1/2} \mathbb{P}_c[G \text{ contains only } \{e_1,e_2\} \text{ and } f(G) = 1] \\ & -3(c/n)^{+1/2} \mathbb{P}_c[G \text{ contains only } \{e_1\} \text{ and } f(G) = 1] \\ & +(c/n)^{+3/2} \mathbb{P}_c[G \cap \{e_1,e_2,e_3\} = \emptyset \text{ and } f(G) = 1]. \end{aligned}$$

*We shall - in a rather imprecise manner - substitute certain quantities by their asymptotic values. We use '$\simeq$' instead of '='.

Because[†]

$$\begin{aligned}
&\mathbb{P}_c\left[G \text{ contains } \{e_1,e_2,e_3\} \text{ and } f(G)=1\right]\\
&=\mathbb{P}_c\left[f(G)=1 \middle|\ G \text{ contains } \{e_1,e_2,e_3\}\right]\\
&\cdot\mathbb{P}_c\left[\ G \text{ contains } \{e_1,e_2,e_3\}\right]\simeq 1\cdot(c/n)^3,
\end{aligned}$$

$$\begin{aligned}
&\mathbb{P}_c\left[G \text{ contains only } \{e_1,e_2\} \text{ and } f(G)=1\right]\\
&=\mathbb{P}_c\left[f(G)=1 \middle|\ G \text{ contains only } \{e_1,e_2\}\right]\\
&\cdot\mathbb{P}_c\left[\ G \text{ contains only } \{e_1,e_2\}\right]\simeq 1/2\cdot(c/n)^2,
\end{aligned}$$

$$\ldots.$$

this simplifies to

$$\begin{aligned}
\hat{f}_{\{e_1,e_2,e_3\}} \simeq &-(c/n)^{-3/2}(c/n)^3\\
&+3(c/n)^{-1/2}\cdot 1/2\cdot(c/n)^2\\
&-3(c/n)^{+1/2}\cdot 1/2\cdot(c/n)\\
&+(c/n)^{+3/2}\cdot 1/2\\
\simeq &-1/2\cdot(c/n)^{3/2}\\
\Rightarrow \hat{f}^2_{\{e_1,e_2,e_3\}} \simeq &(c/n)^3/4.
\end{aligned}$$

Now we know that for all isomorphic triangles the value of $\hat{f}$ must be the same by symmetry. Therefore we get (remember Eq. 2.3):

$$\sum_{H \text{ is a triangle}} \hat{f}^2_H \simeq 1/4\binom{n}{3}(c/n)^3 = 1/4\cdot\varepsilon.$$

$k \geq 2$: In a similar manner we can calculate the other non-zero Fourier-coefficients. We shall get $1/4\cdot\varepsilon^k/k!$.

$$\sum_{\substack{H \text{ is a disjoint union}\\ \text{of } k \text{ triangles}}} \hat{f}^2_H \simeq \sum_k 1/4\cdot\varepsilon^k/k! = 1/2\cdot\sum_k \mathbb{P}_c\left[G \text{ contains } k \text{ triangles}\right].$$

The terms (including the first coefficient $k=0$) sum up to $1/2$ which is, using Parseval's identity 2.1, equal to the sum over *all* squared Fourier-coefficients given by Equation 2.2. Therefore the Fourier-coefficients labelled by disjoint unions of triangles are the only non-zero terms.

[†]Note that the factor $1/2$ stems from the fact that the property '$G$ contains a triangle' is nearly independent of the event that some 'potential' triangle is not 'switched on'. It should be replaced by a correct value in $1/2+o(1)$.

### 2.2.4 3-Colourability (3-SAT) Has a Sharp Threshold.

The spectral techniques discussed above can be employed to show that certain properties have a sharp threshold. In [FK96] Friedgut proved that random 3-SAT has a sharp threshold, and in [AF99] Achlioptas and Friedgut showed this for $k$-colourability. The proofs are essentially indirect proofs and do not quantify the threshold value $c_{\text{crit}}^{(n)}$ nor its asymptotic behaviour.

We will not elaborate much on the random 3-SAT problem here, we refer the reader to [Mol01] and to our Section 7.4.3. Note that 3-SAT is also an $\mathcal{NP}$-complete problem. The idea is to pick $m$ 3-SAT-clauses at random using $n$ variables. The quotient $\alpha = m/n$ plays a role similar to the average degree for random graphs. Again there is a threshold value, possibly depending on $n$. Friedgut was able to show that for $\alpha$ below this value a random 3-SAT formula is a.s. satisfiable, and a.s. unsatisfiable for $\alpha$ above this value. Empirical evidence has been reported that the threshold value should converge to a value close to 4.2.

Now we return to $k$-colourability. Achlioptas/Friedgut used the following theorem from [FK96].

**Fact 6** *From [FK96]. Let $\alpha > 0$. There exist functions $B(\varepsilon, C)$, $b_1(\varepsilon, C)$, $b_2(\varepsilon, C)$* independent of *$n$ such that for all $n$, $c^*$, $C$ and $\varepsilon$, and any monotone graph property $\mathcal{A}$ such that $\alpha < \Pr_{c^*}[A] =: \mu_{c^*} < 1 - \alpha$ and $c^* \cdot \frac{d\mu}{dc}|_{c=c^*} \leq C$ there exists a graph $H$ with no more than $B$ edges such that:*

- *$H$ is balanced, i.e. the average degree of any induced subgraph of $H$ is no larger than the average degree of $H$.*
- *$b_1 < \mathbb{E}_{c^*}[Z_H] < b_2$, where $Z_H$ counts the number of copies of $H$ in $\mathcal{G}_{n,c^*/n}$.*
- *Conditional on the event that some specific copy of $H$ appears in $\mathcal{G}_{n,c^*/n}$,*

$$\Pr_{c^*}[\mathcal{A}|H] > 1 - \varepsilon.$$

We will now sketch - admittedly very vaguely - how Achlioptas and Friedgut (in [AF99]) reach a contradiction assuming that $k$-colourability had a coarse threshold.

a) Intuitively Fact 6 means that for a property with a coarse threshold the 'appearance' of some appropriate rather small subgraph $H$ changes the probability of $k$-colourability significantly.

b) When a property has a coarse threshold, its probability does not change significantly when we 'add' $\delta \cdot n$ edges.

Now 'adding' $\delta \cdot n$ edges will lead to a copy of $H$ to 'appear', due to $H$ being sufficiently small, which by a) should drastically change the probability of $k$-colourability significantly, which it does not, by b). This is the essence of the contradiction leading to their main result stated below.

**Fact 7** *(From [AF99]) Let $f_k(n,c)$ denote the probability $\mathbb{P}_c[\mathcal{A}^{(n)}]$ for the property $\mathcal{A}^{(n)}$ of $k$-colourability. For every integer $k \geq 3$ there exist $d_k^{(n)}$ such that for any $\varepsilon > 0$,*

$$\lim_{n\to\infty} f_k(n, d_k^{(n)} - \varepsilon) = 1 \text{ and } \lim_{n\to\infty} f_k(n, d_k^{(n)} + \varepsilon) = 0.$$

## 2.3 Important Concepts

We will introduce the $k$-core together with some immediately related results, define branching trees, the *extended $k$-core* and the concept *ownership*. We include this material here because it will facilitate reading the next subsections.

### 2.3.1 $k$-Core and its Significance for $k$-Colouring

**Definition 2.3.1** *The $k$-core of a graph $G$ is the union of all induced subgraphs $H$ with minimum degree $\delta(H) = k$.*

**Definition 2.3.2** *A (node-)deletion process is some rule (or protocol) for iteratively removing nodes with degree less than $k$ from $G$ as follows.*

*At 'time' $t$, as long as there are any nodes with degree less than $k$ left in the remainder graph $G_t$, select such a node $v_t$ according to the protocol and remove it from $G_t$, setting $G_{t+1} := G_t \setminus v_t$.*

**Fact 8** *The $k$-core can be algorithmically characterised as the outcome of any deletion process. In particular, the outcome of any deletion process is the $k$-core, irrespective of the details of the specific deletion protocol.*

**Proof** (Of Fact 8.) Denote by $\mathcal{H}(G)$ the collection of all induced subgraphs of minimum degree $k$. Then $C := \bigcup_{H\in\mathcal{H}} H$ is the $k$-core. For convenience let the empty subgraph have minimum degree $k$. Denote by $C'$ the outcome of an arbitrary deletion process consistent with Definition 2.3.2.

$C \subseteq C'$: Let $H$ be any induced subgraph of minimum degree $k$. It cannot be deleted by any deletion process, since otherwise some node in $H$ would have to be the first to be deleted. But since it is the first, it has still degree at least $k$, as the other nodes in H are still in place.

$C' \subseteq C$: Conversely it is obvious that the outcome of any deletion process is a subgraph with minimum degree $k$. □

**Fact 9** *Any graph $G = (V, E)$ with an empty $k$-core can be $k$-coloured in polynomial time.*

**Proof** (Of Fact 9.) When the $k$-core is empty, any (node-)deletion process has removed all nodes in some order. We will consider an arbitrary such deletion process. In the reverse order 'glue' the deleted vertices back on. This yields a sequence of (sub-)graphs. Due to the way in which the deletion process works any re-inserted vertex is connected to at most $k-1$ nodes in the graph reconstructed so far, with the result that it is always possible to extend a proper colouring of the graph reconstructed so far to the next graph in the sequence. □

We will also introduce an *edge-deletion process*. This seems rather artificial when considering only the $k$-core but the significance will become clear when we discuss the *magic subgraph* defined below. When we interpret some graph $G$ as directed we mean that we have replaced all edges by bi-directed arcs. We will frequently omit the word 'directed' when talking about directed edges, that is arcs.

**Definition 2.3.3** *An edge $(u, v)$ is 'bad' if there are less than $k-1$ nodes different from $v$ in the in-neighbourhood of $u$. An edge-deletion process is some rule (or protocol) for iteratively removing 'bad' edges $G$ as follows.*

*At time $t$, as long as there are 'bad' directed edges left in the remainder graph $G_t$, select such an edge $(u, v)$ according to the protocol and remove it from $G_t$, setting $G_{t+1} := G_t \setminus (u, v)$.*

*We shall refer to the outcome of this deletion process as the* extended $k$-core. *It is a directed subgraph of $G$, the* extended $k$-core in the node sense *is the subgraph induced by the nodes 'pointed to' by at least two directed edges.*

**Proposition 2.3.4** *The outcome of the edge-deletion process is the union of all directed subgraphs consisting entirely of 'good' edges. It is independent of the deletion protocol. The $k$-core can be obtained from the outcome of an edge-deletion process as the subgraph induced by the nodes participating in bi-directed edges.*

The proof is analogous to the corresponding proof of Fact 8 (p. 27) for node-deletion. Note that $C$ has to be replaced by the union of all directed subgraphs consisting entirely of 'good' edges ('good' with respect to the given subgraph!).

### 2.3.2 Branching Trees

*Branching trees* are random (possibly infinite) trees that are generated by the following procedure.

- Start with the root and mark it as 'live'.
- While there are live leaves in the tree generated so far, expand a live leaf, chosen according to some protocol (see below). That is generate a random number $X_t$ of fresh nodes distributed according to some integer distribution $\mu_t$, glue them to the leaf being expanded and mark the leaf as dead.

The distribution $\mu_t$ is called the *progeny distribution*. The tree is built up according to a *protocol* determining the order in which live leafs are expanded, such as, for example, breadth first search (BFS). Sometimes the live leaf to be expanded is chosen uniformly at random (note the similarity to Karp's fanning-out process in graphs explained above). Having chosen such a protocol the leafs expanded at time $t$ have progeny $X_t$. Usually $(X_t)_t$ is assumed to be an i.i.d. sequence but it is feasible that $\mu_t$ depends on $t$ and even on the history $x_1, \ldots, x_{t-1}$.

We will mainly consider $\mathrm{Po}_c$ (Poisson) and $\mathrm{Bi}(n, c/n)$ (binomial) progeny distributions. We will frequently discuss trees with *finite radius* $r$, where children are no longer allowed to reproduce when they are at some distance $r$ from the root.

The Branching Tree Connection will be explained in Section 2.4.1 below. The overall idea is that when $\mathcal{G}_{n,c/n}$ random graphs are explored by BFS or some other fanning-out process, small neighbourhoods will closely resemble branching trees since cross edges are rather unlikely, the graphs being as sparse as they are. For a more thorough account of branching trees we refer the reader to [AN72] and [AN97], and to Karp's important paper [Kar90] explaining the appearance of the giant component using branching trees.

### 2.3.3 The Concept of Ownership

We will again fix $k = 3$, for ease of exposition only. For motivation consider a node $v$ in a graph $G$, assuming that the $r$-neighbourhood (w.r.t. the 'shortest path metric') of $v$ is a tree. When $v$ is in the 3-core of $G$ it is adjacent to at least three complete binary trees of depth $r-1$. Otherwise $v$ could be removed by a node deletion process restricted to the interior of the $r$-neighbourhood, contradicting Fact 8. We shall say that '$v$ owns a Cayley tree of depth $r$', or '$v$ is an $r$-Cayley-owner'. We shall state formal definitions and make this

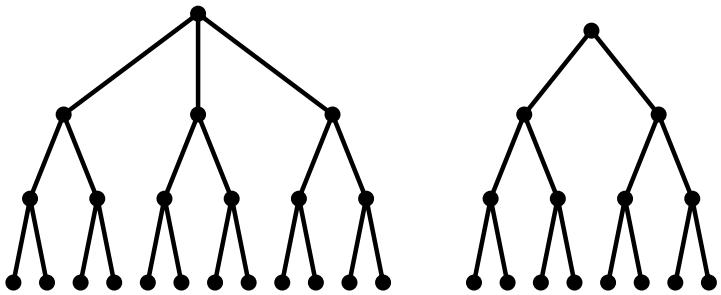

Figure 2.1: A 3-Cayley tree and a 3-binary tree.

precise presently. So being an $r$-Cayley-owner is somehow the 'equivalent' of being in the 3-core on a local basis. Note that in random graphs with average degree $c$ the $r$-neighbourhoods will be trees for most nodes, with high probability, and furthermore closely resemble branching trees. We shall later study ownership in branching trees and shall see that a lot can be learned from this about the 3-core in 'corresponding' random graphs.

**Definition 2.3.5** *An $r$-binary tree is a binary tree of depth $r$. An $r$-Cayley tree is a tree consisting of 3 disjoint complete binary trees of depth $r-1$, their root being connected to some (extra) node $v$. See Figure 2.1.*

We will first define ownership for the special case that $G$ is a tree. This will turn out to be a special case of Definition 2.3.7 below.

**Definition 2.3.6** *(Ownership in Trees, special case of Definition 2.3.7.) Let $v$ be the root of some tree $T$ with depth $r$.*

*The root $v$ is an* $r$-binary owner (w.r.t. $T$)*, if $T$ contains (at least one) complete binary tree of depth $r$ rooted at $v$.*

*The root $v$ is an* $r$-Cayley owner (w.r.t. $T$)*, if $T$ contains (at least one) Cayley-tree of depth $r$ rooted at $v$.*

Note that an $r$-Cayley-owner (w.r.t. $T$) $v$ may own several (possibly overlapping) $r$-Cayley trees. We shall sometimes call the union of those $r$-Cayley trees (rooted at $v$) the *$r$-cone of $v$*, because it is the 'analogue' of the core in trees. Thus the root $v$ of a tree is an $r$-Cayley-owner if and only if its $r$-neighbourhood contains a non-empty $r$-cone.

To include cases where the $r$-neighbourhood is not a tree, we take some node $v \in V(G)$ and decide whether it is an $r$-owner on the basis of the $r$-neighbourhood of $v$ in $G$. The following definition is sufficiently general to

apply to both trees and general graphs. We will also define the concept of *runs* that will be important later.

**Definition 2.3.7** *(Ownership in graphs, generalising Definition 2.3.6.) Let $G$ be some graph.*

1. *Every node $v$ in $G$ is a* 0-binary owner (w.r.t. $G$) *and a* 0-Cayley owner (w.r.t. $G$).

2. *A node $v$ is recursively defined to be an* $r$-binary owner (w.r.t. $G$) *if it is adjacent to at least two $(r-1)$-binary owner w.r.t. $G \setminus v$.*

3. *An* $r$-Cayley owner (w.r.t. $G$) *is a node $v$ that is adjacent to at least three $(r-1)$-binary owners w.r.t. $G \setminus v$.*

4. *An* ($r$-)run (w.r.t. $G$) *is an $r$-Cayley owner (w.r.t. $G$) $v$ that is* not *adjacent to at least three $r$-binary owners w.r.t. $G \setminus v$.*

Note that the 'proofs' for $r$-ownership in the recursive definition above are not necessarily edge-disjoint if the $r$-neighbourhood is not a tree.

We have collected some simple observations in the following Proposition 2.3.8.

**Proposition 2.3.8** *Let $v$ be a node in some graph $G$.*

1. *If the $r$-neighbourhood of $v$ in $G$ is a tree, then $v$ is an $r$-binary(Cayley) owner (w.r.t. $G$), if and only if its $r$-neighbourhood contains an $r$-binary(Cayley) tree.*

2. *Ownership (both $r$-binary and $r$-Cayley, w.r.t. $G$) can be decided upon knowledge of the $r$-neighbourhood of $v$ in $G$.*

3. *Any node in the* 3*-core of $G$ that has a tree-like $r$-neighbourhood is an $r$-Cayley owner (w.r.t $G$). Any node that is not an $r$-Cayley owner (w.r.t. $G$) but has a tree-like $r$-neighbourhood cannot be in the* 3*-core.*

**Proof**

1. Obvious.

2. The $r-1$-neighbourhood w.r.t. $G \setminus v$ of some neighbour $w$ of $v$ is contained in the $r$-neighbourhood of $v$ w.r.t. $G$.

3. We have seen in Fact 8 above, that the deletion protocol may be chosen arbitrarily. The 'tamed' deletion protocol leaves out the complement of the $(r-1)$-neighbourhood as long as possible, that is no node outside the $(r-1)$-neighbourhood may be deleted as long as there are any nodes with degree less than $k$ left in the $(r-1)$-neighbourhood of $v$.

   Clearly, if $v$ has an $r$-neighbourhood that is a tree (w.r.t. $G$), it survives the above 'tamed' deletion process (stopped before it starts attacking the complement of the $(r-1)$-neighbourhood of $v$) if and only if it is an $r$-Cayley owner.

   So if $v$ survives the entire 'tamed' deletion process, due to being in the 3-core, it will in particular survive the 'first phase' and is thus an $r$-Cayley owner.

   Conversely, if it is not an $r$-Cayley owner it will not even survive the 'first phase' and thus not be in the 3-core.

□

To understand why we restricted ourselves to $r$-neighbourhoods that are trees look at the extreme case that $v$ participates in a copy of $K_4$. Then $v$ is not an $r$-owner for $r \geq 2$ according to our definition but it will be in the 3-core. Also observe that an isolated copy of an $r$-Cayley tree will survive the first phase of the 'tamed' deletion process, but nevertheless it is not in the $k$-core.

## 2.4 Sudden Appearance of the Giant $k$-Core

Similarly to the 'appearance of the giant connected component' at $c = 1$ a giant $k$-core appears at some critical value of $c$ which is $3.35(\ldots)$ for $k = 3$. We will start with discussing what is known as the Branching Tree Connection. The first rigorous proof for the appearance of the $k$-core in random graphs by Pittel et.al. (see [PSW96]) will be explained in in Section 2.4.2 below. In Section 2.4.3 we review the important proof of Goerdt and Molloy ([GM00]) for the sudden appearance of a giant $k$-core in faulty random configurations. Our new proof for the appearance of the $k$-core in random graphs (Chapter 6) may essentially be regarded as a generalisation of their proof technique to the $\mathcal{G}_{n,m}$ model.

### 2.4.1 The Branching Tree Connection

The term 'Branching Tree Connection' was coined[‡] by Pittel, Spencer and Wormald in [PSW96], to be reviewed below. Random graphs with constant

[‡]As mentioned before, in [PSW96] the term 'branching process connection' is used.

average degrees locally resemble branching trees. By 'locally' we mean in a sufficiently small neighbourhood w.r.t. the breadth first search (BFS) metric. This is true because cross edges are sufficiently rare. Note that (unfortunately) cross edges are not rare enough to totally disappear for sufficiently large $n$. In the $\mathcal{G}_{n,c/n}$ model the progeny distributions are (quite obviously, close to) Poisson-distributions with parameter $c$, i.e. $\mathrm{Po}_c$. In other models (approximations to) the respective progeny-distributions may be determined analogously.

Before going into more detail we stress the fact that *a priori* the Branching Tree Connection is a heuristic concept. It has frequently been observed that one can learn a lot about 'giant subgraph' phase transitions from studying the 'corresponding' branching-trees and 'corresponding phase transition' therein, instead of the random graphs. Threshold phenomena in branching trees seem to parallel the appearance of giant subgraphs in 'corresponding' sparse random graphs. In particular, the correspondence seems to be so close that numerical values of critical average degree and size of the 'corresponding' subgraphs can be correctly predicted.

Phase transitions in branching trees can be studied rigorously using recurrence equations. Ideas concerning this observation have definitively been around, we have turned them into a little theory (see Chapter 3) providing a comprehensive treatment of phase transitions in branching trees and some novel aspects.

*Some* 'giant subgraph phase transitions' in random graphs have been studied rigorously. Proofs that do not employ the Branching Tree Connection require considerable technical effort. A notable exception is Goerdt and Molloy's result in [GM00] that we shall discuss below. In the special case of the giant component Karp ([Kar90]) gave a proof 'based' on the Branching Tree Connection. But we did not see a way to generalise his proof to other similar situations such as the appearance of the $k$-core. Yet in all cases where 'giant subgraph phase transitions' in random graphs have been analysed the results are such that they could have been 'guessed' from analysing the 'corresponding' branching trees. As to how the 'guessing' works, see the discussion on the giant component in the next paragraph. Somehow the appearance of infinite subtrees in branching trees 'corresponds' to the appearance of giant subgraphs in random graphs. This is the heuristic aspect. The situation cries out for a rigorous theory formalising the Branching Tree Connection. Our work presented in Chapters 5 and 6 shows the progress we have made in making this rigorous. We also have discovered the appearance of a giant subgraph in tripartite models that has not been previously described, *inspired* by the Branching Tree Connection, see Chapter 4.

For example, the appearance of the giant component 'corresponds' to a

branching tree phase transition in the following way. When $c < 1$ the $\mathrm{Po}_c$ branching trees die out with very high probability. For $c > 1$ the branching trees are infinite with a certain probability $q(c)$, we shall say that 'the root owns an infinite subtree'. In the random graph, when $c > 1$ there is a giant component containing $n \cdot q(c)$ nodes.

Thus the random graph apparently behaves as if each node rooted a separate (independent) branching tree, those nodes with an infinite subtree somehow 'merging' into the giant component. Certainly this picture does not at all correctly capture what is going on in the graph, since neighbourhoods are never finite and will more and more overlap the larger their radii become. Still, both the threshold value and the size of the giant subgraph can be correctly predicted by studying branching trees.

For the $k$-core, consider the trace of the 3-core on a (say, treelike) neighbourhood of some node $v$. This subtree will have minimum degree 3 for all interior nodes. This is equivalent to containing a *Cayley-Tree*, that is a complete binary tree with degree 3 at the root $v$. In the 'branching tree world' this leads to a phase transition at $c = 3.35$, the probability that the root owns an infinite Cayley tree being some constant $q'(c)$. And indeed, the Branching Tree Connection appears to work, since the *rigorous* proof of Pittel, Spencer and Wormald ([PSW96]) yields the appearance of a giant $k$-core at $c = 3.35$ its size being sharply concentrated around $q'(c)$.

### 2.4.2 Analysing the Deletion Process

The only existing proof for the sudden appearance of a giant $k$-core in the $\mathcal{G}_{n,m}$-model so far is due to Pittel, Spencer and Wormald ([PSW96]). The overall idea is to analyse the stochastic process of degree sequences induced by the deletion process, rather one version of the deletion process particularly amenable to analysis. By a degree sequence we mean for the moment, slightly abusing notation, the 5-tuple

$$(X_0(G), X_1(G), X_2(G), X_{\geq 3}(G), M(G)),$$

counting the numbers of nodes with degrees $0, 1, 2, \geq 3$, respectively and the number of edges.

Using the differential-equations tool we sketched in Section 2.1, and many other subtle arguments, they can capture the behaviour of the components of the above degree sequence in the course of the deletion process. It turns out that for $c < 3.35(\ldots)$ the process converges to $(n, 0, 0, 0, 0)$ with high probability and to $(n - p(c) \cdot n, 0, 0, p(c) \cdot n, c \cdot (p(c))^2 \cdot n)$, otherwise. The value of $p(c)$ can be 'predicted' from the Branching Tree Connection and it assumes the value $0.27(\ldots)$ for $c$ just above the critical value $3.35(\ldots)$.

Remember that it was Pittel, Spencer and Wormald themselves who brought up the idea of the Branching Tree Connection and asked whether this could lead to a more natural proof. As we know (c.f. Section 2.3) the $k$-core is defined independently of the specific version of the deletion process. Certainly, since the $k$-core can be characterised by deletion processes a detailed analysis of such processes will yield information on the $k$-core. Analysing the full dynamics of the deletion process may turn out to be unnecessarily difficult. Yet Pittel, Spencer and Wormald's proof, involved as it may be, has been the only rigorous proof so far. Empirically the Branching Tree Connection appears to be rather robust against the choice of model, that is it correctly predicts the appearance of, say, the $k$-core for a larger class of models, whereas the existing proofs strongly depend on the specific model, therefore a proof based on the Branching Tree Connection is likely to apply to a wider class of models.

Note that Molloy and Reed ([MR95a] and [MR98]) where able to uniformly treat the appearance of a giant component on graphs that are randomly distributed amongst a large class of pre-defined degree-sequences.

### 2.4.3 Goerdt and Molloy's Result

Very recently Goerdt and Molloy (in [GM00]) have presented an analysis of the appearance of the $k$-core in the model of faulty random configurations. They were the first to give a rigorous proof for the sudden appearance of a giant $k$-core based on the Branching Tree Connection in the model of *random faulty $d$-regular configurations*. Random faulty $d$-regular configurations are generated as follows:

1) Choose a random $d$-regular configuration.
2) Retain each edge independently with probability $p$.

We will try to describe their result in the language developed for our work, trying to point out why it is not merely a straightforward task to generalise their results to $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$. Their main idea is to choose a clever protocol for the deletion process, ingeniously 'by-passing' most of the tedious 'step by step' analysis in [PSW96]. In a first phase consisting of $r - 1 = o(\mathrm{diam}(G))$ rounds iteratively Goerdt and Molloy strip off all nodes that had degrees less than 3 at the beginning of the round, we will refer to this as the *shell-wise deletion process*. At the end of the first phase all nodes with a tree-like neighbourhood that have remained, have had Cayley-trees of radius $r$ in their neighbourhoods, i.e. are $r$-owners in the sense of Section 2.3.3. Thus the expected number of nodes remaining is approximately $p(c) \cdot n$, where $p$ is the value from the Branching Tree Connection for branching trees with

progeny distribution $\mathrm{Bi}(d,p)$. Since the BFS-balls between different nodes do overlap, but 'not too often', they were able to show concentration of the number of remaining nodes using Chebychev's inequality, i.e. by the Second Moment Method. The degree sequence of the remainder configuration at the end of the first phase is extremely close to what would be expected from the Branching Tree Connection, up to $o(n)$ corrections. In particular there are only $o(n)$ nodes with degrees less than $k$, corresponding to runs, in our language. Essentially for further success of their proof strategy, they are able to prove that the remainder configuration is distributed *uniformly* among all configurations having the same degree sequence.

In the second phase (random) single nodes with degrees less than $k$ are deleted. Since the expected increase in the number of 'deficient nodes' having degrees less than $k$ is negative due to the form of the degree sequence, the second phase will terminate with high probability having deleted only $o(n)$ further nodes, this follows from a straightforward 'gambler's ruin' argument. Note that throughout the second phase the degree sequence may only change by $o(n)$ which is immaterial for the negative expected increase in the number of 'deficient nodes'. Again it is crucial that the uniform distribution amongst all configurations with the same degree sequence remains invariant in the course of the second phase.

Our work described in Chapter 5 is similar to Goerdt and Molloy's result concerning the first phase. A substantial part of our further work (Chapter 6) consists in generalising their ingenious proof idea to $\mathcal{G}_{n,m}/\mathcal{G}_{n,p}$ (with average degree $c$). What makes this generalisation a non-trivial task?

1. In random $d$-regular configurations the $r$-neighbourhoods of at most $(n-o(n))$ of the nodes are $d-ary$-trees a.a.s., before the (independent!) edge percolation is performed. For all those nodes the distribution of arbitrary $r$-neighbourhoods equals *exactly* the distribution of $r$-branching trees with progeny distribution $\mathrm{Bi}(d,p)$.

2. Note that in random $d$-regular configurations the sizes of the $r$-neighbourhoods are bounded by $d^r$, deterministically. This remains true after the edge percolation since the degrees are only reduced by this operation. This implies that the $r$-neighbourhood of each node is independent of the $r$-neighbourhoods of all but at most $d^{2r}$ other nodes, deterministically. Therefore it is possible to show that the number of $r$-owners is concentrated around its expected value which is crucial at least for determining the number of runs. Note that in this setting concentration tools relying on the existence of a *dependency graph* could have been employed instead of Chebychev's Inequality, yielding Chernoff-tight concentration.

Generalising the first point to $\mathcal{G}_{n,m}/\mathcal{G}_{n,p}$ with average degree $c$ is more of a technical problem. In $\mathcal{G}_{n,m}/\mathcal{G}_{n,p}$ the $r$-neighbourhoods which are trees are essentially distributed as $r$-branching trees, too, with progeny distribution $\mathrm{Po}_c$. The fact that these distributions are only approximately equal makes calculations more clumsy but not essentially different.

The second point is harder to transfer to $\mathcal{G}_{n,m}/\mathcal{G}_{n,p}$ with average degree $c$. There it is not very hard to prove that the $r$-neighbourhoods of all nodes are bounded by $K^r$, too, for some appropriate constant $K = K(c)$ and a suitable choice of $r = r(c, n)$ with very high probability. However, there is no worst case bound on the sizes of the $r$-neighbourhoods. We can define a 'meta-graph' closely resembling a dependency graph by connecting nodes whose $r$-neighbourhoods overlap by ('meta'-)edges. With very high probability this 'meta-graph' is degree-bounded just as usually required for a dependency graph. It is intuitively obvious that the neighbourhoods are 'independent' for nodes not connected by a 'meta-edge'. This 'independence' should imply concentration of the number of owners. However all concentration tools employing a dependency graph (that we know of) require that the dependency graph, usually together with a degree-bound, is given in advance, i.e. *before* the random graph is chosen.

We have discovered that Goerdt and Molloy's proof for the uniform distribution conditional on the degree sequence carries over to *graphs* in the $\mathcal{G}_{n,m}$-model. We can therefore adapt their proof strategy to the appearance of the $k$-core in the $\mathcal{G}_{n,m}$-model, because the set of owners (to be defined later) is (very close to) what remains after the first phase of shell-wise node deletions.

## 2.5 Explicit Bounds for $k$-Colourability

Remember the result of Achlioptas/Friedgut that $k$-colouring has a sharp threshold. One expects this threshold to converge to an absolute constant independent of $n$, but there is no rigorous proof.

We have seen above (Fact 9 in Section 2.3.1, p. 28) the absence of a $k$-core leads to a trivial $k$-colouring algorithm. From Section 2.4 above we know that the 3-core is a.s. empty for $c < 3.35(\ldots)$. Bollobás conjectured that the a.s. jump in chromatic number might coincide with the appearance of the 3-core.

Things turned out not to be so easy. A first indication was given by Molloy in [Mol96], and Achlioptas and Molloy later showed in [AM97] that $\mathcal{G}_{n,m}$ can be efficiently 3-coloured up to an average degree of $c = 3.85(\ldots)$. We will explain their algorithm in Section 2.5.1 below.

For upper bounds a straightforward application of the First Moment Method counting colourings yields $c_{\text{crit}} < 5.419(\ldots)$. We will review this in Section 2.5.2 together with a series of refinements yielding increasingly tighter upper bounds.

### 2.5.1 Lower Bounds

Achlioptas/Molloy described and analysed an algorithm called *list colouring* that 3-colours a $\mathcal{G}_{n,m}$-graph in linear time with success probability provably greater than some $\varepsilon > 0$, for all $c < 3.85(\ldots)$. In view of the fact that 3-colouring has a sharp threshold (c.f. Section 2.2) this proves that $\mathcal{G}_{n,m}$-graphs are a.s. 3-colourable for $c < 3.85(\ldots)$.

The algorithm works as follows. Associated with each node there is a list initially containing all three colours. The nodes are partitioned in *coloured* nodes, *boundary* nodes and *untouched* nodes. Initially all nodes are untouched. Repeat the following until finished.

- As long as there are uncoloured nodes, pick a node $v$ at random. If there are boundary nodes pick one node $v$ with minimal list length. Otherwise pick $v$ amongst the untouched nodes.
- Colour $v$ with an extra colour if its list is empty. Otherwise colour $v$ with a colour consistent with its list.
- Update all nodes $w$ in the neighbourhood of $v$ by making them boundary nodes if necessary and deleting the newly assigned colour of $v$ from their lists.

Obviously, conditional on success in each step this yields a proper colouring. Note that this is the usual greedy colouring procedure with respect to an ordering of the nodes that is determined on-line.

This algorithm is analysed using Wormald's differential-equations tool (see [Wor95, Wor99] and Section 2.1.3). We partition the boundary nodes into $S_0, S_1, S_2$, according to their list lengths. All other nodes are either contained in the set $C$ of coloured nodes or in the remainder $U$, the untouched nodes.

At time $t = 0$ all nodes are contained in $U$. The time evolution can be analysed using the principle of deferred decisions. At time $t$ we only know the outcome of the random bits corresponding to potential edges within $C^{(t)}$ and the edges connecting $C^{(t)}$ with $S^{(t)}$. When we colour a node $v_{t+1}$ in $S^{(t)}$ or $C^{(t)}$ in the course of the above algorithm the random bits for all potential edges connecting $v_{t+1}$ to $S^{(t)} \setminus v_{t+1}$ and $C^{(t)}$ 'have not yet been looked at', and will be switched on or off independently of the history of the process.

Considering the expected increase (after appropriate scaling) yields a system of differential equations. As long as $c < 3.85$ the solutions are such that $S_1$ remains strictly less than one, for all $t$. Intuitively this means that 'critical nodes' in $S_1$ are processed at a faster rate than they are produced, as long as $c < 3.85$.

### 2.5.2 Upper Bounds

All known upper bounds arise from applications of the First Moment Method. We will review Achlioptas and Molloy [AM99]. We will consider the $\mathcal{G}_{n,m}$-model with $m = r \cdot n$. Thus the average degree is $2r$. Let $\{\alpha_i\}_{i=1}^k$ denote the relative sizes of colour classes in a partition $P : \{V_i\}_{i=1}^k$ of the nodes. There are $M = \binom{n}{2}$ potential edges and

$$T(P) := n^2 \sum_{i<j} \alpha_i \cdot \alpha_j =: \tau n^2$$

potential edges respecting the partition $P$. Thus the probability for the event $C_P$ that $P$ is a proper colouring is given by

$$\mathbb{P}[C_P] = \frac{\binom{T(P)}{m}}{\binom{\binom{n}{2}}{m}} \leq [\frac{T(P)}{\binom{n}{2}}]^m \leq (2\tau)^m O(1).$$

It is straightforward to see that the number $T(P)$ of partitions is maximal when $\alpha_i \equiv 1/k$. The total number of $k$-partitions is $k^n$ and thus the expected number of colourings is bounded above by

$$k^n (\frac{k-1}{k})^m O(1).$$

Upon substituting $m = rn$ this bound converges to zero if

$$(\frac{k-1}{k})^r < 1 \Leftrightarrow r > \frac{\ln k}{\ln k - \ln(k-1)}, \tag{2.4}$$

for $k = 3$ this corresponds to $c > 5.419(\ldots)$.

However, remember our discussion of the First Moment Method in Section 2.1.3, this bound is not necessarily tight. Even for $c < 5.419(\ldots)$ the number of $k$-colourings may be zero with probability converging to one. But if in the rare case that it is greater than zero there are 'very many' colourings, this may cause the expected value to diverge, nevertheless.

Achlioptas and Molloy used an idea that has previously been successfully applied in an analogous situation counting the number of satisfying

assignments of random $k$-SAT instances([KKK96]). Instead of counting *all* colourings they only count *some* colourings. A colouring is *rigid* if and only if for each node $v$, coloured with colour $i$, recolouring $v$ with colour $j > i$ renders the colouring non-proper. If the set of colourings is non-empty, the lexicographically maximum colouring amongst them is rigid. Thus when there are no rigid colourings, there are no colourings at all.

Achlioptas and Molloy calculate a bound on the expected number of rigid colourings that can be approximated numerically to any accuracy. A worst case over all $\{\alpha_i\}_{i=1}^k$ needs to be calculated, in order to see whether a rather complicated term is less than one, just as the term $(2\tau(\{\alpha_i\}_{i=1}^k))^r$ above. Further details concerning their calculations can be found in Section 7.3 below. We reproduce their thresholds for various values of $k$.

| k | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| c | 5.0434 | 9.1722 | 13.8958 | 19.0778 | 24.632 |

Recently Kaporis et. al. ([KKS00]) have been able to even improve on this using tail bounds for the occupancy problem. This was again inspired by a success in the analogous situation for random $k$-SAT instances. The improvement is due to considering more carefully a conjunction of events that occurs in [AM99] when bounding the expected number of rigid colourings. In [AM99] the probability of the conjunction was merely *upper bounded* by the product of the individual events. Since those events are actually negatively correlated, a tighter bound can be achieved using tail bounds from Kamath et.al. in [KMPS95]. Thus the value of 4.989(...) in [KKS00] is the best currently known upper bound for the 3-colouring threshold, apparently it is not easy to get the corresponding numerical values for $k > 3$.

## 2.6 Tripartite Graphs

In Section 2.1.1 we introduced random tripartite graphs with constant average degree. The tripartite case is of independent interest, remember that finding 3-colourings for tripartite graphs is NP-hard.

We will first review a result of Petford and Welsh ([PW89]), empirically describing the *critical slowing down* of a Markov chain on the colourings of random tripartite graphs called *Antivoter Algorithm*.

Later we will report on two rigorously analysed algorithms finding colourings for sparse random tripartite graphs.

Finally the random tripartite model has been employed to show the existence of uniquely 3-colourable graphs with large *girth* (length of the shortest cycle).

### 2.6.1 The Antivoter Phenomenon

The Antivoter Markov chain proposed by Petford and Welsh ([PW89]) samples from the (not necessarily proper) colourings of a graph $G$. It is similar in spirit to the *Gibbs Sampler* Markov chains, studied by statistical physicists.

For a graph $G$ consider the set $\Xi = \{0,1,2\}^n$ of all possible 3-colourings, the set of *states*. The number of 'badly' coloured edges can be interpreted as an energy–function or Hamiltonian, as discussed in Section 1.2. The ground states of minimum energy zero, if they exist, are the proper colourings.

**Antivoter-Algorithm:**

0.) Choose an initial state $\xi_0$ uniformly at random.

1.) While there are monochromatic edges in $\xi_t$:

1a.) When in state $\xi_t$ pick one vertex $i \in V$ that is participating in a monochromatic edge uniformly at random.

1b.) Generate $\xi_{t+1}$ from $\xi_t$ by re-colouring vertex $i$ in colour $\gamma$ with probabilities proportional to $e^{-n_\gamma(\xi_t,i)}$ (where $n_\gamma(\xi,i)$ is the number of neighbours of $i$ coloured in colour $\gamma$ when in state $\xi$).

2.) Return the (correct!) colouring and the number of iterations.

In a little detour we will illustrate the connections between the Antivoter Algorithm and the Gibbs sampler. The Gibbs sampler is a Markov kernel $P$ that is the product of $n$ *local characteristics* $\Pi_i$ with $i$ chosen uniformly at random from the set of vertices:

$$P(\xi,\eta) = \mid V \mid^{-1} \sum_{i \in V} \Pi_i(\xi,\eta) \tag{2.5}$$

$$\Pi_i(\xi,\eta) := \begin{cases} Z_i^{-1} exp(-H(\eta(i)\xi_{V\setminus\{i\}})) & \text{if } x_{V\setminus\{i\}} = y_{V\setminus\{i\}} \\ 0 & \text{otherwise} \end{cases} \tag{2.6}$$

$$Z_i := \sum_{\gamma=0}^{2} \exp(-H(\gamma\xi_{V\setminus\{i\}})) \tag{2.7}$$

Here $(\gamma\xi_{V\setminus\{i\}})$ is a shorthand notation for the state coinciding with $\xi$ on $V \setminus \{i\}$ and with vertex $i$ being coloured in colour $\gamma$. It is straightforward to check that the local characteristics satisfy detailed balance with respect to

$$\xi \mapsto Z_H^{-1} \exp(-H(\xi)), \tag{2.8}$$

the Gibbs measure on $\Xi$.

This carries through to the Gibbs sampler being a convex combination of local characteristics. Since the Gibbs sampler is ergodic, the Gibbs measure is its unique invariant distribution.

Apart from the fact that only nodes with monochromatic edges are recoloured the Antivoter is a Gibbs sampler sampling from the Gibbs measure (2.8) with Hamiltonian function

$$H(\xi) := \frac{1}{2} \sum_{i=1}^{n} n_{\xi(i)}(\xi, i) \tag{2.9}$$

Observe that $H(\xi)$ is simply the number of 'badly' coloured edges in state $\xi$. We end our detour explaining how the Gibbs sampler described above reproduces the updating rule 1b. Assume vertex $i$, presently coloured $\gamma_0$, is to be recoloured with colour $\gamma$ according to the Gibbs updating rule. Let $h_\gamma := H(\gamma\xi_{V\setminus\{i\}})$, $h := H(\xi)$ and $n_\gamma = n_\gamma(\xi, i)$. The number of $\gamma_0$-monochromatic edges is decreased by $n_{\gamma_0}$ and increased by $n_\gamma$, therefore $h_\gamma = h - n_{\gamma_0} + n_\gamma$. Thus the exponential terms in (2.6) are proportional to $e^{-n_\gamma}$ as required in 1b.

Petford and Welsh describe in [PW89] that the expected hitting time appears to be roughly proportional to $n$ for fixed average degree $c$. However, there seems to be a significant slow-down (empirical divergence of the proportionality constant) at a value of $c \approx 5$. This resembles the 'critical slowing down' in the vicinity of phase transitions, as described in Section 1.1.

