Diss. ETH No. 17088

Random Graphs with Structural Constraints

A dissertation submitted to the

Swiss Federal Institute of Technology Zurich

for the degree of

Doctor of Sciences

presented by

Andreas Christian Weissl

Diplom-Informatiker, Technische Universität München

born 18.04.1978

citizen of Germany

accepted on the recommendation of

Prof. Dr. Angelika Steger, examiner

Prof. Dr. Marc Noy, co-examiner

2007

To my parents

Contents

Abstract

The classical way to measure the performance of an algorithm is to consider its worst-case performance. However, in practice this performance measure is often overly pessimistic. This stems from the fact that the worst-case analysis is frequently based on few inputs, which might have a quite artificial structure that does not (or seldom) appear as part of a "typical" input instance.

A natural alternative is to consider the average-case performance of an algorithm, that is, we analyse the performance of an algorithm assuming that the input instances are drawn from all possible instances according to a given probability distribution. In order to analyse the average-case performance of an algorithm and to design algorithms with better average-case performance, it is crucial to understand the properties of a "typical" input instance.

In many real-world scenarios – for example in chip manufacturing and drawing of diagrams – one has to deal with algorithms that take graphs with structural constraints, such as planar graphs, as input. Investigating properties of such constrained graph classes and developing new tools and methods, which help to cope with the difficulty of the dependence of the edges, are central to advance the state of the art in this area of research.

In this thesis, we focus on such constrained graph classes, namely planar graphs with given average degree, cactus graphs, block graphs, and (maximal) $K_{3,3}$-minor-free graphs. We are interested to prove that graphs on n nodes, drawn uniformly at random from the set of all graphs on n nodes of a class with structural constraints, have specific properties with high probability (w.h.p., i.e., with probability tending to 1 as $n \to \infty$). We apply different techniques and provide a new method to obtain results about several graph classes with structural constraints.

Our first result is about planar graphs. Planar graphs are well-known and well-studied combinatorial objects in graph theory. Roughly speaking, a graph is planar if it can be drawn in the plane in such a way that no two edges cross. A *random planar graph* R_n is drawn uniformly at random from the set $\mathcal{P}(n)$ of all simple labelled planar graphs on the node set $\{1, \ldots, n\}$. Here we consider planar graphs with given average degree. More precisely, we are interested in properties of a random planar graph $R_{n,q}$ which is drawn uniformly at random from the class $\mathcal{P}(n, \lfloor qn \rfloor)$ of simple labelled planar graphs with n nodes and $\lfloor qn \rfloor$ edges, where $1 < q < 3$ and the average degree is about $2q$.

We exploit proof techniques of McDiarmid, Steger, and Welsh [**MSW05, MSW06**], to show that for all $1 < q < 3$ the random planar graph $R_{n,q}$ has properties similar to those of a random planar graph R_n. For example, we show that $R_{n,q}$ contains w.h.p. linearly many nodes of each given degree and linearly many node disjoint copies of each given fixed connected planar

graph. Additionally, we investigate the number of automorphisms and we give lower bounds on the maximum node degree and the maximum face size of any plane embedding. We also study the relation between the growth constants of labelled and unlabelled planar graphs with given average degree.

Some of these results were strengthened and many additional results about planar graphs were shown by Giménez and Noy [**GN05b**] using a completely different approach. We also show how our work relates to these newer results.

A key tool of the method is to show that $(|\mathcal{P}(n, \lfloor qn \rfloor)|/n!)^{1/n}$ tends to a limit $\gamma(q)$ as n tends to infinity. We also study the behaviour of this function in more detail. In particular we consider $\gamma(q)$ as q approaches 1 from below and above and 3 from below, and explain the discontinuity as we approach 1 from below by changing scale appropriately.

In our second result we use a new method based on Boltzmann samplers [**DFLS04**] to investigate two moderately complex graph classes with constraints, namely cactus and block graph. A cactus graph is a labelled connected graph, in which each edge is contained in at most one cycle; and a block graph is a labelled connected graph, whose maximal biconnected blocks are cliques.

Additionally, exploiting the (more classical) generating function framework, we asymptotically enumerate the graphs in both classes, derive limit laws for the number of edges, and explore several other properties. Finally, we show how these results can be applied to straightforwardly derive the average-case complexity of two longest path algorithms for both graph classes.

The new approach may be used to investigate properties of "typical" members of constrained graph classes, which seem to be inherently difficult to handle with the analytic combinatorics approach, for instance properties which cannot be directly addressed with (a finite number of) parameters of the generating functions, such as the maximum node degree, or the maximum size of a biconnected block.

Our last result concerns simple labelled $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs, where maximal means that adding any edge to such a graph yields a $K_{3,3}$-minor.

We apply generating function techniques to obtain asymptotic estimates for the number of (maximal) $K_{3,3}$-minor-free graphs. For $K_{3,3}$-minor-free graphs we consider 3-connected, 2-connected, connected and not necessarily connected graphs. We also derive limit laws for several parameters.

Due to Kuratowski's theorem [**Kur30**] planar graphs are $K_{3,3}$- *and* K_5-minor-free. Hence, the class of planar graphs is contained in the class of $K_{3,3}$-minor-free graphs. Due to Wagner's theorem [**Wag37**] the class of triangulations (with the exception of all triangulations on 5 nodes) is contained in the class of maximal $K_{3,3}$-minor-free graphs. Determining the number (of graphs of sub-classes) of planar graphs has attracted considerable attention [**BGW02, GN04, GN05b, BGKN05, BLMK**] in recent years. Giménez and Noy [**GN05b**] obtained precise asymptotic estimates for the number of planar graphs. The asymptotic number of triangulations was given by Tutte [**Tut62**] already in 1962. Investigating how much the

number of planar graphs (triangulations) differs from (maximal) $K_{3,3}$-minor-free graphs is a first important step in examining how "typical" instances of these graph classes will differ.