### 2.6.2 Results for 3-Colouring Tripartite Graphs

With an eye on the Branching Tree Connection, it is at least intuitively clear that the algorithmic *lower bounds* essentially carry over to the $k$-partite case. The fact that the graph is tripartite has little effect locally, as seen from a BFS. For the appearance of the 3-core our techniques explained below work equally well for the tripartite model predicting the same critical $c = 3.35(\ldots)$. We are convinced (and have empirically tested this) that the same holds true for the colouring algorithm of Achlioptas and Molloy. Thus for sufficiently small constant average degree $c$, 3-colouring can be achieved efficiently.

For *upper bounds* we will briefly report on two different algorithms, both rigorously analysed.

The first upper bound is due to Blum and Spencer ([BS95]), who designed and analysed an algorithm for efficiently $k$-colouring $k$-colourable tripartite graphs with average degree as low as $c = c(n) = n^\varepsilon, \varepsilon > 0$. The essence of their method can be explained as follows. Consider two nodes $u$ and $v$ and their common neighbourhood $\Gamma(u) \cap \Gamma(v)$. If $u$ and $v$ are in the same colour

class, the expected number of common neighbours is $2/3 \cdot n \cdot (3/2 \cdot c(n)/n)^2$, and $1/3 \cdot n \cdot (3/2 \cdot c(n)/n)^2$, otherwise. Note that those numbers are the sum of $\theta(n)$ independent indicators (one for each potential edge) and therefore Chernoff bounds apply. Thus for each pair $(u, v)$ we can tell whether they are in the same colour class with a subpolynomially small error. So the presence or absence of all $\theta(n^2)$ potential edges can be a.a.s. decided correctly. Note that counting nodes in the mutual neighbourhood $\Gamma(u) \cap \Gamma(v)$ is tantamount to counting paths of length 2 connecting $u$ and $v$. They refined this basic idea by appropriately considering paths of lengths greater than 2, leading to the aforementioned lower bound of $c(n) = n^{\varepsilon}$.

Blum and Spencer have also considered the *semi-random model*, where an adversary builds up the graph by considering all potential edges in an order of his choice, its choice being reverted with *noise rate* $p$. They refined the above basic idea by using the concept of $k$-*links*. It is convenient to view the colour classes as cliques in the complement graph $(V, K_n \setminus E)$. Nodes $u, v$ in the same clique will share more edges (forming the links for $k = 3$, $K_{k-1}$'s otherwise) within $\Gamma(u) \cap \Gamma(v)$ than nodes in different cliques. Draw a 'meta-edge' between $u, v$ if there are sufficiently many links between them and output the connected components of the 'meta-graph' as a guess for the colour classes. This algorithm efficiently recovers the 3-colouring of the graph with high probability for noise rates as low as $p \geq n^{0.6+\varepsilon}, \varepsilon > 0$.

The second upper bound (Alon and Kahale [AK94]) is based on spectral methods. They were able to prove correctness of their intrinsically efficient algorithm based on calculating eigenvectors of the adjacency matrix of $G$, for some *constant* average degree $c_0$, independent of $n$, improving on the results of Blum and Spencer. The overall idea is that the eigenvectors are a.a.s. sufficiently close to characteristic functions of the colour classes. We refer the reader to the paper for details.

It is apparently hard to 'hide' large planted objects in random graphs, (c.f. the discussion in [FM97]). Using spectral techniques appears to be superior to using 'local information' in the spirit of Blum and Spencer's approach. We have had a similar experience when studying the problem of recovering the cluster structure of 'noisy' clique graphs (joint work with J. Ernst and V. Heun [EHV00]).

### 2.6.3 Uniquely Colourable Graphs of Large Girth

What makes a (random) graph $(k+1)$ chromatic? Remember our discussion of Fourier-Walsh spectra of graph properties. 'Local' properties are characterised by the appearance of *constant* size subgraphs. Indeed, the phase transition of the property 'non-$k$-colourable' can apparently not be explained

by the appearance of small substructures, such as a copy of $K_k$. Small subgraphs that 'explain why the chromatic number is $k$', are rather dense, in fact too dense to appear in a random graph with constant average degree. We feel that our work is relevant in the context of uniquely colourable graphs of large girth, as discussed in [BS76] and [EHK98].

The girth $\gamma(G)$ of a graph $G$ is the length of the shortest cycle. Graphs with large girth are 'locally tree-like', in particular they are very sparse. Note that a uniquely 3-colourable graph becomes 4-chromatic as soon as one edge connecting nodes in the same colour class is inserted.

The existence of uniquely $k$-colourable graphs with arbitrary girth $g$ was proven using the *probabilistic method* by Bollobás and Sauer [BS76]. It came as somewhat of a surprise that there should be 'almost tree-like' graphs exhibiting high chromatic number. Starting with a random tripartite graph with average degree $c(n)$ they proved that it contains a uniquely $k$-colourable subgraph with large girth on $\theta(n)$ nodes with positive probability, as long as $c(n) \geq n^{\varepsilon}, \varepsilon > 0$. Thus such an object exists.

Later Emden-Weinert et.al. [EHK98] improved the above construction and found uniquely 3-colourable graphs of large girth for average degrees down to $c > c_0$, where $c_0$ depends on $k$ but is independent of $n$, more precisely $c_0(k) = k^{10}$. Their focus is on worst case analysis, and they searched for uniquely 3-colourable subgraphs in tripartite random graphs mainly for the sake of finding counterexamples. It may be interesting to note that a new kind of subgraph (to be defined and discussed later, see Chapter 4) suddenly appearing at an average degree of $c = 4.91(\ldots)$ $(k = 3)$ is apparently a.s. 'almost' uniquely 3-colourable, according to our simulations. If this were true one could possibly considerably if not optimally improve on the construction of [EHK98].

# Chapter 3

# Threshold Phenomena in Branching Trees

We ask the reader to re-inspect the definition of branching trees and ownership in Section 2.3.3, pp. 29. In this chapter we shall analyse phase transitions in branching trees with $\mathrm{Po}_c$ progeny distribution, that may be seen as the analogues in *branching trees* of certain phase transitions in *random graphs* with average degree $c$, such as the sudden appearance of a giant $k$-core. We describe a new phase transition in appropriately *coloured* branching trees that has guided us to 'discovering' a new phase transition in $k$-partite random graphs, the appearance of a giant *magic subgraph* (see Chapter 4).

We shall study how the probability of the 'appearance' of certain subtrees, such as binary trees, in a $\mathrm{Po}_c$ branching tree depends on the expected progeny distribution $c$. Such events can be defined recursively, and since disjoint sub-trees are merely independent copies of $\mathrm{Po}_c$ branching trees their probabilities can be quite easily described by *recursive equations*.

We shall later show that the very same recursive equations can be employed in a rigorous analysis of the appearance of *giant subgraphs*, thereby providing structural insight into why the Branching Tree Connection correctly 'predicts' numerical values for appearance and expected sizes of giant subgraphs in random graphs.

For the analogue of the $k$-core (Section 3.1) we have worked out the analysis of the recursive equations in great detail. Most results can be easily 'guessed' from inspecting the plots in Figures 3.1 and 3.2 and the proofs consist in curve discussion essentially at undergraduate level. Therefore the corresponding results for the analogue of the magic subgraph are stated much more concisely, the curve discussions being completely analogous to the sooner case (Section 3.2). In Section 3.3 we present some additional material concerning branching trees and phase transitions therein.

## 3.1 Owners and Runs, Related to the $k$–Core

As the 'analogue' of the appearance of the 3-core in the context of (unrestricted) $\mathrm{Po}_c$-branching trees consider the edge monotone events

$$\mathcal{B} = \text{'the root owns a (complete, unrestricted) binary tree'},$$

and

$$\mathcal{C} = \text{'the root owns a Cayley tree'},$$

i.e. the root has at least three neighbours, each of which root a binary tree.

Denote the probability of the former event by $q(c)$ and the latter by $p(c)$. The probability $q(c)$ can be computed rigorously for branching trees using recursive equations, and $p(c)$ can be easily computed from $q(c)$. We shall presently see that there exists a $c_{\mathrm{crit}} = 3.35(\ldots)$ such that $p(c) \equiv 0$ for all $c < c_{\mathrm{crit}}$ and equal to some $p(c) > 0.2674(\ldots)$ otherwise.

Essentially, what is called the Branching Tree Connection in [PSW96] is the observation that $c_{\mathrm{crit}}$ and $p(c)$ 'happen' to be exactly the critical value and the relative expected size of the 3-core in the random graph.

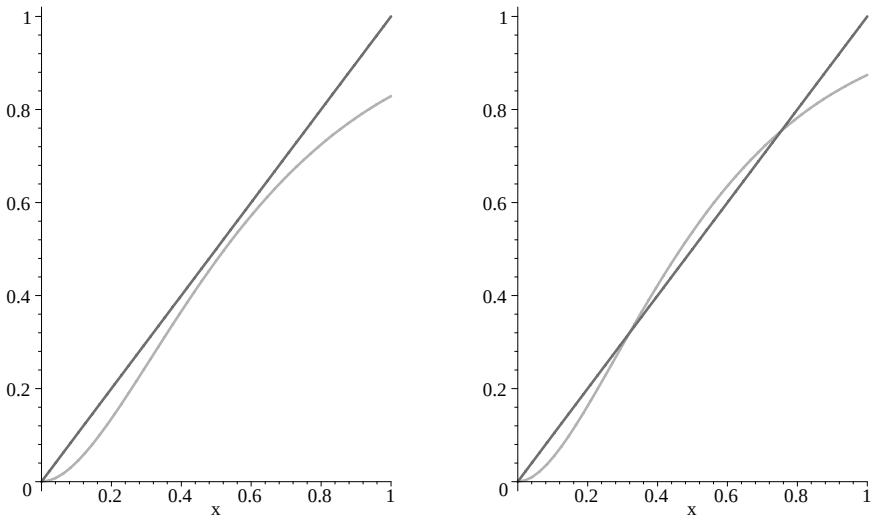


Figure 3.1: The plots show $x \mapsto f_2(3.2, x)$ and $x \mapsto f_2(3.6, x)$. Note that for $c < 3.35(\ldots)$ the only root of $x = f_2(c, x)$ is zero, whereas for $c > 3.35(\ldots)$ the largest root $q(c)$ is positive as can be seen from Figure 3.2 below, portraying $c \mapsto q(c)$. For $k > 3$ similar plots can be drawn.

Here we consider stopped branching trees, i.e. shell-wise fanning out processes discovering the $r$-neighbourhood of the root in some breadth first search (BFS) order. Note that subtrees induced by the children of the root are just independent copies of branching trees with radius $r-1$. The probability $q_r$ of the edge monotone event

$$\mathcal{C}_r = \text{'the root owns a binary tree within its } r\text{-neighbourhood'}$$

can therefore be computed recursively as follows.

$$q_0(c) \equiv 1, \qquad q_{r+1}(c) := f_2(c, q_r(c)) := \underbrace{\mathbb{P}\left[\mathrm{Po}_{cq_r(c)} \geq 2\right]}_{\substack{\text{at least two} \\ \text{children 'fixed'}}}. \tag{3.1}$$

The value of $q(c)$ is the largest root of $x = f_2(c, x)$. The value of this root drastically changes when $c$ becomes larger than $c_{\mathrm{crit}} = 3.35(\ldots)$. For *sub-critical* $c < 3.35(\ldots)$ the only root of $x = f_2(c, x)$ is $x = 0$ whereas for *super-critical* $c > 3.35(\ldots)$ there 'suddenly' appear additional roots greater than zero, c.f. Figures 3.2 and 3.1.

The probability of the edge monotone event

$$\mathcal{C}_r = \text{'the root owns a Cayley tree within its } r\text{-neighbourhood'}$$

is

$$p_r(c) := f_3(c, q_{r-1}(c)) := \mathbb{P}\left[\mathrm{Po}_{cq_{r-1}(c)} \geq 3\right].$$

Finally $p(c) := f_3(c, q(c))$, c.f. Figure 3.2.

We first collect some properties of $f_2$ in Proposition 3.1.1 that are rather obvious from the plots in Figure 3.1 but require a formal proof. We advise the reader to first check the plausibility of the assertions by looking at the plots before reading the proof. Remember that $q(c)$ is the largest root of $x = f_2(c, x)$.

**Proposition 3.1.1** *The function $f_2(c, x)$ is an entire analytic function of $c$ and $x$, strictly increasing in $x$ for $c$ fixed and strictly increasing in $c$ for $x$ fixed. All derivatives (w.r.t. $x$) are bounded on (the compact set) $[0, 1]$.*

*The function $g(c, x) := f_2(c, x) - x$ has a root at $x = 0$, and either none, one or two additional roots in $[0, 1]$.*

*In the first case when $c < 3.35(\ldots)$,*

$$f_2(c, x) < x(1 - \varepsilon(c)),$$

*for all $x \in ]0, 1]$ and for some $\varepsilon(c)$.*

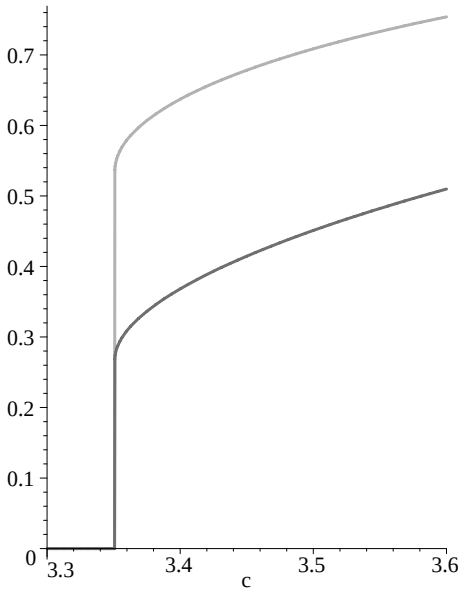


Figure 3.2: The larger curve shows the numerical value of the largest root $q(c)$ of $x = f_2(c, x)$, the second curve shows $p(c) = f_3(c, q(c))$. The latter is the probability that the root of the unbounded branching tree owns an unrestricted Cayley tree (which 'corresponds' to being in the 3-core). Note that $p(c)$ is always slightly larger than $q(c)/2$, for all $c > 3.35(\ldots)$. For $k > 3$ similar plots can be drawn.

*The exceptional second case happens only at one value of c, that can be verified to be* $3.35(\ldots)$*, numerically.*

*In the third case for* $c > 3.35(\ldots)$ *the derivative* $t(c) := \frac{\partial}{\partial x} f_2(c, x)|_{x=q(c)}$ *is strictly less than one, and the curvature of* $f_2(c, x)$ *is negative for* $x$ *in an open interval containing* $[q(c), 1]$*. Thus by Taylor's Theorem we have for all* $x$ *in such an interval*

$$f_2(c, x) \leq q(c) + t(c) \cdot (x - q(c)).$$

*Furthermore, in that third case* $q(c)$ *is strictly increasing and* $C^\infty$ *in* $c$*, and the derivative* $t(c)$ *is equal to* $2 \cdot \mathrm{Po}_{cq(c)}|_{\geq 2}(2)$*, which is strictly decreasing.*

**Proof** Analyticity and monotonicity are easily verified, as well as the 'trivial' root $f_2(c, 0) \equiv 0$.

Note that $g(c, x) := f_2(c, x) - x$ starts at 0 and ends at $-\exp(c) - c\exp(c) < 0$. The second derivative $g''(c, x) = (c^2 - c^3 x) \cdot \exp(-cx)$ of $g$

w.r.t. $x$ starts off with a positive value and changes sign at $x_0(c) := 1/c$. Therefore the first derivative $g' = xc^2 \cdot \exp(-cx) - 1$ is unimodal, starting at $g'(0) = -1$, ending at $g'(1) = -(1 - c^2 \cdot \exp(-c)) < 0$ and attaining its maximum $g'(c, 1/c) = c \cdot \exp(-1) - 1$ at $x_0(c) = 1/c$.

The maximum of $g'(c, 1/c) = c \cdot \exp(-1) - 1$ is greater than zero for all $c > e$. If they exist, $x_1(c)$ denotes the $x$-value where $g'$ changes sign from $-$ to $+$ and $x_2(c)$ the $x$-value where it changes back. Necessarily for $c > e$, $x_1(c) < x_0(c) < x_2(c)$, as $g'$ is continuous.

If $c \leq e$ we are in the first case, as we will presently show. Now $g'(c, x)$ is non-positive. The anti-derivative $g$ is negative, for $g'$ starts at $-1$ and remains less than, say, $-1/2$ within an interval of size $2\varepsilon(c)$ ($g'$ is continuous). Therefore when $c \leq e$, $f_2(c, x) < x(1 - \varepsilon(c))$ for all $x \in [0, 1]$.

From now on we will assume that $c > e$ (note that all three cases are still possible). Now $g'$ increases from $-1$ until $x = x_0(c)$ crossing zero at $x_1(c) < x_0(c)$ then falls again crossing zero at $x_2(c) > x_0(c)$. This implies for $g$ that it falls from zero, attains a local minimum at $x_1(c)$, starting to rise again attaining its maximum at $x_2(c)$. We get $0 = g'(c, x_2(c)) = c^2 x_2(c) \exp(-cx_2(c)) - 1$, knowing that $x_2(c) > 1/c$ and $1/c > 1/e$ Therefore we know that $cx_2(c) = -W_{-1}(-1/c)$. Here $W$ is the Lambert-$W$-function* which is the inverse of $x \mapsto xe^x$. Inserting this expression for $x_2(c)$ into the definition of $g$ yields

$$g(c, x_2(c)) = \frac{cW_{-1}(-1/c) + 1 - W_{-1}(-1/c) + W_{-1}(-1/c)^2}{c \cdot W_{-1}(-1/c)}.$$

We have not succeeded in explicitly solving $g(c, x_2(c)) = 0$ for $c$, even when using the Lambert-$W$-function. Fortunately we can prove that there is only one root and determine it numerically to any accuracy. The derivative of $g(c, x_2(c))$ with respect to $c$ is

$$\frac{-W_{-1}(-1/c)}{c^2},$$

this is positive for all $c \geq e$. For $c = e$ we get $g'(e, x_2(e)) = 1 - 3/e < 0$. Thus there is exactly one value $c_{\text{crit}} = 3.350918871511672732(\ldots)$ such that $g(c_{\text{crit}}, x_2(c_{\text{crit}})) = 0$.

This yields the desired result for $c > 3.35(\ldots)$ (and for $e < c < 3.35(\ldots)$ because the only local maximum of $g$ besides the one at $x = 0$ is strictly

*Just like the more familiar logarithm (ln), which is the inverse of $x \mapsto e^x$, $W$ has *branches* labelled by indices in $\mathbb{Z}$, so $W_{-1}$ is the $(-1)$-st branch of $W$. For an excellent discussion of $W$ we refer the reader to [CGH+96]. Also, some computer-algebra tools, like Maple, can perform computations involving $W$ with the same ease as computations involving ln.

smaller than 1). As $g$ is positive at $x_2(c)$, the largest root $q(c)$ lies (strictly) right of $x_2(c)$. For all $x > x_2(c)$ including $q(c)$ we have negative $g'$ (i.e. $f_2' < 1$), and since $x_2(c) > x_0(c)$ the curvature is indeed negative in $]x_2(c), 1] \supseteq [q(c), 1]$.

Why is $q(c)$ strictly increasing for $c > 3.35(\ldots)$? Monotonicity implies that

$$f_2(c(1+\xi), q(c)) > f_2(c, q(c)) = q(c).$$

By continuity there must be some further root of $f_2(c(1+\xi), x) = x$ strictly larger than $q(c)$ since $f_2(c(1+\xi), 1) < 1$.

Continuity ($C^\infty$) of $c \mapsto q(c)$ follows from the Implicit Function Theorem. $F(c, x) := f_2(c, x) - x$ is $C^\infty$ on $\mathbb{R}^2$ and has a root at some $(c_0, q(c_0))$, furthermore $\frac{\partial}{\partial x} F(c_0, q(c_0)) \neq 0$. Thus there is (at least) a small neighbourhood of $c_0$ such that $c \mapsto q(c)$ is a $C^\infty$ function.

Finally, we know that $cq(c) > cx_2(c) > 1$ from the above discussion for $c > 3.35(\ldots)$ and thus

$$t(c) = \frac{\partial}{\partial x} f_2(c, x)|_{x=q(c)} = c^2 q(c) e^{-cq(c)} = 2 \cdot \mathrm{Po}_{cq(c)}(2)/q(c) = 2 \cdot \mathrm{Po}_{cq(c)|\geq 2}(2)$$

is strictly decreasing, as $a \mapsto ae^a$ is for $a > 1$. □

**Lemma 3.1.2** *In the* $\mathrm{Po}_c$ *branching tree the probability* $p_r(c)$ *of the event* $\mathcal{C}_r$ *'the root is an* $r$*-Cayley owner' converges (w.r.t.* $r$*) to* $p(c)$ *from above at an (at least) exponential rate for all* $c \neq 3.35(\ldots)$*, i.e.*

$$p(c) \leq p_r(c) \leq p(c) + O_{r\to\infty}(t(c)^r).$$

*If* $c < 3.35(\ldots)$ *(sub-critical case)* $p(c)$ *is zero whereas for* $c > 3.35(\ldots)$ *(super-critical case),* $p(c) \geq 0.2674314(\ldots)$*, and it is strictly increasing in* $c$*.*

*For the probability* $q_r(c)$ *of the event* $\mathcal{B}_r$ *'the root is an* $r$*-binary owner' we have*

$$q(c) \leq q_r(c) \leq q(c) + t(c)^r.$$

*If* $c < 3.35(\ldots)$*,* $q(c)$ *is zero whereas for* $c > 3.35(\ldots)$ *(super-critical case),* $q(c) \geq 0.5349949(\ldots)$*, and it is strictly increasing in* $c$*.*

*In the super-critical case* $t(c)$ *can be chosen to be the derivative w.r.t.* $x$ *of* $f_2(c, x)$ *at* $x = q(c)$*.*

**Proof** We first analyse the convergence of the probability $q_r(c)$, of the event $\mathcal{B}_r$ 'the root is an $r$-binary owner', towards the largest root $q(c)$ of $f_2(c, x) = x$ in $[0, 1]$. We will use the results from Proposition 3.1.1.

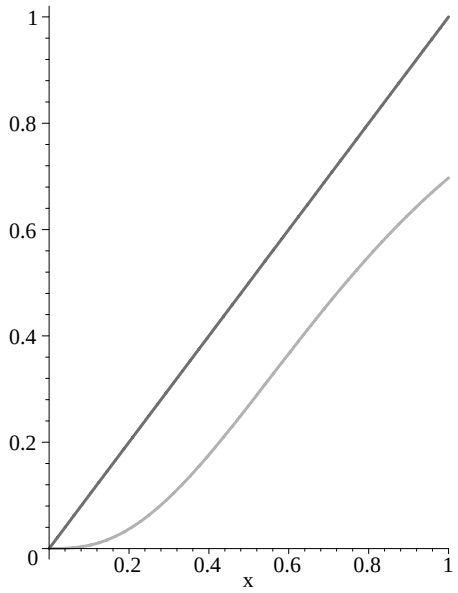


Figure 3.3: A plot of $x \mapsto f_3(3.6, x)$.

If $x \in ]q(c), 1]$ we have $f_2(c, x) \in ]q(c), 1]$, from the monotonicity of $f_2$. By induction over $r$ starting at $q_0(c) = 1$ this implies

$$q_r(c) > q(c).$$

Furthermore, for all $x$ in $]q(c), 1]$

$$\frac{f_2(c, x) - q(c)}{x - q(c)} < t(c).$$

In the sub-critical case we define $t(c) := 1 - \varepsilon(c)$, c.f. Proposition 3.1.1. By induction over $r$ starting at $q_0(c) = 1$ this yields

$$q_r(c) \leq q(c) + (t(c))^r.$$

Figure 3.3 illustrates the fact that $f_3(c, \cdot)$ 'distorts' its input in a smooth and analytic way, mapping errors of $\delta$ to errors of $O(\delta)$, the constants depending only on $c$. Formally $f_3(c, \cdot)$ is analytic and has bounded derivatives, the results will carry over from $q_r(c)$ to $p_r(c)$ by a straightforward application of Taylor's Theorem, i.e.

$$\begin{aligned} & f_3(c, q(c) + (t(c))^r) \\ =& f_3(c, q(c)) + \frac{\partial}{\partial x} f_3(c, x)|_{x=q(c)} \cdot (t(c))^r + \frac{\partial^2}{\partial x^2} f_3(c, x)|_{x=\theta} \cdot (t(c))^{2r}, \end{aligned}$$

where $\theta$ is in $[0,(t(c))^r]$. The error is $O_{r\to\infty}(t(c)^r)$ since all constants are bounded, depending only on $c$. □

The next lemma discusses the probability of runs, see Definition 2.3.7 on p. 31, i.e. nodes that are $r$-owners but not $(r+1)$-owners. This somehow reflects the 'robustness' of ownership further aspects of which shall discuss below.

**Lemma 3.1.3** *In the* $\mathrm{Po}_c$ *branching tree the probability* $\mathbb{P}[\overline{\mathcal{C}_{r+1}} \wedge \mathcal{C}_r]$ *that the root is an* $r$*-Cayley owner but not an* $r+1$*-Cayley owner is* $O_{r\to\infty}(t(c)^r)$*. So is the probability* $\mathbb{P}[\overline{\mathcal{C}_{r+1}}|\mathcal{C}_r]$*.*

**Proof** Select three of the at least $\mathrm{Po}_{cq_{r-1}}|_{\geq 3} \geq 3$ children of an $r$-owner $v$ that are $r-1$-binary owners. Denote by $\mathcal{B}_r$ the event that the root is an $r$-binary owner. For $r-1$-binary-owners we have (c.f. Lemma 3.1.2)

$$\begin{aligned}&\mathbb{P}\left[\mathcal{B}_{r-1} \wedge \overline{\mathcal{B}_r}\right] = \mathbb{P}[\mathcal{B}_{r-1}] - \mathbb{P}[\mathcal{B}_{r-1} \wedge \mathcal{B}_r]\\ &= \mathbb{P}[\mathcal{B}_{r-1}] - \mathbb{P}[\mathcal{B}_r] \leq q(c) + O_{r\to\infty}(t(c)^{r-1}) - q(c) = O_{r\to\infty}(t(c)^r),\end{aligned}$$

i.e. the selected children of $v$ are also $r$-binary owners with probability $1 - O_{r\to\infty}(t(c)^r)$, independently. Even when considering only the three selected children the probability that some child fails to be an $r$-binary owner is therefore still $1 - O_{r\to\infty}(t(c)^r)$. So this is the probability that the root $v$ is a $(r+1)$-Cayley owner conditional on being an $r$-Cayley owner. We remark that $\mathbb{P}[\overline{\mathcal{B}_r} \wedge \mathcal{B}_{r-1}]$ and $\mathbb{P}[\overline{\mathcal{B}_r}|\mathcal{B}_{r-1}]$ differ only by a constant factor disappearing in Landau notation. □

## 3.2 Coloured Owners and Runs, Related to the Magic Subgraph

We will now consider *coloured branching trees*. We 'translate' what a BFS will 'see' in a random tripartite graph to the 'branching tree world'. Whenever new children are generated they are uniformly and independently assigned one of the colours other than the colour of their respective fathers. We will say that the root is a *coloured owner* if it contains a complete binary tree such that every father node has children in all available colours. Such a tree will be called a *coloured binary tree*.

The theory of coloured owners is very similar to the theory of 'ordinary' owners discussed above. We will use the same symbols $f_2, q$ as we did for the 'ordinary' owners.

The probability $q_r$ of the edge monotone event 'the root owns a coloured binary tree within its $r$-neighbourhood' can be computed recursively as follows.

$$q_0(c) \equiv 1, \qquad q_{r+1}(c) := f_2(c, q_r(c)) := \underbrace{\mathbb{P}\left[\mathrm{Po}_{cq_r/2} \geq 1\right]^2}_{\substack{\text{at least one child} \\ \text{of each colour 'fixed'}}}. \tag{3.2}$$

The value of $q(c)$ is the largest root of $x = f_2(c, x)$. The value of this root drastically changes when $c$ becomes larger than $c_{\text{crit}} = 4.91(\ldots)$ (for $k = 3$). For *sub-critical* $c < 4.91(\ldots)$ the only root of $x = f_2(c, x)$ is $x = 0$ whereas for *super-critical* $c > 4.91(\ldots)$ there 'suddenly' appear additional roots greater than zero, c.f. Figures 3.4 and 3.5. There is not really an equivalent to Cayley-owners, so there are no $p$'s.

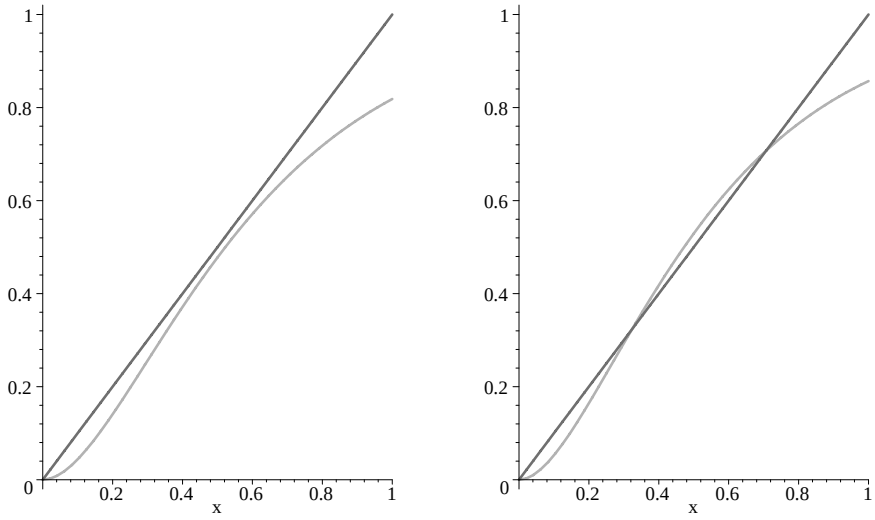


Figure 3.4: 'Coloured case', $k = 3$: The plots show $x \mapsto f_2(4.7, x)$ and $x \mapsto f_2(5.2, x)$. Note that for $c < 4.91(\ldots)$ the only root of $x = f_2(c, x)$ is zero, whereas for $c > 4.91(\ldots)$ the largest root $q(c)$ is positive as can be seen from Figure 3.5 below, portraying $c \mapsto q(c)$. For $k > 3$ similar plots can be drawn.

For larger $k$ Equation 3.2 generalises to:

$$q_0^{(k)}(c) \equiv 1, \qquad q_{r+1}^{(k)}(c) := f_2^{(k)}(c, q_r^{(k)}(c)) := \underbrace{\mathbb{P}\left[\mathrm{Po}_{cq_r^{(k)}/(k-1)} \geq 1\right]^{(k-1)}}_{\substack{\text{at least one child} \\ \text{of each colour 'fixed'}}}.$$

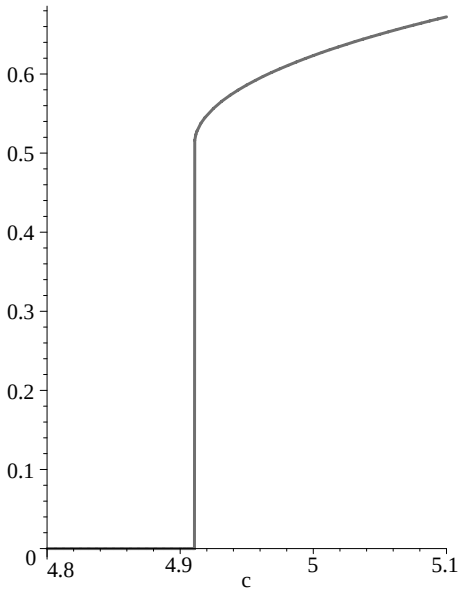


Figure 3.5: 'Coloured case', $k = 3$: Numerical value of the largest root $q(c)$ of $x = f_2(c, x)$, This is the probability that the root of the unbounded branching tree owns an unrestricted *coloured* binary tree (which 'corresponds' to being in the magic subgraph to be defined later). For $k > 3$ similar plots can be drawn. Compare this to the empirically observed sizes of magic subgraphs portrayed in Figure 4.5, p. 75.

In the 'coloured case' we can explicitely compute the values of $c_{\text{crit}}^{(k)}$ for all values of $k \geq 3$. For $k = 3$, solving $\frac{d}{dx}\frac{f_2^{(3)}(c,x)}{x} = 0$ for $x$ yields

$$x_0 = -\frac{2\,W{-}1(-1/2\,e^{-1/2}) + 1}{c},$$

and solving $f_2^{(3)}(c_{\text{crit}}^{(3)}, x_0) = x_0$ for $c_{\text{crit}}^{(3)}$ finally yields

$$\begin{aligned} c_{\text{crit}}^{(3)} &= -\frac{2\,W_{-1}(-1/2\,e^{-1/2}) + 1}{1 - 2\,e^{W_{-1}(-1/2\,e^{-1/2})+1/2} + \left(e^{W_{-1}(-1/2\,e^{-1/2})+1/2}\right)^2} \\ &= 4.910814964568255898751534805 24\ldots. \end{aligned}$$

Along these lines we get the general formula

$$c_{\text{crit}}^{(k)} = -\frac{(k-1)\,W_{-1}(-e^{-(k-1)^{-1}}\,(k-1)^{-1}) + 1}{\left(1 - e^{W_{-1}(-e^{-(k-1)^{-1}}(k-1)^{-1})+(k-1)^{-1}}\right)^{k-1}}$$

For $k = 3, \ldots, 7$ we have filled in the numerical values into the following table.

| $k$ | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $c_{\text{crit}}$ | 4.9108 | 9.2673 | 14.036 | 19.112 | 24.434 |

Although most results in Section 3.1 could have been 'guessed' from the plots in Figures 3.1 and 3.2, we have thoroughly explained how they can be rigorously proved by elementary curve discussion, there. We shall be much more concise here, as the situation is completely analogous, c.f. Figures 3.4 and 3.5.

**Proposition 3.2.1** *Completely analogous to Proposition 3.1.1. Merely replace* $3.35\ldots$ *by* $4.91(\ldots)$*, and* $t(c) := \frac{\partial}{\partial x} f_2(c,x)|_{x=q(c)}$ *takes the value*

$$c \cdot \mathbb{P}\left[\mathrm{Po}_{cq(c)/2} = 0\right] \cdot \mathbb{P}\left[\mathrm{Po}_{cq(c)/2} > 0\right],$$

*being strictly smaller than one.*

**Lemma 3.2.2** *In the coloured* $\mathrm{Po}_c$ *branching tree the probability* $q_r(c)$ *of the event 'the root is a coloured* $r$*-binary owner' converges (w.r.t.* $r$*) to* $q(c)$ *from above at an (at least) exponential rate for all* $c \neq 4.91(\ldots)$*, i.e. for the probability* $q_r(c)$ *of the event* $\mathcal{B}_r$ *'the root is an* $r$*-binary owner' we have*

$$q(c) \leq q_r(c) \leq q(c) + t(c)^r.$$

*If* $c < 4.91(\ldots)$*,* $q(c)$ *is zero whereas for* $c > 4.91(\ldots)$*,* $q(c) \geq 0.5116996(\ldots)$*, and it strictly increasing in* $c$*.*

*In the super-critical case* $t(c)$ *can be chosen to be the derivative w.r.t.* $x$ *of* $f_2(c,x)$ *at* $x = q(c)$*.*

**Lemma 3.2.3** *In the coloured* $\mathrm{Po}_c$ *branching tree the probability* $\mathbb{P}[\overline{\mathcal{B}_{r+1}} \wedge \mathcal{B}_r]$ *that the root is a coloured* $r$*-binary owner but not a coloured* $r+1$*-binary owner is* $O_{r\to\infty}(t(c)^r)$*. So is the probability* $\mathbb{P}[\overline{\mathcal{B}_{r+1}} | \mathcal{B}_r]$*.*

## 3.3 Miscellaneous Results

In the course of our work we have also investigated almost sure bounds on the sizes of branching trees with progeny distributions $\mathrm{Po}_c$ or $\mathrm{Bi}(n, c/n)$ and radius $r(n) = \alpha \log\log n$ for some $\alpha > 0$. Eventually we found that a very coarse estimate based on the maximum degree is sufficient for our purposes, see Proposition 5.1.1 below.

There is also a 'branching tree analogue' for the analysis of the second stage in our new proof for the appearance of the $k$-core, to be discussed below. However, we have not found a rigorous way of employing these observations for proofs in random graphs, as opposed to the material covered in Section 3.1 and 3.2.

We also state some results illustrating that ownership appears to be 'robust' against small perturbations.

### 3.3.1 Expected Size and Upper Bounds

We shall briefly review some rather well known facts about branching trees. The reader is encouraged to re-inspect Section 2.3.2 and references therein. We shall concentrate on the case $c > 1$, i.e. considering super-critical branching trees in the language customary for (Galton–Watson )branching trees. This should not be confused with our use of the words 'sub-critical' and 'super-critical' according to whether $c < c_{\text{crit}}$ or $c > c_{\text{crit}}$. The letter '$c$' will always denote both the average degree of the random graph and the first moment of the progeny distribution of the affiliated branching trees. Let $Z_r$ be the size of the $r$th generation of a branching tree.

$$Z_{r+1} := \sum_{i=1}^{Z_r} X_i, \tag{3.3}$$

the $X_i$ being i.i.d. copies of the progeny distribution $\mu$. It is assumed that $Z_0 \sim \delta_1$ (i.e. $Z_0 \equiv 1$) and $Z_1 = X_1 \sim \mu$.

**Fact 10** *The* expected *size of the $r$th generation is given by*

$$\mathbb{E}[Z_r] = \mathbb{E}[Z_1]^r = c^r.$$

The standard proof uses generation function arguments. We will give a simple and more instructive proof for binomial branching trees.

**Proof** of Fact 10 for $\mu = \text{Bi}(n, c/n)$.

A branching tree with progeny distribution $\text{Bi}(n, c/n)$ and radius $r$ can be constructed by the following procedure:

- Start with the complete $n$–ary tree of radius $r$.
- Perform independent edge percolation with probability $c/n$.
- Delete all nodes that are not connected to the root.

Now, the probability for one of the $n^r$ leafs of the complete $n$–ary tree of radius $r$ to be connected to the root is $(c/n)^r$. Therefore $\mathbb{E}[Z_r] = c^r$. □

A reader familiar with typical applications of Azuma's inequality in the theory of random graphs (see Section 2.1.3) might at first sight conjecture that the relative sizes $W_r := \frac{Z_r}{c^r}$ should be sharply concentrated around their means, asymptotically, i.e. to converge (in an appropriate sense) to a deterministic variable with value $\mathbb{E}[W_r] = c^r/c^r = 1$. This is *not* the case. Nevertheless, martingale techniques may be employed to show that the $W_r$ do converge to a random variable $W$ with $\mathbb{E}[W] = 1$. The distribution of $W$ consists of an absolutely continuous density on ${\mathbb{R}_0}^+$ plus a Dirac measure at 0 (corresponding to the event of extinction) under conditions that are satisfied in our setting. Specifically for $Z_1 \sim \mathrm{Po}_c$ we know that $\mathbb{E}[W] = 1$ and $\mathrm{Var}[W] = \frac{1}{c-1}$. Intuitively, the first few generations determine the 'fluctuation' of the sizes of all later generations. Once the 'population' has reached a certain size the law of large numbers will provide a progression that is sharply concentrated around its expected value but the 'early fluctuations' will persist.

In [KZ95] Karp and Zhang proved tight large deviation bounds for the tails of the $W_r$'s themselves, assuming arbitrary but bounded $Z_1$. Since $\mathrm{Po}_c$ is not bounded we need to take another approach. Moreover, our specific $Z_1$ is not a worst–case instance of a class of distributions as in Karp's proof, thus there is hope for tighter bounds. These bounds are provided by Lemma 3.3.1 and Corollary 3.3.2 below.

**Lemma 3.3.1** *Given $c > 1$ for any $0 < \varepsilon < 0.029 < \exp(-2/3 \cdot \frac{e^{5/3}}{c})$ in a branching tree with $Z_0 \equiv 1$ and $Z_1 \sim \mathrm{Po}_c$*

$$\mathbb{P}[Z_r \geq 1/2 \cdot \log(1/\varepsilon) \cdot (3c)^r] \leq r \cdot \varepsilon,$$

*for all $r \in \mathbb{N}$.*

**Proof** of Lemma 3.3.1.

First note by looking at any proof of the Chernoff bounds for the sum of $n_0$ independent Bernoulli random variables, that we can replace the i.i.d. Bernoulli r.v.s by Poisson r.v.s $\{X_i\}_i$ ($X_i \sim \mathrm{Po}_{c/n_0}$) without essentially altering the proof. The estimate used as an upper bound for the moment generating function of a Bernoulli variable is just the moment generating function of the corresponding Poisson variable. Therefore we know (c.f. [MR95b]) that for $\delta > 1$

$$\mathbb{P}\left[\sum_{i=1}^{t} X_i \geq \delta t c\right] \leq [\frac{e^{\delta-1}}{\delta^\delta}]^{tc}. \tag{3.4}$$

Now we proceed by induction. For $r = 1$ we explicitly show that

$$\mathbb{P}\left[Z_1 \geq 1/2 \cdot \log(1/\varepsilon) \cdot (3c)^1\right] \leq \varepsilon,$$

using

$$\sum_{i=k}^{\infty} \mathrm{Po}_c(i) \leq \frac{c^k}{k!} \sum_{i=0}^{\infty} e^{-c} \cdot \frac{c^i}{i!} = \frac{c^k}{k!}.$$

Substituting $1/2 \cdot \log(1/\varepsilon) \cdot (3c)^1 = 3/2 \cdot c \cdot \log(1/\varepsilon)$ for $k$ and Stirling's formula yield

$$\begin{aligned}\mathbb{P}[\mathrm{Po}_c \geq 3/2 \cdot c \cdot \log(1/\varepsilon)] &\leq \frac{c^{3/2 \cdot c \cdot \log(1/\varepsilon)}}{(3/2 \cdot c \cdot \log(1/\varepsilon)/e)^{3/2 \cdot c \cdot \log(1/\varepsilon)}} \\ &= \varepsilon^{3/2 \cdot c \cdot (\log(-c \log \varepsilon) + \log 3 - \log 2 - 1)}.\end{aligned}$$

This is less than $\varepsilon$ provided $0 < \varepsilon < 0.029 < \exp(-2/3 \cdot \frac{e^{5/3}}{c})$ for all $c > 1$.

For the induction step we use Bayes' formula and the Chernoff bound 3.4 from above. The $Y_i$ are independent $\mathrm{Po}_c$ variables.

$$\begin{aligned}&\mathbb{P}\left[Z_{r+1} \geq \varphi \cdot (\gamma c)^{r+1}\right] = \\ &\mathbb{P}\left[Z_{r+1} \geq \varphi \cdot (\gamma c)^{r+1} | Z_r \geq \varphi \cdot (\gamma c)^r\right] \mathbb{P}\left[Z_r \geq \varphi \cdot (\gamma c)^r\right] \\ +&\mathbb{P}\left[Z_{r+1} \geq \varphi \cdot (\gamma c)^{r+1} | Z_r < \varphi \cdot (\gamma c)^r\right] \mathbb{P}\left[Z_r < \varphi \cdot (\gamma c)^r\right] \\ \leq &\mathbb{P}\left[Z_r \geq \varphi \cdot (\gamma c)^r\right] + \mathbb{P}\left[Z_{r+1} \geq \varphi \cdot (\gamma c)^{r+1} | Z_r < \varphi \cdot (\gamma c)^r\right] \\ \overset{(3.3)}{\leq} &r \cdot \varepsilon + \mathbb{P}\left[\sum_{i=1}^{\varphi \cdot (\gamma c)^r} Y_i \geq \gamma \cdot (\varphi \cdot (\gamma c)^r) \cdot c\right] \overset{(3.4)}{\leq} r \cdot \varepsilon + [\frac{e^{\gamma - 1}}{\gamma^\gamma}]^{\varphi \cdot (\gamma c)^r \cdot c}.\end{aligned}$$

If we choose $\gamma$ such that the square brackets are less than $1/e$, e.g. $\gamma \geq 3$, and $\varphi = 1/2 \cdot \log(1/\varepsilon)$ the last expression is certainly less than $(r+1) \cdot \varepsilon$. □

**Corollary 3.3.2** *Assume $c > 1$ and choose $r = r(n) = \alpha \log\log n$.*

*The size of the $r$-th shell of an $\mathrm{Po}_c$ branching tree stopped at radius $r$ is less than $\log n \cdot (3c)^{r(n)}$ with probability $o(1/n)$. The total number $\lambda(r(n))$ of nodes in such a tree is less than $\log n \cdot (3c)^{r(n)+1}$ also with probability $o(1/n)$.*

*Analogous results hold for $\mathrm{Bi}(n, c/n)$ branching trees stopped at radius $r$ and for BFS–Balls of radius $r$ in a $\mathcal{G}_{n,p}$ graph with average degree $c$.*

**Proof** of Corollary 3.3.2

The statements concerning the sizes of the $r$-shells of the Poisson trees are valid for all sufficiently large $n$ choosing $\varepsilon = \varepsilon(n) = 1/n^2$ in Lemma 3.3.1

above. First we condition on the 'good event', that the sizes of the shells are compliant with our a.s. upper bounds. Summing over all shells yields the total size, and the resulting geometric series can be bounded as stated. Since there are only $r(n)$ 'bad events' each occurring with probability at most $r(n)/n^2$, the 'good event' fails to occur with probability at most $r(n)^2/n^2$ which is still $o(1/n)$ for our choice of $r(n)$.

For binomial trees we re–inspect the proof of Lemma 3.3.1, replacing $\mathrm{Po}_c$ by $\mathrm{Bi}(n, c/n)$. The tail bound for $r = 1$ is correct since we can use Chernoff bounds for Bernoulli random variables. Substituting $t \cdot c$ by $c$ and $\delta$ by $3\varphi(n) = 3/2 \cdot \log(1/\varepsilon(n)) = 3 \cdot \log n$ (c.f. the proof of Lemma 3.3.1 above) in inequality 3.4 yields

$$\mathbb{P}[\mathrm{Bi}(n, c/n) \geq 3 \cdot \varphi(n) \cdot c] \leq [\frac{e^{3\varphi(n)-1}}{(3 \cdot \varphi(n))^{3\varphi(n)}}]^c \leq n^{-3\,c(\log(3)-1+\log(\log(n)))}$$
$$\leq 1/n^2 = \varepsilon(n)$$

for sufficiently large $n$ ($n \geq 6$). For the induction step the Chernoff bounds for Poisson r.v. carry over, since the moment generating function of $\mathrm{Bi}(n, c/n)$ is smaller than the moment generating function of $\mathrm{Po}_c$ ($(1 - c/n + zc/n)^n \leq e^{c(z-1)}$).

In a fanning-out process exploring $r$-neighbourhoods in random graphs, the number of newly-discovered children is at each step stochastically dominated by $\mathrm{Bi}(n, c/n)$.

□

### 3.3.2 Damage Process in Branching Trees

Using *the same* recursive equations as for branching trees, we shall later turn the results discussed in Section 3.1 into a rigorous proof for the appearance of a giant set of $r$-Cayley owners[†] in the random graph, that is 'almost' the $k$-core, choosing $r = r(n)$ appropriately.

By 'almost' we mean that there are a few nodes (i.e. runs) left that do not have degree at least $k$ in the subgraph induced by the $r$-Cayley-owners. Remember that runs were defined as nodes that are $r$-Cayley-owners but not $(r+1)$-Cayley-owners, in Section 2.3. Removing runs may cause nodes that have previously had degree $k$ to have degree $k-1$ and if we continue to delete these nodes recursively we may find a big 'hole' that has been 'sparked' off by removing a single run. We shall call this recursive deletion process a 'damage process'. It may also be interesting to know how large the 'hole sparked off'

[†]A similar result for coloured binary owners seems to be within reach, too.

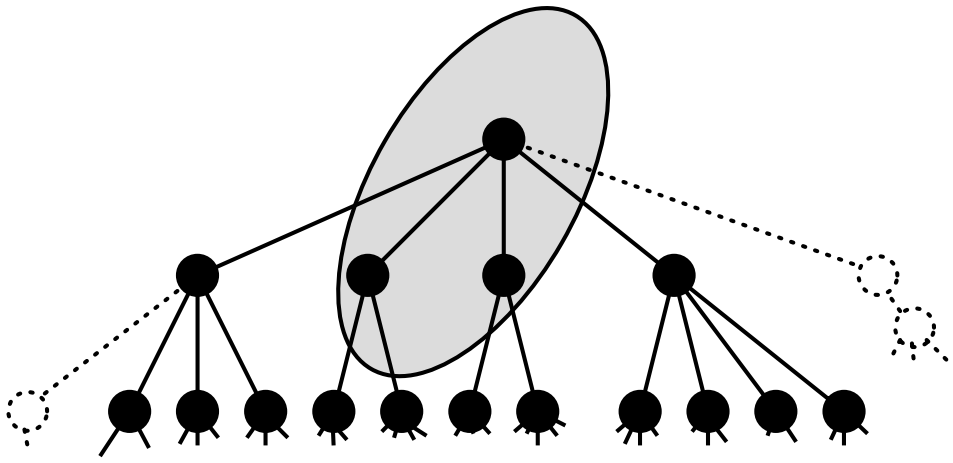

Figure 3.6: The black subtree is the cone, i.e. the union of all Cayley trees proving that the root is a Cayley owner. The subtrees drawn with dotted lines are irrelevant for ownership and hole. The hole induced by the root (shaded area) are all nodes that are removed by the recursive damage process upon removal of the root.

by removing an arbitrary node will be, i.e. to assess the robustness of the aforementioned subgraph against removing single nodes.

We found that there is a branching tree theory for analysing 'damage processes' similar to the ones discussed in the last paragraph. However, we were not able to turn *this* into a rigorous proof for random graphs. It is nevertheless worthwhile to describe our observations.

We state the following definitions only for unrestricted branching trees. See Figure 3.6 for an illustration.

**Definition 3.3.3** *Suppose the root in a* $\mathrm{Po}_c$ *branching tree is a Cayley owner. The* $k$-cone *is the union of all Cayley trees that prove that the root is a Cayley-owner.*

*The* hole induced by $v$ *are all binary owners ('away from' their respective fathers in the tree) that are deleted by a* damage process *started at* $v$*.*

*The damage process starts with deleting the root* $v$ *and recursively deletes all children in the* $k$*-cone that have degree* $(k-1)$ *(w.r.t. the* $k$*-cone!) 'away from' their respective fathers.*

**Corollary 3.3.4** *(of Proposition 3.1.1). The conditional Poisson distribution* $\mathrm{Po}_{cq(c)}|_{\geq 2}$ *for super-critical* $c > 3.35$ *has the following properties.*

$$\pi(c) := \mathrm{Po}_{cq(c)}|_{\geq 2}(2) < (1 - \varepsilon(c))/2.$$

*Moreover,* $\mathrm{Po}_{cq(c)}|_{\geq 2}(2)$ *is strictly decreasing with* $c$*. For general* $k \geq 3$ *we have*

$$\pi^{(k)}(c) := \mathrm{Po}_{cq^{(k)}(c)}|_{\geq k}(k) < (1 - \varepsilon^{(k)}(c))/(k-1).$$

**Proposition 3.3.5** *The* 3*-cone is distributed like an unrestricted* $\mathrm{Po}_{cq(c)}|_{\geq 2}$ *branching tree, where the root plays an exceptional role having at least* $\mathrm{Po}_{cq(c)}|_{\geq 3}$ *children.*

*When the root is a binary owner, the hole induced by the removal of the root is distributed like a* $\mathrm{Bi}(2, \pi(c))$ *branching tree, where* $\pi(c) = t(c)/2 = 1/2 - \varepsilon(c)$*. If the root* $v$ *is not a binary owner, the hole induced by* $v$ *is the union of all holes induced by the children of* $v$ *that are binary owners, each hole being distributed as described above.*

*Similar statements hold when we replace* 3 *by* $k > 3$*.*

**Proof** We shall again use the fact that disjoint induced subtrees are independent copies of the original branching tree.

When the root is a binary owner, we know that it has at least 2 children that are binary owners. Each of the $\mathrm{Po}_c$ children are binary owners independently with probability $q$. Therefore the 3-cone is distributed as a $\mathrm{Po}_{cq(c)}|_{\geq 2}$ branching tree.

Within the 3-cone the damage process will proceed to delete a node, whenever it has *exactly* two children. Those children will be 'exposed' upon removal of their father, i.e. have *exactly* two children themselves, with probability $\mathrm{Po}_{cq(c)}|_{\geq 2}(2) = \pi(c)$. □

Intuitively the holes should die out quickly. This follows from the theory of sub-critical[‡] and is also implicit in the analysis of the giant component for $c < 1$. The proof of the following proposition is based on a straightforward gambler's ruin argument.

**Proposition 3.3.6** *In the (unrestricted)* $\mathrm{Po}_c$ *branching tree, conditional on the root being an owner, the union of the holes induced by the removal of the root* $v$ *contains more than* $(\log n)^{M(c)}$ *owners with probability subpolynomially small in* $n$*, say . Here* $M(c)$ *is a sufficiently large constant.*

[‡]A branching tree with progeny distribution $\mu$ is *sub-critical* branching trees when $\mathbb{E}[\mu] < 1$, which should not to be confused with our usage of 'sub-critical average degrees'.

**Proof** We shall merely sketch the proof and omit some straightforward but tedious details.

The removal of the root $v$ may spark off at most $(\log n)^2$ new holes with probability at least $1 - \exp(-(1+\delta)\log(n)^2)$, for some $\delta > 0$. This is easy to see, for example, see (the proof of) our Proposition 5.1.1 below. Denote by $X_t$, the set of 'live' nodes at time $t$ of the damage process. This can be described by the following gambler's ruin walk starting at $X_0 = (\log n)^2$ and stopped at $X = 0$,

$$\begin{aligned}\mathbb{P}[X_{t+1} = i+1 | X_t = i] &= Bi(2,\pi)(2)\\ \mathbb{P}[X_{t+1} = i | X_t = i] &= Bi(2,\pi)(1)\\ \mathbb{P}[X_{t+1} = i-1 | X_t = i] &= Bi(2,\pi)(0).\end{aligned}$$

We shall be a little sloppy and merely analyse the walk $Y_t$ starting at $Y_0 := X_0$ without self-loops and without the stopping condition. In a more rigorous proof we would have to condition on the number of 'stops' due to self loops being close to their expected values, introducing a constant 'slow-down'. Because of Corollary 3.3.4 we have

$$\begin{aligned}\mathbb{P}[Y_{t+1} = i+1 | Y_t = i] &= (1-\varepsilon)/2\\ \mathbb{P}[Y_{t+1} = i-1 | Y_t = i] &= (1+\varepsilon)/2.\end{aligned}$$

What is the probability that $Y_T$ is still positive after $T = (\log n)^M$ steps? Answer: After $T$ steps $Y_T$ is distributed according to $\mathrm{Bi}(2T,(1-\varepsilon)/2) - T + (\log n)^2$. Using Chernoff bounds yields a subpolynomially small bound, say $\exp(-(\log n)^2))$, provided $M = M(c)$ is a sufficiently large constant:

$$\begin{aligned}\mathbb{P}[X_T \geq 0] &\leq \mathbb{P}[Y_T \geq 0]\\ &= \mathbb{P}\left[\mathrm{Bi}(2T,(1-\varepsilon)/2) \geq T - (\log n)^2\right]\\ &\overset{Z\sim\mathrm{Bi}(2T,(1-\varepsilon)/2)}{=} \mathbb{P}\left[Z - \mathbb{E}[Z] \geq T\varepsilon - (\log n)^2\right]\\ &\overset{\delta:=\varepsilon T-(\log n)^2}{\leq} \mathbb{P}[Z - \mathbb{E}[Z] > \delta]\\ &\overset{\mu=\mathbb{E}[Z]}{\leq} \exp(-\frac{\delta^2}{2\mu(1+\delta/(3\mu))})\\ &= \exp(-3/2\frac{(\varepsilon T - (\log n)^2)^2}{3T - 2\varepsilon T - (\log n)^2}).\end{aligned}$$

□

**Damage Process in Coloured Branching Trees**

We briefly remind the reader that

$$\tilde{t}(c) = c \cdot \mathbb{P}\left[\mathrm{Po}_{c\tilde{q}(c)/2} = 0\right] \cdot \mathbb{P}\left[\mathrm{Po}_{c\tilde{q}(c)/2} > 0\right] < 1,$$

where $\tilde{t}(c)$ denotes the quantity related to the coloured owners (see Proposition 3.2.1, p. 55) can be interpreted as follows. Suppose some coloured binary owner is deleted from the tree. We shall assume arbitrarily that it is red. Then all (of a total of $c$, on average) neighbours that have only 'fixed' children in the colour different from red and their own will no longer be 'fixed'. Therefore $\tilde{t}(c) < 1$ is the expected decrease in the number of 'fixed' nodes, if a 'fixed' node is deleted. It should therefore be possible to discuss cones and holes in a similar manner for coloured owners, as we have done for 'ordinary' owners.

It may be interesting that the quantity $t(c)$, related to the 'ordinary' owners, which is equal to $\mathbb{P}[Po_{cq}|_{\geq 2}(2)]$ can also be written as

$$t(c) = c \cdot \mathbb{P}\left[\mathrm{Po}_{cq} \geq 1\right] < 1,$$

where $\mathbb{P}[\mathrm{Po}_{cq} \geq 1]$ is the probability for being a 'unary owner'.

### 3.3.3 Expected Number of Cayley Trees, Robustness

Let $Z$ be the random variable that counts the number of Cayley trees in a branching tree. We have proven the drastic change in the probability $\mathbb{P}_c[Z = 0]$ when $c$ crosses $c_{\mathrm{crit}}$. Now we will demonstrate the surprising fact that $\mathbb{E}_c[Z]$ is perfectly 'well-behaved', it is analytic in $c$. Remember the constructive characterisation of $\mathrm{Bi}(n, c/n)$ branching trees stopped at radius $r$ in our proof of Fact 10. From this $\mathbb{E}[Z]$ can be computed by counting trees.

We shall need the following (obvious) observation.

**Proposition 3.3.7** *A* Cayley tree *of radius $r$ has $32^{r-1} - 2$ inner nodes including the root, $32^{r-1}$ leaves and $3 \cdot 2^r - 3$ edges.*

Note that when embedding a Cayley tree in the complete $n$–ary tree we have $\binom{n}{3}$ choices at the root and we have $\binom{n}{2}$ choices for all the other inner nodes.

**Proposition 3.3.8**

$$\begin{aligned}\mathbb{E}_c\left[Z\right] &= \lim_n \binom{n}{3} \cdot \binom{n}{2}^{|V_{inner}|-1} \cdot (c/n)^{|E|} \\ &= \lim_n \frac{n^{3+2(|V_{inner}|-1)}}{6 \cdot 2^{|V_{inner}|-1}} \cdot (c/n)^{|V|-1} = 8/6 \cdot 1/c^3 \cdot (c/\sqrt{2})^{3 \cdot 2^r}. \qquad (3.5)\end{aligned}$$

Formula 3.5 is perfectly analytic in $c$, nothing appears to happen at $c = 4.91(\ldots)$. We will try to shed some light on this seemingly paradoxical result below, by explaining it as a 'jackpot phenomenon'. Return to Proposition 3.3.8 and compute the expected number of Cayley trees conditional on $Z \neq 0$.

**Proposition 3.3.9** *Let $p_i = \mathbb{P}_c[Z = i]$ and $\tilde{p}_i = \mathbb{P}_c[Z = i | \, Z \neq 0]$. $\tilde{Z}$ is the random variable that counts Cayley trees conditional on $Z \neq 0$. Assume that $\mathbb{E}_c[Z]$ and $p_0 = \mathbb{P}_c[Z = 0]$ are known. Then*

$$\tilde{p}_i = \begin{cases} 0 & i = 0 \\ \frac{p_i}{1-p_0} & i > 0 \end{cases} \qquad \text{and} \qquad \mathbb{E}_c\left[\tilde{Z}\right] = \frac{\mathbb{E}_c[Z]}{1 - \mathbb{P}_c[Z = 0]}. \tag{3.6}$$

The denominator is nothing but $p_r(c)$. Qualitatively assuming a 'jackpot phenomenon' implies that in the super-critical case there will be many different Cayley trees conditional on the existence of at least one, and in the sub-critical case there must be even more to provide an expected value smooth in $c$. The typical 3-cone, the union of all Cayley trees, will therefore have many more than $32^{r-1}$ leaves at level $r$. Intuitively the cone is 'robust', or 'bushy'.

'Robustness' of the cone can also be demonstrated more directly by the following argument. We know that the cone is itself a branching tree with $\mathrm{Po}_{cq}|_{\geq 2}$ progeny distribution. Now we may introduce further errors in two different ways.

a) Either we perform 'node percolation' at the outer shell at radius $r$, that is tantamount to defining $q_0 = 1 - \varepsilon$ instead of $q_0 = 1$. Clearly, the recursive equation (describing probabilities in the cone) for binary ownership in the presence of this percolation is given by

$$f(c, x) = \sum_{k=2}^{\infty} \mathrm{Po}_{cq}|_{\geq 2}(k) \cdot \mathbb{P}\left[\mathrm{Bi}(k, x) \geq 2\right].$$

It has a fixed point at $x = 1$ and convergence of the $q_r$ will be like

$$1 - O(\varepsilon t^r) \leq q_r \leq 1,$$

for some $0 < t < 1$. We implicitly have assumed that $\varepsilon$ is $o(1)$, or small and constant, in the worst case.

b) Or, more strongly, we may 'disown' each owner with probability $\varepsilon$, independently in each level. The recursive equation for binary ownership is then given by

$$f(c, x) = \sum_{k=2}^{\infty} \mathrm{Po}_{cq}|_{\geq 2}(k) \cdot \mathbb{P}\left[\mathrm{Bi}(k, x(1 - \varepsilon)) \geq 2\right].$$

For $\varepsilon = 0$ the recursive equation of a) is reproduced. Increasing $\varepsilon$ has the effect that the fixed point shifts from 1 to $1 - \delta\varepsilon$, the dependence on $\varepsilon$ is (locally) linear due to analyticity, this also holds for $t$. Thus for sufficiently small $\varepsilon = o(1)$ the $q_r$ will converge with the 'usual' rate of convergence to some value converging to 1.

# Chapter 4

# The Magic Subgraph

We were inspired by the Branching Tree Connection in searching for some subgraph 'corresponding' to the coloured owners (Section 3.2) in an analogous way as the Cayley owners 'correspond' to the $k$-core. We actually found a 'corresponding' subgraph. Similarly to the extended $k$-core it can be characterised by a *directed-edge* deletion process. The *magic subgraph* is thus well-defined, in any graph *with respect to* a given $k$-colouring, moreover it can be found in linear time, algorithmically.

We have collected strong empirical evidence (see Section 4.3) that in tripartite random graphs there 'suddenly appears a giant magic subgraph' at an average degree of $c_{BT} = 4.91\ldots$. The same seems to be true fore $k > 3$, the critical $c$-values being the ones from the Branching Tree Connection for coloured owners, Moreover, magic subgraphs found in random tripartite graphs are 'almost uniquely colourable', empirically. However, magic subgraphs are not alway uniquely colourable, deterministically.

Being 'almost uniquely colourable' would make the magic subgraph a candidate giant subgraph for explaining the jump in chromatic number, which is certainly one of the most central challenges in the field, see [Mol01]. Unfortunately, it is only defined with respect to a specific, 'built-in' $k$-partition. We shall further discuss this in Chapter 7. We have demonstrated above (see Section 2.6) that there are serious results and open questions related to random tripartite graphs. It may well be that the Antivoter Phenomenon turns out to be a 'critical slowing down' due* to the 'sudden appearance of the magic subgraph'. Also, magic subgraphs may be interesting in the context of finding uniquely colourable graphs with large girth (see Section 2.6.3).

*Note that we do not claim anything rigorous on a general mechanism of the 'critical slowing down', see our comments in Section 1.2.

## 4.1 Motivation and Definition

When looking for something similar to the $k$-core 'corresponding' to the coloured owners of Section 3.2, the naive idea that jumps to mind is modifying the node deletion process characterising the $k$-core by now deleting nodes with a monochromatic neighbourhood.

However, this does not seem to yield anything exciting. We observed no discontinuity in the size of the remaining subgraph at the critical value $c = 4.91\ldots$ ($k = 3$), as might be expected considering the Branching Tree Connection and the results on coloured owners from the last chapter. We can furthermore give an explicit example proving that the 'naive' node deletion process as described above is indeed 'weaker' than the deletion process leading to the magic subgraph. Consider a $C_4$ with an additional edge. This graph is uniquely 3-colourable, and every node has a non-monochromatic neighbourhood. Thus it is robust against the above 'naive' deletion process. Yet, as we will presently see, it is not robust against the deletion process leading to the magic subgraph. This example graph is just a special case of the (uniquely colourable) graphs defined in the proof of Lemma 4.2.6 below.

The breakthrough was devising a *directed-edge deletion process*, similar to the seemingly artifical one introduced for the extended $k$-core in Section 2.3.1.

Remember the intuition that a coloured owner $u$ is 'fixed', in the unrestricted branching tree. Let us assume that $u$ is red and has a blue child $v$ and a green child $w$, both of them being 'fixed'. Suppose we re-colour $u$ with, say, blue. Then the edge $\{u, v\}$ becomes monochromatic, which we do not like. If we decide to fix this by re-colouring $v$, recursively and 'away from $u$', we will spark off an infinite chain of re-colouring actions. By the (recursive) definition of coloured owners this is true *no matter what colours we try to choose as replacement colours.*

If, on the other hand $u$ is not a coloured owner, we *can* (cleverly) choose replacement colours such that we will need to perform only a finite number of re-colouring actions. At some stage in the re-colouring process we will meet a node $\hat{u}$ that has only fixed children of *one* of the available colours, other than the colour of $\hat{u}$. Note that $\hat{u}$ does *not* need to have a monochromatic neighbourhood, we only require that the 'fixed' part of its neighbourhood *except $\hat{u}$'s parent* be monochromatic.

It is therefore more natural to look at directed edges $(u, v)$. If $\Gamma^-(u) \setminus v$ contains ('fixed') nodes of each available colour, then we consider the edge $(u, v)$ to be a 'warning' not to recolour $v$ and to fix it by re-colouring $u$, and so on recursively 'away from $v$', in the way we have described above.

Therefore we make the following definition.

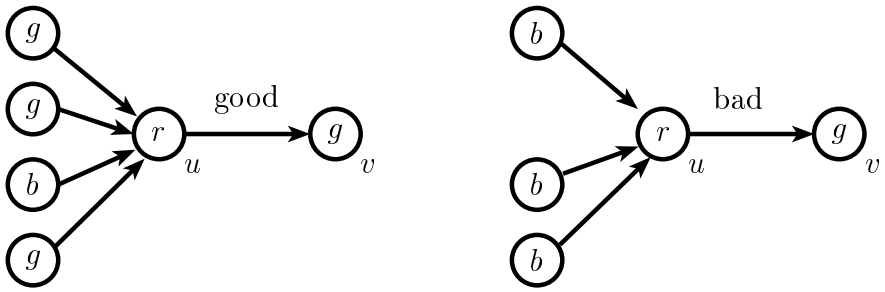


Figure 4.1: A directed edge $(u, v)$ is *bad*, if $\Gamma^-(u) \setminus v$ is monochromatic, and *good* otherwise. Note that there may be more directed edges, but $\Gamma^-(u) \setminus v$ is completely portrayed.

**Definition 4.1.1** *Let $D = (V, A)$ be some directed tripartite graph. A directed edge $(u, v) \in A$ is* bad*, if $\Gamma^-(u) \setminus v$ is monochromatic, and* good *otherwise. An edge is good or bad with respect to $A$. A subgraph $(V, A')$ is* good*, if it consists entirely of good edges with respect to $A'$.*

*The* magic subgraph *of $D = (V, A)$ is the union of all good directed subgraphs. Simple graphs are interpreted as directed graphs by replacing undirected edges by anti-parallel directed edges.*

Note that unions of good subgraphs are good themselves, obviously.

**Remark 4.1.2** *The above can be generalised to the case $k > 3$. In that case an edge $(u, v)$ is good if the in-neighbourhood contains nodes of all $k - 1$ available colours. We have chosen to state the definitions only for the case $k = 3$ in order to avoid complicating our notation.*

*Here and in what follows we will always assume that $V$ is partitioned into $k$ 'built in' colour classes, not necessarily balanced, even if not explicitly stated in our notation.*

## 4.2 Properties of the Magic Subgraph

The magic subgraph is in many ways similar to the $k$-core, see Section 2.3.1. Remember that we have shown that the $k$-core may also be characterised by a directed-edge deletion process. We have seen empirically, that the magic subgraphs found in random graphs are 'almost uniquely $k$-colourable'. We

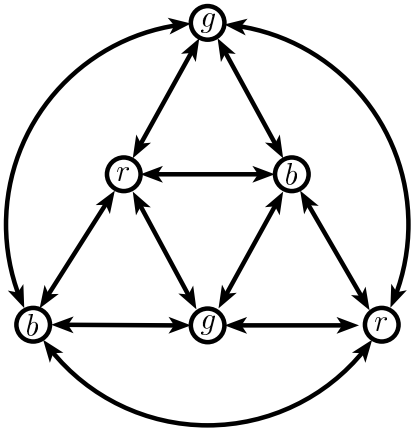


Figure 4.2: The tripartite *globe graph* is its own magic subgraph. It is uniquely 3-colourable, but note Figure 4.3.

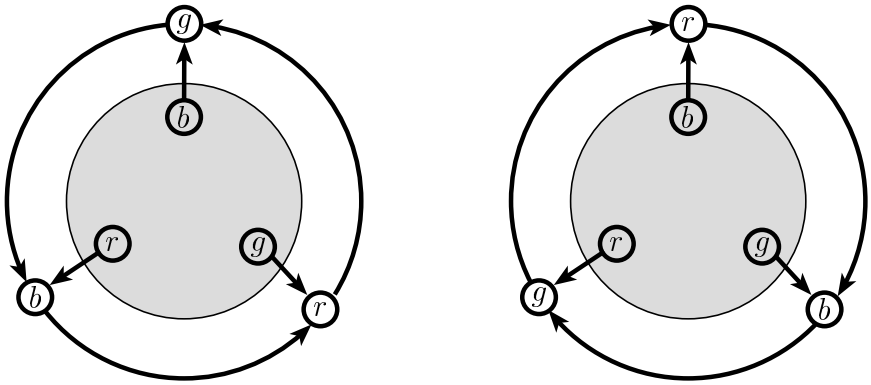


Figure 4.3: The colouring of a (simple) tripartite graph restricted to the magic subgraph is not unique, in general. The grey area stands for some appropriate graph, for example the graph from Figure 4.2.

therefore have collected some fairly simple statements illuminating some interconnections between uniquely $k$-colourable graphs, $k$-cores and magic subgraphs.

Again, the magic subgraph can be characterised by deletion processes.

**Definition 4.2.1** *An edge deletion process is some rule (or protocol) for iteratively removing bad edges from $D$ (with respect to the remainder graph).*

*At 'time' $t$, as long as there are any bad edges left in the remainder graph $D_t$, select such an edge $e_t$ according to the protocol and remove it from $D_t$, setting $D_{t+1} := D_t \setminus e_t$.*

**Proposition 4.2.2** *The magic subgraph can be algorithmically characterised as the outcome of any edge deletion process. In particular, the outcome of any edge deletion process is the magic subgraph, irrespective of the details of the specific deletion protocol.*

**Proof** Denote by $\mathcal{G}(D)$ the collection of all good subgraphs. Then $M := \bigcup_{D' \in \mathcal{G}(D)} D'$ is the magic subgraph. The empty subgraph is good. Denote by $M'$ the outcome of an arbitrary deletion process consistent with Definition 4.2.1.

$M \subseteq M'$: Let $D'$ be any good subgraph. It cannot be deleted by any deletion process, since otherwise some edge in $D'$ would have to be the first to be deleted. But since it is the first, it is still good with respect to the remainder graph, due to all other edges in $D'$ still being there.

$M' \subseteq M$: Conversely, it is obvious that the outcome of any deletion process consists entirely of good edges. □

**Definition 4.2.3** *The* magic subgraph in the node sense *of some simple graph $G$ is the (simple) subgraph induced in $G$ by the nodes that do not have monochromatic in-neighbourhoods in the magic subgraph $M$ of $G$.*

The smallest (as far as we know) tripartite example graph containing a non-trivial magic subgraph is the *globe graph* portrayed in Figure 4.2. It is uniquely 3-colourable, but this is not generally true for magic subgraphs, at least deterministically. Figure 4.3 shows a counterexample.

**Proposition 4.2.4** *Subgraphs of tripartite simple graphs induced by magic subgraphs are not necessarily uniquely 3-colourable.*

Yet, in Section 4.3 we shall present empirical evidence, that the colourings of random tripartite graphs, restricted to the respective magic subgraphs (in the node sense) appear to be 'almost unique'.

There is one assertion that we can prove deterministically for magic subgraphs.

**Lemma 4.2.5** *Let $D = (V, A)$ be the magic subgraph of some simple graph $G$. For each pair of colour classes the subgraph induced by bi-directed edges in $D$ on those two colour classes has minimum degree 2. The statement generalises to $k > 3$, 'minimum degree 2' having to be replaced by 'minimum degree $k - 1$'.*

**Proof** We start with $k = 3$. Assume the adjacent nodes $u$ and $v$ are coloured red and green, respectively and that there is a bi-directed edge between $u$ and $v$, i.e. $(u, v) \in A$ and $(v, u) \in A$. Then there must exist nodes $w_1, w_2, u'v'$ together with the arcs $(w_1, u), (w_2, v), (v', u), (u', v)$ as shown in Figure 4.4, otherwise the arcs connecting $u$ and $v$ would not be good. Therefore the arcs $(u, v')$ and $(v, u')$ (dotted arrows in Figure 4.4) must be present, too. For $(u, v')$ this is because it was there in the original graph $G$, and the presence of the arcs $(w_1, u)$ and $(v, u)$ proves that it cannot have been deleted by the deletion process. The same holds true for the arc $(v, u')$, by symmetry.

For $k > 3$ the proof is analogous. □

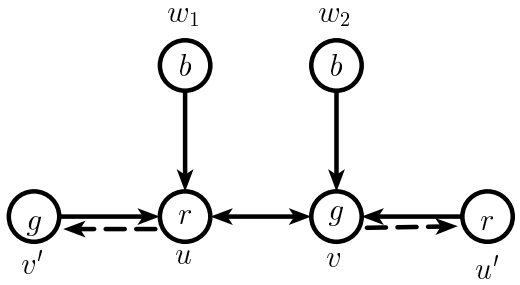


Figure 4.4: Illustration for the proof of Lemma 4.2.5.

Conversely, one may ask whether uniquely colourable graphs always contain magic subgraphs. The answer is no, in general.

**Lemma 4.2.6** *There are uniquely $k$-colourable graphs of arbitrary size having an empty magic subgraph.*

**Proof** (of Lemma 4.2.6) Start with an edge $\{u, v\}$, $u$ coloured red and $v$ coloured green. Recursively attach 'cherries' $w_t$ to the graph, where $w_t$ is attached to the graph $G_t$ so far with exactly two edges, connecting $w_t$ with nodes of two different colours. The colour of $w_t$ is uniquely determined by the colours of its two neighbours in $G_t$. Such graphs are uniquely colourable,

but their magic subgraphs are empty. The latter, if not obvious, is implied by Corollary 4.2.9 b) below. □

Note that if we replace the 'starting edge' in the proof of Lemma 4.2.6 by a non-empty magic subgraph (such as the globe graph from Figure 4.2), all the 'cherries' *will* be in the magic subgraph (in the node sense).

We may further ask whether uniquely colourable graphs that do contain a $k$-core need to also contain non-empty magic subgraphs (note that the counterexamples in the proof of Lemma 4.2.6 are *not* of that type). However the *pyramid* (join of circle $C_4$ with an isolated vertex) is a counterexample. This counterexample does *not* readily generalise to $k > 3$.

**Proposition 4.2.7** *There exist uniquely 3-colourable graphs with a non-empty 3-core but an empty magic subgraph.*

Let us now return to the directed-edge deletion process that characterises the extended $k$-core, as described in Section 2.3.1. The analogue of the magic subgraph in the node sense is the subgraph induced by all vertices with in-degree at least two, with respect to the directed graph surviving the deletion process. Remember that this was called the extended $k$-core in Definition 2.3.3.

**Lemma 4.2.8** *A non-empty extended $k$-core implies a non empty $k$-core in the classical sense.*

**Proof** (of Lemma 4.2.8) We will only discuss the case $k = 3$, generalisation to $k > 3$ is straightforward. Note that as soon as there is a node $v$ with in-degree at least 3 there will be a non-empty 3-core. Clearly, all in-edges of $v$ must be bi-directed edges, bi-directed edges 'never end' (Lemma 4.2.5), and thus at least the vertices contained in the 1-in-neighbourhood of $v$ will be in the 3-core.

Thus if there is no 3-core, the in-degree of all nodes in the extended 3-core is exactly two. We will presently show that if there is no 3-core, the extended 3-core is acyclic. This yields a contradiction, since there exists no (non-empty) acyclic directed graph with in-degree exactly two.

Assume that there is a cycle in the extended 3-core and consider any node $v$ on the cycle and the (directed) edges $(u, v)$, $(v, w)$ on that cycle. Since $(v, w)$ has not been removed by the (3-core, directed-edge) deletion process there must exist an edge $(u', v)$. Now consider the time at which the deletion process has deleted the first edge in the opposite cycle, say $(v, u)$. At that time the edges $(u', v)$ and $(w, v)$ are still present, a contradiction. In other words, when there are cycles there are also double edges, implying a classical 3-core which we have assumed to be absent. □

**Corollary 4.2.9** *a) The magic subgraph is contained in the extended $k$-core.*
*b) Any magic subgraph contains a non-empty 'classical' $k$-core.*

**Proof** ad a) We choose a specific protocol for the magic-subgraph- deletion process, remember that any will do according to Proposition 4.2.2. First delete all edges that are bad due to having in-degree less than $k-1$, recursively. At that stage we have found the extended $k$-core. The magic subgraph deletion process may now carry on deleting further edges that are bad in the 'magic subgraph sense'.

ad b) Apparently what is left by the magic subgraph directed-edge deletion process consists entirely of directed edges that are also good with respect to the $k$-core node deletion process. It forms an extended $k$-core. Because of Lemma 4.2.8 it also contains a $k$-core in the 'classical' sense. □

## 4.3 Empirical Results

In this section we present some experimental results. The 'sudden appearance of a magic subgraph' at $c = 4.91\ldots$ in tripartite random graphs is presumably the most interesting novelty. In Section 4.3.1 we shall show how the sizes of magic subgraphs (in the node sense) in random tripartite graphs depend on $c$ and $n$ - in full accordance with what would be expected heuristically from the Branching Tree Connection.

It is not hard to see that magic subgraphs are not uniquely 3-colourable, deterministically, c.f. Section 4.2. Still magic subgraphs of random tripartite graphs do seem to be 'almost' uniquely 3-colourable. In Section 4.3.2 we will present striking empirical evidence for this observation. But counterexamples do occur. We have included 'screen-shots' of a randomly-generated magic subgraph that behaves like the counterexample portrayed in Figure 4.3, leading to Proposition 4.2.4.

Finally, in Section 4.3.3, we shall state and discuss various observations.

### 4.3.1 Sudden Appearance of a Giant Magic Subgraph

Figure 4.5 shows the relative sizes of 'real life' magic subgraphs in the node sense at various average degrees for $n = 3000, 30000$ and $300000$ nodes, respectively. They are the outcome of the directed-edge deletion process applied to random tripartite graphs with average degree $c$. Remember that the magic subgraph in the node sense is the set of nodes having a non-monochromatic in-neighbourhood, w.r.t. the magic subgraph. Intuitively, it

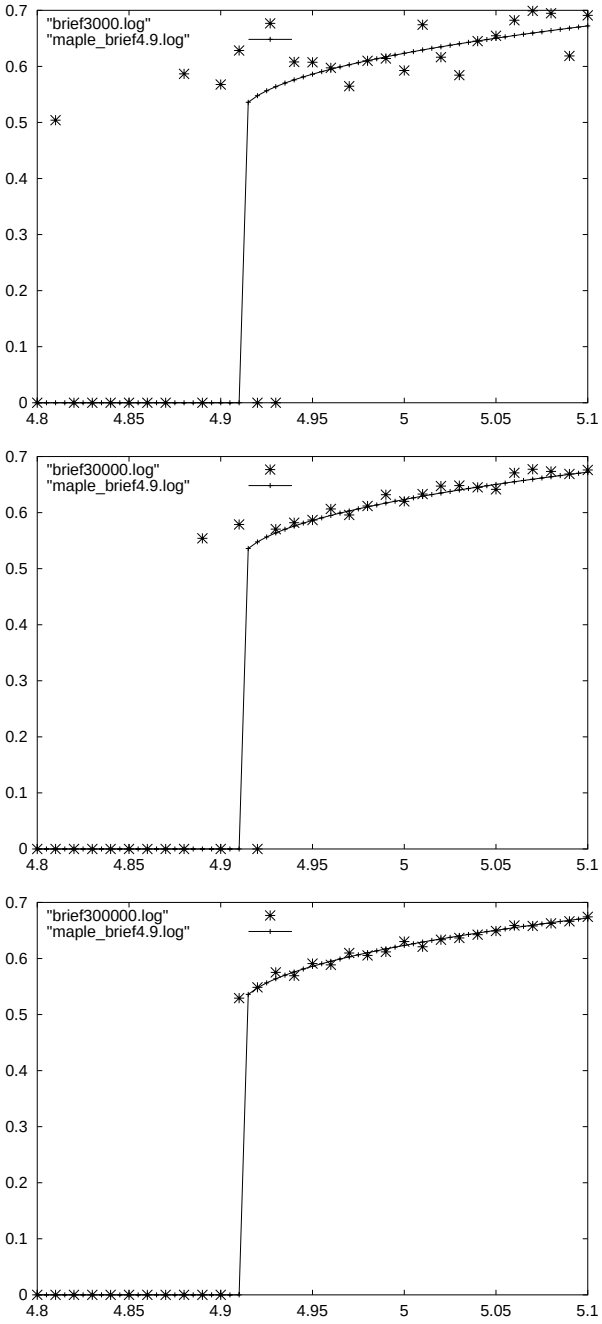


Figure 4.5: Relative sizes of empirically observed magic subgraphs for $c \in [4.8, 5.1]$, $n = 3000, 30000, 300000$. We have also plotted the value predicted by the Branching Tree Connection, see Figure 3.5, p. 54.

'corresponds' to the set of coloured binary-owners. We have also plotted the value $q(c)$, see Section 3.2, which is the expected (relative) size according to the Branching Tree Connection.

The reader may be interested in the following details concerning our implementation.

- Random tripartite graphs with average degree $c$ are not generated according to the $\mathcal{G}_{n,3,p}$ model but to the following $\mathcal{G}_{n,3,m}$ model, using a procedure similar to the one which generates $\mathcal{G}_{n,m}$ graphs (c.f. Section 2.1.1, p. 12):

  Partition the set of nodes into three 'colour classes' of equal size. As long as there are less than $m = \frac{c \cdot n}{2}$ edges repeatedly choose a pair $(u, v)$ of nodes in different 'colour-classes' uniformly at random. When the edge $\{u, v\}$ is not yet present in the graph generated so far incorporate it, otherwise ignore $(u, v)$ and retry.

  This procedure will generate a random graph in (an expected number of) $O(n)$ steps as opposed to the corresponding $\mathcal{G}_{n,3,p}$ procedure that has to consider each of the $\theta(n^2)$ 'potential' edges. Finding 'translation tools' between $\mathcal{G}_{n,3,p}$ and $\mathcal{G}_{n,3,m}$ generalising the tools explained in Section 2.1.1, p. 13, should be a straightforward task.

- The deletion process finding the magic subgraph is carried out exactly as described in Definition 4.2.1, p. 70.

- We have generated only one magic subgraph for each value of $n$ and $c$. We could instead have generated several magic subgraphs for each $(c, n)$ in order to plot error bars (based on a variance estimator). We have not done so because it seems plausible that fluctuations will not change much for closely adjacent $c$-values. Looking at the 'fluctuations' in the plots therefore immediately gives an impression of the variance.

### 4.3.2 Magic Subgraph 'Almost Uniquely Colourable'?

Before describing our experiments and presenting the results we shall make a few remarks on motivation.

#### Qualitative Considerations

Remember the intuition behind the definition of the magic subgraph. In the branching tree world, the (root) node is 'fixed' if recolouring to a different colour will spark off an 'infinite chain of repair steps' when trying to 'repair

the colouring away from the root'. As a rule of thumb infinite subtrees translate to giant ($\theta(n)$ nodes) subgraphs in the random graph world. Conversely, consider a uniquely colourable (giant) subgraph. Recolour one selected node from red to green, say. Then we need to change the whole formerly green colour class into red, when trying to 'repair the colouring away from the selected node'. In the light of these admittedly very hand-waving considerations it seems plausible to conjecture that the magic subgraph of a random tripartite graph might be uniquely colourable.

If this were true computing magic subgraphs of random tripartite graphs might lead to a new method for generating uniquely colourable graphs of large girth (c.f. Section 2.6.3). More importantly, adding a single edge between nodes of the same colour in uniquely 3-colourable graphs increases the chromatic number by one. This will happen a.s. after adding $o(n)$ random edges *no longer respecting the colour classes*. Thus the appearance of a giant uniquely colourable subgraph could provide an appealing explanation for the jump in chromatic number in the non-tripartite case. However, we do not know how to translate results for tripartite random graphs to results concerning non-tripartite random graphs.