For 3-connected $K_{3,3}$-minor-free graphs, the change from planar graphs can be easily described: it follows from a theorem of Wagner [**Wag37**] that the set of 3-connected $K_{3,3}$-minor-free graphs consists of all 3-connected planar graphs and the complete graph on 5 nodes. Thus, on this connectivity level our graph class differs only in the existence of one additional graph from planar graphs. But as we shall show, adding K_5 to the set of 3-connected graphs yields a slightly larger exponential growth rate for 2-connected, connected, and not necessarily connected $K_{3,3}$-minor-free graphs. It also slightly changes other parameters, for instance the expected number of edges in a random $K_{3,3}$-minor-free graph. For maximal $K_{3,3}$-minor-free graphs the growth rate also increases slightly compared to triangulations.

To establish these results for $K_{3,3}$-minor-free graphs, we follow the approach taken for planar graphs [**BGW02, GN05b**], i.e., we follow a well-known decomposition along the connectivity structure of a graph and translate this decomposition into relations of our generating functions. This is possible as the decomposition for $K_{3,3}$-minor-free graphs, which is due to Wagner [**Wag37**], fits well into this framework. For maximal $K_{3,3}$-minor-free graphs the situation is different, as the decomposition, which is again due to Wagner, has further constraints. The functional equations for the generating functions of *edge-rooted* maximal graphs are easy to obtain but in order to go to unrooted graphs, special integration techniques based on rational parametrisation of rational curves are needed.

Zusammenfassung

Die klassische Art die Güte eines Algorithmus zu messen, ist seine Worst-case Performance zu betrachten. In der Praxis zeigt sich jedoch, dass dieses Gütemaß oft übermäßig pessimistisch ist. Der Hauptgrund dafür ist, dass die Worst-case Analyse häufig auf wenigen Eingaben beruht, die eine sehr künstliche Struktur besitzen, welche nicht (oder nur selten) in einer „typischen“ Eingabeinstanz auftritt.

Eine natürliche Alternative stellt die Average-case Analyse dar. Dabei wird die Güte eines Algorithmus unter der Annahme analysiert, dass die Eingabeinstanzen zufällig aus allen möglichen Instanzen entsprechend einer gegebenen Wahrscheinlichkeitsverteilung gezogen werden. Für die Average-case Analyse von Algorithmen und die Entwicklung von Algorithmen mit besserer Average-case Güte ist es entscheidend, zu verstehen, welche Eigenschaften „typische“ Eingabeinstanzen besitzen.

In vielen praktischen Anwendungen – zum Beispiel im Bereich der Chipherstellung und beim Zeichnen von Diagrammen – beschäftigt man sich mit Algorithmen, die Graphen mit strukturellen Einschränkungen, wie z.B. planare Graphen, als Eingabe entgegennehmen. Die Untersuchung von Eigenschaften solcher Graphklassen mit strukturellen Einschränkungen, sowie die Entwicklung neuer Werkzeuge und Methoden, die helfen die Schwierigkeiten, die durch die Abhängigkeit der Kanten entstehen, zu überwinden oder zu umgehen, sind von zentraler Bedeutung für den Fortschritt in diesem Bereich der Forschung.

In dieser Arbeit konzentrieren wir uns auf solche Graphklassen mit strukturellen Einschränkungen und zwar auf planare Graphen mit gegebenem Durchschnittsgrad, Kaktusgraphen, Blockgraphen und (maximale) $K_{3,3}$-Minor-freie Graphen. Wir werden beweisen, dass ein Graph auf n Knoten, der zufällig und gleichverteilt aus allen Graphen auf n Knoten einer Graphklasse mit strukturellen Einschränkungen gezogen wird, bestimmte Eigenschaften mit hoher Wahrscheinlichkeit (d.h. mit Wahrscheinlichkeit gegen 1 strebend für $n \to \infty$) hat. Wir wenden verschiedene Beweistechniken an und stellen eine neue Methode zur Verfügung, um Resultate über mehrere Graphklassen mit strukturellen Einschränkungen zu erhalten.

In unserem ersten Resultat betrachten wir planare Graphen. Planare Graphen sind kombinatorische Objekte, die in der Graphentheorie seit langem intensiv untersucht werden. Grob gesprochen ist ein Graph planar, wenn er in der Ebene gezeichnet werden kann, ohne dass sich zwei Kanten kreuzen. Ein zufälliger planarer Graph R_n wird zufällig und gleichverteilt aus der Menge $\mathcal{P}(n)$ aller einfachen nummerierten planaren Graphen auf der Knotenmenge $\{1, \ldots, n\}$ gezogen. Hier betrachten wir planare Graphen mit gegebenem Durchschnittsgrad. Genauer interessieren wir uns für Eigenschaften eines zufälligen planaren Graphen $R_{n,q}$, der zufällig und gleichverteilt aus der Klasse $\mathcal{P}(n, \lfloor qn \rfloor)$ aller einfachen nummerierten planaren

Graphen auf n Knoten mit $\lfloor qn \rfloor$ Kanten gezogen wird, wobei $1 < q < 3$ und der Durchschnittsgrad ungefähr $2q$ ist.