### How to Generate Random Colourings

We can generate random colourings using the Antivoter algorithm described in Section 2.6.1. It is highly plausible to assume that a proper colouring returned by the Antivoter is distributed (very close) to uniformly at random amongst all proper colourings, since the invariant distribution of the Antivoter is the (Boltzmann-)Gibbs-distribution on the set of all colourings of a graph (see 1.2 and 2.6.1). Even if the mixing time of the Antivoter was much longer than the first hitting time of a proper colouring, the fact that it starts from a completely random i.i.d. colouring should prevent the Antivoter from introducing any 'external' bias towards colourings that are somehow correlated with the built in colouring. In other words, when the colourings produced by the Antivoter are somehow correlated with the 'built in' colouring, this should only be due to the internal structure of the graph and not to the sampling algorithm.

### Algorithm and Results

We have generated random tripartite graphs as described above for $n = 3000, 9000$ and $c = 4.30, 4.35, \ldots, 5.45$. For each graph we computed the magic subgraph and generated two different, independent (*bona fide*) random colourings using the Antivoter algorithm. Each time we counted the overlap

| $c$ | $T$ | $q$ | $r$ | $Rr$ | $Gr$ | $Br$ | $r\|_{MS}$ | $Rr\|_{MS}$ | $Gr\|_{MS}$ | $Br\|_{MS}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 4.30 | 160 | 0 | 33 | 31 | 34 | 35 | - | - | - | - |
| 4.35 | 730 | 0 | 33 | 38 | 27 | 35 | - | - | - | - |
| 4.40 | 240 | 0 | 33 | 33 | 38 | 29 | - | - | - | - |
| 4.45 | 920 | 0 | 34 | 32 | 24 | 44 | - | - | - | - |
| 4.50 | * | 0 | * | * | * | * | * | * | * | * |
| 4.55 | 310 | 0 | 34 | 14 | 70 | 16 | - | - | - | - |
| 4.60 | 460 | 0 | 34 | 77 | 14 | 9 | - | - | - | - |
| 4.65 | 82 | 0 | 33 | 8 | 81 | 10 | - | - | - | - |
| 4.70 | 760 | 0 | 34 | 78 | 10 | 1 | - | - | - | - |
| 4.75 | 1100 | 0 | 33 | 8 | 85 | 7 | - | - | - | - |
| 4.80 | 120 | 0 | 34 | 8 | 10 | 83 | - | - | - | - |
| 4.85 | 400 | 0 | 33 | 9 | 6 | 84 | - | - | - | - |
| 4.90 | 430 | 0 | 33 | 9 | 7 | 84 | - | - | - | - |
| 4.95 | 160 | 62 | 34 | 87 | 7 | 7 | 21 | 100 | 0 | 0 |
| 5.00 | 40 | 64 | 33 | 7 | 87 | 6 | 21 | 0 | 100 | 0 |
| $5.00^+$ | 92 | 60 | 34 | 7 | 5 | 87 | 20 | $\simeq 0$ | $\simeq 0$ | 99.003 |
| 5.05 | 46 | 63 | 34 | 5 | 88 | 7 | 21 | 0 | 100 | 0 |
| 5.10 | 73 | 67 | 34 | 89 | 4 | 7 | 23 | 100 | 0 | 0 |
| 5.15 | 73 | 69 | 33 | 6 | 88 | 6 | 22 | 0 | 100 | 0 |
| 5.20 | 210 | 65 | 33 | 5 | 88 | 6 | 21 | 0 | 100 | 0 |
| 5.25 | 16 | 74 | 33 | 90 | 5 | 5 | 25 | 100 | 0 | 0 |
| 5.30 | 18 | 72 | 34 | 6 | 6 | 88 | 24 | 0 | 0 | 100 |
| 5.35 | 36 | 75 | 34 | 4 | 91 | 5 | 25 | 0 | 100 | 0 |
| 5.40 | 40 | 77 | 34 | 93 | 3 | 3 | 26 | 100 | 0 | 0 |
| 5.45 | 17 | 75 | 33 | 5 | 4 | 91 | 25 | 0 | 0 | 100 |

'*': Antivoter stopped unsuccessfully. More than $9 \times 10^7$ iterations.
'-': Magic subgraph empty.
'+': Repeated 9 times until we saw deviations of the colourings restricted to the magic subgraph.
All columns in % except $c$ (average degree, no units) and $T$ (hitting time of Antivoter, in units of $n \ln n$).

Table 4.1: The magic subgraph appears to be almost uniquely 3-colourable, $n = 3000$. In this experiment the colourings restricted to the magic subgraph were always identical up to permutation of the colour classes. It took 9 repetitions of the experiment for $c = 5.0$ to actually see a (slight) deviation, reported in the extra line of the table. Further explanations see Table 4.3.

| $c$ | $T$ | $q$ | $r$ | $Rr$ | $Gr$ | $Br$ | $r\|_{MS}$ | $Rr\|_{MS}$ | $Gr\|_{MS}$ | $Br\|_{MS}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 4.30 | 50 | 0 | 34 | 32 | 33 | 35 | - | - | - | - |
| 4.35 | 120 | 0 | 34 | 35 | 31 | 33 | - | - | - | - |
| 4.40 | * | 0 | * | * | * | * | * | * | * | * |
| 4.45 | * | 0 | * | * | * | * | * | * | * | * |
| 4.50 | * | 0 | * | * | * | * | * | * | * | * |
| 4.55 | * | 0 | * | * | * | * | * | * | * | * |
| 4.60 | * | 0 | * | * | * | * | * | * | * | * |
| 4.65 | * | 0 | * | * | * | * | * | * | * | * |
| 4.70 | * | 0 | * | * | * | * | * | * | * | * |
| 4.75 | 210 | 0 | 34 | 9 | 8 | 83 | - | - | - | - |
| 4.80 | 370 | 0 | 33 | 9 | 8 | 83 | - | - | - | - |
| 4.85 | 890 | 0 | 33 | 85 | 7 | 7 | - | - | - | - |
| 4.90 | 400 | 0 | 33 | 7 | 7 | 85 | - | - | - | - |
| 4.95 | 520 | 61 | 33 | 7 | 87 | 6 | 21 | $\simeq 0$ | 99.62 | $\simeq 0$ |
| 5.00 | 240 | 63 | 34 | 6 | 89 | 5 | 21 | 0 | 100 | 0 |
| 5.05 | 260 | 64 | 33 | 89 | 5 | 6 | 21 | 100 | 0 | 0 |
| 5.10 | 63 | 64 | 33 | 7 | 87 | 6 | 21 | 0 | 100 | 0 |
| 5.15 | 110 | 71 | 33 | 5 | 91 | 4 | 24 | $\simeq 0$ | 99.95 | $\simeq 0$ |
| 5.20 | 160 | 71 | 33 | 90 | 5 | 6 | 23 | 100 | 0 | 0 |
| 5.25 | 51 | 73 | 33 | 6 | 89 | 5 | 24 | c | c | c |
| 5.30 | 63 | 74 | 33 | 5 | 90 | 5 | 25 | $\simeq 0$ | 99.72 | $\simeq 0$ |
| 5.35 | 57 | 75 | 34 | 4 | 4 | 92 | 25 | 0 | 0 | 100 |
| 5.40 | 40 | 77 | 33 | 5 | 4 | 91 | 26 | 0 | 0 | 100 |
| 5.45 | 58 | 76 | 33 | 91 | 5 | 4 | 25 | 100 | 0 | 0 |

'*': Antivoter stopped unsuccessfully. More than $9 \times 10^7$ iterations.
'-': Magic subgraph empty.

All columns in % except $c$ (average degree, no units) and $T$ (hitting time of Antivoter, in units of $n \ln n$).

Table 4.2: The magic subgraph appears to be almost uniquely 3-colourable, $n = 9000$. The colourings restricted to the magic subgraph are frequently identical up to permutations of the colour classes. If not, the deviations are restricted to a tiny fraction of the nodes. Further explanations see Table 4.3.

| | |
|---|---|
| $c$ | Average degree, no units. |
| $T$ | Hitting time of the (first) Antivoter, in multiples of $n \ln n$. |
| $q$ | Relative size of the magic subgraph, in %, rounded. |
| $r$ | Relative size of the (first) red colour class, in %, rounded. |
| $Rr$ | Size of the overlap of the new colour class $R$ with the (first) red class $r$ relative to the size of $r$, in %, rounded. $Gr, Br$ analogous. |
| $r\|_{MS}$ | Relative size of the (first) red colour class restricted to the magic subgraph, in %, rounded. |
| $Rr\|_{MS}]$ | Size of the overlap of the new colour class $R$ with class $r\|_{MS}$ relative to the size of $r\|_{MS}$, in %, rounded. Yet whenever we write 100% in this column, *we mean it.* $Gr\|_{MS}, Br\|_{MS}$ analogous. |

Table 4.3: Explanation for the columns in Tables 4.1 and 4.2.

of the 'new' colour classes $\{R, G, B\}$ of the second colouring with the colour class $r$ of the second colouring, as well as those quantities restricted to the respective magic subgraph. The data presented in Tables 4.1 and 4.2 (p. 78 and p. 79, respectively) clearly show that those two colourings either coincide completely or occasionally differ in only a tiny fraction of the nodes, up to globally relabelling the colour classes. They are rather self-explanatory, Table 4.3 contains information on the columns in Tables 4.1 and 4.2.

We shall discuss observations concerning those counterexamples in more detail below.

### Structure of Counterexamples

We have randomly generated counterexamples, i.e. magic subgraphs that are assigned non-equivalent colourings in independent runs of the Antivoter, and inspected them in detail in an attempt to guess the structure implied by the fact that there are more than one colourings of the magic subgraph. However, we did not get very far. All we can say is that two colourings of a given magic subgraph seem to differ by 'shifts' along odd directed cycles, and on some other nodes that are 'somehow influenced' by those cycles, not unlike the small uni-cyclic connected components before the advent of the giant component. However striking the empirical evidence, there is a lot to be done on the theoretical side.

The situation reminds us of Section 7 in [Mol01] where a giant 'bi-linked collection' $B$, somewhat similar to the magic subgraph, is conjectured to contain a uniquely colourable subgraph $B'$ with $|B'| = |B| - o(n)$.

In Figures 4.6, 4.7, 4.8 (pages 81, 82, 83, respectively) we show various layouts of a specific magic subgraph found in a graph on $n = 60$ nodes which has more than one colouring. This specific example resembles the counterexample in Figure 4.3.

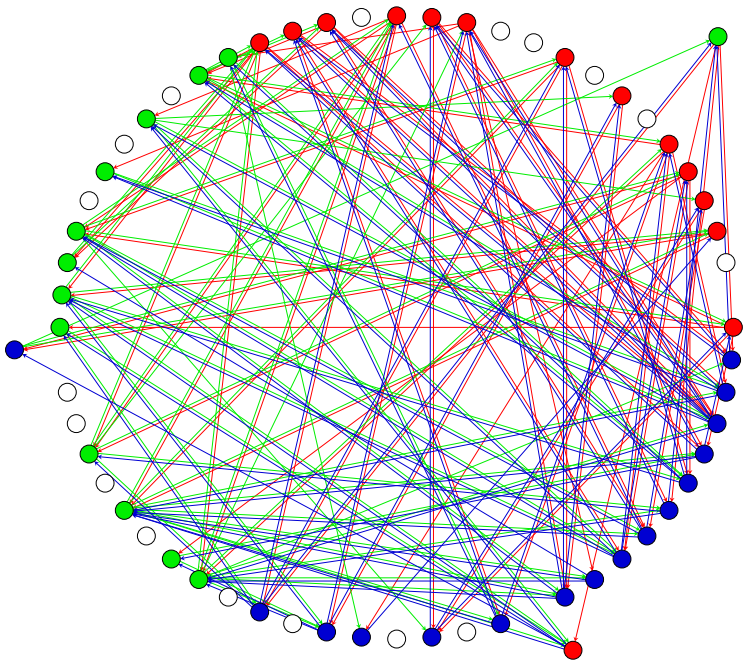

Figure 4.6: Example magic subgraph (see text), Layout I. Note that the 'built-in' in colouring is reproduced on all but 3 nodes.

### 4.3.3 Miscellaneous Observations

We shall state various observations that we have made performing our simulations.

- Apparently the hitting time of the Antivoter depends on the 'temperature' $T$ (for fixed $(n, c)$ in a nice, 'concave' way), whereas Petford and

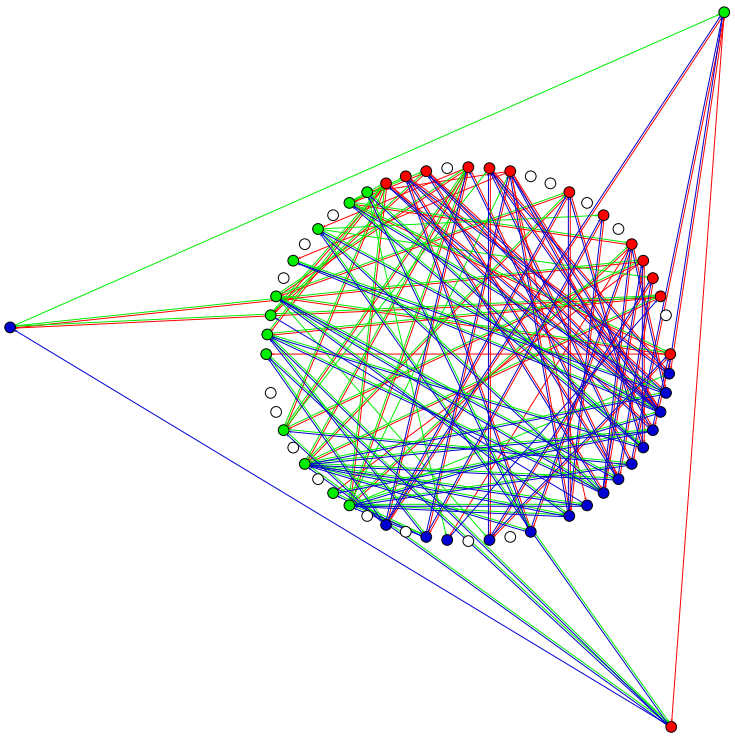

Figure 4.7: Example magic subgraph (see text), Layout II. The 'wrongly-coloured' node lie on a directed triangle, like in the counterexample given in Figure 4.3.

Welsh in [PW89] merely reported that the precise value of 'temperature' does not affect the 'critical slowing down', qualitatively. We have not checked whether there is a 'universal' optimal choice $T$ or whether the optimal $T$ should depend on $c$, possibly even on $n$, but have merely hand-tuned this parameter to obtain reasonable running times.

- We have not thoroughly studied the case $k > 3$. Some simulations made for $k = 4$ suggest that the situation looks very similar to the case $k = 3$, apart from different numerical values, obviously.

- When looking at the data in Tables 4.1 and 4.2 for $c$-values *just before* the appearance of the magic subgraph two observations jump to mind. Firstly, the worst case running times of the Antivoter appear to occur in this regime. Secondly, the colourings seem to be 'correlated', even

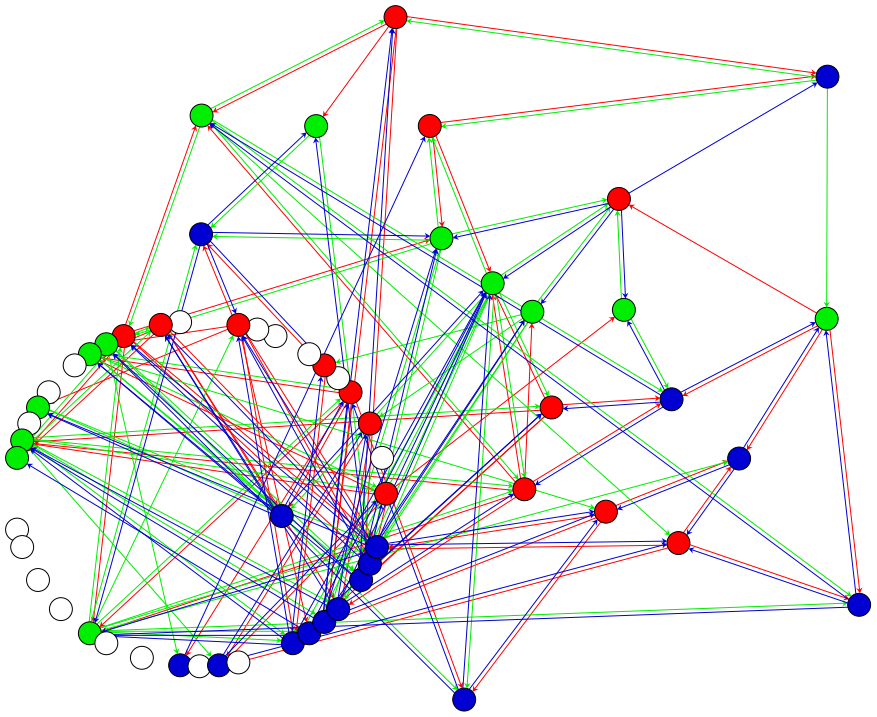

Figure 4.8: Example magic subgraph (see text), Layout III. The neighbourhood of the 'wrongly-coloured triangle' (top right corner).

though there is no magic subgraph present yet. We shall present some ideas concerning this below.

- We have not included plots for the appearance of the extended 3-core nor of the 'classical' 3-core, as they look exactly the same as the plots in Figure 4.5, apart from the fact that numerical values are different, obviously. Note that we can fully analyse the appearance of the 'classical' 3-core but - at the time being - not the appearance of the closely related extended 3-core nor of the magic subgraph. The fact that in all cases the empirical sizes, even for finite, moderately large values of $n$, so closely follow the Branching Tree Connection suggests that generalising the proof for the 'classical' 3-core should be possible.

- Unsurprisingly in the light of the Branching Tree Connection, i.e. from a 'local' perspective, both the appearance of the 3-core appears at a

critical average degree $c = 3.35(\ldots)$ in the tripartite model, too.

Moreover, we have implemented the list-colouring algorithm of Achlioptas and Molloy (see Section 2.5.1). Both in the tripartite and in the non-tripartite model it seems to fail only when $c \geq 3.85(\ldots)$ and the performance is completely comparable for both models.

- We have also investigated, and observed, the appearance of a giant magic subgraph with respect to a random i.i.d.-colouring, this scenario is discussed separately in Section 7.4.1.

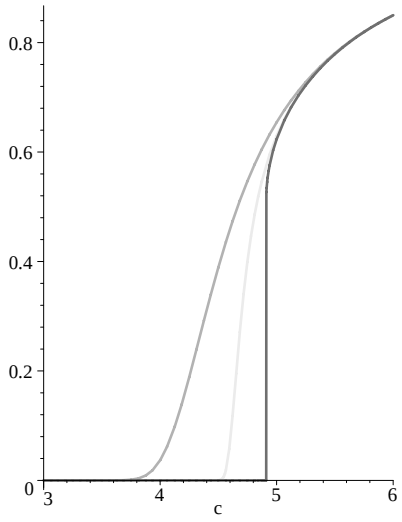


Figure 4.9: Plots of $c \mapsto q_8(c)$ and $c \mapsto q_{16}(c)$ as well as $c \mapsto q_\infty(c)$.

We conclude with an informal discussion of the aforementioned 'correlations' between independent colourings for slightly sub-critical values of $c$ and the fact that the Antivoter appears to diverge *just before* a magic subgraph has appeared.

Idea I: One may conjecture that another (possibly uniquely colourable) subgraph appears *before* the magic subgraph. 'Correlations' between colourings may appear somehow 'blurred' when we do not restrict ourselves to the new subgraph but to the total graph, just as can be observed in the supercritical case (columns 5, 6 and 7 in Tables 4.1 and 4.2). The reader might also like to consult our observations laid out in Section 7.4.1 on p. 148.

Idea II: In Figure 4.9 we have plotted $c \mapsto q_8(c)$ and $c \mapsto q_{16}(c)$ compared to $c \mapsto q_\infty(c)$. Even though $q_r(c)$ converges to $q_\infty(c)$ (w.r.t. $r$) this convergence is by no means uniform. The plots show that there will be many coloured $r$-owners, for moderately small constant values of $r$, even for $c$-values way below $4.91(\ldots)$. Now the intuition for being an $r$-owner is that not even a 'clever' backtracking algorithm *restricted to the interior of the $r$-neighbourhood $v$* may find a colouring such that the root $v$ has a different colour than before. This will even more apply to the Antivoter which is essentially 'greedy with a random backtracking facility', intuitively. Thus it may already be hard for the Antivoter to colour the set of $r$-owners (for moderately small constant values of $r$), which has size $\theta(n)$ even for sub-critical values of $c$. Note that the diameter of 'our' random graphs is $\simeq \log n$ and it seems hard to discern 'global' owners from, say, 10-owners for moderately sized $n$, anyway. In the light of what we have just said it is even more surprising that the relative sizes of magic subgraphs found by the edge-deletion process follow $q_\infty(c)$ so closely, as illustrated above.

# Chapter 5

# Relating Branching Trees and Random Graphs

Both the empirical results and the fact that sparse random graphs locally resemble branching trees suggest that there should be a rigorous proof employing the ideas behind the Branching Tree Connection, both for the sudden appearance of the $k$-core and the magic subgraph. As described in Section 2.4.3 Goerdt and Molloy were the first to achieve this goal by proving the appearance of the $k$-core in random faulty configurations. One essential ingredient of their proof is predicting the degree sequence of the subgraph induced by the set we called $r$-owners with reasonable accuracy.

This part of Goerdt and Molloy's proof does not readily extend to $\mathcal{G}_{n,p}$ and $\mathcal{G}_{n,m}$, mainly because both the sizes of BFS-balls and their mutual overlaps are not *deterministically* bounded. Sufficient concentration of the degree sequence of the set of $r$-owners follows from an application of Chebychev's inequality only in *the degree bounded* model of faulty configurations. In $\mathcal{G}_{n,p}$, when considering sums of indicator variables (indexed by nodes) depending on the respective the $r$-neighbourhoods, none of the standard concentration tools seems to directly apply. We have developed a concentration tool for *semi-local* functions of random graphs, implying a corresponding concentration result in the $\mathcal{G}_{n,p}$ models. A second but less critical point is that $r$-neighbourhoods in random faulty configurations *are* $\mathrm{Bi}(d,p,\cdot)$ branching trees, conditional on the $r$-neighbourhood in the underlying $d$-regular graph being a tree. In $\mathcal{G}_{n,c/n}$ the $r$-neighbourhoods are merely *similar to* $\mathrm{Po}_c$-branching trees but corrections have to be taken into account for finite $n$.

Combining our results with the proof strategy of Goerdt and Molloy, together with our observation that uniform distribution conditional on the degree sequence remains invariant also for *graphs*, will finally yield a new proof for the appearance of the $k$-core in the $\mathcal{G}_{n,m}$ model, see Chapter 6.

## 5.1 Ownership and Concentration of Semi-Local Functions

In this section we will show that quantities like the number of $r(n)$-owners are sharply concentrated around their means, always considering the $\mathcal{G}_{n,c/n}$ model. For regular faulty configurations it is possible to show reasonable concentration using Chebychev's inequality, see Goerdt/Molloy [GM00], but as explained above this does not work in $\mathcal{G}_{n,c/n}$.

We first state some rather trivial facts concerning the $r$-neighbourhoods. When $r = r(n)$, we define $\lambda(n) := (\log n)^{2r(n)}$, a rather crude upper bound on the size of the $r(n)$-neighbourhoods.

**Proposition 5.1.1** *For sufficiently large $n$, the maximum degree of a graph $G \in \mathcal{G}_{n,c/n}$ is greater than $(\log n)^2$ with a probability which is subpolynomially small in $n$, say $\exp(-(\log n)^2)$. Thus, when $r = r(n)$ there exist $r$-neighbourhoods containing more than $\lambda(n) = (\log n)^{2r(n)}$ nodes, with subpolynomially probability $\exp(-(\log n)^2)$, only.*

**Proof** The degree of any node is a $\mathrm{Bi}(n, c/n)$ random variable which is greater than $(\log n)^2$ with probability at most

$$\exp(-1/2 \cdot \frac{\log(n)^4}{c(1 + 1/3\log(n)^2/c)}),$$

by Chernoff. Note that the negative exponent is $\sim 3/2 \cdot (\log n)^2$. Since there are only $n$ 'bad events' with all but subpolynomially small probability bounded by $\exp(-(\log n)^2)$ the maximum degree of the graph is $(\log n)^2$. □

**Remark 5.1.2** *When some function like $\exp(-(\log n)^2)$ is* subpolynomially small*, it goes to zero even when multiplied by a constant number of arbitrary (fixed degree) polynomials in $n$.*

*Instead of having used Chernoff Bounds in the proof of Proposition 5.1.1 we could have calculated the tail probability more explicitly, using Stirling's formula, yielding essentially the same result. However, re-inspecting the proof of the Chernoff bounds yields that any upper bound on the tail of the $\mathrm{Po}_c$-distribution is automatically an upper bound for the tail of the 'corresponding' $\mathrm{Bi}(n, c/n)$-distribution.*

**Corollary 5.1.3** *A node contains a cycle in its $r(n)$-neighbourhood with probability $O(\lambda(n)^2/n)$.*

**Proof** Because of Proposition 5.1.1 there are at most $\binom{\lambda(n)}{2} \leq \lambda(n)^2$ potential cross edges available. The probability that at least one cross edge is switched on is therefore less than $1-(1-c/n)^{\lambda(n)^2/2}$. For sufficiently large $n$ this is at most $1-\exp(-c\lambda(n)^2/n) \leq c\lambda(n)^2/n$. □

As explained on p. 14 we shall order the $\binom{n}{2}$ 'random bits' standing for the potential edges being switched on or off into $n$ 'groups' $X_1, \ldots, X_n$. The $X_i$ describe outcome of the random bits for all potential edges connecting node $i$ with the nodes $[i-1]$ discovered 'so far' in the course of vertex exposure. Each function of $G$ will depend on the variables $X_1, \ldots, X_n$, just as required in the conditions of the Simple Concentration Bound, see p. 18.

We consider functions of $G$ of the form $f(G) := \sum_{v \in V} Z_v^{(r)}$, where the $Z_v^{(r)}$ are 0/1 random variables depending only on the $r$-neighbourhood of the respective node $v$. Such an $f$ will be called *$r$-semi-local*.

**Definition 5.1.4** *Let $f = f(G)$ be a sum of indicator variables, labelled by the nodes $v$ of a graph $G$:*

$$f(G) := \sum_{v \in V} Z_v^{(r)}.$$

*If all the indicator-variables $Z_v^{(r)}$ can be decided upon knowledge of the $r$-neighbourhood of $v$ in $G$, we call $f$* $r$-semi-local.

Intuitively speaking the number $f(G)$ is the sum of 'nearly independent' 0/1 random variables and may therefore be expected to be sharply concentrated. The $Z_v^{(r)}$ are 'nearly independent', because for each $v$ all $w$'s outside the $2r$-neighbourhood around $v$ will have disjoint proofs for $Z_w^{(r)}$ being zero or one. Since with high probability all $r(n)$-neighbourhoods are small, that is $(\log n)^{\alpha(c)M(c)}$, each $Z_v^{(r)}$ is therefore 'independent' of all but $(\log n)^{\alpha(c)M(c)}$ of the $Z_w^{(r)}$'s. Yet this does not imply the existence of a suitable dependency graph, since we do not know for sure how many of the $Z_w^{(r)}$'s are independent of $Z_v^{(r)}$ *in advance*. Further note that the $Z_v^{(r)}$'s are slightly positively correlated if the events '$Z_v^{(r)} = 1$' are edge monotonous.

We aim at applying the Simple Concentration Bound (Fact 4). As mentioned above, in our case the $X_i$ will be the respective random bits for all potential edges connecting node $i$ to $[i-1]$, and $f(\mathbf{X})$ is simply $f(G)$, $G$ being encoded by the $X_i$.

The crucial point is that $f(X)$ would be concentrated, if $f$ satisfied a Lipschitz condition with reasonably small Lipschitz-constant. But this is true only for most graphs (w.r.t. the $\mathcal{G}_{n,c/n}$ model), ruling out a straightforward application of the Simple Concentration Bound.

As an example let (the semi-local function) $f$ be the number of all nodes with degree greater 0 and consider the (very unlikely) case that $G$ is a star, centred at node number $n$. Then 'flipping' the last coordinate $X_n$ will change the value of $f$ by $n$, in the worst case, and a Lipschitz constant of $n$ does not yield concentration. Admittedly, we know that with very high probability the maximum degree is $(\log n)^2$ and conditional on that event flipping any coordinate $X_i$ could alter the value of $f$ by at most that value. Note that conditioning on the 'typical' maximum degree would not help, since then the coordinates would no longer be independent, and the Simple Concentration Bound would not apply.

Yet we can prove the following '*Semi-Local Lemma.*'

**Lemma 5.1.5** (Semi-Local Lemma) *Suppose $f$ is $r(n)$-semi-local, for some $r(n) \leq \alpha(c) \log\log n$. Then the random variable $f(G)$ lies outside*

$$[\mathbb{E}[f(G)] - n^{1/2+\varepsilon}, \mathbb{E}[f(G)] + n^{1/2+\varepsilon}]$$

*with probability subpolynomially small in $n$, say* $\exp(-(\log n)^2)$.

**Proof**

We consider the random variables $Y_i$, the $X_i$ conditional on the event $\bar{B}_0$ that no vertex has no more than $(\log n)^2$ neighbours in the graph discovered so far, when 'glued' to the graph in the course of vertex exposure. Note that $\bar{B}_0$ is a *product event* and the $Y_i$ are still independent, since we can tell for each $X_i$ separately, whether more than $(\log n)^2$ bits are switched on. (If we had conditioned on some bound on the total vertex degree we would have lost independence of the $X_i$!)

In order to achieve a Lipschitz condition we define $\tilde{f}$ to be the number of owners in an appropriately *truncated* graph, i.e. $\tilde{f} = f \circ T$ . A given graph $G = (x_1, \ldots, x_n)$ gets truncated by the following procedure defining $T$.

For each node $v$ in $G$ with degree greater than $(\log n)^2$ consider all adjacent edges $\{v, w_1\}, \{v, w_2\}, \ldots$ in increasing order of the labels of the $w_i$. Mark all but the first $(\log n)^2$ edges for deletion (but do not delete them yet). The remainder graph $T(G)$ consists of the edges that were not marked for deletion and has maximum degree $(\log n)^2$ by construction.

What is the Lipschitz constant of $\tilde{f}$? We consider two graphs $G$ and $G'$ differing only in the $k$-th component. I.e. when partitioning the edges in $E_1, E_2, \ldots, E_n$ and $E'_1, E'_2, \ldots, E'_n$, only $E_k \neq E'_k$. By our condition $\bar{B}_0$ the symmetric difference between $E(G)$ and $E(G')$ will contain at most $2(\log n)^2$ edges. We will show that $T(G)$ and $T(G')$ differ only by $6(\log n)^2$ edges.

First of all introduce a sequence of intermediate graphs

$$G_0 = G, G_1, \ldots G_k, G_{k+1}, G_{k+k'} = G'$$

where $k, k' \leq (\log n)^2$, first switching off all edges in $E_k$ and then switching on all edges in $E'_k$ one after the other. Two consecutive graphs $G_i$ and $G_{i+1}$ will differ in exactly one edge $e_i = \{u, v\}$. For convenience we assume that $e \notin E(G_i)$ but $e \in E(G_{i+1})$, but the argument is obviously symmetric.

If two graphs $G_i$ and $G_{i+1}$ differ only by a single edge, by how many edges will $T(G_i)$ and $T(G_{i+1})$ differ in the worst case? After inserting edge $e$, node $u$ may have degree $\geq (\log n)^2 + 1$ (in $G_{i+1}$!) and either $e$ or some other edge adjacent to $u$ may be removed by $T$. The same holds true for the other endpoint $v$ of $e$. Thus in the worst case the symmetric difference is 3, that is the case when both $u$ and $v$ lose an edge other than $e$ upon truncation (in $G_{i+1}$!). This worst case is illustrated in Figure 5.1.

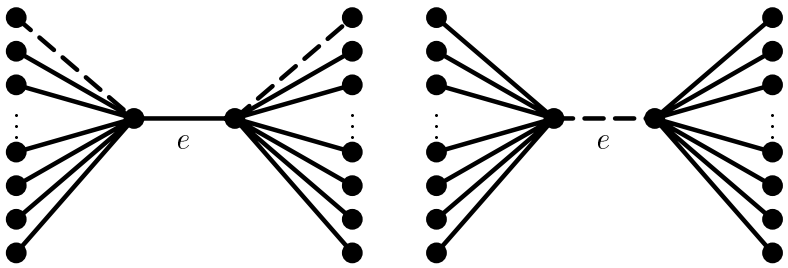


Figure 5.1: The worst case effect of presence or absence of edge $e$ on the truncated versions $T(G_i)$ and $T(G_{i+1})$ of two graphs $G_i$ and $G_{i+1}$ differing only in a single edge $e$. The right hand picture shows the $(\log n)^2$ neighbours with minimum label of the endpoints of $e$ in $G_i$. If $e$ is switched on, only the dotted edges (if at all present) may possibly now be marked for deletion by the truncation operation. Conversely, if $e$ is switched off, only (possibly) the dotted edges will no longer be marked for deletion. Thus the symmetric difference $|T(G_i) \oplus T(G_{i+1})|$ is at most 3.

We thus know that the symmetric difference between $T(G)$ and $T(G')$ is at most $6(\log n)^2$. Changing one of these edges $e = \{u, v\}$ may affect the status of ownership only for nodes lying in the $r$-neighbourhood of $e$ which is contained in the $(r+1)$-neighbourhood of any one of the respective endpoints. This is illustrated in Figure 5.2.

Therefore the Lipschitz constant is bounded,

$$\left| \tilde{f}(G) - \tilde{f}(G') \right| \leq 6(\log n)^2 (\log n)^{2(1+\alpha(c) \log\log n)}.$$

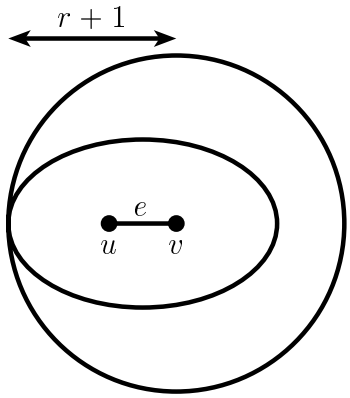


Figure 5.2: The $r$-fattening of edge $e$ is contained in the $(r+1)$-neighbourhood of node $v$. Flipping $e$ may affect $r$-ownership only for nodes in the $r$-fattening of $e$.

Putting things together we get the subpolynomially small tail bound

$$\begin{aligned}&\mathbb{P}\left[\left|\tilde{f}(\mathbf{X})-\mathbb{E}\left[\tilde{f}(\mathbf{X})\right]\right|\geq n^{1+\varepsilon/2}|\bar{B}_0\right]\\&\leq 2\exp(-\frac{2n^{1+\varepsilon}}{n(6(\log n)^2(\log n)^{2(1+\alpha(c)\log\log n)})^2})\end{aligned}\tag{5.1}$$

from the simple concentration bound for $\tilde{f}(X)$. Note that $\varepsilon$ has been replaced by $\varepsilon/2$ in this bound on $\tilde{f}(X)$ for reasons which will become obvious below.

It remains to show that this bound yields also a bound for $f(X)$, as stated in the assertion.

All probabilities above were conditional on $\bar{B}_0$. We want to use the good event $\bar{B}_1$ 'no vertex has total degree greater than $(\log n)^2$'. Note that $\bar{B}_1$ implies $\bar{B}_0$ and that $\mathbb{P}[B_1 \cap \bar{B}_0]$ is subpolynomially small by Proposition 5.1.1.

Let us briefly summarise the remainder of the proof in simple words. We have two functions $f$ and $\tilde{f}$ that coincide on the event $\bar{B}_1$ which occurs with very high probability. In the subpolynomially rare event $B_1$ they may differ, but they can differ by no more than $n$, in the worst case. Therefore it should be intuitively clear that $f$ is concentrated around the same value as $\tilde{f}$.

The distribution of $\tilde{f}(X)$ can be decomposed into

$$\mathcal{L}_{\tilde{f}(X)} = \mathbb{P}\left[\bar{B}_1\right]\mathcal{L}_{\tilde{f}(X)|\bar{B}_1} + \mathbb{P}\left[B_1\cap\bar{B}_0\right]\mathcal{L}_{\tilde{f}(X)|B_1\cap\bar{B}_0} = \mathbb{P}\left[\bar{B}_1\right]\mu + \mathbb{P}\left[B_1\cap\bar{B}_0\right]\tilde{\delta},$$

and the distribution of $f(X)$ into

$$\mathcal{L}_{f(X)} = \mathbb{P}\left[\bar{B}_1\right]\mathcal{L}_{f(X)|\bar{B}_1} + \mathbb{P}\left[B_1\cap\bar{B}_0\right]\mathcal{L}_{f(X)|B_1\cap\bar{B}_0} = \mathbb{P}\left[\bar{B}_1\right]\mu + \mathbb{P}\left[B_1\cap\bar{B}_0\right]\delta.$$

By construction of $\tilde{f}$ the first two terms already denoted by the same letter '$\mu$' are identical.

Therefore $\mathbb{E}[f(X)]$ and $\mathbb{E}[\tilde{f}(X)]$ differ by at most $2n\mathbb{P}[B_1 \cap \bar{B}_0]$, because

$$\left|\mathbb{E}\left[f(X) - \tilde{f}(X)\right]\right| = \mathbb{P}\left[B_1 \cap \bar{B}_0\right] \sum_{x=0}^{n} x \cdot \left|\delta(x) - \tilde{\delta}(x)\right| \leq 2n\mathbb{P}\left[B_1 \cap \bar{B}_0\right].$$

This is subpolynomially small by Proposition 5.1.1, and because $\mathbb{E}[f(X)]$ and $\mathbb{E}[\tilde{f}(X)]$ differ so little the desired event

$$[n] \setminus [\mathbb{E}\left[f(\mathbf{X})\right] - n^{1/2+\varepsilon}, \mathbb{E}\left[f(\mathbf{X})\right] + n^{1/2+\varepsilon}]$$

is contained in

$$[n] \setminus [\mathbb{E}\left[\tilde{f}(\mathbf{X})\right] - n^{1/2+\varepsilon/2}, \mathbb{E}\left[\tilde{f}(\mathbf{X})\right] + n^{1/2+\varepsilon/2}],$$

for $n$ large enough. Therefore bounding the probability of $f(X)$ hitting the second event will suffice. Note that

$$\sum_{x=0}^{n} \left|\delta(x) - \tilde{\delta}(x)\right| \leq 2\mathbb{P}\left[B_1 \cap \bar{B}_0\right]$$

implies a subpolynomially small difference $\exp(-(\log n)^2)$, because of Proposition 5.1.1) in variation norm* between the distributions of $f(X)$ and $\tilde{f}(X)$, and thus it is equivalent to bound the probability of $\tilde{f}(X)$ hitting the second event instead, up to a subpolynomially small error. The latter probability is bounded by Inequality (5.1) above. This completes the proof. □

## 5.2 Appearance of 'Almost the $k$-Core'

We start with a remark on notation.

**Remark 5.2.1** *Intuitively, the effect of small perturbations (in the inputs) on the outcome of analytic functions is small. We will write*

$$a = b \pm \delta,$$

*whenever $a$ is contained in the interval $[b - \delta, b + \delta]$. The notation*

$$f(a \pm \delta) = f(a) \pm \gamma$$

*See our Remark 5.3.4, p. 105.

*means that whenever $x$ is in $[a-\delta, a+\delta]$, $f(x)$ is in $[f(a)-\gamma, f(a)+\gamma]$. Note that whenever the derivative $f'$ is uniformly bounded by some constant $C$ we may choose $\gamma$ to be $C \cdot \delta$, by Taylor's Theorem.*

*In the applications below, both the values $a, b, \ldots$ and the perturbations $\gamma, \delta, \ldots$ will be functions of $n$.*

Remember Definition 2.3.7 (p. 31) of the various concepts of ownership, and that we may decide $r$-ownership upon knowing the $r$-neighbourhood. Thus the number of $r(n)$-owners is always $r(n)$-semi-local and, provided that $r(n) \leq \alpha(c) \log\log n$, the Semi-Local Lemma 5.1.5 applies.

Both in the graph and the branching tree $v$ is an $r$-binary owner, if at least two of its children are $(r-1)$-binary-owners in $G \setminus v$. When considering branching trees we used the fact that the respective subtrees induced by the children of a node $v$ were *independent* copies of the original branching trees. Moreover, using this independence we were able to calculate various probabilities such as of $v$ being an $r$-Cayley owner, being a run, and the conditional number of children of $v$ that were themselves $r$–binary owners.

Apart from the fact that Poisson progeny distributions should be replaced by binomial progeny distributions, which is more of a technical issue, the only reason that prevents us from computing the aforementioned probabilities by the same set of arguments in the random graph is lack of independence. However, the Semi-Local Lemma 5.1.5 allows us to calculate those probabilities in the random graph by essentially (up to negligible errors) the same calculations as we did in branching trees.

The following Lemma 5.2.2 shows why.

**Lemma 5.2.2** *Assume that $c$ is contained in some interval $C$ not containing $3.35\ldots$. Take a 'test particle' $v$ that is connected to each node of a 'remainder graph' $G$ on $n'(n) := n - \lambda_0(n)$ nodes, independently with probability $c/(n-1)$. Suppose that there is some 'marked' subset $V_0$ of the nodes of $G$ containing*

$$|V_0| \leq bn'(n) \pm \delta(n)$$

*nodes.*

*Here $0 \leq \lambda_0(n) \leq \Lambda_0(n) = o(n)$, and $0 \leq \delta(n) \leq \Delta(n) = \omega(1)$. We further assume that $\lambda_0(n) < \delta(n)$, and that $\Delta(n) = o(n)$.*

*Then for sufficiently large $n$ (depending on $\Lambda_0(n)$ and $\Delta(n)$),*

*a) $v$ is adjacent to at least two marked nodes with probability*

$$f_2(c,b) \pm T_2(C)\Delta(n)/n,$$

*b) $v$ is adjacent to at least three marked nodes with probability*

$$f_3(c,b) \pm T_3(C)\Delta(n)/n,$$

*c) conditional on $v$ being adjacent to at least two marked nodes, $v$ is adjacent to exactly two marked nodes with probability*

$$\mathrm{Po}_{cb}|_{\geq 2}(2) \pm T_0(C)\Delta(n)/n,$$

*where $T_0(C), T_1(C)$ and $T_2(C)$ are constants depending only on $C$.*

Note that the internal structure of $G$ is irrelevant. All randomness involved stems from the $n'(n)$ independent random bits connecting the 'test particle' to $G$.

Before we will prove this intuitively rather obvious lemma, we will quantify the intuition that it 'does not matter asymptotically' whether we use $\mathrm{Po}_{\ldots}$ or $\mathrm{Bi}(\ldots)$ in the 'sandwiching' Proposition 5.2.3 below.

**Proposition 5.2.3** *For all $\xi(n) = \omega(1/n)$ and $n$ sufficiently large*

$$\mathbb{P}\left[\mathrm{Po}_{c-\xi(n)} \geq 2\right] \leq \mathbb{P}\left[\mathrm{Bi}(n,c/n) \geq 2\right] \leq \mathbb{P}\left[\mathrm{Po}_{c+\xi(n)} \geq 2\right],$$

$$\mathbb{P}\left[\mathrm{Po}_{c-\xi(n)} \geq 3\right] \leq \mathbb{P}\left[\mathrm{Bi}(n,c/n) \geq 3\right] \leq \mathbb{P}\left[\mathrm{Po}_{c+\xi(n)} \geq 3\right].$$

$$\mathrm{Po}_{c+\xi(n)}|_{\geq 2}(2) \leq \mathrm{Bi}(n,c/n)|_{\geq 2}(2) \leq \mathrm{Po}_{c-\xi(n)}|_{\geq 2}(2).$$

*Furthermore $c$ may even depend slightly on $n$, as long as $c(n)$ converges to some $c_0$.*

**Proof** By definition

$$\begin{aligned}
\mathbb{P}\left[\mathrm{Bi}(n,c/n) \geq 2\right] &:= 1 - (1-c/n)^n - c(1-c/n)^{n-1} \\
\mathbb{P}\left[\mathrm{Po}_c \geq 2\right] &:= 1 - e^{-c} - ce^{-c} \\
\mathbb{P}\left[\mathrm{Bi}(n,c/n) \geq 3\right] &:= 1 - (1-c/n)^n - c(1-c/n)^{n-1} \\
&\quad - \frac{c^2}{2}(1-1/n)(1-c/n)^{n-2} \\
\mathbb{P}\left[\mathrm{Po}_c \geq 3\right] &:= 1 - e^{-c} - ce^{-c} - c^2/2e^{-c} \\
\mathrm{Po}_c|_{\geq 2}(2) &:= \frac{c^2/2e^{-c}}{1 - e^{-c} - ce^{-c}} \\
\mathrm{Bi}(n,c/n)|_{\geq 2}(2) &:= \frac{c^2/2(1-1/n)(1-c/n)^{n-2}}{1 - (1-c/n)^n - c(1-c/n)^{n-1}}.
\end{aligned}$$

In view of these formulas it suffices to show that

$$e^{-(c-\xi)} \stackrel{\text{a}}{\geq} (1-c/n)^{n-2} \geq \begin{cases} (1-c/n)^n \\ (1-c/n)^{n-1} \\ (1-1/n)(1-c/n)^{n-2} \end{cases}$$

and

$$e^{-(c+\xi)} \stackrel{\text{b}}{\leq} (1-1/n)(1-c/n)^n \leq \begin{cases} (1-c/n)^n \\ (1-c/n)^{n-1} \\ (1-1/n)(1-c/n)^{n-2} \end{cases}.$$

Inequalities a) and b) can be shown to be true by straightforward calculations using the well known estimates $\exp(-x) \geq 1-x$ and $\exp(-x-x^2) \leq 1-x$, the latter holding true for sufficiently small $x$. Note that in both cases below it is important that $\xi(n) = \omega(1/n)$.

Inequality a) follows from

$$\begin{aligned} e^{-(c-\xi)} &= (e^{-c/n})^n e^{\xi} \geq (1-c/n)^{n-2} e^{\xi} (1-c/n)^2 \\ &\geq (1-c/n)^{n-2} \underbrace{e^{\xi - 2c/n - 2c^2/n^2}}_{\to\infty, \text{ for } \xi(n)=\omega(1/n)} \geq (1-c/n)^{n-2}, \end{aligned}$$

for $n$ sufficiently large.

Inequality b) follows from

$$\begin{aligned} e^{-c(1+\xi)} &= (e^{-c/n-c^2/n^2-1/n^2-1/n^3+c^2/n^2+1/n^2+1/n^3-\xi/n})^n \\ &= e^{-1/n-1/n^2}(e^{-c/n-c^2/n^2})^n \underbrace{(e^{c^2/n^2+1/n^2+1/n^3-\xi/n})^n}_{\to 0, \text{ for } \xi(n)=\omega(1/n)} \\ &\leq (1-1/n)(1-c/n)^n, \end{aligned}$$

for $n$ sufficiently large.

Re-inspecting the proof yields that it does not matter if $c$ depends on $n$, as long as $c(n)$ converges. □

**Proof** (Of Lemma 5.2.2.) The following calculations hold for sufficiently large $n$. It would be difficult to explicitly specify just how large $n$ should be, but we will be content with the fact that it could be determined in principle. Note that the answer will depend on the behaviour of $\lambda_0(n)$ and $\delta(n)$, and it is clear that the worst case is determined by $\Lambda_0(n)$ and $\Delta(n)$.

Ad a): We use Proposition 5.2.3.

$$\begin{aligned}
&\mathbb{P}[\text{At least two marked children}]\\
\leq&\mathbb{P}\left[\mathrm{Bi}(n'(n), \frac{(bn'(n)+\delta(n))\cdot c/(n-1)}{n'(n)}) \geq 2\right]\\
\leq&\mathbb{P}\left[\mathrm{Po}_{(bn'(n)+\delta(n))\cdot c/(n-1)+\xi(n)} \geq 2\right]\\
\overset{\xi:=c\delta/(2n)}{\leq}&\mathbb{P}\left[\mathrm{Po}_{cb+2c\delta(n)/(n-1)} \geq 2\right]\\
=&f_2(c, b+2\delta(n)/(n-1))\\
\leq&f_2(c,b)+T_2(C)\delta(n)/n,
\end{aligned}$$

where $T_2(C)$ is some universal constant depending only on $c$. Up to constants it is the maximum of $\frac{\partial}{\partial x} f_2(c,x)$. The last inequality is a straightforward application of Taylor's Theorem.

By an essentially symmetrical argument we get the lower bound as follows.

$$\begin{aligned}
&\mathbb{P}[\text{At least two marked children}]\\
\geq&\mathbb{P}\left[\mathrm{Bi}(n'(n), \frac{(bn'(n)-\delta(n))\cdot c/(n-1)}{n'(n)}) \geq 2\right]\\
\geq&\mathbb{P}\left[\mathrm{Po}_{(bn'(n)-\delta(n))\cdot c/(n-1)-\xi(n)} \geq 2\right]\\
\overset{\xi:=c(\lambda_0+\delta)/n}{\geq}&\mathbb{P}\left[\mathrm{Po}_{cb-c(4\delta(n))/(n-1)} \geq 2\right]\\
=&f_2(c, b+(4\delta(n))/(n-1))\\
\geq&f_2(c,b)+T_2(C)\delta(n)/n,
\end{aligned}$$

remembering that $\lambda_0 < \delta$.

Ad b): Completely analogous to a).

Ad c): Very similar to a). We use again Proposition 5.2.3.

$$\begin{aligned}
&\mathbb{P}[\text{Exactly two} \mid \text{at least two marked children}]\\
\leq&\mathrm{Bi}(n'(n), \frac{(bn'(n)-\delta(n))\cdot c/(n-1)}{n'(n)})|_{\geq 2}(2)\\
\leq&\mathrm{Po}_{(bn'(n)-\delta(n))\cdot c/(n-1)-\xi(n)}|_{\geq 2}(2)\\
\overset{\xi:=c(\lambda_0+\delta)/n}{\leq}&\mathrm{Po}_{cb-c(\lambda_0(n)+\delta(n))/(n-1)}|_{\geq 2}(2)\\
=&\mathrm{Po}_{c(b+(\lambda_0(n)+\delta(n))/(n-1))}|_{\geq 2}(2)\\
\leq&\mathrm{Po}_{cb}|_{\geq 2}(2)+T_0(C)\delta(n)/n,
\end{aligned}$$

where $T_0(C)$ is again some universal constant depending only on $c$, up to constants it is the largest value of $\frac{\partial}{\partial x}\mathrm{Po}_{cx}(2)/f_2(c,x)$. Remember that $\lambda_0 < \delta$.

Again the lower bound follows from an essentially symmetric argument.

□

**Corollary 5.2.4** *(Of Proposition 3.1.1.) For any $r$ and any fixed $c_0 \neq 3.35\ldots$ the function $c \mapsto q_r(c)$ is analytic and strictly increasing in $c$, and its derivative is uniformly (in $r$) bounded by $\tilde{T}(C)$ in some interval $C$ containing $c_0$ but not $3.35\ldots$, depending only on the choice of the interval $C$.*

**Proof** By its definition as the $r$-fold application to 1 of the monotonous and entire analytic function $f_2(c,\cdot)$, $c \mapsto q_r(c)$ is monotonous and entire analytic.

We will show that $\frac{d}{dc}q_r(c)$ is uniformly (in $r$) bounded in some neighbourhood $C$ of $c_0$.

For the super-critical case we proceed as follows. Note that $\frac{d}{dc}q_0(c) \equiv 0$. Now, aiming at a recursively defined upper bound,

$$\begin{aligned}\frac{d}{dc}q_r(c) &= \frac{\partial}{\partial c}f_2(c,q_{r-1}(c)) + \frac{\partial}{\partial x}f_2(c,q_{r-1}(c))\frac{d}{dc}q_{r-1}(c)\\ &= q_{r-1}(c)(cq_{r-1}(c))\exp(-cq_{r-1}(c)) + \tilde{t}(r-1,c)\frac{d}{dc}q_{r-1}(c)\\ &\leq 1/e + \tilde{t}(r-1,c)\frac{d}{dc}q_{r-1}(c).\end{aligned}$$

We know from Proposition 3.1.1 that $\tilde{t}(r-1,c)$ is always less than $t(c) < 1$. Thus the above indeed establishes a recursive equation for an upper bound on $\frac{d}{dc}q_r(c)$ which is solved by

$$r \mapsto \exp(-1)(1-t(c)^r)/(1-t(c)).$$

This converges from below to some continuous function in $c$ and therefore $\frac{d}{dc}q_r(c)$ is uniformly bounded on some interval $C$ by some constant $\tilde{T}(C)$.

In the sub-critical case $q_r(c)$ converges to 0. After some number of steps $\Delta(C)$, depending only on the right boundary of the interval $C$, for all $c \in C$, $q_r(c)$ will have reached a value where $f_2(c,\cdot)$ has derivative less than one, for all $c \in C$. From that point on we may argue as above. For the (w.r.t. $r$) constant number of functions $q_1,\ldots,q_{\Delta(C)}$ there will be some separate bounds as their derivatives are continuous on the compact set $C$. Thus we may choose $\tilde{T}(C)$ to be the maximum of those $\Delta(C)+1$ bounds. □

**Lemma 5.2.5** *Remember that* $r(n) := \alpha(c)\log\log n$ *and assume that* $\rho \leq r(n)$. *In a* $\mathcal{G}_{n,c/(n-1)}$ *graph when* $c \neq 3.35\ldots$ *the number of* $\rho$*-binary owners is sharply concentrated around* $n \cdot q_\rho(c)$. *More precisely, it is* not *equal to*

$$q_\rho(c)n \pm (4T(C))^\rho n^{1/2+\varepsilon}$$

*with probability subpolynomially small in* $n$, *say* $\exp(-(\log n)^2)$. *Here* $C$ *is some interval containing* $c$ *but not* $3.35\ldots$, *and* $T(C) > \max\{1, 1/c\}$ *is larger than the maximum of* $\frac{\partial}{\partial x} f_2(c,x)$ *on* $C \times [0,1]$.

Before presenting a proof we will show that the statement of Lemma 5.2.5 is actually more general than it seems.

**Corollary 5.2.6** *Suppose that* $0 \leq n - n' = \lambda_0(n) = o(n)$. *The statement of Lemma 5.2.5 implies that the number of* $\rho$*-binary owners in* $\mathcal{G}_{n',c/(n-1)}$ *is* not *equal to*

$$n'q_\rho(c) \pm (n'c\tilde{T}(C)\frac{\lambda_0(n)}{n-1} + (4T(C))^\rho n^{1/2+\varepsilon})$$

*with probability subpolynomially small in* $n$, *say* $\exp(-(\log n)^2)$. *The constant* $\tilde{T}(C)$ *stems from Corollary 5.2.4, and* $n$ *needs to be sufficiently large.*

**Proof** (Of Corollary 5.2.6.) For fixed $\rho$ the statement of Lemma 5.2.5 above for $\mathcal{G}_{n,c/(n-1)}$ replaced by $\mathcal{G}_{n',c'/(n'-1)}$ reads that the number of $\rho$-binary owners in $\mathcal{G}_{n',c'/(n'-1)}$ lies outside

$$[q_\rho(c')n' + (4T(C))^\rho (n')^{1/2+\varepsilon}, q_\rho(c')n' - (4T(C))^\rho (n')^{1/2+\varepsilon}],$$

with probability subpolynomially small in $n'$. Now $\mathcal{G}_{n',c'/(n'-1)}$ is nothing but $\mathcal{G}_{n',c/(n-1)}$, setting $c' = c(1 - \lambda_0(n)/(n-1))$.

We have seen in Corollary 5.2.4 that $c \mapsto q_\rho(c)$ is analytic and its derivative on an appropriately chosen interval $C$ containing $c$ but not $3.35\ldots$ is uniformly bounded by some constant $\tilde{T}(C)$. For sufficiently large $n$, $c'$ will be contained in $C$. Therefore by Taylor's Theorem $q(c')$ deviates from $q(c)$ by less than $c\tilde{T}(C)\lambda_0(n)/n$.

Finally observe that $n' \sim n$ and thus the probability is subpolynomially small in $n$, too. □

**Proof** (Of Lemma 5.2.5.) Denote the probability that a node is a $\rho$-binary owner in $\mathcal{G}_{n,c/(n-1)}$ by $\hat{q}_\rho$.

The basic idea is that concentration of the number of $(\rho-1)$-binary owners enables us to calculate the probability that a node is a $\rho$-binary owner to be close to the value expected in branching trees. Simply 'glue' a 'test particle'

to the graph, and see whether it is adjacent to at least two $(\rho-1)$-binary owners. This gets us the expected number of $\rho$-binary owners. Provided that $\rho$ is still sufficiently small, the number of $\rho$-binary owners is again a semi-local function and we may apply Lemma 5.1.5 proving that it is also concentrated around what it should be. We will now turn this idea into an inductive proof over $\rho = 0, \dots, r(n) = \alpha(c)\log\log n$.
$\rho = 0$: There are $n = nq_0(c)$ 0-binary owners.
$\rho - 1 \rightsquigarrow \rho$: We assume for the induction step that the statement is true for $\rho - 1$ choosing $n' := n - 1$. We consider (without loss of generality) the last node $v_n$ and the remainder graph $G \setminus v_n$ being an $\mathcal{G}_{n-1,c/(n-1)}$ random graph. Each of the potential $n-1$ edges connecting $v_n$ to $G \setminus v_n$ is switched on i.i.d. with probability $c/(n-1)$. Node $v_n$ is a $\rho$-binary owner in $\mathcal{G}_{n,c/(n-1)}$ if and only if it is connected to at least two $(\rho-1)$-binary owners in $G \setminus v_n$, i.e. in $\mathcal{G}_{n-1,c/(n-1)}$. We thus may apply an argument as in Lemma 5.2.2, where the marked set is the set of $(\rho-1)$-binary owners in $G \setminus v_n$.

Now we calculate the probability $\hat{q}_\rho$ conditional on $G \setminus v_n$ having the right number of owners.

$$
\begin{aligned}
& q_\rho(c) - 3/c \cdot T(C)(4T(C))^{\rho-1} n^{1/2+\varepsilon}/n \\
\overset{\text{a}}{\leq} & f_2(c, q_{\rho-1}(c) - 3/c \cdot (4T(C))^{\rho-1} n^{1/2+\varepsilon}/n) \\
= & \mathbb{P}\left[\mathrm{Po}_{cq_{\rho-1}(c) - 3\cdot(4T(C))^{\rho-1} n^{1/2+\varepsilon}/n} \geq 2\right] \\
\overset{\text{b}}{\leq} & \mathbb{P}\left[\mathrm{Bi}(n-1, cq_{\rho-1}(c)/(n-1) - 2\cdot(4T(C))^{\rho-1} n^{1/2+\varepsilon}/(n-1)^2) \geq 2\right] \\
\leq & \mathbb{P}\left[\mathrm{Bi}(n-1, c/(n-1)\frac{q_{\rho-1}(c)(n-1) - c\tilde{T}(C) - (4T(C))^{\rho-1} n^{1/2+\varepsilon}}{n-1}) \geq 2\right] \\
\overset{\text{c}}{\leq} & \hat{q}_\rho(c).
\end{aligned}
$$

- Inequality a) follows from Taylor's Theorem, using the fact that $0 \leq \frac{\partial}{\partial x} f_2(c,x) \leq T(C)$.
- Inequality b) follows from Proposition 5.2.3 .
- Inequality c) follows from the assumption using Corollary 5.2.6 for $(\rho-1)$ and $\lambda_0(n) = 1$.

For the last argument it is crucial that the number of $\rho - 1$ owners is at least

$$
q_{\rho-1}(c)(n-1) - (n-1)c\tilde{T}(C)\underbrace{\lambda_0(n)}_{=1}/(n-1) - (4T(C))^{\rho-1} n^{1/2+\varepsilon}
$$

and at most

$$q_{\rho-1}(c)(n-1) + (n-1)c\tilde{T}(C)\underbrace{\lambda_0(n)}_{=1}/(n-1) + (4T(C))^{\rho-1}n^{1/2+\varepsilon}.$$

Symmetrically,

$$\begin{aligned}
&\hat{q}_\rho(c)\\
\overset{c}{\leq}&\mathbb{P}\left[\mathrm{Bi}(n-1, c/(n-1)\frac{q_{\rho-1}(c)(n-1) + c\tilde{T}(C) + (4T(C))^{\rho-1}n^{1/2+\varepsilon}}{n-1}) \geq 2\right]\\
\leq&\mathbb{P}\left[\mathrm{Bi}(n-1, cq_{\rho-1}(c)/(n-1) + 2\cdot(4T(C))^{\rho-1}n^{1/2+\varepsilon}/(n-1)^2) \geq 2\right]\\
\overset{b}{\leq}&\mathbb{P}\left[\mathrm{Po}_{cq_{\rho-1}(c)+3\cdot(4T(C))^{\rho-1}n^{1/2+\varepsilon}/n} \geq 2\right]\\
=&f_2(c, q_{\rho-1}(c) + 3/c\cdot(4T(C))^{\rho-1}n^{1/2+\varepsilon}/n)\\
\overset{a}{\leq}&q_\rho(c) + 3/c\cdot T(C)(4T(C))^{\rho-1}n^{1/2+\varepsilon}/n.
\end{aligned}$$

The condition fails to be true with probability only subpolynomially small in $n$. If it does, the probability of being a $\rho$-binary owner may be anything (in $[0,1]$, obviously) but this yields only a subpolynomially small correction.

Now we may apply Lemma 5.1.5. The expected number of $\rho$-binary owners is now 'blurred' by another additive term $n^{1/2+\varepsilon}$, except for some subpolynomially small probability, and thus in the worst case

$$q_\rho(c)n - 4\cdot(4T(C))^{\rho-1}n^{1/2+\varepsilon} \leq \hat{q}_{\rho-1}(c)n \leq q_\rho(c)n + 4\cdot(4T(C))^{\rho-1}n^{1/2+\varepsilon}.$$

with probability one minus some subpolynomially small correction. □

Together with Lemma 5.2.2 this implies that the numbers of owners and runs are concentrated around the values expected from $\mathrm{Po}_c$-branching trees.

**Lemma 5.2.7** *Let $G \in \mathcal{G}_{n-\lambda_0(n),c/(n-1)}$, where $\lambda_0(n) = O(n^{1/2})$, $n' := n - \lambda_0(n)$. Then the numbers of the various kinds of owners are concentrated around the values expected from the branching trees with an error less than some $n^{1/2+\varepsilon'}$.*

*Assume again that $0 \leq \rho \leq r(n)$ With probability one minus some subpolynomially small error, say $\exp(-(\log n)^2)$ ...*

*... the number of $\rho$–binary owners is*

$$n'q_\rho(c) \pm n^{1/2+\varepsilon'},$$

*... the number of $(r(n)-1)$–binary owners is*

$$n'q(c)\pm, n^{1/2+\varepsilon'},$$

*... the number of $\rho$–Cayley owners is*

$$n'q_\rho(c) \pm n^{1/2+\varepsilon'},$$

*... the number of $r(n)$-Cayley owners is*

$$n'p(c) \pm n^{1/2+\varepsilon'}$$

*... the number of $r(n)$-binary owners that are adjacent to exactly two binary owners is*

$$n'\mathrm{Po}_{cq_{r(n)}}|_{\geq 2}(2) \pm n^{-1/2+\varepsilon'}.$$

*Finally, the number of runs is*

$$O(n/(\log n)^{\alpha(c)M(c)}),$$

*almost surely.*

**Proof** We can obviously overestimate errors of the form $K^{r(n)}n^{-1/2+\varepsilon}$ by the more expressive form $n^{-1/2+\varepsilon'}$. In all cases below we will select such an $\varepsilon'$ and take the maximum of these as the referent of the lemma.

- Lemma 5.2.5 and Corollary 5.2.6 directly imply the first part for $\rho$-binary owners, where $0 \leq \rho \leq r(n)$.
- The second statement follows from Lemma 3.1.2 on branching trees.
- For $r(n)$-Cayley owners we apply Lemma 5.2.2, where the marked set are the $(r-1)$–binary owners. Another application of Lemma 5.1.5 yields concentration of the $r$–Cayley owners around $n'p_{\rho(c)}$ and the third statement.
- The fourth statement follows again from Lemma 3.1.2.
- For the fifth statement we know, using Lemma 5.2.2 once more the marked set being the $r$–binary owners in the remainder graph, that an $r$–binary owner is adjacent to exactly two binary owners with probability

  $$[\mathrm{Po}_{cq_{r(n)}}|_{\geq 2}(2) - T_0(c)n^{-1/2+\varepsilon''}, \mathrm{Po}_{cq_{r(n)}}|_{\geq 2}(2) + T_0(c)n^{-1/2+\varepsilon''}].$$

  Applying Lemma 5.1.5 yields the desired result.

For runs, we may again argue as in Lemma 3.1.3. The $(r-1)$-binary owners that are not $r$-binary owners are sharply concentrated around $O(t(c)^{r(n)})$, whereas there are $O(n)$ $(r-1)$-binary owners. Being an $r$-Cayley owner, the node $v$ is adjacent to at least three $(r-1)$-binary owners in $G \setminus v$. Select three arbitrary children. The probability that at least one of those children is a 'pure' $(r-1)$-binary owner is in $O(t(c)^r(n))$, and by Markov's inequality the total number of runs is less than $\log n \cdot O(t(c)^r(n))$ with probability $O(1/\log n)$. □

## 5.3 Degree Sequence of the $r$-Cayley-Owners, $\mathcal{G}_{n,p}$ vs. $\mathcal{G}_{n,m}$

We have by now already 'translated' the branching tree results to random graphs. The set of owners induces a subgraph of the random graph, obviously very closely related to the $k$-core, and the proof of its sudden appearance is based on rigorously translating the branching tree ideas to random graphs. However the set of owners is not a subgraph of minimal degree $k$, since there are runs. Considering the gambler's ruin argument in the branching tree world it is intuitively clear that those runs do not matter a lot, but proving this requires further analysis. Following the strategy of Goerdt and Molloy we will later show that (essentially) the set of $r$-owners is distributed uniformly conditional on its degree sequence. Given this and knowing the degree sequence reasonably accurately the gambler's ruin argument can be made rigorous. However, this will only work in the $\mathcal{G}_{n,m}$ model.

We finish this chapter by calculating the degree sequence of the graph induced by the set of owners fairly accurately and translating those results to the $\mathcal{G}_{n,m}$ model. Also, the number of nodes with a non-tree-like $r$-neighbourhood is small in the $\mathcal{G}_{n,m}$ model, too. We shall make use of the 'translation tools' Fact 1 and Fact 2 (see Section 2.1.1).

In view of Lemma 5.2.2 the degree-sequence of the set of $r$-owners should be pretty close to what we expect it to be from the Branching Tree Connection, heuristically, that is close to

$$d_i^{(0)} = \begin{cases} n \cdot \mathrm{Po}_{cq}(i) & i \geq 3 \\ 0 & 0 < i < 3 \\ n \cdot \mathbb{P}[\mathrm{Po}_{cq} < 3] & i = 0 \end{cases} \tag{5.2}$$

In order to facilitate the translation from $\mathcal{G}_{n,p}$ to $\mathcal{G}_{n,m}$ we are also interested in the *integrated degree sequence*, that is $\{D_i\}_{i \in \mathbb{N}_0}$, where $D_i$ counts the number

of nodes having at least degree $i$. We expect it to be close to

$$D_i^{(0)} = \begin{cases} n \cdot \mathbb{P}[\mathrm{Po}_{cq}(i) \geq i] & i \geq 3 \\ n - n \cdot \mathbb{P}[\mathrm{Po}_{cq} < 3] & 0 < i < 3 \\ n & i = 0 \end{cases} \tag{5.3}$$

Observe that knowing the $D_i$ within some small margins of errors implies knowledge of the $d_i$ within essentially the same margins of errors, because

$$d_i = D_i - D_{i+1}.$$

We shall say that an integrated degree sequence $\{D_i\}_{i\in\mathbb{N}_0}$ is *$\varepsilon$-close* to some 'idealised' integrated degree sequence $\{D_i^0\}_{i\in\mathbb{N}_0}$ if

$$D_i^0 - \varepsilon \leq D_i \leq D_i^0 + \varepsilon,$$

for all $i$. Note that $D_i, D_i^0$ and $\varepsilon$ can and will depend on $n$.

**Lemma 5.3.1** *Consider any $\varepsilon = \varepsilon(n)$. The (sequence of) event(s)*

$$\mathcal{E} = \mathcal{E}^{(n)} = \{G \in \mathcal{G}_n | \textit{ integrated degree sequence of } G \textit{ is } \varepsilon\textit{-close to } \{D_i^0\}_{i\in\mathbb{N}_0}\}$$

*is convex.*

**Proof** Take $F \subseteq G \subseteq H$, and assume that $F, H \in \mathcal{E}$. Then

$$D_i(F) \leq D_i(G) \leq D_i(H),$$

because any node that is adjacent to $i$ $r$-owners in some graph will be adjacent to at least $i$ $r$-owners after adding some edges, i.e. in any super-graph. But since both $D_i(F)$ and $D_i(H)$ are contained within $[D_i^0 - \varepsilon, D_i^0 + \varepsilon]$, so is $D_i(G)$. □

**Lemma 5.3.2** *The integrated degree sequence $\{D_i\}_{i\in\mathbb{N}_0}$ of the set of $r$-owners is $\varepsilon$-close to $\{D_i^0\}_{i\in\mathbb{N}_0}$ for some $\varepsilon(n) = O(n/\log(n)^{\alpha(c)M(c)})$, with probability approaching one. This holds true in the $\mathcal{G}_{n,p}$ model and in the corresponding $\mathcal{G}_{n,m}$ model.*

Before proving Lemma 5.3.2 we will provide some little tools.

**Proposition 5.3.3** *The derivative of $f_k(c,x) = \mathbb{P}[\mathrm{Po}_{cx} \geq k]$ is uniformly bounded by $c$, that is*

$$0 \leq \frac{d}{dx} f_k(c,x) \leq c,$$

*for all $k \in \mathbb{N}_0, x \in \mathbb{R}$.*

**Proof** Consider $A(k) := \max_{x \in \mathbb{R}} \frac{d}{dx} f_k(c,x)$. Now straightforward curve discussion yields that $\frac{d}{dx} f_k(c,x)$ is maximised at $\frac{k-1}{c}$, and there it takes the value

$$\frac{(k-1)^{k-1} e^{-(k-1)} c}{(k-1)!} \leq c,$$

because $(k-1)^{k-1}/(k-1)! \leq \sum_{i=1}^{\infty} (k-1)^i/i! = e^{k-1}$. Thus $A(k) \leq c$, and the proof is complete. □

We quote a powerful statement from Janson's paper [Jan94]:

**Fact 11** *(Straightforward consequence of Inequality 11 in [Jan94].)*

$$d_{TV}(\mathrm{Bi}_{n,c/n}, \mathrm{Po}_c) \leq \frac{c}{n}.$$

**Remark 5.3.4** *Remember that the total variation distance $d_{TV}(\mu,\nu)$ between probability measures $\mu$ and $\nu$ is defined as*

$$d_{TV}(\mu,\nu) := \sup_A |\mu(A) - \nu(A)|,$$

*where the* sup *runs over all possible events. Thus it is a worst case bound.*

**Proof** (of Lemma 5.3.2). We shall only discuss $\mathcal{G}_{n,p}$. Note that the good event is convex according to Lemma 5.3.1. It is then easy to translate the result to $\mathcal{G}_{n,m}$ using Fact 1.

We know from Lemma 5.2.5 that the set of $r-1$ binary-owners has size $nq \pm \delta(n)$ and the set of $r-1$ binary owners that are not $r$-binary owners has size $\gamma(n) = O(n/\log(n)^{\alpha(c)M(c)})$, both with all but a subpolynomially small probability. We shall merely make use of the fact that $\delta(n)$ and $\gamma(n)$ are some functions in $o(n)$.

We shall first calculate the expected values of the $D_i$ by using the 'test-particle approach'. Concentration will follow from Lemma 5.1.5. Since we only have to handle a polynomial number of events (note that all $D_i$ are certainly 0 for $i \geq n$!) all the $D_i$ will remain within their respective 'confidence intervals' simultaneously, with probability tending to one.

$i \geq 3$: We consider a 'test particle' $v$, connected with $n$ potential edges to the remainder graph, each edge being switched on with probability $c/n$. We know that the number of $r(!)$-binary owners in the remainder graph is

$nq \pm \delta(n)$ with all but a subpolynomially small probability. Therefore

$$\begin{aligned}
&\mathbb{P}[\text{`}v \text{ adjacent to} \geq i \text{ } r\text{-binary-owners'}|\text{`\# } r\text{-binary-owners is } nq \pm \delta(n)\text{'}] \\
&= \mathbb{P}\left[\mathrm{Bi}(nq \pm \delta(n), \frac{c}{n}\frac{nq \pm \delta(n)}{nq \pm \delta(n)}) \geq i\right] \\
&\overset{\text{Fact 11}}{=} \mathbb{P}\left[\mathrm{Po}_{c(q\pm\delta(n)/n)} \geq i\right] \pm \underbrace{\frac{\frac{c(nq\pm\delta(n))}{n}}{nq \pm \delta(n)}}_{=c/n} \\
&\overset{\text{Proposition 5.3.3, Taylor}}{=} \mathbb{P}[\mathrm{Po}_{cq} \geq i] \pm 2c \cdot \delta(n)/n.
\end{aligned}$$

Now, conditional on the event '# $r$-binary-owners is not $nq \pm \delta(n)$' the probability of $v$ being adjacent to at least $i$ $r$-binary-owners may be anything, but this event only happens with a subpolynomially small probability. Therefore

$$\mathbb{E}[D_i] = n\mathbb{P}[\mathrm{Po}_{cq} \geq i] \pm 3c \cdot \delta(n).$$

Finally, we can use Lemma 5.1.5, to see that '$D_i = n\mathbb{P}[\mathrm{Po}_{cq} \geq i] \pm 4c \cdot \delta(n)$' holds with probability one minus some subpolynomially small correction.

$i < 3$: We will only give a rather crude but sufficient argument. It is clear that only $n - D_3$ nodes have degree less than 3. Those that have degree 1 or 2 are runs, and we know there are only $\gamma(n)$ of them around. Thus $d_0 = n - S_3 - \gamma$. Remember that $d_i = D_i - D_{i+1}$, and that $D_0 \equiv n$. We this implies that

$$\begin{aligned}
D_1 &= D_0 - d_0 = D_3 \pm \gamma = n\mathbb{P}[\mathrm{Po}_{cq} \geq 3] \pm 2\gamma. \\
D_2 &= D_1 - d_1 = D_3 \pm 3\gamma.
\end{aligned}$$

(We have - realistically - assumed that $\gamma$ is $\omega(\delta)$.)

Thus we may choose $\varepsilon(n) = 3\gamma(n)$. □

Observe that nodes with a *treelike* $r$-neighbourhood will survive $(r-1)$ steps of shell-wise deletion if and only if they are $r$-Cayley owners. Since the expected number of nodes with a non-tree-like $r$-neighbourhood is (bounded above by) $\log(n)^{2r(n)}$ (c.f. Proposition 5.1.3) we have essentially predicted the outcome of the shell-wise deletion process. Using Markov's inequality we know that there are $o(n)$ nodes with a non-tree-like $r$-neighbourhood, a.a.s. Using Fact 2 we can easily translate this to $\mathcal{G}_{n,m}$.

**Proposition 5.3.5** *The number of nodes whose $r$-neighbourhood is not a tree is less than $n^{1/2+\varepsilon}$, with probability converging to one in the $\mathcal{G}_{n,m}$ model with average degree $c$.*

**Proof** The probability that a node has a non-tree-like $r$-neighbourhood is $\log(n)^{2r(n)}/n$ by Proposition 5.1.3, in $\mathcal{G}_{n,p}$. By Fact 2 this probability increases by at most a factor of $\sqrt{m}$, that is $\sqrt{c/2 \cdot n}$. Now, using Markov's inequality for a sum of indicators $\sum_{v \in V} X_v$ we get

$$\mathbb{P}\left[\sum_{v \in V} X_v \geq t\right] \leq \frac{\mathbb{E}[\sum_{v \in V} X_v]}{t}.$$

Setting $t := n^{1/2+\varepsilon}$ yields

$$\mathbb{P}\left[\text{'more than } n^{1/2+\varepsilon} \text{ nodes with a non-tree-like } r\text{-neighbourhood'}\right] \leq \frac{\log(n)^{2r(n)}\sqrt{c/2 \cdot n}}{n^{1/2+\varepsilon}} \overset{n \to \infty}{\to} 0.$$

□

# Chapter 6

# Sudden Appearance of the $k$-Core

The result of Goerdt and Molloy concerning the $k$-core for random faulty configurations consists of two essential ingredients.

The first is proving the existence of a sub-configuration that has almost the right degree sequence - for $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ we have parallelled this with our results on the set of owners described in Chapter 5. It remains to show that we have also correctly predicted the degree sequence after $(r-1)$ rounds of the shell-wise deletion process. This is tantamount to by-passing most of the tedious step by step analysis of the single node deletion process in [PSW96].

The second ingredient consists in proving that both shell-wise and single-node deletion processes leave invariant the uniform distribution conditional on the given degree sequence. There are only $o(n)$ nodes (or classes) with a non-treelike neighbourhood, both inrandom graphs and faulty configurations. All other nodes will survive $(r-1)$ rounds of the shell-wise deletion process if and only if they are $r$-owners. From that point on it is straightforward to show that the single-node deletion process terminates after $o(n)$ steps using a gambler's ruin argument, provided both the degree sequence and uniform distribution conditional on that degree sequence are known.

It is fairly easy to see that that there may be no $k$-core in the sub-critical case (Section 6.1), using an edge-density argument as Goerdt and Molloy did. In the super-critical case (Section 6.2) we study the invariance of uniform distribution under the various deletion processes (Section 6.2.1), prove that the degree sequence after $(r-1)$ rounds of edge deletion is essentially the same as the degree sequence of the $r$-owners (Section 6.2.1) and show that the single node process removing remaining runs will a.s. terminate after $o(n)$ steps (Section 6.2.1). We shall state and prove the new theorem for the sudden appearance of the $k$-core in the $\mathcal{G}_{n,m}$ model in Section 6.3.

## 6.1 Sub-Critical Case

Before we can complete the proof of the main theorem, we need to discuss the sub-critical case ($c < 3.35\ldots$). We argue similar to as Goerdt and Molloy did in [GM00].

**Lemma 6.1.1** *Almost surely a graph in $\mathcal{G}_{n,c/n}$ contains no induced subgraph with minimum degree 3 and $o(n)$ nodes.*

**Proof** Consider all subsets $S$ of size $s \leq \varepsilon n$, for some $\varepsilon$ that will be fixed below.

Denote by $X_S$ the indicator random variable of the event '$S$ induces a subgraph with minimum degree three'. We will show that

$$\begin{aligned}\mathbb{E}\left[\sum_{|S|\leq\varepsilon n} X_S\right] &= \sum_{s=2}^{\varepsilon n}\mathbb{E}\left[\sum_{|S|=s} X_S\right] \\ &= \sum_{s=2}^{\varepsilon n}\binom{n}{s}\mathbb{P}\left[X_S = 1 \wedge |S| = s\right] \qquad (6.1)\end{aligned}$$

converges to zero which proves the lemma via the First Moment Method.

$$\begin{aligned}&\binom{n}{s}\mathbb{P}\left[X_S = 1 \wedge |S| = s\right] \leq \binom{n}{s}\binom{\binom{s}{2}}{3/2\cdot s}(c/n)^{3/2\cdot s} \\ &\leq (en/s)^s (es/3)^{3/2\cdot s}(c/n)^{3/2\cdot s} \\ &\leq (1/27\cdot e^5 s c^3/n)^{1/2\cdot s} \\ &\leq (6sc^3/n)^{1/2\cdot s} \\ &=: f(s).\end{aligned}$$

Now

$$\sum_{s=2}^{\varepsilon n} f(s) = \sum_{s=2}^{\varepsilon n}\left[(6sc^3/n)^{1/2}\right]^s =: \sum_{s=2}^{\varepsilon n}\left[(As/n)^{1/2}\right]^s.$$

Choosing $\varepsilon \leq 1/(2Ae) = 1/(12c^3e)$ we get $f(2) = 2A/n$ and for $s \geq 3$

$$\begin{aligned}&f(s)/f(s-1) \\ =&(As/n)^{(s-1)/2}\sqrt{As/n}(A(s-1)/n)^{-(s-1)/2} \\ =&((s+1)/s)^{s/2}\sqrt{As/n} \overset{s/n\leq\varepsilon}{\leq} e^{1/2}\sqrt{1/(2e)} \\ =&1/\sqrt{2}.\end{aligned}$$

Therefore

$$\sum_{s=2}^{\varepsilon n} f(s) \leq 2A/n \cdot \sum_{k=0}^{\infty} (1/\sqrt{2})^k = O(1/n).$$

□

## 6.2 Super-Critical Case

### 6.2.1 Deletion Processes and Uniform Distribution

Remember that the $k$-core can be found by any suitable protocol of node deletion. Following the overall proof strategy of Goerdt and Molloy we will in a *first phase* delete bad nodes shell by shell. Applying this shell-wise deletion process $(r-1)$ times will yield a subgraph that is essentially the set of $r$-Cayley owners. That is, we can predict the degree sequence after $(r-1)$ rounds of edge deletion sufficiently accurate.

All but $o(n)$ nodes will have degree at least $k$. Moreover, the form of the predicted degree sequence is such that we can show that the single node deletion process applied to the remainder graph (in a *second phase*) will a.s. die out removing only $o(n)$ nodes using a straightforward gambler's ruin argument.

For the proof to work as sketched above it is essential to know that the uniform distribution, conditional on the given degree sequence, remains invariant during each step of both shell-wise and single-node deletion.

The deletion process can be seen as a mapping

$$\Delta : G \mapsto \Delta(G).$$

In the case of shell-wise deletion all bad nodes (having degree less than $k$) are deprived of their incident edges. For single-node deletion we uniformly pick a random bad vertex and deprive it of its incident edges, if there are no bad vertices we do nothing. Note that in the latter case $\Delta$ is a *random* mapping.

Let us denote the mapping $G \mapsto \{d_i(G)\}_{i \in \mathbb{N}_0}$ by the letter $\delta$. We will sometimes denote degree sequences by the letter '$\delta$', too, which is sloppy but should not lead to confusion.

**Proposition 6.2.1** *Let $G$ be distributed according to the $\mathcal{G}_{n,m}$-model and condition on the event $\mathcal{E}$ = 'G has degree sequence $\{d_i\}_{i \in \mathbb{N}_0}$', where $\sum_i i d_i = 2m$. Then $G$ conditional on $\mathcal{E}$ is distributed uniformly amongst all graphs with degree-sequence $\{d_i\}_{i \in \mathbb{N}_0}$.*

**Proof** Note that $\mathcal{E} \subseteq G_n$ is a *subset* of the set of all graphs on $n$ nodes with exactly $m$ edges. Considering the $\mathcal{G}_{n,m}$-model is tantamount to considering the uniform distribution on the set of all such graphs. But uniform distribution on some discrete probability space $\Omega$ obviously induces uniform distribution on any subset of $\Omega$. □

We will call a degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$ *relevant* if it can be reached by a series of shell-wise or single-node deletion steps, i.e. there is a $G$ such that for some $\ell$,

$$\delta(\Delta^{(\ell)}(G)) = \{d_i\}_{i\in\mathbb{N}_0}\,.$$

Two degree sequences $\{d_i\}_{i\in\mathbb{N}_0}$ and $\{d'_i\}_{i\in\mathbb{N}_0}$ are *consistent* if there is at least one graph $G$ with $\delta(G) = \{d_i\}_{i\in\mathbb{N}_0}$, such that

$$\delta(\Delta(G)) = \{d'_i\}_{i\in\mathbb{N}_0}\,.$$

In what follows we will always assume all degree sequences to be relevant and pairs of degree sequences to be consistent, as this will quite obviously be the case for all degree sequences encountered in the course of the deletion process(es).

We will presently show that both deletion processes leave invariant the uniform distribution conditional on the (present) degree sequence. The essence of the proof idea is to take two arbitrary graphs with degree sequence $\{d'_i\}_{i\in\mathbb{N}_0}$ and to prove that their predecessor sets under $\Delta$ amongst all graphs with degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$ are disjoint and of the same cardinality. Note that the disjointness is evident since $\Delta$ is a mapping.

### Shell-Wise Deletion

**Lemma 6.2.2** *For all consistent pairs $(\{d_i\}_{i\in\mathbb{N}_0}, \{d'_i\}_{i\in\mathbb{N}_0})$ and for all relevant graphs $G'$ (such that $\delta(G') = \{d'_i\}_{i\in\mathbb{N}_0}$) the number of predecessors $G \in \Delta^{-1}[G']$ under the shell wise deletion process $\Delta$ (such that $\delta(G) = \{d_i\}_{i\in\mathbb{N}_0}$), depends only on $(\{d_i\}_{i\in\mathbb{N}_0}, \{d'_i\}_{i\in\mathbb{N}_0})$, not on the specific choice of $G'$.*

### Proof

**Semi-formal version of the proof:** Consider two arbitrary graphs $G'_1$ and $G'_2$ which both have degree sequence $\{d'_i\}_{i\in\mathbb{N}_0}$. Choose an arbitrary relabelling $\pi = \pi_{G'_1,G'_2} \in S_n$ of the nodes, preserving the degrees, i.e. $\deg_{G'_1}(v) = \deg_{G'_2}(\pi(v))$. Let $G_1$ be an arbitrary predecessor of $G'_1$ under $\Delta$. Then

$$G_2 := G'_2 \uplus \pi[\underbrace{G_1 \setminus G'_1}_{=:H}]$$

is a predecessor of $G_2'$. In other words $H \mapsto \pi[H]$ induces a bijection between the (disjoint, see above) sets of predecessor graphs. It remains to justify using the symbol '$\uplus$', for disjoint union.