Wir nutzen Beweistechniken von McDiarmid, Steger und Welsh [**MSW05, MSW06**], um für alle $1 < q < 3$ zu zeigen, dass der zufällige planare Graph $R_{n,q}$ ähnliche Eigenschaften hat wie der zufällige planare Graph R_n. So zeigen wir unter anderem, dass $R_{n,q}$ mit hoher Wahrscheinlichkeit linear viele Knoten von jedem festen gegebenen Grad und linear viele knotendisjunkte Kopien jedes festen zusammenhängenden planaren Graphen enthält. Desweiteren untersuchen wir die Anzahl der Automorphismen. Außerdem geben wir untere Schranken für den maximalen Knotengrad und die maximale Größe eines Gebiets in jeder planaren Einbettung an. Zudem untersuchen wir die Beziehung zwischen der Wachstumskonstante von nummerierten und nicht nummerierten planaren Graphen mit gegebenem Durchschnittsgrad.

Giménez and Noy [**GN05b**] haben einige dieser Resultate verstärkt und neue Ergebnisse gezeigt, wobei sie einen gänzlich anderer Ansatz verwendet haben. Wir zeigen auch auf, welcher Bezug zwischen unserer Arbeit und diesen neueren Resultaten besteht.

Wesentlich ist bei der verwendeten Methode, zu zeigen, dass $(|\mathcal{P}(n, \lfloor qn \rfloor)|/n!)^{1/n}$ für n gegen unendlich gegen einen Grenzwert $\gamma(q)$ strebt. Wir untersuchen das Verhalten dieser Funktion im Detail. Insbesondere betrachten wir $\gamma(q)$ für q gegen 1 von unten und gegen 3 von oben strebend und erklären die Diskontinuität, wenn wir uns 1 von unten nähern, indem wir die Skalierung entsprechend anpassen.

In unserem zweiten Resultat verwenden wir eine neue Methode basierend auf Boltzmann Samplern [**DFLS04**], um zwei komplexere Graphklassen mit strukturellen Einschränkungen zu untersuchen und zwar Kaktus- und Blockgraphen. Ein Kaktusgraph ist ein nummerierter zusammenhängender Graph, in dem jede Kante in höchstens einem Kreis enthalten ist und ein Blockgraph ist ein nummerierter zusammenhängender Graph, in dem jeder maximale zweifach zusammenhängende Block eine Clique ist.

Außerdem zählen wir asymptotisch, unter Verwendung des (klassischen) Ansatzes mit erzeugenden Funktionen, die Graphen beider Klassen, leiten ein Grenzwertgesetz für die Anzahl Kanten her und untersuchen mehrere andere Eigenschaften. Zuletzt zeigen wir, wie diese Resultate angewendet werden können, um in direkter Weise die Average-case Komplexität zweier Längste-Pfade Algorithmen für beide Klassen herzuleiten.

Der neue Ansatz kann verwendet werden, um Eigenschaften „typischer“ Mitglieder von Graphklassen mit strukturellen Einschränkungen zu untersuchen, die inhärent schwierig mit dem Ansatz der analytischen Kombinatorik handhabbar zu sein scheinen, z.B. Eigenschaften, die nicht direkt mit (einer endlichen Anzahl von) Parametern der erzeugenden Funktionen beschrieben werden können. Beispiele sind der maximale Knotengrad oder die maximale Größe eines zweifachzusammenhängenden Blocks in einem Graph.

Unser letztes Resultat beschäftigt sich mit einfachen nummerierten $K_{3,3}$-Minor-freien und maximalen $K_{3,3}$-Minor-freien Graphen, wobei maximal bedeutet, dass das Hinzufügen einer beliebigen Kante zu einem solchen Graphen zur Bildung eines $K_{3,3}$-Minor führt.

Wir wenden Techniken mit erzeugenden Funktionen an, um die asymptotische Anzahl (maximaler) $K_{3,3}$-Minor-freier Graphen zu erhalten. Für $K_{3,3}$-Minor-freie Graphen betrachten wir dreifachzusammenhängende, zweifachzusammenhängende, zusammenhängende und nicht notwendigerweise zusammenhängende Graphen. Außerdem leiten wir Grenzwertgesetze für mehrere Parameter her.

Nach Kuratowskis Theorem [**Kur30**] sind planare Graphen $K_{3,3}$- *und* K_5-Minor-frei. Daher ist die Klasse der planaren Graphen in der Klasse der $K_{3,3}$-Minor-freien Graphen enthalten. Nach Wagners Theorem [**Wag37**], ist die Klasse der Triangulierungen (mit Ausnahme aller Triangulierungen auf 5 Knoten) in der Klasse der maximalen $K_{3,3}$-Minor-freien Graphen enthalten. Der Bestimmung der Anzahl (der Graphen von Subklassen) planarer Graphen kam in den letzten Jahren beträchtliche Beachtung zu [**BGW02, GN04, GN05b, BGKN05, BLMK**]. Giménez und Noy haben asymptotische Näherungen für die Anzahl planarer Graphen hergeleitet [**GN05b**]. Die asymptotische Anzahl von Triangulierungen wurde von Tutte [**Tut62**] bereits 1962 hergeleitet. Zu untersuchen, wieviel sich die Anzahl planarer Graphen (Triangulierungen) von der Anzahl (maximaler) $K_{3,3}$-Minor-freier Graphen unterscheidet ist ein erster wichtiger Schritt, um zu erforschen, wie sich typische Instanzen dieser Graphklassen unterscheiden.