We shall say that $H$ 'fits on' $G_1'$ if no edge in $H$ has been present in $G_1'$, i.e. adding $H$ to $G_1'$ will not introduce multiple edges. The essential observation is that when $H$ 'fits on' $G_1'$ we know that $\pi[H]$ 'fits on' $G_2'$ since each edge in $H$ is incident to at least one node with degree 0 in $G_1'$ and $\pi$ preserves degrees.

Note that we know that the set of $H$'s is never empty because of relevance and consistency of the degree sequences involved.

**Entirely formal proof:** Choose $\pi = \pi_{G_1',G_2'} \in S_n$ preserving degrees as above. Denote by

$$\begin{aligned}\pi^* :&\mathcal{G}_n \to \mathcal{G}_n\\ &\{\{u,v\},\ldots\} \mapsto \{\{\pi(u),\pi(v)\},\ldots\}\end{aligned}$$

the bijection between the set of subgraphs induced by $\pi$ (we encode an element of $\mathcal{G}_n$ by the subset of edges 'switched on').

We shall show that

$$\begin{aligned}\Phi :&\Delta^{-1}[G_1'] \to \Delta^{-1}[G_2']\\ &G_1 \mapsto G_2' \uplus \pi^*(G_1 \setminus G_1')\end{aligned}$$

and

$$\begin{aligned}\Psi :&\Delta^{-1}[G_2'] \to \Delta^{-1}[G_1']\\ &G_2 \mapsto G_1' \uplus (\pi^*)^{-1}(G_2 \setminus G_2')\end{aligned}$$

are well defined mappings between $\Delta^{-1}[G_1']$ and $\Delta^{-1}[G_2']$, and that

$$\Phi \circ \Psi^{-1} = \mathbf{id}_{\Delta^{-1}[G_1']} \text{ and } \Psi \circ \Phi^{-1} = \mathbf{id}_{\Delta^{-1}[G_2']}.$$

The latter is clear by construction of $\Phi$ and $\Psi$. By symmetry it will suffice to show that $\Phi$ is well defined. We need to show that for every $G_1 \in \Delta^{-1}[G_1']$ there is exactly one $G_2 \in \Delta^{-1}[G_2']$ such that $G_2 = G_2' \uplus \pi^*(G_1 \setminus G_1')$. Keep in mind that the $\uplus$ is justified as we have argued above, thus $G_2$ is unique. It remains to show that $G_2 \in \Delta^{-1}[G_2']$. Now $\Delta(G_2)$ is contained in $G_2'$ because all edges in $\pi^*(G_1 \setminus G_1')$ will be incident to nodes with degree at most $k-1$ by construction. Conversely, $G_2'$ is contained in $\Delta(G_2)$, because assuming otherwise would yield the following contradiction. By construction, $G_2$ has the same degree-sequence as $G_1$ implying that there are $d_1 + d_2 + \ldots + d_{k-1}$ nodes that are bad in $G_2$, too. On the other hand, all nodes that had degree

0 in $G_1'$ and are bad in $G_1$ are mapped to those $d_1 + d_2 + \ldots + d_{k-1}$ nodes in $G_2$. Thus no other edge in $G_2$ will be deleted than the edges that correspond to edges in $H$.

□

This proof also intuitively 'explains' why it is important to remove *all* edges incident to bad nodes. Otherwise, if some difference graph $H$ 'fits on $G_1'$, it might be that $\pi[H]$ introduces some multiple edges when 'added' to $G_2'$. By a simple induction we get the following corollary.

**Corollary 6.2.3** *Suppose an initial graph $G_0$ is chosen uniformly at random from all configurations with degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$. Then the graphs $G^{(\ell)}$ derived from $G_0$ by an $\ell$-fold application of the shell-wise deletion process, with degree sequences $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, are distributed uniformly conditional on $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, respectively.*

**Remark 6.2.4** *We may replace the $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$ by the integrated degree sequence $\{D_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, the discrete antiderivative of the degree sequence, i.e.*

$$D_i = \sum_{j\geq i} d_j.$$

*Any graph has $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$ when it has the corresponding $\{D_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, provided $D_i - D_{i+1} = d_i$, for all $i \geq 0$.*

### Single-Node Deletion

**Lemma 6.2.5** *Consider a consistent pair $(\{d_i\}_{i\in\mathbb{N}_0}, \{d_i'\}_{i\in\mathbb{N}_0})$. Denote by $\Delta$ the action of the (randomised!) single-node deletion process, and assume that $G$ is distributed uniformly amongst all graphs with degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$.*

*Then $\Delta(G)$ is distributed uniformly amongst all graphs with degree sequence $\{d_i'\}_{i\in\mathbb{N}_0}$, conditional on having this degree sequence.*

**Proof** Fairly similar to the proof of Lemma 6.2.2.

From inspecting $\{d_i\}_{i\in\mathbb{N}_0}$ and $\{d_i'\}_{i\in\mathbb{N}_0}$ we can easily determine the degree $0 < \ell < k$ of the bad node that is going to be deleted - it is simply equal to the number of edges 'before' minus the number of edges 'after'.

We shall consider the *$\ell$-strip* $\mathcal{S}_\ell$ consisting of all pairs $(G, v)$, where $G$ has degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$ and $v$ is contained in the set of bad nodes with degree $\ell$, i.e. in $B_\ell(G)$ consisting of all nodes with $\deg_G(v) = \ell$.

Since we assume $G$ to be distributed uniformly amongst all graphs with degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$, and the single node deletion process picks bad nodes uniformly at random, any pair $(G, v)$ in $\mathcal{S}_\ell$ is equally likely.

Instead of considering the randomised (single node) deletion operation $\Delta$ we consider the deterministic mapping $\Delta_\ell$ from $\mathcal{S}_\ell$ to the set of graphs with degree sequence $\{d'_i\}_{i\in\mathbb{N}_0}$. The pair $(G, v)$ is mapped to $\Delta_v(G) = G \setminus v$. Here $\Delta_v$ is the deterministic deletion of node $v$.

It is clear that the randomised operation $\Delta$ produces uniform distribution on the set of graphs with degree sequence $\{d'_i\}_{i\in\mathbb{N}_0}$ if the cardinality of $\Delta_\ell^{-1}[G']$ does not depend on the specific choice of $G'$.

Now

$$\left|\Delta_\ell^{-1}[G']\right| = \sum_{\{v_0 \,|\, \deg_{G'}(v_0)=0\}} \left|\{(G, v_0) \in \mathcal{S}_\ell \,|\, \Delta_{v_0}(G) = G'\}\right|.$$

Clearly the number of summands depends only on $\{d'_i\}_{i\in\mathbb{N}_0}$.

It remains to show that the summands are equal for all $(G'_1, v_1), (G'_2, v_2)$, where $v_1, v_2$ are arbitrary choices of nodes such that $\deg_{G'_1}(v_1) = 0$ and $\deg_{G'_2}(v_2) = 0$. We shall choose a suitable node-relabelling $\pi = \pi_{G'_1,G'_2,v_1,v_2} \in S_n$ preserving degrees and identifying $v_1$ and $v_2$.

Again, as in the proof of Lemma 6.2.2, any graph in $\Delta_{v_1}^{-1}[G'_1]$ can be written as $(G'_1 \uplus H, v_1)$, and $(G'_2 \uplus \pi[H], \pi(v_1))$ is a corresponding predecessor in $\Delta_{v_2}^{-1}[G'_2]$. Thus all summands are enumerated by the set of predecessor graphs of, say, $G'_1$ and the have identical cardinality for all $(G'_i, v_i)$. It is again crucial that the $v_i$ are *isolated nodes* in the graphs $G'_i$. Note that all difference graphs $H$ are stars centred at the respective $v_i$'s with $\ell$ edges.

□

**Corollary 6.2.6** *Suppose an initial graph $G_0$ is chosen uniformly at random from all configurations with degree sequence $\{d_i\}_{i\in\mathbb{N}_0}$. Then the graphs $G^{(\ell)}$ derived from $G_0$ by an $\ell$-fold application of the single-node deletion process, with degree sequences $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, are distributed uniformly conditional on $\{d_i^{(\ell)}\}_{i\in\mathbb{N}_0}$, respectively.*

### 6.2.2 Degree Sequence after Deleting $(r-1)$ Shells

We shall fist observe that the degree-sequence of the graph after $(r-1)$ steps of shell-wise deletion is $\varepsilon$-close to the degree sequence of the set of owners.

**Lemma 6.2.7** *Let $G$ be some graph distributed according to the $\mathcal{G}_{n,m}$ model with average degree $c$. A.a.s., the degree-sequence of the subgraph $G_\mathcal{S}$ induced*

*by $\mathcal{S}$, the set of nodes surviving $(r-1)$ steps of shell wise deletion with a non-zero degree,is $\varepsilon$-close to the degree sequence of the subgraph $G_{\mathcal{O}}$ induced by set $\mathcal{O}$ of $r$-owners. We may choose $\varepsilon(n) = n^{1/2+\varepsilon'}$, where $\varepsilon'$ is larger than the $\varepsilon$ of Proposition 5.3.5.*

**Proof** Denote by $\mathcal{T}$ the set of nodes that have a tree-like $r$-neighbourhood w.r.t. the given graph $G$.

We shall first prove that

$$\mathcal{O} \cap \mathcal{T} \subseteq \mathcal{S} \subseteq \mathcal{O} \cup \overline{\mathcal{T}}. \tag{6.2}$$

The first inclusion is true since any $r$-owner with a tree-like $r$-neighbourhood will survive $(r-1)$ rounds of edge deletion with degree at least 3. The second inclusion is tantamount to the contraposition of the following implication: if a node is an $r$-non-owner (i.e. in $\overline{\mathcal{O}}$) and has a tree-like $r$-neighbourhood it will certainly *not* survive $(r-1)$ steps of shell wise deletion with a non-zero degree, see Proposition 2.3.8.

We shall next show that the (integrated) degree-sequences of the corresponding induced subgraphs of $G$ are monotonous. We shall write $\deg_{\mathcal{A}}(v)$ to denote the degree of $v$ with respect to the subgraph $G_{\mathcal{A}}$ of $G$ induced by the (node-)set $\mathcal{A}$. Clearly, $\deg_{\mathcal{O}\cap\mathcal{T}}(v) \leq \deg_{\mathcal{S}}(v)$.

Now let $\underline{X}_v^i$ be an 'indicator' that is 1 if $\deg_{\mathcal{O}\cap\mathcal{T}}(v) \geq i$ and 0 otherwise and $Y_v^i$ be an 'indicator' that is 1 if $\deg_{\mathcal{S}}(v) \geq i$ and 0 otherwise. Summing over all nodes yields

$$D_i(G_{\mathcal{O}\cap\mathcal{T}}) = \sum_v \underline{X}_v^i \leq \sum_v Y_v^i = D_i(G_{\mathcal{S}}).$$

In a completely analogous way we can show that

$$D_i(G_{\mathcal{S}}) = \sum_v Y_v^i \leq \sum_v \overline{X}_v^i = D_i(G_{\mathcal{O}\cup\overline{\mathcal{T}}}).$$

It remains to show that $D_i(G_{\mathcal{O}\cap\mathcal{T}})$ and $D_i(G_{\mathcal{O}\cup\overline{\mathcal{T}}})$ remain reasonably close to $D_i(G_{\mathcal{O}})$, respectively. We shall prove the following inequality for arbitrary disjoint subsets $\mathcal{A}, \mathcal{B}$ below ($\mathcal{C} := \mathcal{A} \uplus \mathcal{B}$).

$$D_i(G_{\mathcal{A}}) \geq D_i(G_{\mathcal{A}\uplus\mathcal{B}}) - (\Delta(G)+1)\,|\mathcal{B}|\,. \tag{6.3}$$

Substituting

i) $\mathcal{A}_1 := \mathcal{O} \cap \mathcal{T}, \mathcal{B}_1 := \mathcal{O} \setminus (\mathcal{O} \cap \mathcal{T})$ and

ii) $\mathcal{A}_2 := \mathcal{O}, \mathcal{B}_2 := (\mathcal{O} \cup \overline{\mathcal{T}}) \setminus \mathcal{O}$

in Inequality 6.3 actually yields that

$$\begin{aligned} & D_i(G_{\mathcal{O}}) - (\Delta(G)+1)\,|\mathcal{B}_1| \\ & \leq D_i(G_{\mathcal{O}\cap\mathcal{T}}) \leq D_i(G_{\mathcal{S}}) \leq D_i(G_{\mathcal{O}\cup\overline{\mathcal{T}}}) \\ & \leq D_i(G_{\mathcal{O}}) + (\Delta(G)+1)\,|\mathcal{B}_2|\,. \end{aligned}$$

Observe that in either case $\mathcal{B}_i$ is contained in (the small set) $\overline{\mathcal{T}}$ and the fact that a.a.s. the degrees in $G$ are bounded*. Indeed, the maximum degree is less than $(\log n)^2$, a.a.s., in the $\mathcal{G}_{n,m}$ model, too. This is an easy consequence of Proposition 5.1.1 and Fact 2. Remember that $\overline{\mathcal{T}}$ contains at most $n^{1/2+\varepsilon}$ nodes, a.a.s. by Proposition 5.3.5.

Therefore we may choose $\varepsilon(n) = n^{1/2+\varepsilon'}$, with an appropriately large $\varepsilon'$ (independent of $n$!).

It remains to prove Inequality 6.3.

$$\begin{aligned} D_i(G_{\mathcal{A}\uplus\mathcal{B}}) &= \sum_v C_v^i \\ &= \sum_{v\in\mathcal{A}} C_v^i + \underbrace{\sum_{v\in\mathcal{B}} C_v^i}_{\leq|\mathcal{B}|} \\ &\leq \sum_{v\in\mathcal{A}} A_v^i + (\Delta(G)+1)\,|\mathcal{B}| \\ &= D_i(G_{\mathcal{A}}) + (\Delta(G)+1)\,|\mathcal{B}|\,. \end{aligned}$$

Here $\Delta(G)$ denotes the maximum degree in $G$, $A_v^i$ is indicating that $v$ has degree $\geq i$ in $G_{\mathcal{A}}$ and $C_v^i$ is indicating that $v$ has degree $\geq i$ in $G_{\mathcal{C}}$. The inequality sign holds because every node in $\mathcal{B}$ can contribute to the degree (w.r.t. $\mathcal{C}$) of at most $\Delta(G)$ nodes in $\mathcal{A}$; in the worst case the degrees (w.r.t. $\mathcal{C}$) of all affected nodes in $\mathcal{A}$ drop below $i$, for all other nodes in $\mathcal{A}$ we have $C_v^i = A_v^i$. Observe that the $C_v^i = 1$ but $A_v^i = 0$ only happens when $v$ is affected, i.e. adjacent to some node in $\mathcal{B}$. □

Combining this with the result on the degree sequence of the set of owners stated in Lemma 5.3.2 it is now easy to predict the degree sequence after $(r-1)$ steps of shell-wise deletion.

*Note that without some reasonable degree bound the statement would not be very useful even when $\mathcal{B}$ is very small, as can be seen from the following counterexample: Let $G$ be a star on $n+1$ nodes, centred at node 1, choose $\mathcal{B} := \{1\}$ and $\mathcal{A} := [n+1] \setminus \mathcal{B}$. Then $D_1(G_{\mathcal{A}}) = 0$ but $D_1(G_{\mathcal{A}\uplus\mathcal{B}}) = n$. I.e. Inequality 6.3 is sharp.

**Corollary 6.2.8** *Let $G$ be distributed according to the $\mathcal{G}_{n,m}$ model with average degree $c$. A.a.s., the degree-sequence of the subgraph $G_{\mathcal{S}}$ induced by $\mathcal{S}$, the set of nodes surviving $(r-1)$ steps of shell wise deletion with a non-zero degree,is $\varepsilon$-close to the degree sequence $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$ expected from the Branching Tree Connection. Here $\varepsilon = \varepsilon(n) = O(n/\log(n)^{\alpha(c)M(c)})$.*

### 6.2.3 Single Node Deletion, Gambler's Ruin

In Corollary 6.2.8 above we have - for $\mathcal{G}_{n,m}$ - determined the degree sequence $\{d_i(G)\}_{i\in\mathbb{N}_0}$ of the graph remaining after an $(r-1)$-fold application of the shell-wise deletion process, for the super-critical case $c > 3.35\ldots$. It is a.a.s. $\varepsilon$-close to the degree-sequence $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$ that we would intuitively expect from the Branching Tree Connection. Furthermore, we know that the graph is distributed uniformly amongst all graphs having degree sequence $\{d_i(G)\}_{i\in\mathbb{N}_0}$.

It remains to show that under these circumstances iteratively deleting remaining bad nodes until none are left will a.a.s. terminate after $o(n)$ steps.

**Lemma 6.2.9** *Suppose some graph $G$ is distributed uniformly amongst all graphs with degree sequence $\{d_i(G)\}_{i\in\mathbb{N}_0}$, and that this degree sequence is $\varepsilon$-close to $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$, for some $\varepsilon = \varepsilon(n) = O(n/\log(n)^{\alpha(c)M(c)})$.*

*Then a.a.s. we will be left with a linear sized $k$-core after $o(n)$ applications of the single node deletion process. Moreover, the degree sequence of the $k$-core found will be $\varepsilon'$-close to $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$, for some $\varepsilon' = \varepsilon'(n) = o(n)$*

Before proving this essential lemma we shall start a little excursion trying to point out that it states merely the random graph analogue to the damage process in trees discussed above (Section 3.3.2). We shall also give an impression how the branching tree ideas relate to the various models available, that is $\mathcal{G}_{n,p}$, $\mathcal{G}_{n,m}$ and random configurations. Any reader not interested in top-level, semi-intuitive ramifications is encouraged to skip this and to continue with the proof of Lemma 6.2.9 on pp. 121.

The generic run $v$ has at most 2 neighbours (again restricting ourselves to $k = 3$). Using the 'test-particle approach' (see Section 5.2, Lemma 5.2.2) it is not hard to show that in the $\mathcal{G}_{n,p}$ model the number of neighbours of $v$ having degree exactly three is distributed as $\mathrm{Bi}(2,\pi)$, up to $o(1)$ corrections; we shall write $\mathrm{Bi}(2,\pi) + o(1)$. Remember that $\pi = 1/2 - \varepsilon(c)$ is the percentage of the binary owners that are adjacent to exactly two binary owners. When we remove $v$ from the set of $r$-Cayley-owners we shall thus produce $\mathrm{Bi}(2,\pi) + o(1)$ new bad nodes. If we could iterate this argument for the

newly created runs, recursively, the number $Y_t$ of bad nodes would be clearly captured by a gambler's ruin process with drift $-\varepsilon(c)/2$, starting at some $Y_0 = \theta(n/\log(n)^{\alpha(c)M(c)})$. Such a gambler's ruin process is well known to a.s. die out after $o(n)$ steps, c.f. the proof of Lemma 6.2.9 below.

We have amply demonstrated above how to translate calculations valid in branching trees to the $\mathcal{G}_{n,p}$ model, taking into account all those annoying errors that do not matter in the end. In particular we can show that the *first* run to be deleted will spark off $\mathrm{Bi}(2,\pi)+o(1)$ new bad nodes, as sketched in the last paragraph. With considerable more effort we shall show below that this is also true in the $\mathcal{G}_{n,m}$ model. So why did not we stay in the $\mathcal{G}_{n,p}$ model, where we have our machinery working?

- It is not sufficient to merely know the degree sequences of the subgraphs involved, as encoded in Lemma 5.3.2. Otherwise an adversary could 'pile up' many ($\theta(n)$) of the $\theta(n)$ nodes with degree exactly 3 around a single run. In that deterministic worst case the single node process would quite clearly *not stop* after $o(n)$ steps. Note that - at least intuitively - assuming uniform distribution amongst all graphs with the given degree sequence would turn such a worst case scenario highly improbable. In the $\mathcal{G}_{n,p}$ model it may be intuitively clear that all subgraphs with the given degree sequence are equally likely, but we lack a formal proof. Thus we have to work in the $\mathcal{G}_{n,m}$ model, where we have such a proof, see Section 6.2.1, Corollary 6.2.3. However, for the first (few) run(s) we could still prove that the number of newly created bad nodes is distributed according to $\mathrm{Bi}(2,\pi)+o(1)$) in the $\mathcal{G}_{n,p}$ model, as explained above.

- We also need to know that the uniform distribution (conditional on the degree sequence) remains invariant in the course of successive single node deletion steps. Otherwise we would not know that the number of newly created bad nodes is $\mathrm{Bi}(2,\pi)+o(n)$, each time afresh, and we could not describe the evolution of $Y_t$ by a gambler's ruin. Obviously 'having looked at' only a few nodes in the $\mathcal{G}_{n,p}$ model does not matter a lot - but it *does* matter. To see this, consider the 'remainder graph' at the last but one step of the single node deletion process. It will deterministically not have any bad nodes at all. Therefore it can not be distributed as a $\mathcal{G}_{n-o(n),p}$ random graph, because such a graph would have to have $\theta(n/\log(n)^{\alpha(c)M(c)})$ runs (w.r.t. the remainder graph!) with high probability, according to Lemma 5.2.7. This seemingly paradox result can be intuitively explained by observing that runs w.r.t. the remainder graph are not necessarily runs with respect to the total graph.

Thus the principle of deferred decisions does not apply when we consider more than the first run, and we can presently not control the complicated dependencies in the $\mathcal{G}_{n,p}$ model.

For those reasons we have translated our results achieved in $\mathcal{G}_{n,p}$ to $\mathcal{G}_{n,m}$ where we can adapt the proof strategy of G/M. Note that it would be highly desirable to devise a proof entirely within the $\mathcal{G}_{n,p}$ model, avoiding the analysis of deletion processes altogether. We shall continue this discussion in Chapter 7.

It is thus essential to translate the 'test-particle' approach to the $\mathcal{G}_{n,m}$ model. Goerdt and Molloy, working in the model of random faulty configurations, have demonstrated how to go about it, we shall briefly describe the central formula and how it 'miraculously' yields the same result as the 'test-particle' approach in the $\mathcal{G}_{n,p}$ model. The equivalent of the 'test-particle' approach in random configurations is not too hard to show, essentially as we may employ random matchings, see the characterisation of random configurations in Section 2.1.1, p. 13. However, we do not work with random configurations but with random graphs (both distributed uniformly under all objects with a given degree sequence). It is standard practice to prove results in the random graph model by proving them in the 'corresponding' configuration model. We have alluded to the 'translation tools' available in Section 2.1.1. It appears to be difficult (although the result is intuitively obvious) to translate Goerdt and Molloy's result concerning the change in the number of bad nodes $Y_t$ from the configuration model to the 'corresponding' random graph model *step by step*. We consider the following strategy feasible but tedious: Analyse the gambler's ruin in the 'corresponding' configuration model and translate the bad event 'gambler's ruin does not terminate after $o(n)$ steps' back into the random graph model.

Instead we have found a direct proof for random ($\mathcal{G}_{n,m}$) graphs based on a switching argument.

Before we end our excursion we shall explain how the damage process in branching trees translates to the different models, $\mathcal{G}_{n,p}$ on one side and $\mathcal{G}_{n,m}$/random configurations on the other side, and why the seemingly different central formulae yield the same value $\mathrm{Bi}(2,\pi)+o(1)$.

$\mathcal{G}_{n,p}$ Assume that the degree sequence of the *binary owners* in the remainder graph is ($\varepsilon$-close to) $\beta = \{b_i\}_{i=0}^{\infty}$, where $b_1 \simeq 0$ and $b_0 = n - \sum_{j=1}^{\infty} b_j$. Clearly, attaching a test-particle with one edge to the binary owners in the remainder graph will hit a binary owner with degree *exactly* two (w.r.t. the remainder graph!) with probability

$$\simeq \frac{b_2}{\sum_{i=2}^{\infty} b_i}.$$

$\mathcal{G}_{n,m}$ We shall *here* discuss only the configuration model. In the proof of Lemma 6.2.9 below we shall prove that the same formula holds for the 'corresponding' graphs, too. Assume that the degree sequence of the *Cayley-owners* is $\varepsilon$-close to $\delta = \{d_i\}_{i=0}^{\infty}$, where $d_1, d_2 \simeq 0$ and $d_0 = n - \sum_{j=1}^{\infty} d_j$. Essentially we pick a node $v$ with degree 1 and discern cases according to what kind of class the the edge emanating from $v$ ends. Obviously, to hit a class of degree exactly 3 (w.r.t. the set of Cayley owners altogether!) is proportional to both $d_i$ and $i$, yielding the formula

$$\simeq \frac{3 \cdot d_3}{\sum_{i=3}^{\infty} i \cdot d_i}. \tag{6.4}$$

Similar results hold when the test-particle has degree greater than one.

These formulae appear to be different. But note that in the case of Cayley- and binary-owners we have $b_i \simeq n \cdot \mathrm{Po}_{cq}(i)$, for $i \geq 2$ and $d_i \simeq n \cdot \mathrm{Po}_{cq}(i)$, for $i \geq 3$. Observe that

$$\frac{b_2}{\sum_{i=2}^{\infty} b_i} \simeq \frac{\mathrm{Po}_{cq}(2)}{\sum_{i=2}^{\infty} \mathrm{Po}_{cq}(i)} = \pi(c)$$

and that

$$\frac{3 \cdot d_3}{\sum_{i=3}^{\infty} i \cdot d_i} \simeq \frac{3 \cdot \mathrm{Po}_{cq}(3)}{\sum_{i=3}^{\infty} i \cdot \mathrm{Po}_{cq}(i)} = \pi(c),$$

too. It is straightforward to compute that those formulae yield the same value $\pi(c)$ using the characterisation $q = \mathbb{P}[\mathrm{Po}_{cq} \geq 2]$. However, this equality seems to reflect some kind of 'equivalence' of dealing with Cayley-owners and dealing with binary-owners that we do not understand well enough. Note that we *can* fully analyse the sudden appearance of the 'classical' 3-core (essentially the set of Cayley-owners) adapting the proof strategy of Goerdt and Molloy whereas we *can not* fully analyse the sudden appearance of the extended 3-core (see Definition 2.3.3 on p. 28, essentially the equivalent to the binary-owners). Yet we can prove the 'appearance of a giant set of binary-owners' (see Section 3.1) which relate to the extended 3-core in a seemingly completely analogous way as the Cayley owners relate to the classical 3-core. This apparent analogy is further backed by simulations using the respective deletion processes, strongly suggesting that the size of the extended 3-core is 'obeying' the Branching Tree Connection in the very same way as the size of the 'classical' 3-core does. Note that the much sought for proof strategy avoiding the analysis of the deletion process altogether, as mentioned above, could most likely be adapted to both situations, easily.

**Proof** (Of Lemma 6.2.9.) We shall show further down in this proof that at all 'relevant times' $t$ the increase $Y_t - Y_{t-1}$ in the number of bad nodes is

stochastically dominated by a $\mathrm{Bi}(2, 1/2 - \kappa(c)) - 1$ random variable, initially assuming uniform distribution under a degree sequence $\delta_0$ that is $\varepsilon$-close to the degree sequence expected from the Branching Tree Connection, for some $\varepsilon = \varepsilon(n) = K_1 \cdot n/\log(n)^{\alpha(c)M(c)}$, c.f. Corollary 6.2.8.

Note that the conditional uniform distribution remains invariant due to Corollary 6.2.6, p. 115.

By 'relevant times' $t \leq T$ we mean that we stop the processes $Y$ and $X$ at time $T$. In other words we shall show that the random process $(Y_t)_t$ is stochastically dominated by a gambler's ruin process $(X_t)_t$, having increments distributed according to $X_t - X_{t-1} \sim \mathrm{Bi}(2, 1/2 - \kappa(c)) - 1$. The gambler's ruin process is well known to a.s die out quickly (see Feller [Fel68]). More formally, we know that the expected hitting time of the origin 0 is proportional to $X_0/\kappa$ ([Fel68], XIV). Since $X_0 = Y_0 = K_2 \cdot n/\log(n)^{\alpha(c)M(c)}$, we know that the hitting time of the origin 0 is at most, say, $T := K_2 \cdot n/\log(n)^{\alpha(c)M(c)} \cdot \log n$, a.a.s., using Markov's Inequality.

Note that at each time $t \leq T$, i.e. after having deleted $t \leq T$ nodes, the degree sequence of the remainder graph will still be $K_3 \cdot n/\log(n)^{\alpha(c)M'(c)}$-close to to the degree sequence expected from the Branching Tree Connection, which follows from an argument similar to the argument in the proof of Lemma 6.2.7, using Inequality 6.3 (see p. 116).

We next turn to proving the aforementioned stochastic dominance. Unfortunately it is necessary to discuss several cases. However, it will turn out that in all cases the number of newly created bad nodes is stochastically dominated by $\mathrm{Bi}(2, 1/2 - \kappa(c))$. The generic case is the following:

a) The (bad) node $v$ to be removed has degree 2 and it is adjacent to two other good nodes.

Apart from this there are a few 'atypical' cases.

b) The (bad) node $v$ to be removed has degree 2 and it is adjacent to one good node and one bad node.

c) The (bad) node $v$ to be removed has degree 2 and it is adjacent to two bad nodes.

d) The (bad) node $v$ to be removed has degree 1 and it is adjacent to a good node.

e) The (bad) node $v$ to be removed has degree 1 and it is adjacent to a bad node.

Note that a run that has only one neighbour left in the set of Cayley-owners will occur much less frequently than the 'generic run' with two neighbours. It should be clear, that case a) is indeed not only the generic case but the worst case. Only in case a) it is possible to create more bad nodes than we remove. In cases c) and e) the increase in bad nodes is -1, deterministically, and cases b) and d) are essentially the same.

We shall start with discussing case d), explaining the switching argument, as we go along. This is a rigorous version of the informal discussion on p. 121. There are two possibilities. Either the bad node $v$ is adjacent to a good node with degree exactly three, or it is adjacent to a good node with degree greater than three. We thus expect the number of newly created bad nodes to be distributed according to $\mathrm{Bi}(1,\pi)+o(1)$. Now we need to prove this formally. We should really discern cases for all pairs $(i,j)$ that the bad node $v$ to be removed is node number $i$ and that its only neighbour is node number $j$. We shall merely discuss the case that $i=1$ and $j=2$. This is sufficient because all cases are equivalent due to symmetry against relabelling nodes. Assuming that the original graph had degree-sequence $\delta$, we shall denote by $\delta_{1,i}$ the degree sequence arising from $\delta$ upon 'moving' one node with degree 1 (i.e. our node number 1) to the set of nodes with degree 0 and one node with degree $i$ (i.e. our node number 2) to the set of nodes with degree $i-1$. Denote by $\mathcal{G}^{\delta}$ the set of all graphs having degree sequence $\delta$, where additionally node 1 is adjacent to node 2. By $\mathcal{G}^{\delta_{1,i}}$ we shall denote all graphs in $\mathcal{G}^{\delta}$ where node 2 has degree $i-1$ and node 1 has degree 0. Clearly.

$$\left|\mathcal{G}^{\delta}\right| = \sum_{i\geq 1} \left|\mathcal{G}^{\delta_{1,i}}\right|.$$

Now the number of newly created bad nodes may either be 0 or 1. We shall compute the probability that it is 1. Since we have uniform distribution on all graphs in $\mathcal{G}^{\delta}$, we have

$$\mathbb{P}\left[\text{one new bad node created}\right] = \frac{\left|\mathcal{G}^{\delta_{1,3}}\right|}{\sum_{i\geq 1}\left|\mathcal{G}^{\delta_{1,i}}\right|}. \tag{6.5}$$

Our goal is to show that this probability is close to the value expected in Formula 6.4. In order to see this, we shall use a switching argument. We shall show that

$$\frac{\left|\mathcal{G}^{\delta_{1,i}}\right|}{\left|\mathcal{G}^{\delta_{1,3}}\right|} \simeq \frac{i\cdot d_i}{3\cdot d_3}. \tag{6.6}$$

This will clearly lead towards Formula 6.4, when substituting this in Equation 6.5, as the $d_i$ are close to the values expected from the Branching Tree Connection. How does one prove a statement like Equation 6.6? Consider a bipartite graph between some sets $A$ and $B$ (we shall choose $A := \mathcal{G}^{\delta_{1,i}}$ and $B := \mathcal{G}^{\delta_{1,3}}$ presently). We can count the number of edges in the bipartite graph in two ways:

$$\sum_{v\in A}\deg(v) = \sum_{w\in B}\deg(w).$$

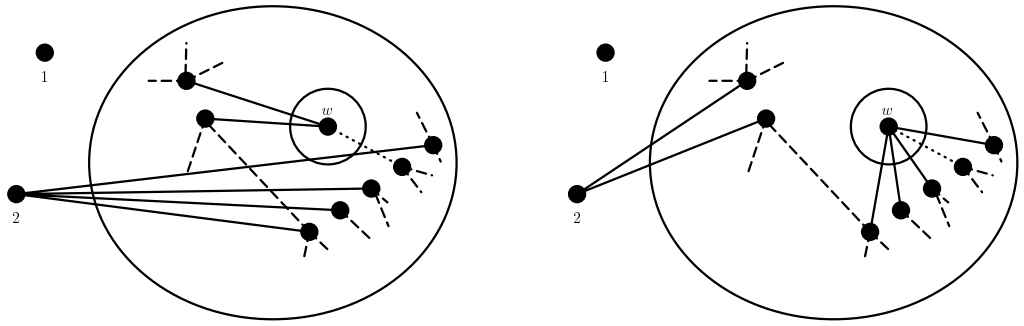


Figure 6.1: Switching between a graph in $A := \mathcal{G}^{\delta_{1,5}}$ (left) and a graph in $B := \mathcal{G}^{\delta_{1,3}}$ (right). The edges drawn with full lines are the ones to be switched. The dotted edge is the one that is selected to remain incident to $w$. Further edges that may be present are drawn with dashed lines.

Note that when all nodes in $A$ have degree $b$ and all nodes in $B$ have degree $a$ this reduces to

$$|A| \cdot b = |B| \cdot a \Leftrightarrow \frac{|A|}{|B|} = \frac{a}{b}.$$

Even when the degrees are not exactly equal to $a$ and $b$, respectively, we may produce approximate formulae along those lines. The difficult part is to construct an appropriate bipartite graph, i.e. to appropriately pair the elements of $A$ with elements of $B$. We shall from now on refer to the bipartite graph as a 'meta-graph', simply to avoid confusion due to the fact that the elements of $A := \mathcal{G}^{\delta_{1,i}}$ and $B := \mathcal{G}^{\delta_{1,3}}$ are graphs themselves.

We consider a meta-node $G \in A$ and determine its meta-neighbours in $B$. Remember that node 2 has degree $a-1$ in $G$. We shall also select a node $w$ with degree 3. Now we shall 'switch' the edges incident to node 2 with the edges incident to node $w$ in the following way, see Figure 6.1. One of the edges incident to $w$ remains, the other edges get attached to node 2. All edges that were originally incident to node 2 get attached to node $w$. Note that there are usually 3 ways to switch, depending on how the remaining edge incident to $w$ is selected. We shall not allow a switching in (the rare) case we introduce self-loops or multiple edges, or do not get a graph in $B$ otherwise. Conversely we consider a meta-node $G \in B$ and determine its meta-neighbours correspondingly. Note that we have actually defined a simple bipartite (meta-)graph this way, because the fact that each allowed switching can be reversed proves that each (a priori directed) meta-edge from $A$ to $B$ has a corresponding 'partner' from $B$ to $A$. So what do we know about

the degrees in the bipartite meta graphs? At first sight it seems that each graph in $A := \mathcal{G}^{\delta_{1,i}}$ has indeed exactly $3 \cdot d_3$ meta-neighbours in $B := \mathcal{G}^{\delta_{1,3}}$ and each graph in $B$ has exactly $i$ meta-neighbours in $A$, but this is not quite true. If we switch with nodes $w$ in the 2-neighbourhood of node 2 we might not end up with a graph in $A$, or $B$, respectively. Fortunately we know that all graphs considered are degree-bounded by $\Delta(G) = (\log n)^2$. In the worst case we do not know what will happen for $(\log n)^4$ candidates $w$. Thus the meta-degrees are only determined up to an error in the following way.

$$\deg(G) = 3 \cdot (d_3 \pm (\log n)^4), \text{ for } G \in A,$$

and

$$\deg(G) = i \cdot (d_i \pm (\log n)^4), \text{ for } G \in B.$$

Alas, the $d_i$ are only known up to additive errors of $\pm K_3 \cdot n/\log(n)^{\alpha(c)M'(c)}$, anyway. We may thus replace those errors by $\pm K_4 \cdot n/\log(n)^{\alpha(c)M'(c)}$ and ignore the $\pm(\log n)^4)$ in the following. Now the number of edges in the meta-graph may be written as $\sum_{G \in A} \deg(G) = \sum_{G' \in B} \deg(G')$. Substituting the degree sequence (within the aforementioned additive errors of size $\tilde{\varepsilon} := K_4 \cdot n/\log(n)^{\alpha(c)M'(c)}$) yields

$$|A| \cdot 3 \cdot (n \cdot Po_{cq}(3) \pm \tilde{\varepsilon}) = |B| \cdot i \cdot (n \cdot i \cdot Po_{cq}(i) \pm \tilde{\varepsilon}),$$

for $i \geq 4$,

$$|A| \cdot 3 \cdot (n \cdot Po_{cq}(3) \pm \tilde{\varepsilon}) = |B| \cdot 2 \cdot (0 \pm \tilde{\varepsilon}),$$

for $i = 2$,

$$|A| \cdot 3 \cdot (n \cdot Po_{cq}(3) \pm \tilde{\varepsilon}) = |B| \cdot 1 \cdot (0 \pm \tilde{\varepsilon}),$$

for $i = 1$. Therefore,

$$\begin{aligned}
&\mathbb{P}[\text{one new bad node created}] = \frac{\left|\mathcal{G}^{\delta_{1,3}}\right|}{\sum_{i \geq 1} \left|\mathcal{G}^{\delta_{1,i}}\right|} \\
&= \frac{1}{\sum_{i \geq 1} \frac{\left|\mathcal{G}^{\delta_{1,i}}\right|}{\left|\mathcal{G}^{\delta_{1,3}}\right|}} = \frac{1}{\frac{\left|\mathcal{G}^{\delta_{1,1}}\right|}{\left|\mathcal{G}^{\delta_{1,3}}\right|} + \frac{\left|\mathcal{G}^{\delta_{1,2}}\right|}{\left|\mathcal{G}^{\delta_{1,3}}\right|} + 1 + \sum_{i \geq 4} \frac{\left|\mathcal{G}^{\delta_{1,i}}\right|}{\left|\mathcal{G}^{\delta_{1,3}}\right|}} \\
&= \frac{3 \cdot Po_{cq}(3) \pm \tilde{\varepsilon}/n}{\left(\sum_{i \geq 3} i \cdot Po_{cq}(i) \pm \tilde{\varepsilon}/n\right) \pm O(\tilde{\varepsilon}/n)} = \dots
\end{aligned}$$

Due to the degree bound the summation over the $i$ extends only to $\Delta(G) = (\log n)^2$ and we may continue

$$\begin{aligned}
&= \frac{3 \cdot Po_{cq}(3) \pm \tilde{\varepsilon}/n}{\sum_{i=3}^{\infty} i \cdot Po_{cq}(i) - \left(\sum_{(\log(n))^2+1}^{\infty} i \cdot Po_{cq}(i)\right) \pm \left(\sum_{i=3}^{(\log(n))^2} \tilde{\varepsilon}/n\right) \pm O(\tilde{\varepsilon}/n)} \\
&= \frac{3 \cdot Po_{cq}(3) \pm \tilde{\varepsilon}/n}{\sum_{i=3}^{\infty} i \cdot Po_{cq}(i) \pm O((\log n)^3 \cdot \tilde{\varepsilon}/n)} = \pi(c) + o(1).
\end{aligned}$$

Note that

$$\sum_{(\log(n))^2+1}^{\infty} i \cdot Po_{cq}(i) = \sum_{(\log(n))^2+1}^{\infty} i \cdot e^{-cq}\frac{(cq)^i}{i!} = cq \sum_{(\log(n))^2}^{\infty} e^{-cq}\frac{(cq)^i}{i!},$$

which is subpolynomially small by Lemma 5.1.1. Also,

$$\sum_{i=3}^{(\log(n))^2} \tilde{\varepsilon}/n \leq \log(n)^3 \cdot \tilde{\varepsilon}/n = o(1),$$

for sufficiently large $\alpha = \alpha(c)$.

We have lost track on the rate of convergence on the way, and a careless use of Landau symbols is potentially dangerous. Fortunately, in this proof there are only a constant (independent of $n$) number of different rates of convergence to be considered, so we can take the slowest amongst them as a worst cat rate.

We have now rigorously discussed case d), remember that for cases c) and e) there is nothing to be done.

The cases a) and b) can be treated in an analogous way, but we shall not discuss them as detailed as case d). First of all, we exclude all cases where nodes 2 and 3 have distance less than 3 in the graph with node 1 deleted, this will lead to a $o(1)$ correction, even if we make the worst case assumption that in this case we have always produced two bad nodes in this case.

We shall merely sketch a proof for the sub-case of a) that node 1 is adjacent to nodes 2 and 3, both of them being good. We need to prove that

$$\begin{aligned}&\mathbb{P}\left[\text{two new bad nodes created}\right]\\ &= \frac{\left|\mathcal{G}^{\delta_{1,3,3}}\right|}{\sum_{\substack{i\geq 1\\ j\geq 1}} \left|\mathcal{G}^{\delta_{1,i,j}}\right|} \stackrel{!}{=} \frac{(3 \cdot \mathrm{Po}_{cq}(3))^2}{\sum_{i,j=3}^{\infty} i \cdot j \cdot \mathrm{Po}_{cq}(i) \cdot \mathrm{Po}_{cq}(i)} + o(1) = \pi(c)^2 + o(1)\end{aligned}$$

This can be shown using

$$\frac{\left|\mathcal{G}^{\delta_{1,i,j}}\right|}{\left|\mathcal{G}^{\delta_{1,3,3}}\right|} = \frac{i \cdot j \cdot d_i \cdot d_j}{9 \cdot d_3^2} + \text{ corrections as in d)},$$

which follows from a switching argument as in case d). Essentially, choose two nodes with degrees $i$, $j$ sufficiently far apart from each other and from nodes 2 and 3, and switch all but one of their incident edges with the edges incident to nodes 2 and 3, respectively. □

## 6.3 Sudden Appearance of the $k$-Core, New Proof

We are finally ready to state and prove the Main Theorem 1, giving the new proof of the sudden appearance of a giant $k$-core in $\mathcal{G}_{n,m}$-graphs with average degree $c$.

**Theorem 1** *Let $G$ be distributed according to the $\mathcal{G}_{n,m}$ model with average degree $c$. Denote by $c_{crit}$ the critical value for ownership in branching trees, which is $3.35(\ldots)$ for $k = 3$.*

1. *For sub-critical average degrees $c < c_{crit}$ the $k$-core is empty, with probability approaching one.*

2. *For super-critical average degrees $c > c_{crit}$ there is a giant $k$-core with probability approaching one. It consists of the set of $r(n)$-Cayley owners of $G$ minus $o(n)$ nodes. Moreover, the degree sequence of the $k$-core found will be $\varepsilon'$-close to $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$, for some $\varepsilon' = \varepsilon'(n) = o(n)$.*

*Here $\{d_i^{(0)}(G)\}_{i\in\mathbb{N}_0}$ is the degree sequence 'predicted' by the Branching Tree Connection, see Equation 5.2, p. 103.*

**Proof** (Of the Main Theorem 1.)
1. Inequality 6.2 (p. 116) states that the set $\mathcal{S}$ surviving $(r-1)$ steps of the shell-wise deletion process is contained in the set $\mathcal{O} \cup \overline{\mathcal{T}}$, where $\mathcal{O}$ are the $r$-owners and $\overline{\mathcal{T}}$ are the nodes with an $r$-neighbourhood that is *not a tree*. Lemma 5.3.2 and Proposition 5.3.5 imply that $|\mathcal{S}| = o(n)$. Clearly, $\mathcal{S}$ is a super-set of the $k$-core, since the shell wise-deletion process is a specific deletion protocol stopped (shortly, one should think) before it has found the $k$-core. Now Lemma 6.1.1 shows that there exists no subgraph with minimum degree ($\geq 3$) containing $o(n)$ nodes, with probability approaching one, and thus there there exists no $k$-core, with probability approaching one.

2. The supercritical case is a straightforward consequence of Corollary 6.2.8 on p. 118, stating that the degree sequence after $(r-1)$ rounds of shell-wise deletion is right, Corollary 6.2.3 on p. 114, implying that we have uniform distribution under this degree sequence, and Lemma 6.2.9 on p. 118 proving that the subsequent single node degree sequence will die out after at $o(n)$ steps.

□

# Chapter 7

# Loose Threads

In this chapter we discuss various results and insights we have collected during our studies. The material is incomplete and has (apart from Section 7.1) a heuristic, or even speculative flavour. Still, it may reveal insights valuable to anybody pursuing further studies in this field.

We have given a new proof for the sudden appearance of a giant $k$-core avoiding most of the tedious step by step analysis of the deletion process and discovered a new type of giant subgraph, the magic subgraph, both motivated by the Branching Tree Connection. The reader may ask why we have not included a proof for the sudden appearance of extended $k$-core or magic subgraph using our new techniques. Section 7.1 explains to what extent the proofs can be adapted to directed subgraphs such as the extended $k$-core or the magic subgraph and why we feel unable to generalise the proof for the invariance of uniform distribution which is essential for the analysis of the second phase, i.e. for removing the remaining runs.

In Section 7.2 we explain why we could not completely avoid analysis of the deletion processes characterising $k$-core and related subgraphs.

Even though the tripartite case deserves attention independently of non-tripartite models it is natural to ask whether there is some phenomenon similar to the sudden appearance of a giant magic subgraph in non-tripartite graphs. In particular, what can be learned about the $k$-colourability threshold? We shall discuss this in Section 7.3.

Finally, in Section 7.4 we present some further recursive equations predicting critical values in branching trees. They may be a starting point for finding new giant subgraphs in the 'corresponding' random graphs, just as the recursive equations connected to coloured owners (see Section 3.2) eventually lead to 'discovering' the magic subgraph. We also include such a recursive equation possibly related to random $k$-SAT formulae.

## 7.1 Sudden Appearance of Extended $k$-Core and Magic Subgraph

Note that the new proof for the $k$-core is conceptually simple but writing it down in a rigorous way involves a lot of technical effort mainly due to having to carefully discuss the errors terms appearing in the calculations due to finite $n$.

Similarly as in branching trees, ownership in graphs is defined recursively, and the numbers of owners are concentrated by the Semi-Local Lemma 5.1.5. We may thus compute the expected number of owners in graphs essentially using the very same recursive equations that are describing and explaining the 'corresponding' phase transitions in branching trees.

It is natural to seek for generalisations of the existing proof to the appearance of the extended $k$-core and of the magic subgraph, as well as the 'classical' $k$-core in the $k$-partite model, sticking closely to analysing the two phases (shell-wise deletion followed by single node deletion) which turned out successful for the 'classical' $k$-core.

### Sources of errors to be taken into account

The aforementioned technical effort arises for two reasons.

In order to show that the errors do not matter asymptotically we need to write them down, and to discuss them carefully. Some errors stem from the deviations between $\mathrm{Po}_c$ and $\mathrm{Bi}(n, c/n)$ distributions. Sizes of certain subsets are concentrated around their expected values but not *equal* to the expected values, introducing a further source of errors.

Because we need to analyse the single node deletion process in the second phase and show that it behaves like a gambler's ruin random walk, we need to translate the results achieved for the $\mathcal{G}_{n,p}$ model to the $\mathcal{G}_{n,m}$ model because only there we can analyse the deletion process. This is an additional source of errors and forces us to go through the whole 'translation machinery'.

We shall report below in Section 7.2 on some ideas how having to translate from $\mathcal{G}_{n,p}$ to $\mathcal{G}_{n,m}$ could possibly be avoided.

### Modifications necessary for $k$-partite models

When re-inspecting the proofs in Chapters 5 and 6 it seems highly plausible to believe that comparable results may also be shown for the 'corresponding' $k$-partite models. Here is a list of changes for $k = 3$.

- $\mathrm{Bi}(2/3 \cdot n, 3/2 \cdot c/n)$ will have to be approximated by $\mathrm{Po}_c$, instead of $\mathrm{Bi}(n, c/n)$.

- Vertex exposure will have to be reformulated for the tripartite case, and the proof for the Semi-Local Lemma 5.1.5 needs to be adapted accordingly.
- We need to translate from $\mathcal{G}_{n,3,p}$ to $\mathcal{G}_{n,3,m}$. There is no doubt that there are 'translation tools' closely analogous to the ones for translating from $\mathcal{G}_{n,p}$ to $\mathcal{G}_{n,m}$. However, they would need to be written down.

**Conjecture 1** *All statements concerning the (generalised, see below) degree-sequence of the extended and 'classical' $k$-core that we proved (or can be proved) for $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ with average degree $c$ are also valid for the tripartite models $\mathcal{G}_{n,3,p}/\mathcal{G}_{n,3,m}$ with average degree $c$, with possibly slightly modified error terms.*

### Generalised degree sequence

What should be the analogue of the degree sequence, when talking about directed edge deletion processes? There are certainly several possibilities, we have singled out a choice that appeared amongst the most sensible. Yet we were not able to fully generalise the analysis of the invariance of uniform distribution during the course of the deletion processes (see below). Hence we shall only sketch what seems to be feasible to show. We shall mainly discuss the scenario of the extended $k$-core (Definition 2.3.3, p. 28) since we expect a further generalisation to the magic subgraph to be relatively easy - the hard bit seems to be analysing the *directed-edge* deletion process properly.

Remember that the extended $k$-core (in the node sense) consists of all nodes 'pointed to' by at least two 'good' directed edges. Intuitively, and consistently with our empirical observation, this is essentially the set of binary owners.

**Definition 7.1.1** *The* generalised degree sequence *is a family $\{d_I\}_{I\in\mathcal{I}}$ of non-negative numbers labelled by an index $I = (i, j, k)$. The numbers $d_I$ count the number of nodes in some digraph with a specific* 1*-neighbourhood, having $i$ in/out-edges, $j$ 'pure' in-edges and $k$ 'pure' out-edges.*

*Instead of regarding $I$ as a triple it is more convenient to interpret it as a 'picture' of the (undirected)* 1*-neighbourhood of a node in a digraph, ignoring the labels.*

*There is a natural partial ordering on $\mathcal{I}$, namely $J \succeq I$ whenever $I$ can be obtained from $J$ by deleting some directed edges.*

*Finally we can define the* integrated generalised degree sequence $\{D_I\}_{I\in\mathcal{I}}$ *by*

$$D_I := \sum_{J\succeq I} d_J.$$

Note that $\{d_I\}_{I\in\mathcal{I}}$ is related to $\{D_I\}_{I\in\mathcal{I}}$ in the same way as multivariate (discrete) probability-distributions are related to the corresponding multivariate distribution functions.

The above definition satisfies one important necessary 'constraint', that all other alternative definitions should also comply to. When we choose an undirected random graph initially and make it bi-directed, it will be distributed uniformly amongst all digraphs with a certain generalised degree sequence, where the only non-zero $d_i$ are labelled by indices $I$ of the type $(d_i, 0, 0)$, $d_i$ being the 'normal' degree sequence of the undirected graph. This will provide the basis of any inductive proof concerning uniform distribution conditional on a generalised degree sequence.

Next we have to specify the deletion protocol. A node will be bad, if there are out-edges bad in its 1-neighbourhood. Note that when the generalised degree sequence $\{d_I\}_{I\in\mathcal{I}}$ is such that there are no more bad nodes, it is the degree sequence of an extended $k$-core, possibly empty. We shall 'purge' the bad edges in a 'node-wise' fashion. Whenever a node is bad, remove all emanating out-edges that are bad. For the shell-wise deletion process, we will check on all nodes presently bad and determine which edges would have to be deleted from their out-neighbourhoods, in 'single-node mode'. Then we remove the union of all those edges.

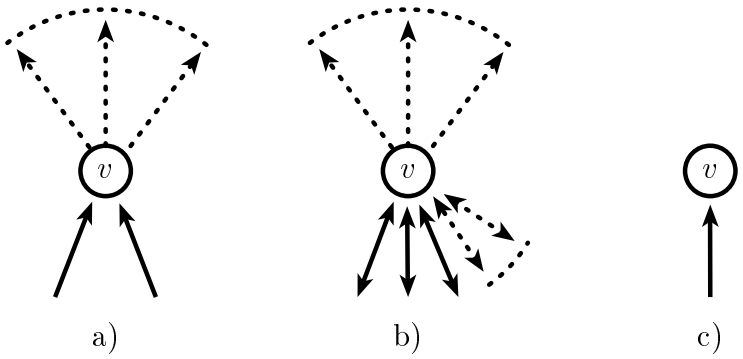


Figure 7.1: All cases of possible 'pictures' $I$ that can occur (i.e. have non-zero $d_i$) in the generalised degree sequence $\{d_I\}_{I\in\mathcal{I}}$ of the extended 3-core, apart from the empty 'picture'. See text for a discussion.

Can we, like in the case of the 'classical' $k$-core predict the degree-sequence after $r-1$ rounds of shell-wise deletion? We believe that the answer is yes, but we give only a hand-waving derivation. Remember the results from

Chapter 5, i.e. the 'test-particle' approach. There are binary owners, Cayley owners and also *unary owners*, the latter are nodes adjacent to exactly one binary owner. Whenever a 'test-particle' $v$ is attached to the remainder graph it will be connected to $\mathrm{Po}_{cu}$ unary owners, where

$$u := \mathbb{P}\left[\mathrm{Po}_{cq} = 1\right],$$

and to $\mathrm{Po}_{cq}$ binary owners. We have portrayed all cases that can happen in Figure 7.1 and will calculate the expected $d_I$ below.
a) 'Picture' $I = (0, 2, k)$, $k \geq 0$ has probability $\mathbb{P}[\mathrm{Po}_{cq} = 2] \cdot \mathbb{P}[\mathrm{Po}_{cu} = k]$.
b) 'Picture' $I = (i, 0, k)$, $i \geq 3$, $k \geq 0$ has probability $\mathbb{P}[\mathrm{Po}_{cq} = i] \cdot \mathbb{P}[\mathrm{Po}_{cu} = k]$.
c) 'Picture' $I = (0, 1, 0)$ has probability $\mathbb{P}[\mathrm{Po}_{cq} = 1] (= u)$.

**Conjecture 2** *The generalised degree sequence of the digraph obtained after $r - 1$ rounds of shell-wise edge deletion is $\varepsilon -$ close to the generalised degree sequence obtained from the 'branching tree' probabilities just calculated.*

*Moreover, in the case of the magic subgraph an analogous statement, concerning an even more complicated but straightforward generalisation* of the degree sequence, holds true.*

### Analysing the deletion process in directed graphs is difficult

We now turn to explaining why we failed in generalising the proof strategy of Goerdt and Molloy to the *directed* subgraphs extended $k$-core and magic subgraph.

We remind the reader of the essential ideas that worked for the 'classical' $k$-core. In the $\mathcal{G}_{n,m}$ model it is obvious that conditional on some specific degree sequence the graph $G$ will be distributed uniformly amongst all graphs with this specific degree sequence. In order to show that this conditional degree sequence remains invariant in the course of both shell-wise and single-node deletion process, we had to essentially show the following.

Given two degree sequences $\delta$ ('before') and $\delta'$ ('after'), count the pre-images of each graph† $G'$ under the action of the deletion process $\Delta$ and show that their sizes are independent of the specific choice of $G'$ amongst all graphs with degree sequence $\delta'$. More specifically, we showed that for any two graphs $G'_1$ and $G'_2$ we can explicitely state a bijection between the pre-images $\Delta^{-1}[G'_1]$ and $\Delta^{-1}[G'_2]$. This bijection was induced by a relabelling $\pi_{G'_1, G'_2}$ of the nodes preserving $\delta'$. Whenever $G'_1 \uplus H$ is a predecessor of $G'_1$,

*Re-reading the discussion on p. 63 may help when wondering what the typical extended degree sequence should look like in the coloured case.

†Even when considering digraphs we shall use the letter '$G$'.

$G_2' \uplus \pi[H]$ is a predecessor of $G_2'$. Recall that we demanded that both $\delta$ and $\delta'$ were relevant (i.e. they may occur in the course of some deletion process) and consistent (i.e. the set of predecessors is not empty). When discussing generalised degree sequences, relevance implies that certain pictures that may not appear due to the way the deletion process works.

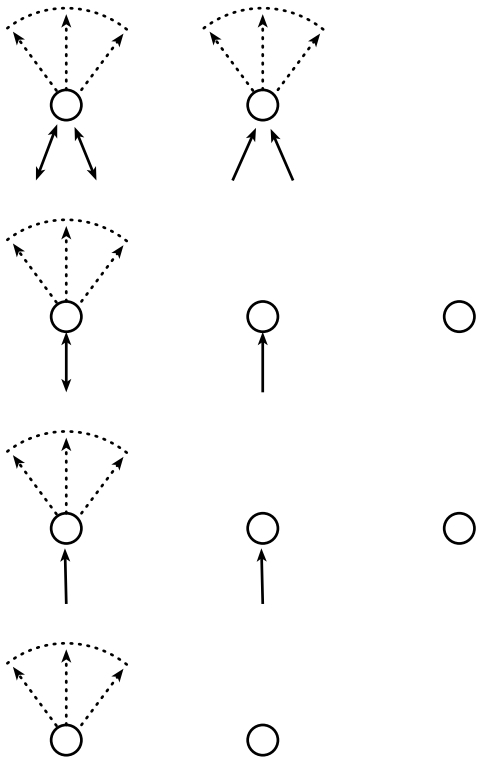

Figure 7.2: All cases of bad 'pictures' that may appear, and how they are dealt with by the deletion process. See text for a discussion.

It appears natural to attempt choosing a relabelling $\pi_{G_1',G_2'}$ of the nodes, now preserving the generalised degree sequence $\delta'$. However, here the problems arise that we feel unable to overcome at present. Note that there are several ways for a node of being bad, portrayed in Figure 7.2. We will also make the deletion process a little more 'aggressive' in that pictures of type c) (c.f. Figure 7.1) will be reduced to the empty picture at once, since we are only interested in nodes 'pointed to' by at least two directed edges.

The 'bad picture' portrayed in the first row of Figure 7.2 we shall call an

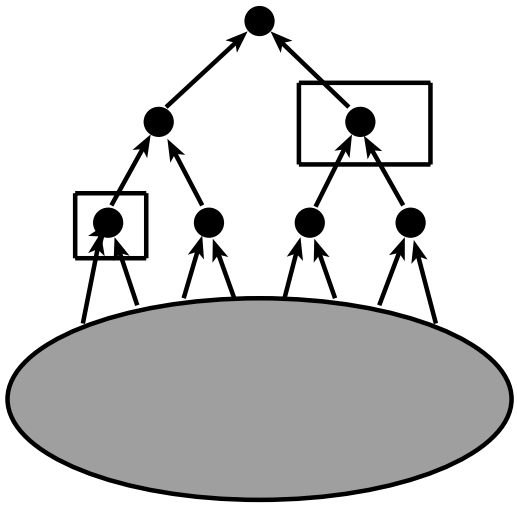

Figure 7.3: Two purged cherries (rectangular boxes) that are not 'equivalent'. One cherry has been 'freshly' purged, whereas the other has been purged in a previous step of shell-wise deletion. 'Non-equivalent' cherries must not be identified by $\pi_{G_1',G_2'}$! The grey area stands for some subgraph robust against the deletion process, e.g. a $K_8$.

*un-purged cherry*, and the picture it is turned into by the deletion process a *purged cherry*. Now the aforementioned relabelling $\pi_{G_1',G_2'}$ will identify purged cherries with purged cherries. But looking at Figure 7.3 shows that this will in general lead to problems. Some cherries have been purged 'some while ago' in the course of the deletion process and they should not be 'un-purged' when looking for predecessors under the given graph, whereas other purged cherries should. It is certainly no longer true that whenever $G_1' \uplus H$ is a predecessor of $G_1'$, $G_2' \uplus \pi[H]$ is a predecessor of $G_2'$.

Matters are even worse. There is a counterexample of two graphs having the same generalised degree sequence $\delta'$ but having (respectively) unique predecessors under the shell wise deletion process with different generalised degree sequences, $\delta_1$ and $\delta_2$. Thus the number of predecessors with degree sequence $\delta_1$ is 0 for one graph, say $G_1'$ and 1 for $G_2'$. The counterexample is shown in Figure 7.4. In the graph on the left the shell-wise deletion process will have purged the cherries in both 'cranes' in a symmetric way until there are no bad edges left. Thus for the left graph, reverting the deletion process should make all four cherries 'at the basis of the cranes' un-purged. Whereas for the graph on the right, the smaller 'crane' on the right hand side will have

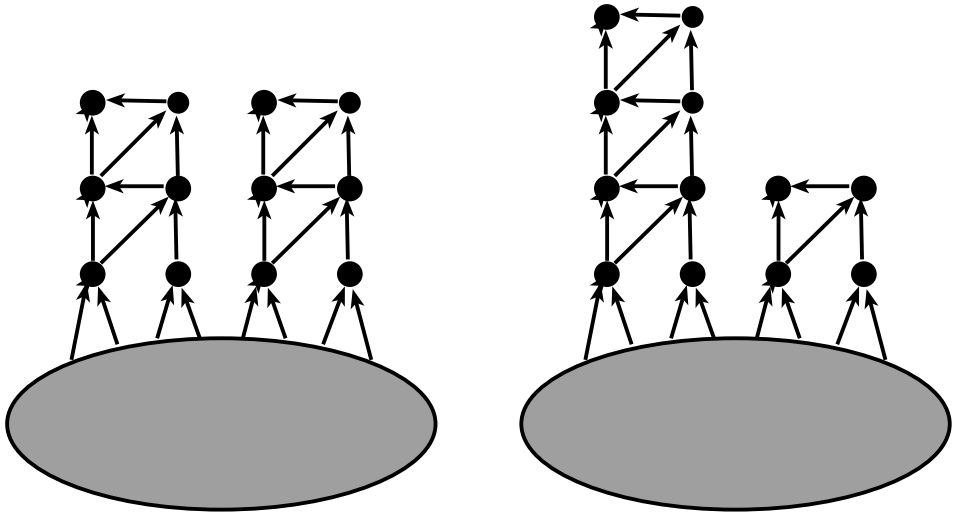

Figure 7.4: Two graphs with the same generalised degree sequence having (respectively) unique but *different* predecessors under the shell-wise deletion process. The grey area stands again for some subgraph robust against the deletion process, e.g. a $K_8$. Note by increasing the 'height of the cranes' this example yields a hole class of counterexamples.

been purged before the shell-wise deletion process has terminated. Thus for the right graph, reverting the deletion process should make only the two cherries 'at the basis of the left crane' un-purged.

At this point we are stuck because we can not answer the following questions. Are the graphs obtained by a repeated application of the shell-wise edge deletion process distributed *uniformly* amongst all graphs with a specific generalised degree sequence? Does only the *proof strategy* fail, that worked so well for the undirected case? Is the *graph invariant* 'generalised degree sequence' (as defined above) the wrong invariant amongst several other candidates for our purposes? Did we choose the 'wrong' deletion protocol, in the sense that it does not lead to the results we expected?

## 7.2 Attempts to Avoid Analysis of the Deletion Process

The 'classical' $k$-core and other subgraphs such as the the extended $k$-core and the magic subgraph can be characterised by deletion processes. Clev-

erly analysing cleverly chosen protocols of the deletion process will therefore teach us something about the respective subgraphs, indirectly. Analysing those processes is possible, but not easy, and seems to introduce unnecessary overhead. Intuitively, the subgraphs 'do not care' about the deletion process and how we analyse it. We feel that the 'ideal proof' could get away without any reference to deletion processes.

We have discussed (in Section 7.1) to what extent we can generalise the proof strategy successful for the 'classical' $k$-core to the extended $k$-core and the magic subgraph. Our proof, adapting Goerdt and Molloys proof strategy to the $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model, is simpler than the original proof of Pittel et.al. ([PSW96]) in that *most of* the tedious analysis of the single node deletion process is replaced by studying the set of owners instead. But we do not avoid deletion processes *altogether*. Even this remainder makes the proof far more clumsy than intuitively necessary. Moreover, we were unable to carry over the proof strategy to the case of (directed) edge deletion processes, as far as the treatment of remaining runs is concerned. This is possibly 'merely' a technical issue but we would be happy to avoid it. We expect a proof for the 'classical' $k$-core not having to rely on deletion processes at all to translate to the scenario of directed subgraphs (essentially corresponding to the binary owners) in a fairly straightforward manner.

We could settle with a weaker result, the 'sudden appearance of giant subgraph that has only $o(n)$ runs', which can be proved without any reference to deletion processes whatsoever (Section 5.2). This subgraph is simply the set of $r$-owners. Note that this result gives a good deal of structural insight why a phase transition in branching trees corresponds to a phase transition in random graphs, and why the numerical values coincide. Yet the set of owners is only an approximation to the $k$-core based on 'semi-local information' which itself is actually a 'global phenomenon'.

Ultimately we want to prove the sudden appearance of the $k$-core and similar structures and not approximations. How can we 'get rid' of the runs if not by analysing the single node deletion process as we have done for the 'classical' $k$-core? In other words, can we exploit the fact that the radii $r(n)$ (on which the 'semi-local' decisions rely) grow with $n$ sufficiently well to capture the 'globalness' of the $k$-core and related subgraphs?

### Arguing similarly as in the sub-critical case

This approach appears to be the most promising and would lead to a short and elegant proof, if successful.

Remember that in the sub-critical case we knew that there were only $o(n)$ nodes that might form a non-empty $k$-core. This result followed from

predicting the degree sequence of the $r$-owners. Now *without* having to refer to any deletion processes we could indirectly argue that those $o(n)$ nodes are highly unlikely to form or contain a $k$-core, essentially considering the edge-density of any induced subgraph containing less than $\varepsilon \cdot n$ nodes, for some sufficiently small $\varepsilon$ (following Goerdt and Molloy).

In the super-critical case we need to dis-prove that more than $o(n)$ nodes will be deleted from the set of $r$-owners. We have done this using the gambler's ruin argument, but is there a simpler proof?

Intuitively, consider an adversary that is allowed to pick a degree-sequence as predicted in Lemma 5.3.2. Then he builds a graph obeying this degree sequence in a worst case fashion. There are $\theta(n)$ nodes that have degree exactly 3 and some $o(n)$ nodes that are runs. Presumably the adversary would take (at least) one run and 'pile up' many ($\theta(n)$) nodes with degrees exactly 3 into a neighbourhood of this run. Upon removal of the run all those 'piled up' nodes will be removed by the deletion process.

However, it seems very unlikely that such 'giant piles' of nodes having degree exactly 3 will occur in a *random* setting.

How could this insight be formalised other than by analysing the gambler's ruin? We think that assuming that more than $o(n)$ nodes, say $\varepsilon \cdot n$ for some sufficiently small $\varepsilon$, are deleted from the set of owners implies the existence of some subgraph that is unlikely to occur.

Possibly such an argument turns out to be simple, but we were not successful in finding one. In view of the sub-critical case the first idea that jumps to mind is to analyse the edge density in the subgraph induced by the $r$-owners that were deleted. Essentially there are only two edges per node, and the edges connecting the deleted nodes with the remaining owners must not be counted. Thus the edge density seems - at first sight - indeed 'atypically low'. It is certainly bounded above by 2. But note that we are considering subgraphs that are induced by $\varepsilon \cdot n$ nodes, only. Here the expected number of edges should be

$$\binom{\varepsilon \cdot n}{2} \cdot c/n \simeq \varepsilon^2 \cdot n/2 \cdot c,$$

and, accordingly, the expected edge density in the induced subgraphs should be roughly equal to

$$\varepsilon \cdot c/2.$$

But this does *not* imply a contradiction to the observation that the edge density is upper bounded by 2 in the induced subgraphs, the smaller $\varepsilon$ the less there is a problem. Let us contrast this to the analysis of the sub-critical case. Here we know that the edge density of the induced subgraphs is *lower*

*bounded* by 3 and there is a contradiction, the smaller $\varepsilon$ the further away is the expected edge density from the lower bound 3.

Finally note that there is a faint hope that the specific structure of the degree sequence is such that building up 'piles' (as explained above) is impossible for the adversary and more generally that we can deduce the deletion process to terminate after $o(n)$ steps *deterministically*. But we neither know whether this can be true nor do we have any ideas how we should prove such a statement.

### Increasing $r(n)$ is too naive

We first present a 'naive' approach that does not work. Remember that runs can be shown to be rare, because $\rho$-owners fail to be $(\rho+1)$-owners with probability $O(t^\rho)$, only. Now choosing $\rho := r(n)$ for $r(n)$ growing with $n$ will lead to increasingly smaller probabilities. When we could choose $r(n)$ such that the aforementioned probability was $o(1/n)$, the expected number of runs would converge to zero and there would be none a.a.s. by Markov's Inequality. Clearly we would have to substitute $\rho := r(n) = \omega(\log n)$ in order to get $O(t^\rho)$ small enough. It is well known that the diameter of a random graph with constant average degree is roughly equal to $\log n$. Therefore the $r$-neighbourhoods will resemble branching trees only for $r(n) = o(\log n)$ - when choosing $r(n)$ larger than the diameter, the branching tree analogy will be definitely lost. More formally, we can re-inspect the proof of the Semi-Local Lemma 5.1.5 and see that the inductive proof would break down when choosing $r(n)$ considerably larger than we have done.

There is a last informal observation to be stated in that context. The aforementioned problems are in some sense due to the fact that we restrict ourselves to neighbourhoods with fixed radii when determining whether nodes are owners, non-owners or runs. Within our framework, we need to do that, because otherwise we could no longer employ the recursive equations.

Intuitively, the status of the root is very 'sensitive' against changes on the boundary of the $r$-neighbourhood whenever its cone contains a binary tree (not necessarily complete) 'connecting' the root with the $r$-boundary, such that every interior node has degree *exactly* 3.

We have seen in the context of branching trees that the part of the cone that has degree exactly 3 can be described by a sub-critical branching tree with progeny distribution $\mathrm{Bi}(2, \pi(c))$. Such branching trees will die out after only $(\log(n))^{\hat{K}(c)}$ nodes are discovered, yet this will not always happen in a 'balanced' way, i.e. these sub-critical branching trees will with very high probability die out after having 'generated' $(\log(n))^{\hat{K}(c)}$ nodes but not always within a radius of order $\log\log n$. If that happens, it will be necessary to look

'beyond the horizon' of $r(n)$. We therefore believe that ownership etc. should probably be determined not only on the basis of the $r(n)$-neighbourhood, but that in the rare cases described above it would be necessary to explore the graph beyond the $r$-th shell of the BFS, starting from those nodes on the boundary that are part of a sub-critical binary tree. We have not found a way of turning this intuition into something rigorous, but note that the 'proofs' for ownership etc. would presumably still be based on the outcome of some $(\log(n))^{\tilde{K}(c)}$ edges. So, the vaguely suggested 'improved' concept of ownership could presumably still be decided 'locally', in some new appropriate sense.

### Extending the 'test particle' approach?

What we present next does not work, too, but it is more sophisticated and there may be faint hope that something can be shown along those lines, eventually. In Section 5.2 we have successfully applied the 'test particle' approach. Apart from having taken into account errors due to finite $n$ we essentially showed that the probability for being an owner is governed by the same recurrence equations as in branching trees.

Now, in branching trees we were also able to describe the evolution of the holes induced by runs. Those holes are branching trees themselves with progeny distributions $\mathrm{Bi}(2, \pi(c))$. The expected progeny is less then 1 (see Section 3.3.2) - such branching trees are well known to die out quickly. So there is something like a Branching Tree Connection for the evolution of the hole. Obviously the gambler's ruin argument, whenever we can get it to work(!), rigorously proves that this intuition is correct.

We furthermore know that, at least in branching trees, the probability of the event 'the root is an $r$-owner' is robust against small perturbations (see 3.3.3). So intuitively one might expect that the impact of 'knowing' about some nodes already deleted on the set of owners in the remainder graph should somehow be small.

Positively speaking, we have pursued the following approach somewhat similar to analysing the gambler's ruin process arising from the single node deletion process, but using only the principle of deferred decisions. The idea is to describe the evolution of the hole $H_t$ by a fanning out process, which hopefully should die out quickly.

We shall at first be sloppy with the concept of ownership and clarify problems arising from this afterwards. For ease of exposition, we shall restrict ourselves to the 'generic case', i.e. that each run is adjacent to exactly two binary owners, and we shall assume that the hole $H_t$, i.e. the set of nodes that have been removed so far, is a union of trees.

Suppose we have removed $o(n)$ nodes from the set of owners. Each node $v$ in the hole $H_t$ was an owner w.r.t. the original graph. If $v$ was a run originally, it was adjacent to exactly two owners (generic case!). Or it was 'made bad' because one father in the set of nodes already deleted was removed, in which case it is adjacent to exactly two other owners. In particular, each node that is still bad in the boundary of the hole $H_t$ is adjacent to exactly two binary-owner nodes in the remainder graph.

What do we know about the remainder graph? Essentially all we know is that there are at least some $2\,|H_t|$ nodes in the remainder graph that are binary owners. However, we know that the set of binary-owners has linear size, anyway, the size being sharply concentrated. Thus conditioning on having at least $2\,|H_t|$ binary owners is conditioning on an event that fails to happen with sub-polynomially small probability.

One may be tempted to assume (but this will turn out to be wrong!) that information on how nodes in the boundary of $H_t$ are connected to those owners in the remainder graph is essentially an issue of how the edges connecting $H_t$ with the remainder graph are selected. The owners reached by edges from $H_t$ will be Cayley owners w.r.t. the remainder graph with probability strictly larger than $1/2$. And $H_{t+1}$ should therefore be smaller than $H_t$ by a factor strictly less than one, on average.

So it seems as if an analysis of the evolution of $H_t$ could be possible. However, when it is true that that the remainder graph is essentially a 'fresh' $\mathcal{G}_{n,p}$, as long as only $o(n)$ nodes are removed, it should contain $\theta(t^{r(n)})$ runs (c.f. Lemma 5.2.7). So how should the bad nodes die out in the process?

What is wrong? The paradoxical conclusion is due to the aforementioned sloppiness. An owner w.r.t. the remainder graph is not the same as an owner w.r.t. the entire graph and we need to study the latter. More specifically, the ideas laid out in the last but one paragraph are incorrect. It is correct that the set of owners w.r.t. an *arbitrary* remainder graph on $n - o(n)$ nodes is concentrated around what it should be, independently of the outcome of the edges connecting $H_t$ with the remainder graph, see Lemma 5.2.7. But the Principle of Deferred Decisions may *not be applied*, as ownership w.r.t. the original graph depends on the random edges within the remainder graph *and* on the edges connecting $H_t$ with the remainder graph.

## 7.3 Comparing $k$-Partite and Non-$k$-Partite Models

Guided by the intuition that lead to defining the magic subgraph and in view of the empirically observed 'almost unique colourability' one might conjecture that the magic subgraph is a *candidate* for a subgraph explaining the jump in chromatic number. Remember that the appearance of a giant uniquely 3-colourable subgraph would a.a.s. make the chromatic number jump after adding only $o(n)$ edges, c.f. the discussion in [Mol01].

Unfortunately, the magic subgraph is defined *w.r.t. some given tripartition*, and it is is not at all clear *which* tripartition should be used in the non-tripartite model, in the presence of several proper colourings. We start with a couple of (confusing) observations.

**Proposition 7.3.1** *The $\mathcal{G}_{n,c/n}$-model, conditional on respecting some* specific *(balanced) colouring, is equivalent to the $\mathcal{G}_{n,3,c\cdot c/n}$ model. I.e. the average degree is reduced to $2/3 \cdot c$ by conditioning on a specific colouring.*

**Proof** We may assume that the colour classes consist in the node sets $[1, n/3]$, $[n/3+1, 2\cdot n/3]$ and $[2\cdot n/3+1, n]$. Obviously, this partition induces a proper colouring if and only if there are no edges within the colour classes. Since one third of the potential edges in the non-tripartite model lie within a colour class we have just discounted $1/3$ of the potential edges without any restriction on the other potential edges. Thus the average degree is reduced to $2/3 \cdot c$. □

Thus Proposition 7.3.1 suggests that conditioning on 3-colourability in the non-tripartite case might reduce the 'effective' average degree. On the other hand, it seems plausible to assume that whenever $c$ is strictly below the (unknown) threshold in the non-tripartite model, the graph is 3-colourable with very high probability. If the probability of failure was sub-polynomially small, the sum of indicators counting the number of edges should not be considerably changed, in other words conditioning on the existence of *some* proper colouring should not affect the average degree (nor any other 'reasonable event'), for sub-critical values of $c$. We have tried to discern 'cases', applying Proposition 7.3.1 to all feasible (balanced?!) colourings. But those 'cases' do not form a partition of the probability space and we were unable to perform any rigorous calculations. We remind the reader that the 3-core appears at the same critical average degree both in the the tripartite and in the non-tripartite model which is certainly backed by empirical evidence. Moreover, it is very plausible that our proofs for the appearance of the giant

$k$-core can be adapted to the tripartite model without changing the critical value (we are merely careful not to claim anything that we have not fully checked).

We thus do not even intuitively know whether the tripartite and the non-tripartite model should be considered comparable, as far as critical average degrees for the jump in chromatic number are concerned. Still it is a natural *ad hoc* assumption that the jump in the chromatic number from $k$ to $(k+1)$ in the non-$k$-partite model 'corresponds' to a 'dying out' of the $k$-colourings different from the 'built in' colouring in the $k$-partite model, and that the critical average degrees might be the same in both models. We find it worthwhile to discuss the issue here.

We have given empirical and (partial) theoretical evidence for the sudden appearance of a giant magic subgraph in random tripartite graphs at an average degree $c_{BT}$ (the values from the Branching Tree Connection), and this seems to generalise to $k > 3$. Let us assume this to be true for the moment. We encourage the reader to re-read the discussion of Achlioptas and Molloys upper bound $c_{AM}$ on the critical average degree for the jump in chromatic number in Section 2.5.2.

If the assumption described in the last but one paragraph were true the magic subgraph in the $k$-partite model should better not appear *before* the point where there are no more proper $k$-colourings. Those critical points were upper bounded by Achlioptas and Molloy using the First Moment Method.

In the light of the above considerations it is tempting to compare the numerical values for the appearance of the magic subgraph in the $k$-partite model and the upper bound on the colourability threshold in the non-$k$-partite model.

| $k$ | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $c_{AM}$ | 5.0434 | 9.1722 | 13.8958 | 19.0778 | 24.632 |
| $c_{BT}$ | 4.9108 | 9.267 | 14.035 | 19.112 | 24.434 |

Observe that for $k = 4, 5, 6$ there are $c$-values such that there are a.s. no colourings left in the non-$k$-partite model, whereas a magic subgraph has not yet appeared in the $k$-partite model with the same average degree $c$.

It is therefore natural to ask whether the (First Moment) upper bound from the non-$k$-partite model applies to the $k$-partite case as well. To achieve a fair basis for comparison, we have repeated the arguments of Achlioptas and Molloy for the $k$-partite model, see Lemma 7.3.2 below. We remark that in order to achieve a fair basis of comparison one could also switch entirely to the non-tripartite model. But there it is unclear with respect to which colouring the magic subgraph should be defined, see above.

Note that the significance of the rigid colourings in the $k$-partite case is different than in the non-$k$-partite case. In the non-$k$-partite case a non-empty set of (proper) colouring implies the existence of at least one rigid colouring and therefore it suffices to show that the expected number of rigid colourings converges to zero. However, in the $k$-partite case the set of proper colourings is never empty by construction and so is the set of rigid colourings. When arguing that the jump in the chromatic number in the non-$k$-partite model should 'correspond' to the numbers of proper colourings jumping from many to one in the $k$-partite model, one should thus possibly only count 'orthogonal' rigid colourings, 'orthogonal' being defined somehow appropriately, see our discussion at the end of this section.

**Lemma 7.3.2** *The upper bound $c_{AM}$ on the number of rigid colourings in the non-$k$-partite case is* not *the same as the analogous upper bound in the $k$-partite case. Specifically, for $k = 4$ it is strictly larger than the branching tree value $c_{BT}$ at which a magic subgraph should appear.*

**Proof** (of Lemma 7.3.2) In [AM99] Achlioptas and Molloy count the *rigid* colourings only. The crucial quantity in [AM99] is the function

$$f : (c, \alpha_1, \ldots, \alpha_k) \mapsto \frac{\left[2\sum_{i<j}\alpha_i\alpha_j\right]^{c/2}}{\prod_i (\alpha_i)^{\alpha_i}} \prod_{j>1}\left(1 - \exp\left(-\alpha_j \frac{c/2}{2\sum_{i<j}\alpha_i\alpha_j}\right)\right)^{\sum_{i<j}\alpha_i},$$

since the expected number of rigid colourings is bounded above by

$$\left[\sup_{\alpha_1,\alpha_2,\alpha_3,\alpha_4} f(c, \alpha_1, \alpha_2, \alpha_3, \alpha_4)\right]^{n+o(n)}.$$

By mimicking their analysis for the $k$-partite case we get an upper bound

$$\left[\sup_{(\alpha_{ij})_{i,j}} F(c, (\alpha_{ij})_{i,j})\right]^{n+o(n)},$$

where

$$F : (c, (\alpha_{ij})_{i,j}) \mapsto \frac{\left[\frac{2k}{k-1}\tau((\alpha_{ij})_{i,j})\right]^{c/2}}{\prod_{i,j}(k\alpha_{ij})^{\alpha_{ij}}} \prod_i \prod_{j<k} \prod_{j'>j} \left(1 - \exp\left(-\sum_{i'\neq i} \alpha_{i'j'} \frac{c/2}{\tau((\alpha_{ij})_{i,j})}\right)\right)^{\alpha_{ij}},$$

here (as above) all indices in sums and products are in $[k]$ and

$$\tau((\alpha_{ij})_{i,j}) := \sum_{i<i'} \sum_{j \neq j'} \alpha_{ij}\alpha_{i'j'}.$$

We include detailed calculations below. From now on we assume that $k = 4$. Whenever the suprema are less than one the expected number of rigid colourings converges to zero and there are no rigid colourings a.a.s. by the First Moment Method. By substituting $\alpha_{ij} := \alpha_j/4$ the function $F$ reduces to $f$, therefore

$$\sup_{\alpha_1,\alpha_2,\alpha_3,\alpha_4} f(c,\alpha_1,\alpha_2,\alpha_3,\alpha_4) \leq \sup_{(\alpha_{ij})_{i,j}} F(c,(\alpha_{ij})_{i,j}).$$

Moreover, for $c = c_{AM} = 9.172$ and even for $c = c_{BT} = 9.267$ there are $(\alpha_{ij})_{i,j})$ such that $F(c,(\alpha_{ij})_{i,j}))$ is definitely greater than one. Choose the following parametrisation:

$$x \mapsto \begin{pmatrix} 1/4-x & x/3 & x/3 & x/3 \\ x/3 & 1/4-x & x/3 & x/3 \\ x/3 & x/3 & 1/4-x & x/3 \\ x/3 & x/3 & x/3 & 1/4-x \end{pmatrix} =: \alpha(x). \tag{7.1}$$

We have plotted $x \mapsto F(c,\alpha(x))$ for the aforementioned values of $c$ in Figure 7.5. From that plot it is evident that choosing, say, $x = 0.05$, yields an $F$-value strictly greater than one.

It remains to state details concerning the upper bound on the expected numbers of colourings in the $k$-partite model, we proceed analogously as Achlioptas and Molloy in [AM99].

We will encode all possible partitions of the nodes in 'built in' and new colour classes by $k \times k$-matrices $(\alpha_{ij})_{i,j}$. The rows correspond to the 'built in' colouring and thus the row sums are all $1/k$, respectively. Entry $\alpha_{ij}$ describes the percentage of nodes that were coloured by colour $i$ in the 'built in' colouring and are coloured $j$ by the new colouring encoded by $\alpha$.

Let

$$T(\alpha) := n^2\tau(\alpha) := n^2 \sum_{i<i'} \sum_{j \neq j'} \alpha_{ij}\alpha_{i'j'}$$

be the number of proper colourings consistent with $\alpha$.