Für dreifachzusammenhängende $K_{3,3}$-Minor-freie Graphen ist der Unterschied zu planaren Graphen einfach zu beschreiben: es folgt aus einem Theorem von Wagner [**Wag37**], dass sich die Menge der dreifachzusammenhängenden $K_{3,3}$-Minor-freien Graphen aus allen dreifachzusammenhängenden planaren Graphen und dem vollständigen Graphen auf 5 Knoten zusammensetzt. Somit unterscheidet sich unsere Graphklasse auf dieser Konnektivitätsstufe nur in der Existenz eines zusätzlichen Graphen von planaren Graphen. Wie wir aber zeigen, führt das Hinzufügen des K_5 zu dreifachzusammenhängenden Graphen zu einer geringfügig größeren exponentiellen Wachstumsrate für zweifachzusammenhängende, zusammenhängende und nicht notwendigerweise zusammenhängenden $K_{3,3}$-Minor-freie Graphen. Es ändern sich auch weitere Parameter geringfügig, wie z.B. die erwartete Anzahl von Kanten in einem zufälligen $K_{3,3}$-Minor-freien Graph. Für maximale $K_{3,3}$-Minor-freie Graphen ändert sich die Wachstumsrate ebenfalls geringfügig verglichen mit Triangulierungen.

Um diese Ergebnisse für $K_{3,3}$-Minor-freie Graphen herzuleiten, verwenden wir den Ansatz, der bei planaren Graphen [**BGW02, GN05b**] angewandt wurde, d.h. wir folgen einer wohl bekannten Zerlegung eines Graphen entlang seiner Konnektivitätsstruktur und übersetzen diese in Beziehungen unserer erzeugenden Funktionen. Dies ist möglich, da die Zerlegung für $K_{3,3}$-Minor-freie Graphen, die auf Wagner [**Wag37**] zurückgeht, gut in dieses Rahmenwerk passt. Für maximale $K_{3,3}$-Minor-freie Graphen ist die Situation verschieden, da die Zerlegung, die wiederum auf Wagner zurückgeht, zusätzlichen Beschränkungen unterliegt. Die funktionalen Gleichungen für die erzeugenden Funktionen von an Kanten gewurzelten maximalen Graphen sind zwar einfach herzuleiten, aber um zu ungewurzelten Graphen zu gelangen sind besondere Integrationstechniken notwendig, die auf der rationalen Parametrisierung rationaler Kurven basieren.

Acknowledgements

First and foremost, I wish to express my deep gratitude to my supervisor Angelika Steger. With her interesting lectures and challenging projects she attracted my attention to theoretical computer science and brought me closer to this area. I owe her the opportunity to do research in one of the most exciting areas of computer science. Her ambition to solve problems inspired and encouraged me on my way to finish this work. I am also grateful for the freedom to work on many different topics.

Many thanks go to Marc Noy for introducing me to the world of generating functions and analytic combinatorics. I am honoured that he is the co-examiner of this thesis.

I sincerely appreciate the collaboration with my coauthors Stefanie Gerke, Omer Giménez, Martin Marciniszyn, Colin McDiarmid, Marc Noy, Konstantinos Panagiotou, Jan Remy, Justus Schwartz, Reto Spöhel, and Angelika Steger. I had interesting discussions with many guests in our research group, particularly memorable ones with Mihyun Kang and Eric Fusy.

In particular, I want to thank Stefanie for her support, countless discussions, endless hours of proof-reading, her incredible patience, the good humour, and many, many cookies. I am particularly obliged to Konstantinos Panagiotou for being such a good friend, for his patience, and for making research and solving problems so much fun. I will never forget the fun, weird discussions, and endless calculations in the evenings.

The past years within Angelika's research group were extremely enjoyable. I would like to thank my colleagues for their invaluable help and the good time we spent together. Special thanks go to my office mates Stephan Micklitz in Munich and Martin Marciniszyn in Zurich. Also, many thanks to Martin for his advice how to master the administrative issues of the Ph.D. program at ETH and his excellent LaTeX support. I also want to thank Justus Schwartz for many enjoyable coding sessions and his technical support when faced with Linux problems. I thank the whole group (and former members) – Nicla Bernasconi, Joshua Cooper, Stefanie Gerke, Manuel Huber, Florian Jug, Julian Lorenz, Martin Marciniszyn, Konstantinos Panagiotou, Jan Remy, Justus Schwartz, Alex Souza, Reto Spöhel, Angelika Steger – for the lively, amicable, and motivating atmosphere, the entertaining cooking and game sessions in the evening, and for many contested matches at the "Töggelichaschte".

I also want to thank my friends for their support and encouragement while working on my thesis. I especially want to thank Jutta Kiener, Manuel Huber, Stephan Micklitz, and Chris Hodges.

Finally, I want to thank my family and especially my parents for their love and invaluable support at any time.

Andreas Weißl

Zurich, 15 December 2006

CHAPTER 1

Introduction

1.1. Average Case Analysis

The analysis (and development) of algorithms is a fundamental and well-studied area of computer science. The classical way to measure the performance of an algorithm, such as the running time or memory consumption, is to consider the *worst-case* performance of an algorithm, i.e., to analyse the behaviour of an algorithm for the most ill-behaved input. A clear advantage of this approach and probably the main reason why research focused on worst-case analysis in the past decades, is that worst-case bounds guarantee that under all circumstances the algorithm will need at most the estimated quantity of the considered resources.

However, the worst-case performance measure also has a major drawback. In practice it turns out that this performance measure is often overly pessimistic. This might result in favouring an algorithm which is, despite of its theoretical superiority, not very efficient in practice. The difference of the theoretical prediction and the observed behaviour stems from the fact that the worst-case analysis is often based on few inputs, which might have a quite artificial structure that does not (or seldom) appear as part of a "typical" input instance.

The sorting algorithm QuickSort is probably the best-known example, where one can observe this kind of behaviour. The algorithm selects a *pivot* element from the sequence to be sorted, divides the sequence into elements that are smaller and elements that are larger or equal than the pivot element, and recursively proceeds on these two sub-sequences; finally, the recursion stops on a trivial sequence of length at most two. It can be shown that for every deterministic pivot selecting strategy there exist input sequences of length n that yield a running time of $\Theta(n^2)$. Compared to other sorting algorithms, as for instance MergeSort, which have a worst-case running time of $\mathcal{O}(n \log n)$, QuickSort appears to be an inefficient and thus "bad" algorithm. Contrary to that, in practice QuickSort is considered to be one of the best comparison based sorting algorithms. This empirical observation was supported by a probabilistic analysis (see e.g., Motwani and Raghavan [**MR95**]). If all input sequences are considered to be equally probable, then the probability for the worst-case behaviour to occur is very unlikely and a good running time of $\mathcal{O}(n \log n)$ can be expected. From an algorithm design perspective this *average-case* result yields that we may want to randomise the pivot selecting strategy or to permute the input sequence appropriately to obtain a good average-case performance, even if no assumptions can be made about the distribution of input sequences.

Furthermore, while for sorting there also exist efficient algorithms in a worst-case sense, for most combinatorial optimisation problems the situation is different. In the 1970's Cook,

Karp, Garey, and Johnson [**Coo71, GJ79, Kar72**] developed with their seminal work on $\mathcal{NP}$-completeness a rich set of tools, which allow to characterise the solvability of a specific problem. If we prove that a problem is $\mathcal{NP}$-hard, we classify it as being very difficult to solve, meaning that the existence of a polynomial (worst-case) time algorithm is very unlikely; hence, such a proof usually terminates the search for a polynomial time algorithm. From a practical point of view, this is rather dissatisfying: many important real-world problems, which have enormous economic potential and impact and whose solution is hence inevitable, fall into this class. Nevertheless, it may be possible to design polynomial-time algorithms that calculate good solutions in the "typical" case. That is, we are looking for algorithms with good expected running time.

From the above discussion it follows that a natural alternative to the worst-case performance measure is to consider the average-case performance of an algorithm; that is, we analyse the performance of an algorithm assuming that the input instances are drawn from all possible instances according to a certain probability distribution.

A plausible approach, taken in this thesis, is to consider all input instances of a given size to be equally likely. Although at first sight this might appear to be a quite strong restriction, it is often the only reasonable choice: in practice we frequently don't have enough information about the distribution of the input instances. Moreover, the lack of tools and techniques to perform an average-case analysis – the example of QuickSort above is one of the rare cases where one could carry out an average-case analysis – leads to study such "simplified" models first.

1.2. An Example - Graph Colouring

From the discussion before, it is clear that, in order to analyse the average-case performance of an algorithm, we have to choose a probability distribution on the class of possible inputs. If we consider graph algorithms and if we assume a uniform distribution over all labelled graphs on n nodes, it is easy to see that this model is equivalent to analysing the behaviour of the algorithm in question with respect to *random graphs*. The theory of random graphs started in the middle of the last century with the pioneering work of Erdős and Rényi. In the widely known $G(n,p)$ model each edge in graph G on n nodes is present with probability p independently of all other edges; for a very good introduction to the theory of random graphs see [**Bol85**] and [**JŁR00**].

In order to analyse the average-case performance of graph algorithms and to improve them, many algorithms working on random graphs were investigated. Let us briefly sketch such an algorithm for determining the chromatic number $\chi(G)$ of a given graph G – that is the minimum number of colours needed to colour the nodes of a graph such that no adjacent nodes have the same colour – a problem which is $\mathcal{NP}$-hard in general. This problem is important in graph theory itself, but also appears frequently as a core problem within numerous practical applications. In 1984 Wilf [**Wil84**] already presented an algorithm, which decides for every given (positive) integer l and graph G in constant expected time whether $\chi(G) \leq l$ – assuming that all graphs are equally likely. If $\chi(G) \leq l$ the algorithm also determines a colouring with exactly $\chi(G)$ colours. The reason why this result holds is that almost all graphs contain a

large number of $(l+1)$-cliques and hence are not l-colourable. With very high probability it suffices to investigate only a small part of the given graph until one finds a $(l+1)$-clique and therefore a certificate for the fact that G is not l-colourable. As the cases that such a clique is not found are so unlikely, one can afford to use an exhaustive search over all possible l-colourings to obtain a proper colouring or to determine that the graph is not l-colourable. Thus, exploiting characteristics of the underlying graph model yields a constant expected time algorithm.

1.3. Graphs with Structural Constraints

Similar to the example in the previous section, it turns out that in general for the analysis of the average-case performance and the design of algorithms with better average-case performance it is crucial to understand the properties of a "typical" input instance. Hence, for graph algorithms we have to investigate which properties a graph, drawn uniformly at random from the set of all graphs on n nodes within the considered class, has.

Although the random graph model $G(n,p)$ introduced in Section 1.2, has led to a fruitful branch in combinatorics and made the average-case analysis of many graph algorithms possible, in many real-world applications it does not appropriately model realistic input instances. In real-world scenarios, one rather has to deal with graphs with structural constraints, such as planar graphs. Numerous applications with such inputs exist, for example chip manufacturing, drawing of diagrams and many more. Unfortunately, many optimisation problems remain $\mathcal{NP}$-hard even if the input is restricted to graph classes with structural constraints. For example, the problem to decide whether one needs three or four colours to colour a planar graph is still $\mathcal{NP}$-hard (from the celebrated four-colour theorem [**AH77, AHK77, AH89**] it follows that we don't need more than four colours).