Now a colouring encoded by $\alpha$ is proper with probability

$$\mathbb{P}[C_\alpha] = \frac{T(\alpha)^m}{(\binom{n}{2} - k\binom{n/k}{2})^m} = (\frac{2k}{k-1}\tau(\alpha))^m.$$

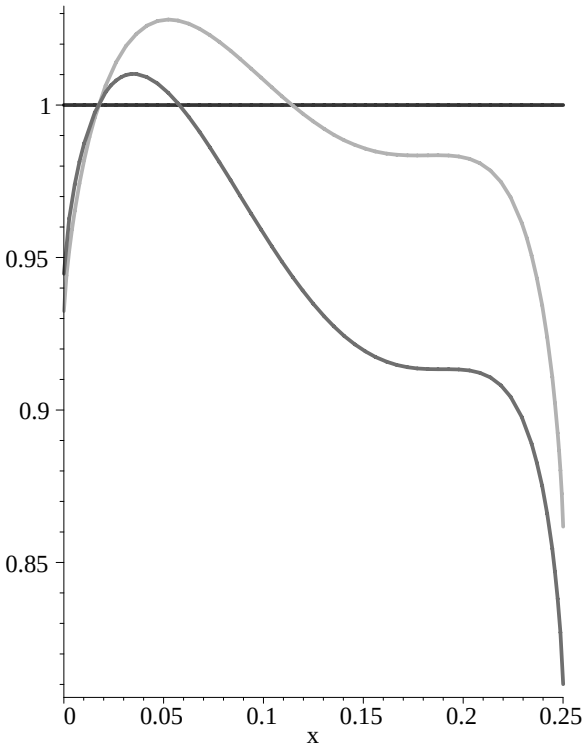


Figure 7.5: Plot of $F(c,\alpha)$ (see Eq. 7.2) along the curve $x \mapsto \alpha(x)$ defined in Eq. 7.1, for $c_{AM} = 9.172$ and $c_{BT} = 9.267$ .

Next we need the probability $\mathbb{P}[R_\alpha | C_\alpha]$ that a colouring is rigid conditional on being proper.

$$\mathbb{P}[R_\alpha | C_\alpha] \leq \prod_i \prod_{j<k} \prod_{j'>j} \left(1 - \exp\left(-\sum_{i' \neq i} \alpha_{i'j'} \frac{c/2}{\tau(\alpha)}\right)\right)^{\alpha_{ij} n}.$$

Thus the expected number of rigid colourings is

$$\mathbb{E}[R] = \sum_\alpha \mathbb{P}[R_\alpha] \leq \sup_\alpha \prod_i \binom{n/k}{\alpha_{i1}, \ldots, \alpha_{ik}} \mathbb{P}[R_\alpha] \cdot n^{k^2},$$

since there are at most $n^{k^2}$ ways to partition $n$ into $k^2$ ordered summands.

Therefore we get

$$\mathbb{E}[R] \leq \sup_\alpha F(c,\alpha)^{n+o(n)},$$

where

$$F(c,\alpha)=\frac{\left[\frac{2k}{k-1}\tau(\alpha)\right]^{c/2}}{\prod_{i,j}(k\alpha_{ij})^{\alpha_{ij}}}\prod_{i}\prod_{j<k}\prod_{j'>j}\left(1-\exp\left(-\sum_{i'\neq i}\alpha_{i'j'}\frac{c/2}{\tau(\alpha)}\right)\right)^{\alpha_{ij}}. \quad (7.2)$$

□

Note that in the above proof colourings parametrised with the matrix $\alpha(0.05)$ may certainly regarded 'orthogonal', in the sense that they do not coincide with the 'built' in colouring. One may regard the colourings whose $\alpha$-matrix maximises $F$ as the typical colourings. They will presumably be even 'more orthogonal' than the ones restricted to the curve $x \mapsto \alpha(x)$.

The above discussion is not satisfactory. In our eyes the problems arising may reflect the essential weakness of trying to determine the colourability threshold via the First Moment Method. The appearance of a giant uniquely colourable subgraph is certainly an appealing feasible mechanism suggesting an explanation for the jump in chromatic number, c.f. [Mol01]. Yet, even in the presence of such a giant uniquely colourable subgraph there may still be many different colourings of the graph as a whole. The unique colouring of the giant subgraph may be extended to the nodes not in the giant subgraphs in possibly exponentially many ways, the colours of those node are still 'flexible'.

From this point of view the refinements of the naive application of the First Moment Method (see p. 39) may be regarded as an attempt to (partially) reduce the 'flexibility' of nodes in the complement of a uniquely colourable subgraph in order to reduce the number of 'extensions' of the colouring restricted to the uniquely colourable subgraph. As a simple example consider the $\simeq n \cdot \mathrm{Po}_c(0)$ isolated nodes. They certainly introduce an exponential factor in the number of colourings, and by restricting ourselves to rigid colourings this 'flexibility'-effect (as well as other more complicated effects) is excluded. First Moment upper bounds counting colourings are *never wrong*, but they may systematically overestimate the threshold due to the 'jackpot phenomena' as discussed. Moreover, it seems very hard to completely eliminate the 'flexibility' without explicitely referring to the uniquely colourable subgraph, provided such a subgraph was indeed the 'true reason' for the phase transition.

So, what is the conclusion of this section? Our Lemma 7.3.2 merely gives an indication that the First Moment Bounds derived for the non-$k$-partite model do not carry over in a straightforward way to the $k$-partite model. The thresholds may have different numerical values in the two models, even though some subgraph *similar* to the magic subgraph may still be responsible for the threshold-phenomenon in the non-$k$-partite model.

## 7.4 Further Recurrence Equations

This section contains a heuristic discussion of further recursive equations indicating phase transition in branching trees. We have decided to include the material at this point.

The theory of phase transition in sparse random graphs may benefit from studying recurrence equations for branching trees in two ways. We have seen that the concept of ownership can be transfered to the 'corresponding' graphs and that we can devise a new rigorous proof for the appearance of a giant $k$-core based on the very same recurrence equations playing the central role for branching trees. The 'discovery' of the magic subgraph as the 'analogue' to the coloured owners illustrates the second benefit. Motivated by a new threshold phenomenon in branching trees we set out to find - and found - a 'corresponding' giant subgraph that actually appears in graphs. Apparently, when the root is 'fixed' in the branching tree with probability $p$ there seems to appear a 'corresponding' giant subgraph of size $p \cdot n$ in the random graph.

We therefore consider it worthwhile to report on several other recursive equations that exhibit threshold phenomena. They may serve as starting points when looking for new giant subgraphs playing some role for understanding phase transitions.

### 7.4.1 Heuristic: Arbitrarily Coloured Graphs

Consider the analogue of what a BFS 'sees' in a *arbitrarily coloured* random graph. By arbitrarily coloured we mean that all vertices get some colour, uniformly and independently, thus the colouring will usually not be proper. This resembles the starting state of the Antivoter Algorithm, see Section 2.6.1.

We consider a vertex 'fixed' when it has at least two fixed children of different colours. This yields a recursive equation associated to

$$\begin{aligned} f(c,x) = & \underbrace{\left(1-e^{-1/3\,xc}\right)}_{\text{at least one red child fixed}} \left(1-\left(e^{-1/3\,xc}\right)^2\right) + \underbrace{e^{-1/3\,xc}}_{\text{no red child fixed}} \left(1-e^{-1/3\,xc}\right)^2 \\ & = 1-3\,e^{-2/3\,xc}+2\,e^{-xc}, \end{aligned}$$

with a critical value at $c = 4.1546(\ldots)$.

Indeed, we have empirically observed the sudden appearance of giant magic subgraphs of random *non-k-partite* graphs at this critical value, defined *with respect to arbitrary random colourings*, see Figure 7.6.

However, as opposed to the magic subgraph in the tripartite case *this* magic subgraph does not appear to be 'almost uniquely' colourable, empirically. Neither does its absence imply some (obvious) efficient colouring

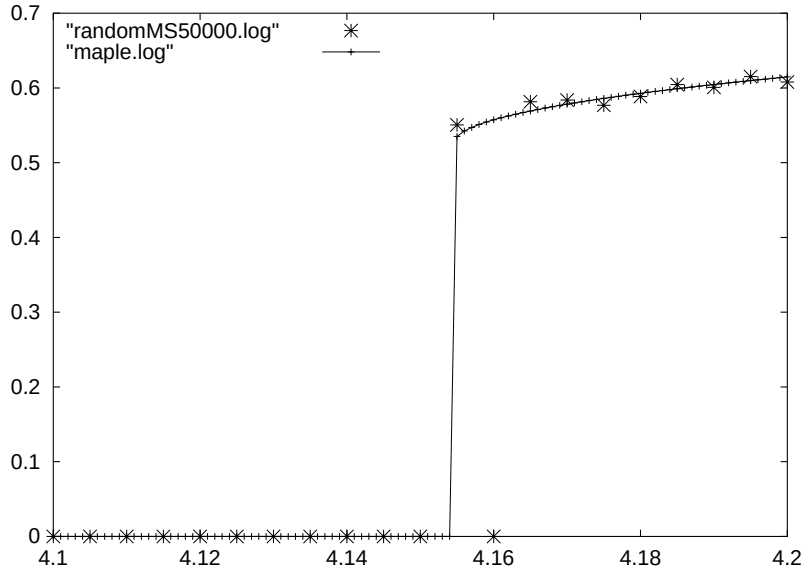


Figure 7.6: Magic subgraph with respect to an i.i.d. colouring. Value from the Branching Tree Connection and empirical result of a deletion process on a random graph with $n = 50000$ nodes.

algorithm like in the case of the absence of a $k$-core, c.f. Section 2.5.1. It may deserve attention when further investigating the 'correlation' effect described on p. 84 in Section 4.3.

### 7.4.2 Heuristic: Uniquely Colourable Subgraphs

The phase transition in coloured branching trees as explained in Section 3.2 does actually appear to translate to the appearance of a giant subgraph *in the tripartite model*. Moreover there is empirical evidence that this subgraph (the magic subgraph) may be uniquely 3-colourable or contain a giant uniquely colourable subgraph. Remember (see Chapter 1) that the advent of a uniquely 3-colourable subgraph would immediately imply an upper bound for the 3-colourability threshold *in the non-tripartite model*. It is by no means clear whether the thresholds for the tripartite and the non-tripartite model are comparable (see Section 7.3). In particular, since the magic subgraph is defined with respect to a 'built in' tripartition it is not well-defined in non-$k$-partite graphs.

In the non-$k$-partite model one feasible mechanism for explaining the sudden jump in chromatic number goes as follows: In the course of vertex exposure there suddenly appears a giant uniquely $k$-colourable graph, implying chromatic number $k+1$ after only $o(n)$ further steps. Since we know that below the (unknown) threshold the random graph is $k$-colourable with high probability it seems therefore plausible to assume that the graph 'generated so far' contains a certain proportion of 'fixed' nodes, and condition on the event that the newly attached vertex has 'fixed' neighbours only in $k-1$ of the colour classes, which we further assume to be of roughly equal size. I.e., we assume that some unknown subgraph of size $q \cdot n$ induced by 'fixed' nodes is uniquely colourable. We shall ask for the probability that the newly attached node is 'fixed', i.e. becomes a member of this uniquely colourable subgraph itself.

This leads to the following recursive equation:

$$q = f(c,q) = \frac{k \cdot \overbrace{\mathbb{P}\left[\mathrm{Po}_{cq/k} > 0\right]^{k-1} \cdot \mathbb{P}\left[\mathrm{Po}_{cq/k} = 0\right]}^{\text{adjacent to fixed nodes in all but one colour}}}{\underbrace{1 - \mathbb{P}\left[\mathrm{Po}_{cq/k} = 0\right]^{k}}_{\text{not adjacent to fixed nodes in all colours}}}.$$

For $k = 3, \ldots, 7$ this leads to the following critical values, stated in the second row in the table below.

| k | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| c | 4.851 | 8.891 | 13.205 | 17.753 | 22.496 |
| $c_{AM}$ | 5.0434 | 9.1722 | 13.8958 | 19.0778 | 24.632 |
| $c_{BT}$ | 4.9108 | 9.267 | 14.035 | 19.112 | 24.434 |

.

Note that those critical values are *slightly smaller* than the values for the $k$-partite model (last row). In particular, they are consistent for all $k = 3, 4, \ldots, 7$ with all presently known upper bounds on the $k$-colouring threshold (third row).

### 7.4.3 Heuristic: Uniquely Satisfiable Subformulae

In the last subsection, where we studied - very heuristically - what may be the 'branching tree analogue' of a uniquely colourable subgraph in the 'corresponding' random graph. We may also ask for the analogue of a uniquely satisfiable formula in random 3-SAT with $n$ variables and $m = \alpha \cdot n$ 3-clauses.

Instead of properly defining things we shall merely give an example for a 'random' (rather a generic example of a) formula $F$ avoiding clumsy notation.

$$F = \underbrace{(X_1 \vee \bar{X}_2 \vee X_{23}) \wedge (\bar{X}_3 \vee \bar{X}_9 \vee X_{12}) \wedge \ldots \wedge (X_{18} \vee \bar{X}_{27} \vee X_{37})}_{m \text{ 3-clauses}},$$

i.e. such 3-SAT formulae are obtained by collecting $m$ random 3-clauses [‡].

We can view the 3-clauses as *randomly coloured* hyper-edges in a random 3-uniform hyper-graph. The nodes correspond to *variables* that are elements from the set $\{X_1, \ldots X_n\}$. The 'colour' is a number in $\{0, 1, \ldots, 7\}$ describing whether the variables in the 3-clause are negated or not:

Order the variables in each clause (according to the canonical ordering) and interpret the 'colour' as a bit-string $(b_1, b_2, b_3)$. The $\ell$-th variable in a clause becomes a negated literal if and only if $b_\ell = 0$. Thus the first clause in the above example formula corresponds to the hyper-edge $\{X_1, X_2, X_3\}$ with colour $1 \cdot 2^0 + 0 \cdot 2^1 + 1 \cdot 2^2 = 6$.

Again we shall assume that those hyper-graphs locally look like 'trees'. We will not make this precise and we shall assume that there is some analogue of vertex exposure. It is not hard to see that we should expect the new variable $X$ 'glued to the formula' (remember our 'test-particle' approach) to participate im roughly $\mathrm{Po}_{3\alpha}$ clauses. So $3\alpha$ is something like 'average degree'. The generic case will be that those clauses intersect only in node $X$. The variable $X$ will be 'fixed' if it participates in at least one clause $\{X, Y, Z\}$ such that both 'endpoints' $Y, Z$ are fixed and the colouring of $\{X, Y, Z\}$ is such that the value of $X$ is indeed uniquely determined. It is straightforward to see that the latter happens for 1/4-th of all possible 'colours', that is truth-assignments of the variables.

We therefore get the recurrence equation

$$q = f(\alpha, q) := \sum_{i=1}^{\infty} \mathrm{Po}_{3\alpha}(i) \cdot \underbrace{(1 - (1 - q^2/4)^i)}_{\substack{\text{fixed by at least} \\ \text{one hyper-edge}}} = 1 - e^{-3/4 \cdot \alpha \cdot q^2}.$$

Analysing this recurrence equation yields a critical value at

$$\alpha = 3.273(\ldots).$$

Note that this value is consistent with the presently best lower bound

$$\alpha = 3.26(\ldots).$$

by Achlioptas and Sorkin (as reported by [Mol01]), but it is far away from the conjectured (according to [Mol01]) value of $\alpha \stackrel{?}{\simeq} 4.2$.

[‡] We do not pay any attention to whether we sample with or without replacement, at this heuristic level. This resembles $\mathcal{G}_{n,m}$, there are also models analogous to $\mathcal{G}_{n,p}$ and random multigraphs.

# Chapter 8

# Discussion and Open Questions

We shall first report on what we consider our most important achievements, and then make suggestions for future research based on the insights gained during our studies, in particular discussing how our results and new proof techniques might be further improved.

## 8.1 Central Results

We have investigated aspects of the heuristic 'analogy' between phase transitions in branching trees and 'corresponding' random graphs, referred to as the Branching Tree Connection.

We showed that essentially the same calculations that describe phase transitions in branching trees can be translated into rigorous proofs for random graphs, by giving a new proof for the sudden appearance of the giant $k$-core in the $\mathcal{G}_{n,m}$ model.

Also, the search for new giant subgraphs that may play an important role for understanding phase transitions in random graphs can apparently be guided by first studying new phase transitions in branching trees, which is a relatively easy task, and then identifying the 'corresponding' subgraphs in random graphs, as we have demonstrated for the new magic subgraph.

### New proof for the $k$-core

In Chapter 5 we proved the sudden appearance of a giant subgraph induced by the set of $r(n)$-Cayley owners in the $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model. This subgraph has minimum degree $k$ for all but $o(n)$ nodes and is closely related to the giant $k$-core which has been analysed in [PSW96] by another proof technique analysing the dynamics of the deletion process.

Our rigorous proof employs the same recursive equations that are the key for analysing the *intuitively* 'corresponding' phase transition in branching trees (as discussed in Section 3.1) and we are thus providing a structural explanation for why the heuristic predictions suggested by the Branching Tree Connection are so surprisingly accurate.

Clearly, Goerdt and Molloy should be honoured for being the first to employ recursive equations related to 'corresponding branching trees' for rigorously proving the sudden appearance of the $k$-core in random structures closely related to random graphs, namely random faulty configurations, in [GM00]. We have explained their results in Section 2.4.3, also trying to point out why it is a non-trivial task to generalise their result to the important $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model, which we have done. Our contribution here mainly consists in the 'test particle approach', combined with the 'Semi-Local' Lemma 5.1.5, the proof of which requires somewhat more than an application of standard concentration tools based on a degree-bounded dependency graph, see Sections 5.1 and 5.2.

The aforementioned subgraph of $r(n)$-Cayley owners is only 'almost' the $k$-core. Again, Goerdt and Molloy have pioneered our proof for the sudden appearance of a giant $k$-core in the $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model (see Chapter 6), by showing that a single-node deletion process applied to the set of $r(n)$-Cayley owners will a.a.s die out after only $o(n)$ steps, in the model of random faulty configurations. Our proof essentially follows the route laid out in [GM00]. Yet mimicking Goerdt and Molloy's proof in the $\mathcal{G}_{n,p}/\mathcal{G}_{n,m}$ model required some additional work on a more or less technical level. In particular, we showed for the $\mathcal{G}_{n,m}$ model that uniform distribution remains invariant under the shell-wise deletion process, and that at each step of the single node deletion process the expected increase in the number of 'bad' nodes is negative, both as in the model of random faulty configurations. Negative expected increase is slightly harder to show in the $\mathcal{G}_{n,m}$ model, we used a switching argument.

### Directed subgraphs, extended $k$-core and magic subgraph

Our second central achievement is the 'discovery' and formal definition of the magic subgraph, together with the empirical observations that its appearance and size follow the values suggested by the Branching Tree Connection in a completely analogous way as the $k$-core does, and that magic subgraphs *of random $k$-partite graphs* seem to be 'almost uniquely colourable'. See Chapter 4.

We were motivated by the branching tree result as laid out in Section 3.2 in searching for a subgraph that 'corresponds' to the set of coloured owners

just as the 'ordinary' owners correspond to the $k$-core. Surprisingly, the 'corresponding' magic subgraph turned out to be a *directed* subgraph, and in retrospect we saw that there is an analogous directed subgraph closely related to the $k$-core, that is the extended $k$-core.

Intuitively, the directed subgraphs correspond to the binary owners, either coloured or 'ordinary', whereas the 'classical' $k$-core corresponds to the Cayley owners. Remember that when defining Cayley owners we had to make an exception, taking into account the 'special role' of the root, which was not necessary for binary owners. Further note that, using our results laid out in Section 4.2, it is a trivial* task to deduce the appearance of a giant 'classical' $k$ core from the appearance of a giant extended $k$-core. Finally, talking about 'coloured Cayley owners' does not seem to make sense, and thus there is no 'classical' magic subgraph, possibly reflecting the fact that the magic subgraph can *not* be characterised by a node-deletion process.

Considering the arguments laid out in the last paragraph, we think that *analysing directed subgraphs* appears to be the most natural approach, mainly because of the similarity between (extended) $k$-core and the magic subgraph seems best reflected in this framework. Having said this we feel even more dissatisfied that we were not even able to fully generalise the proof for the sudden appearance of the giant $k$-core to a proof of the sudden appearance of a giant extended $k$-core. This is essentially due to the fact that we could not prove the invariance of uniform distribution under what seems to be a natural generalisation of the graph invariant 'degree sequence' to directed graphs, in the course of the deletion processes. We have discussed this issue in Section 7.1 and we will highlight important ideas concerning this problem below.

## 8.2 Future Research

We shall now discuss what can or should to be done next, also highlighting the most important conclusions that can be drawn from our investigations described in Chapter 7.

### Avoiding the analysis of the deletion process

It should by now be clear that our proofs would become a lot shorter if we did not have to translate from the $\mathcal{G}_{n,p}$ model to the $\mathcal{G}_{n,m}$ model, which we have only done in order to be able to analyse the single node deletion process

*The 'classical' $k$-core is induced by the set of nodes with in-degree at least $k$ w.r.t. the extended $k$-core. See the proof of Lemma 4.2.8

in the second phase. Moreover, once in the $\mathcal{G}_{n,m}$ model, there was still a lot of work to be done in order to show the invariance of uniform distribution under the deletion processes for the 'classical' $k$-core and, see pp. 131 in Section 7.2, for the extended $k$-core and magic subgraph we do not even know how to generalise the proof. So, can we avoid analysing the deletion process?

The most promising approach appears to be finding an indirect argument, similar to the one used in the sub-critical case for the $k$-core. We know the degree sequence of the set induced by the $r$-owners and we know how the single node deletion process works. Assuming that we need to delete more than $o(n)$ nodes might deterministically imply that there is some subgraph contained in the graph which can be shown to be unlikely to appear using the First Moment Method. We have discussed this idea on pp. 137 in Section 7.2.

The second best approach may consist in trying to 'translate' the branching tree version of the damage process (see Section 3.3.2) to random graphs. Further material concerning this can be found on pp. 140 in Section 7.2.

There may be a very faint hope that we could give up on considering $r$-neighbourhoods only, but still only retain some 'semi-local proofs of ownership', see pp. 139 in Section 7.2.

### Analysing the deletion process in the right way

The reader is encouraged to cross-check with Section 7.1. We have failed in analysing the invariance of the uniform distribution under the generalised degree sequence $\{d_I\}_{I\in\mathcal{I}}$ during the course of the deletion processes.

Formally, we have considered a *graph invariant*[†], here the generalised degree sequence, and attempted to show that invariance of the uniform distribution conditional on whatever value the graph invariant takes, remains invariant during repeated applications of some deletion protocol. Note that we are free to both choose practically any graph invariant and practically any deletion protocol, as long as we can prove what we want in the end. The existence of these two 'degrees of freedom' makes finding the right proof hard, if there is one at all.

Naturally, we have tried to single out 'sensible' invariants and deletion protocols, we have discussed this in detail in Section 7.1. We shall demonstrate in the next paragraph that there are still many alternatives left, that are also 'sensible'.

When considering the generalised degree sequence we classify graphs according to how many of its nodes have 1-neighbourhoods 'matching' a cer-

[†]In the strict mathematical sense: a mapping from the set of graphs to some other set such that isomorphic graphs are mapped to the same object

tain 'picture'[‡] $I$. We could use more complicated 'pictures', say, of the 2-neighbourhood. Apart from the fact that this leads to 'combinatorial explosion' we think that the counterexample portrayed in Figure 7.4 also rules out this possibility. Since we are dealing with a directed-edge deletion process, really, we might also consider classifying graphs according how many of its (directed, undirected?) *edges* 'match certain pictures'.

It may further be possible that both choices (deletion protocol and invariant) were 'right', and only our attempt to generalise the proof for the invariance of the uniform distribution was 'wrong', or that we do not even need uniform distribution amongst all directed graphs with a given invariant, in order to prove that the second phase a.a.s. terminates after $o(n)$ steps.

Having depicted this scenario in rather dark colours, we remind the reader of how strongly the empirical results and the 'branching tree analogue' (c.f. Section 3.3.2) suggest that the directed subgraphs in question 'behave completely analogous' to the $k$-core. Quite possibly there is just a simple idea lacking, but lacking it is.

### Specify 'almost uniquely colourable'

We have shown in Chapter 4, that the magic subgraph is not uniquely colourable, deterministically, but have presented empirical evidence, that magic subgraphs as found in random tripartite graphs are 'almost uniquely colourable'.

More precisely, we have observed, that two different random colourings found by independent runs of the Antivoter algorithm coincided on the magic subgraph (w.r.t. the 'built in' 3-partition) either *completely*, or differed occasionally only on a tiny set of nodes. Moreover, the nodes in the magic subgraph on which the colourings differed appeared to lie in small neighbourhoods of small cycles and the colours of those nodes appeared to be 'shifted along these cycles' in a specific way that we are unable to further describe.

The following questions may be a starting point for theoretical investigations in order to make sense of those vaguely stated observations.

- How do magic subgraphs with respect to different proper colourings of the same graph relate to each other?
- Do they, when they are magic subgraphs of a random graph *w.r.t. different colourings*, (almost) always 'coincide' in essentially the same node set, as suggested to be true by our simulations?

[‡]Remember that the indices $I$ of the generalised degree sequence may be regarded as unlabelled 'photographs' of feasible 1-neighbourhoods in directed graphs, see pp. 131

- If they do not 'essentially coincide', does this imply the existence of some subgraph that can be shown to be unlikely to occur in random graphs?

Finally one may ask whether generating magic subgraphs can be employed for generating uniquely colourable graphs of large girth, as discussed in Section 2.6.3.

### $k$-Partite vs. non-$k$-Partite

The magic subgraph is only well-defined *with respect to* a 'built-in' tripartition. What can we learn about the non-tripartite case?

We have discussed this question in Section 7.3. Yet, essentially we do not know, to what extent the $k$-partite and the non-$k$-partite model with the same average degree should be regarded comparable, in particular with respect to predicting critical average degrees.

If the critical degrees *were* comparable, there would be a 'contradiction', since the chromatic number would have jumped in the non-$k$-partite model at average degrees at which in the $k$-partite model no magic subgraph has 'yet' appeared for $k = 4, 5, 6$ which follows from the First Moment Bounds calculated in [AM99].

We have reproduced the following table from Section 7.3

| k | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|
| $c_{\text{ad hoc}}$ | 4.851 | 8.891 | 13.205 | 17.753 | 22.496 |
| $c_{AM}$ | 5.0434 | 9.1722 | 13.8958 | 19.0778 | 24.632 |
| $c_{BT}$ | 4.9108 | 9.267 | 14.035 | 19.112 | 24.434 |

The last line shows the critical values for the appearance of a giant magic subgraph in the $k$-partite model. The last but one line shows the upper bound on the $k$-colouring threshold from [AM99], for the non-$k$-partite model. The second line shows values predicted by a heuristic and fairly *ad hoc* guess, when a uniquely colourable subgraph in the non-$k$-partite model might appear, see Section 7.4.2 and below.

In the non-$k$-partite case, the advent of a giant subgraph that is uniquely $k$–colourable yields an upper bound on the $k$-colouring threshold, as a.s. adding $o(n)$ additional edges will destroy[§] $k$-colourability. But according to Molloy (see Questions 7.2 in [Mol01]) it is still feasible, though 'unlikely', that the chromatic number might jump *before* the appearance of a giant uniquely colourable subgraph that is not a triangle. The aforementioned

[§]A.s. some edge will connect nodes in the same colour class of the unique colouring.

'contradiction' could thus also be a hint that this 'unlikely' scenario is indeed true, for certain values of $k$.

A brief aside: In Chapter 4 we showed that there are large uniquely colourable graphs with an empty magic subgraph, that are 'built up from a triangle by adding cherries', see Lemma 4.2.6. Should they also be excluded from the statement of Question 7.2 in [Mol01]?

Since we have not presented a rigorous proof for the appearance of a giant magic subgraph it may be interesting to know that empirically, at least for the case $k = 4$, the 'branching tree value' seems to correctly predict the appearance of a giant magic subgraph.

All in all, we do not know enough. Certainly, answering the questions stated on p. 157 should help, because we do not know which magic subgraph is the 'right' one, in the presence of more than one ('canonical') colourings, and how those different magic subgraphs relate to each other.

It is also feasible that another subgraph similar to the magic subgraph appears in the non-$k$-partite model at smaller average degree. Possibly at the value we have referred to as '*ad hoc*' above, but this is purely speculative.

### Finding new subgraphs or sub-formulae?

We knew about the phase transition related to coloured owners *in branching trees* (Section 3.1) a long time before we actually 'found' the magic subgraph in random graphs. At least in this particular case we have demonstrated that phase transitions in branching trees may 'announce' the sudden appearance of certain giant subgraphs in 'corresponding' random graphs, the subgraphs being well defined objects.

We have discussed above that the advent of a giant uniquely colourable subgraph in the non-$k$-partite model would be a very interesting phenomenon to describe. Also, the magic subgraph seems somewhat resemblant the rather recent concept of a *backbone* in random SAT formulae, see [BBC+99], which is closely related to uniquely satisfiable sub-formulae. In Section 7.4.2 and Section 7.4.3 we have discussed recursive equations that *may* be interpreted as describing the appearance of a giant uniquely colourable subgraph in the non-$k$-partite model of random graphs with average degree $c$ and that of a giant uniquely satisfiable subgraph in random instances of 3-SAT with average clauses to variable ration $\alpha$, respectively.

This is just an educated guess, but note that the predicted values are consistent with all known upper and lower bounds on the respective thresholds.

See the second row ($c_{\text{ad hoc}}$) in the table on p. 158 for $k$-colouring. For 3-SAT the heuristic predicts a value of $\alpha_{\text{crit}} = 3.273(\ldots)$, whereas the best

known lower bound is 3.26(...). However empirical evidence has been reported that $\alpha_{\text{crit}} \simeq 4.2$.

We do not see any way of performing simulations that might confirm the appearance of such subgraphs/sub-formulae.

We think that one might be able to study *random satisfiable* 3-SAT *formulae* with average clauses to variable ratio $\alpha$ empirically, by implementing a hyper-edge deletion process analogous to the edge deletion characterising the magic subgraph. It would be certainly interesting to know what would happen.

In Section 7.4.2 we have described yet another branching tree heuristic, i.e. recursive equations, supported by empirical evidence in this case, that in the non-tripartite model there appears a giant magic subgraph w.r.t. an i.i.d. 3-colouring, not necessary proper, at $c = 4.16(\ldots)$. But we do not know what to make of this observation. Neither does its absence seem to imply an efficient colouring algorithm nor does it appear to be 'almost uniquely colourable'.

### Algorithmic implications

In the tripartite case we can imagine that the 'critical slowing down' of the Antivoter Algorithm is 'indicating' the advent of the magic subgraph. However, we are far from understanding even intuitively, why the advent of the magic subgraph should 'confuse' the Antivoter Algorithm, and even further from proving anything.

Yet, note our observations stated in Section 4.3. Apparently the hitting time of the Antivoter 'peaks' slightly *before* the magic subgraph appears. Moreover, we have noticed 'correlations' between colourings produced by independent runs of the Antivoter, for critical average degrees strictly less than 4.91(...). It may well be that the critical slowing down of the Antivoter and the appearance of the magic subgraph are not causally related. Then again, it is hard to extrapolate the empirical results performed on graphs with as little as 3000 or 6000 nodes. We have informally discussed this issue at the end of Section 4.3.

There are two interesting recent publications that we want to mention in this context. In [DF01] Dyer and Frieze are apparently able to show rapid mixing of the so called *Glauber dynamics* on random graphs, which is nothing but the Gibbs sampler ($\simeq$ Antivoter) described in Section 2.6.1 in *random graphs*. The other paper is [ABM01] by Achlioptas, Beame and Molloy. There it is shown that a certain backtracking algorithm takes exponential time to find a satisfying assignment for a random SAT instance with average degree *below* the 'generally accepted' threshold value.

### Abstraction

The reason that prevented us from writing down Conjecture 1 and Conjecture 2 as proper lemmas was the fact that otherwise we would have had to repeatedly go through our proofs laid out in Chapters 5 and 6 line by line, adapting things to a very similar but different scenario (e.g. random tripartite with average degree $c$, instead of random non-tripartite with average degree $c$), without essentially adding any new ideas.

One might want to consider defining a class of 'sparse random combinatorial structures admitting a Branching Tree Connection' that contains as special instances sparse random graphs, $k$-partite or not, sparse random configurations and possibly even random $k$-SAT-formulae. So, for example, there would be only one 'Generalised Semi-Local Lemma' to be shown, which should yield similar but different versions that seem plausible to hold for each specific random structure as corollaries.

### Fourier Analysis

The sharpness of both the threshold for $k$-colourability in random graphs and for random $k$-SAT was proved using Fourier analysis, see our Section 2.2. We have explained that the proof is indirect and therefore does not predict the threshold value, merely the assumption of a coarse threshold can be disproved.

In 2.2.3 we have sketched, following [FK96], how the Fourier-Walsh transform of the *local* property '$G$ contains a triangle' is asymptotically concentrated on the Fourier-coefficients that are (a disjoint union of) triangles. We have the intuition, possibly unjustified, that this reflects the fact that one only has to 'look at triangles' in order to decide the property.

We know that properties such as '$G$ contains a giant $k$-core' are *non-local*. Yet we have in some sense demonstrated that it 'suffices to look at small neighbourhoods of radius $r = r(n)$', which mostly are tree-like, in order to determine the apparently closely related property '$G$ contains a giant set of $r(n)$-owners'.

We should definitely leave this to the experts, but we dare to ask: is it feasible that the Fourier-Walsh transform of the property '$G$ contains a giant set of $r(n)$-owners', or even the property '$G$ contains a giant $k$-core', can be asymptotically well approximated by considering only Fourier-coefficients that are trees of radius $r(n)$? Could this lead to explicitely computing the threshold value for certain properties?

# Bibliography


[ABM01] D. Achlioptas, P. Beame, and M. Molloy. A sharp threshold in proof complexity. In *Proceedings of the 33th Annual Symposium on the Theory of Computing*, pages 337–346, New York, July 2001. ACM Press.

[AF99] D. Achlioptas and E. Friedgut. A sharp threshold for k-colorability. *Random Structures and Algorithms*, 14, 1999.

[AK94] N. Alon and N. Kahale. A spectral technique for coloring random 3-colorable graphs. In *Proceedings of the 26th Annual Symposium on the Theory of Computing*, pages 346–355, New York, May 1994. ACM Press.

[Alo98] N. Alon. Spectral techniques in graph algorithms. *Lecture Notes in Computer Science*, 1380:206–215, 1998.

[AM97] D. Achlioptas and M. Molloy. Analysis of a list–colouring algorithm on a random graph. In *Proceedings of the 38th Annual Symposium on Foundations of Computer Science*, pages 204–212, 1997.

[AM99] D. Achlioptas and M. Molloy. Almost all graphs with $2.522n$ edges are not 3-colorable. *The Electronic Journal of Combinatorics*, 6:#R29, 1999. `http://www.combinatorics.org/`.

[AN72] K.B. Athreya and P.E. Ney. *Branching Processes.* Springer, Heidelberg, 1972.

[AN97] K.B. Athreya and P.E. Ney. *Classical and Modern Branching Processes.* Springer, Heidelberg, 1997.

[AS92] N. Alon and J. H. Spencer. *The Probabilistic Method.* John Wiley and Sons, New York, 1992, 1992.

[Bax89] R.J. Baxter. *Exactly solved models in statistical mechanics.* Academic Press, London, 1989.

[BBC+99] B. Bollobás, C. Borgs, J. Chayes, J.H. Kim, and D.B. Wilson. The scaling window of the 2-Sat transition. Technical report, MS Research, 1999.

[BD97] R. Bubley and M. Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *38th Annual Symposium on Foundations of Computer Science*, Miami Beach, Florida, 20–22 October 1997. IEEE.

[Bol85] B. Bollobás. *Random Graphs.* Academic Press, London, 1985.

[BS76] B. Bollobás and N. Sauer. Uniquely colourable graphs with large girth. *Can. J. Math.*, 28:1340–1344, 1976.

[BS95] A. Blum and J. Spencer. Coloring random and semi-random $k$-colorable graphs. *Journal of Algorithms*, 19:204–234, 1995.

[BW99a] G. Brightwell and P. Winkler. Graph homomorphisms and phase transitions. *J. Comb. Theory (Series B)*, 77:415–435, 1999.

[BW99b] G. Brightwell and P. Winkler. Random colorings of a Cayley tree. Submitted., 1999.

[BW00] G. Brightwell and P. Winkler. Gibbs measures and dismantlable graphs. *J. Comb. Theory (Series B)*, 78:141–169, 2000.

[CGH+96] R.M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W function. *Advances in Computational Mathematics*, 5:329–359, 1996.

[DF01] M. Dyer and A. Frieze. Randomly colouring graphs with lower bounds on girth and maximum degree. Preprint, URL: `http://www.math.cmu.edu/~af1p/papers.html`, 2001.

[EHK98] T. Emden-Weinert, S. Hougardy, and B. Kreuter. Uniquely colourable graphs and hardness of colouring graphs of large girth. *Comb. Probab. Comput.*, 7(4):375–386, 1998.

[EHV00] J. Ernst, V. Heun, and U. Voll. A spectral approach to clustering. TU München, Germany, 2000.

[ER60] P. Erdős and A. Rény. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.

[Fel68] W. Feller. *An introduction to probability theory and its applications. Volume I.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Chichester-New York-Brisbane-Toronto-Singapore, 1968.

[FJ96] U. Feige and Kilian J. *Zero* knowledge and the chromatic number. In *Proc. of the 32$^{nd}$ IEEE FOCS*, 1996.

[FK96] E. Friedgut and G. Kalai. Every monotone graph property has a sharp threshold. *Proc. Amer. Math. Soc.*, 124:2993–3002, 1996.

[FM97] A. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Structures and Algorithms*, 10:2–42, 1997.

[GJ97] V.K. Gore and M. Jerrum. The Swendsen-Wang process does not always mix rapidly. In *Proceedings of the Eight Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 674–681. ACM/SIAM Press, 1997.

[GM00] A. Goerdt and M. Molloy. Analysis of edge deletion processes on faulty random regular graphs. In *LATIN 2000*, volume 1776 of *Lect. Notes Comput. Sci.*, pages 38–47, Berlin, 2000. Springer.

[HMRR98] M. Habib, C. McDiarmid, J. Ramirez-Alfonsin, and B. Reed. *Probabilistic Methods for Algorithmic Discrete Mathematics.* Springer, Heidelberg, 1998.

[Jan94] S. Janson. Coupling and poisson approximation. *Acta Appl. Math.*, 34:7–15, 1994.

[JKŁP93] S. Janson, D. E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4(3):233–358, 1993.

[JŁR00] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs.* Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley, New York, 2000.

[Kar90] R.M. Karp. The transitive closure of a random digraph. *Random Structures and Algorithms*, 1(1):73–92, 1990.

[KKK96] L.M. Kirousis, E. Kranakis, and D. Krizanc. Approximating the unsatisfiability threshold of random formulas. In *ESA: Annual European Symposium on Algorithms*, 1996.

[KKS00] A.C. Kaporis, L.M. Kirousis, and Y.C. Stamatiou. A note on the non-colorability threshold of a random graph. *Electron. J. Comb.*, 7(1), 2000. Research paper R29.

[KMPS95] A. Kamath, R. Motwani, K. Palem, and P. Spirakis. Tail bounds for occupancy and the satisfiability threshold conjecture. *Random Structures and Algorithms*, 7, 1995.

[KZ95] R.M. Karp and Y. Zhang. Bounded branching process and AND/OR tree evaluation. *Random Structures and Algorithms*, 7(2):97–18, 1995.

[Mol96] M. Molloy. A gap between the appearances of a k-core and a (k+1)-chromatic graph. *Random Structures and Algorithms*, 8, 1996.

[Mol01] M. Molloy. Thresholds for colourability and satisfiability in random graphs and boolean formulae. In *Surveys in Combinatorics*, pages 165–197, 2001.

[MR95a] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures and Algorithms*, 2(3):161–179, 1995.

[MR95b] R. Motwani and P. Raghavan. *Randomized Algorithms*. Press Syndicate of the University of Cambridge, Cambridge, 1995.

[MR98] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Comb. Probab. Comput.*, 7(3):295–305, 1998.

[MR99] M. Molloy and B. Reed. Critical subgraphs of a random graph. *The Electronic Journal of Combinatorics*, 6:#R35, 1999. `http://www.combinatorics.org/`.

[PSW96] B. Pittel, J. Spencer, and N. Wormald. Sudden emergence of a giant $k$-core in a random graph. *Journal of Combinatorial Theory Series B*, 67:111–151, 1996.

[PW89] A.D. Petford and D.J.A. Welsh. A randomised 3-colouring algorithm. *Discrete Mathematics*, 74:253–261, 1989.

[Rue78] D. Ruelle. *Thermodynamic formalism*, volume 5 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1978.

[Tal98] M. Talagrand. Huge random structures and mean field models for spin glasses. *Doc. Math., J. DMV, Extra Vol. ICM Berlin 1998*, I:507–536, 1998.

[Wor95] N.C. Wormald. Differential equations for random processes and random graphs. *Annals of Applied Probability*, 5:1217–1235, 1995.

[Wor99] N.C. Wormald. The differential equation method for random graph processes and greedy algorithms. In M. Karonski and H.J. Proemel, editors, *Lectures on Approximation and Randomized Algorithms*, pages 73–155. PWN, Warsaw, 1999.

# List of Symbols and Abbreviations

| | |
|---|---|
| $G = (V, E)$ | Graph with edge set $E$ and node set $V$ |
| $m$, $n$, $c$ | Usually, $\|E\| = m$ edges and $\|V\| = n$ vertices, average degree is $c = 2 \cdot m/n$ |
| $(d_i(G))_{i \geq 0}$ | Degree sequence of $G$ |
| $(D_i(G))_{i \geq 0}$ | Integrated degree sequence of $G$, $D_i := \sum_{j \geq i} d_j$ |
| $\chi(G), \gamma(G)$ | Chromatic number of $G$, girth of $G$ |
| $K_n, C_n$ | Complete graph on $n$ vertices, cycle on $n$ vertices |
| $\mathcal{G}_n$ | Set of all graphs on $n$ nodes |
| $\mathcal{G}_{n,p}$, $\mathcal{G}_{n,m}$ | Random graphs, see Subsection 2.1.1 |
| $\mathbb{P}[\mathcal{A}]$, $\mathbb{P}_c[\mathcal{A}]$ | Prob. of event $\mathcal{A}$ on random graphs with average degree $c$ |
| $X = X^{(n)}$, $Y$ | Integer random variables, possibly depending on $n$ |
| $\mathbb{E}[X]$, $\mathbb{E}_c[X]$ | Expectation or first moment of $X$ |
| $\mathrm{Po}_c(\cdot), \mathrm{Bi}(n, c, \cdot)$ | Poisson and binomial distributions |
| (a.)a.s. | (Asymptotically) almost surely |
| i.i.d. | Identically and independently distributed |
| $O$, $o$, $\theta$, $\sim$ | Landau symbols ($n \to \infty$, usually, sometimes $r \to \infty$) |
| $W_{-1}$ | (Branch of the) Lambert-$W$-function, see [CGH$^+$96] |
| $W_r$ | Relative size $Z_r/\mathbb{E}[\mu]^r$ of the $r$-th shell in a $\mu$-branching tree. |