Performing an average-case analysis for an algorithm, which takes constrained graphs as input, turns out to be rather difficult. This has several reasons, the main two being the following. First, there does not exist such a rich set of tools and methods to prove structural properties for "typical" members of these classes, as for example for the traditional random graph $G(n,p)$. Second, we have to cope with the difficulty of the dependence of the edges which makes the analysis in many cases technically quite involved.

Hence, investigating properties of constrained graph classes and developing new tools and methods to analyse such graph classes are central to advance the state of the art in this area of research.

1.4. Overview

In this thesis, we focus on graphs with structural constraints as input instances, for example planar graphs. More precisely, we are interested to prove that graphs on n nodes, drawn uniformly at random from the set of all graphs on n nodes of a class with structural constraints, have specific properties with high probability (w.h.p., i.e., with probability tending to 1 as $n \to \infty$). To achieve this, there exist several methods; central to all is the enumeration of the

graphs in the class in interest. In this thesis, we apply such techniques and provide a new method to obtain results about several graph classes with structural constraints.

In the first part, we use a purely combinatorial approach, following McDiarmid, Steger, and Welsh [**MSW06**], to investigate planar graphs with given average degree. In the second part, we use analytic combinatorics and generating function techniques to derive the asymptotic number of cactus graphs, block graphs, $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs as well as to investigate several of their properties. We also present a new method based on the framework of Boltzmann samplers – recently introduced by Duchon, Flajolet, Louchard, and Schaeffer [**DFLS04**] – which may be used as an alternative to known methods to show properties of "typical" members of a class under consideration. The new approach seems especially useful for certain parameters which are difficult (or impossible) to handle with known methods. We illustrate how this approach works on cactus and block graphs. In the sequel, we give a more detailed overview over the subsequent chapters.

Chapter 2 - Tools and Techniques

In this chapter we briefly describe the two major approaches in literature for analysing constrained graph classes, namely the combinatorial and the analytic combinatorics approach. Moreover, we give a brief introduction into a framework for sampling combinatorial objects, which we will use in a subsequent chapter to develop a new method for investigating structural properties of random constrained graphs. We briefly present the main idea of this new approach, too. The main focus of this chapter is on the exposition of the ideas behind the different methods and on providing theorems from literature as a basis for the later chapters.

Chapter 3 - Planar Graphs with given Average Degree

Planar graphs are well-known and well-studied combinatorial objects in graph theory. Roughly speaking, a graph is planar if it can be drawn in the plane in such a way that no two edges cross. A random planar graph R_n is a simple labelled planar graph that is drawn uniformly at random from the set $\mathcal{P}(n)$ of all simple planar graphs on the node set $\{1, \ldots, n\}$.

In this chapter we consider planar graphs with given average degree. More precisely, we are interested in properties of a random planar graph $R_{n,q}$ which is drawn uniformly at random from the class $\mathcal{P}(n, \lfloor qn \rfloor)$ of simple labelled planar graphs with n nodes and $\lfloor qn \rfloor$ edges, where $1 < q < 3$.

We show that for all $1 < q < 3$ the random planar graph $R_{n,q}$ has properties similar to those of a random planar graph R_n, which was investigated in [**DVW96, MSW05, GN05b**]. For example, we show that $R_{n,q}$ contains w.h.p. linearly many nodes of each given degree and linearly many node disjoint copies of each given fixed connected planar graph. Additionally, we investigate the number of automorphisms and we give lower bounds on the maximum node degree and the maximum face size of any plane embedding. We also study the relation between the growth constants of labelled and unlabelled planar graphs with given average degree.

Some of these results were strengthened and many additional results about planar graphs were shown by Giménez and Noy [**GN05b**] using a completely different approach. We also show how our work relates to these newer results.

The proof techniques exploited in this chapter are purely combinatorial and follow McDiarmid, Steger, and Welsh [**MSW05, MSW06**]. A key tool of the method is to show that $(|\mathcal{P}(n, \lfloor qn \rfloor)|/n!)^{1/n}$ tends to a limit $\gamma(q)$ as n tends to infinity. We study the behaviour of this function in more detail. In particular we consider $\gamma(q)$ as q approaches 1 from below and above and 3 from below, and explain the discontinuity as we approach 1 from below by changing scale appropriately.

Chapter 4 - Cactus and Block Graphs

Recently, the usage of generating functions has led to immense progress in the enumeration and the understanding of properties of graph classes with structural constraints, such as planar graphs. However, it seems to be inherently difficult to investigate properties of "typical" members of those classes, which cannot be directly addressed with (a finite number of) parameters of the generating functions, such as the maximum node degree, or the maximum size of a biconnected block.

In this chapter we address this problem. We propose a new method, that is based on the analysis of the behaviour of Boltzmann sampler algorithms, and may be used to obtain precise estimates for the maximum degree and maximum size of a biconnected block of a "typical" member in the class in question.

In contrast to Chapter 3, here we use analytic combinatorics. The main focus, however, is on the newly developed method based on Boltzmann samplers. We illustrate how our method works on two graph classes, namely cactus graphs and block graphs, two moderately complex graph classes with constraints. A cactus graph is a labelled connected graph, in which each edge is contained in at most one cycle; and a block graph is a labelled connected graph, whose maximal biconnected blocks are cliques.

Additionally, exploiting the (more classical) generating function framework, we asymptotically enumerate the graphs in those classes, derive limit laws for the number of edges, and explore several other properties. Finally, we show how these results can be applied to straightforwardly derive the average-case complexity of two longest path algorithms for both graph classes.

Chapter 5 - $K_{3,3}$-minor-free Graphs

In the last chapter we consider simple labelled $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs, where maximal means that adding any edge to such a graph yields a $K_{3,3}$-minor.

Due to Kuratowski's theorem [**Kur30**] planar graphs are $K_{3,3}$- *and* K_5-minor-free. Hence, the class of planar graphs is contained in the class of $K_{3,3}$-minor-free graphs. Due to Wagner's theorem [**Wag37**] the class of triangulations (with the exception of all triangulations on 5 nodes) is contained in the class of maximal $K_{3,3}$-minor-free graphs. Determining the number (of graphs of sub-classes) of planar graphs has attracted considerable attention [**BGW02,**

GN04, GN05b, BGKN05, BLMK] in recent years. Giménez and Noy [**GN05b**] obtained precise asymptotic estimates for the number of planar graphs. The asymptotic number of triangulations was given by Tutte [**Tut62**] already in 1962. Investigating how much the number of planar graphs (triangulations) differs from (maximal) $K_{3,3}$-minor-free graphs is a first important step in examining how "typical" instances of these graph classes will differ.

We apply generating function techniques to obtain asymptotic estimates for the number of (maximal) $K_{3,3}$-minor-free graphs. For $K_{3,3}$-minor-free graphs we consider 3-connected, 2-connected, connected and not necessarily connected graphs.

For 3-connected $K_{3,3}$-minor-free graphs, the change from planar graphs can be easily described: it follows from a theorem of Wagner [**Wag37**] that the set of 3-connected $K_{3,3}$-minor-free graphs consists of all 3-connected planar graphs and the complete graph on 5 nodes. Thus, on this connectivity level our graph class differs only in the existence of one additional graph from planar graphs. But as we shall show, adding K_5 to the set of 3-connected graphs yields a slightly larger exponential growth rate for 2-connected, connected, and not necessarily connected $K_{3,3}$-minor-free graphs. It also slightly changes other parameters, for instance the expected number of edges in a random $K_{3,3}$-minor-free graph. For maximal $K_{3,3}$-minor-free graphs the growth rate also increases slightly compared to triangulations.

To establish these results for $K_{3,3}$-minor-free graphs, we follow the approach taken for planar graphs [**BGW02, GN05b**]: we use a well-known decomposition along the connectivity structure of a graph, i.e., into connected, 2-connected and 3-connected components, and translate this decomposition into relations of our generating functions. This is possible as the decomposition for $K_{3,3}$-minor-free graphs which is due to Wagner [**Wag37**] fits well into this framework. Then we use singularity analysis to obtain asymptotic estimates and limit laws for several parameters from these equations.

For maximal $K_{3,3}$-minor-free graphs the situation is different, as the decomposition which is again due to Wagner has further constraints. The functional equations for the generating functions of *edge-rooted* maximal graphs are easy to obtain but in order to go to unrooted graphs, special integration techniques based on rational parametrisation of rational curves are needed.

This thesis is based on the papers [**GMSW05, GMSW07**], [**GGNW**] and [**PW**].

CHAPTER 2

Tools and Techniques

In this chapter we give a brief overview over the techniques and tools which we will use in the subsequent chapters for our analysis of random graphs with structural constraints. The purpose of the exposition is twofold. Firstly, we want to provide the main ideas of the different methods and not the (in depth) technical details. We give references to the literature and to related parts of this thesis for further details. Secondly, we state the fundamental theorems (without proof) which our work relies on and which we will use later without further reference.

2.1. Combinatorial Approach

In this section, we briefly describe the main ideas behind the combinatorial approach following McDiarmid, Steger, and Welsh [**MSW05, MSW06**]. We will use this method in Chapter 3 to deduce properties of random planar graphs with given average degree; there, we will see that we have to slightly adapt the overall approach to make it applicable for this particular graph class.

Let $\mathcal{G}$ denote the class of labelled graphs we are interested in and let $|\mathcal{G}_n|$ denote the number of graphs on n nodes in $\mathcal{G}$. To deduce properties of a graph G_n drawn uniformly at random from $\mathcal{G}_n$, it is central to enumerate the graphs in $\mathcal{G}_n$. The main technical tool of the combinatorial approach is to prove that there exists a constant γ – called the growth constant of $\mathcal{G}$ – such that

$$\left(\frac{|\mathcal{G}_n|}{n!}\right)^{\frac{1}{n}} \to \gamma \quad \text{as } n \to \infty. \tag{2.1}$$

For the combinatorial approach the existence of such a growth constant suffices to prove a rather general result. More precisely, we want to prove that there exist a linear number of node disjoint copies of any fixed graph $H \in \mathcal{G}$ in G_n with high probability. To prove this, we proceed roughly as follows (for details, see the proof of Theorem 3.21 in Section 3.5): we assume the contrary, i.e. that the probability that G_n contains linearly many copies of H is small. Under this assumption, we show that we can construct – by attaching appropriate graphs to a given graph G – more graphs in $\mathcal{G}_n$ than (2.1) allows, which yields the desired contradiction. The key point in the counting argument is that we can show that each graph is not constructed too often, since G had few appearances of H. This rather general result can then be used to prove further structural properties of a random graph G_n.

To show (2.1), we apply Fekete's Lemma (see e.g. [**LW92**, Lemma 11.6]) which states the following: if $g : \mathbb{N} \to \mathbb{R}^+$ is a function such that $g(i+j) \geq g(i) \cdot g(j)$ for all $i, j \in \mathbb{N}$, and $c = \sup_n g(n)^{\frac{1}{n}} < \infty$, then $g(n)^{\frac{1}{n}} \to c$ as $n \to \infty$. The idea is to set $g(n) := |\mathcal{G}_n| \cdot \text{poly}(n)/n!$, where $\text{poly}(n)^{\frac{1}{n}} \to 1$ as $n \to \infty$. To prove that $\sup_n g(n)^{\frac{1}{n}}$ exists and is finite we can often use the existence of a growth constant of a super class of $\mathcal{G}$. Now, to prove $g(i+j) \geq g(i) \cdot g(j)$, we

consider all graphs in $\mathcal{G}$ on $i+j$ nodes on the left hand side and each tuple of *connected* graphs on i and j nodes respectively on the right hand side. If we can show that an appropriate portion of all graphs in $\mathcal{G}_n$ is connected, then the inequality is fulfilled, as we consider only graphs with two components on the right hand side. To show that the proportion of connected graphs in $\mathcal{G}_n$ is appropriate, is the second important ingredient of this approach. McDiarmid, Steger, and Welsh proved that this holds for a wide class of graphs using a Markov chain (see [**MSW05**] and [**MSW06**] for details). We will see later that their result cannot be directly applied to the graph class we consider in Chapter 3, which makes it harder to deal with.

2.2. Analytic Combinatorics

In this section we very briefly describe the idea behind the analytic combinatorics approach based on generating functions. A kind of revival of this method started with its successful application for planar graphs by Giménez and Noy [**GN05b**]. For precise estimates on the number of graphs in a given class and certain parameters this approach is quite powerful; for instance, Giménez and Noy used this method to determine the expected number of edges of a random planar graph, a problem which was open for quite some time [**DVW96, GM04, OPT03, BGH03**].

Here, we only give a brief overview of the method and state the main theorems that we will need later. Note, that the exposition in the sequel follows closely [**FS06**], but focuses on the application of the method for graphs with constraints. For a detailed description of the topic we refer the reader to [**FS06**] (and for an introduction to [**Wil90**]), where one can also find the proofs.

We will apply the method described here in Chapter 5 to obtain an estimate on the number of $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs. We will also apply it in Chapter 4. Moreover, it is the basis for the Boltzmann sampler framework of Duchon, Flajolet, Louchard, and Schaeffer [**DFLS04**], which we will briefly describe in Section 2.3 and which we will need later in Chapter 4 as well.

Before we can explain the method, we need some basic definitions from complex analysis. This is necessary as, roughly speaking, the method relies on two different views on a power series. On the one hand we view them and calculate with them as formal power series, which we use to encode information of the combinatorial objects we are interested in. On the other hand, if we want to get estimates for the coefficients, then we need to change our point of view and take the complex analysis part into account.

Definition 2.1 (Definition IV.1 of [**FS06**]). A function $f(z)$ defined over a region Ω is analytic at a point $z_0 \in \Omega$ if, for z in some open disc centred at z_0 and contained in Ω, it is representable by a convergent power series expansion

$$f(z) = \sum_{n \geq 0} c_n (z - z_0)^n. \tag{2.2}$$

A function is analytic in a region Ω iff it is analytic at every point of Ω.

It can be shown that analyticity and complex differentiability are equivalent notions, i.e., a function is analytic in a region Ω if and only if it is complex-differentiable in Ω.

Let $f(z)$ be an analytic function defined over the interior region Ω determined by a simple closed curve γ, and let z_0 be a point of the bounding curve γ. If there exists an analytic function $f^*(z)$ defined over some open set Ω^* containing z_0 and such that $f^*(z) = f(z)$ in $\Omega^* \cap \Omega$, one says that f is *analytically continuable* at z_0 and that f^* is an *immediate analytic continuation* of f.

With these notions, we can define what is commonly called a singularity.

DEFINITION 2.2 (Definition IV.4 of [**FS06**]). Given a function f defined in the region interior to the simple closed curve γ, a point z_0 on the boundary (γ) of the region is a *singular point* or a *singularity* if f is not analytically continuable at z_0.

Now, we are ready to explain how the analytic combinatorics approach works in more detail.

As in the combinatorial approach (see the previous section), the enumeration of the objects in interest is a key tool (and actually the main application of this method). Let $\mathcal{G}$ denote the class of labelled graphs we are interested in. We define its exponential generating function $G(x) = \sum_n g_n x^n/n!$, where $g_n = |\mathcal{G}_n|$ denotes the number of graphs in $\mathcal{G}$ on n nodes. (For unlabelled graphs, we consider the corresponding ordinary generating function $G(x) = \sum_n g_n x^n$.) For a generating function $G(x)$ we denote its n-th coefficient by $g_n := [x^n]G(x)$.

We say that a number sequence $\{a_n\}$ is of *exponential order* K^n or has exponential growth rate K (in the combinatorial approach this corresponds to the growth constant, see Section 2.1), in short form $a_n \bowtie K^n$ if and only if $\limsup |a_n|^{\frac{1}{n}} = K$.

One can prove that the exponential growth rate of the coefficients of a power series such as $G(x)$ is determined by its singularities on the boundary of the disc of convergence, the so called *dominant* singularities. For ease of exposition (and as we will need just this case in the subsequent sections), we assume in the following that there is only *one* dominant singularity.

THEOREM 2.3 (Theorem IV.7 of [**FS06**]). *If $f(z)$, $f : \mathbb{C} \to \mathbb{C}$, is analytic at 0 and R is the modulus of a singularity nearest to the origin in the sense that*

$$R := \sup\{r \geq 0 \mid f \text{ is analytic in } |z| < r\},$$

then the coefficient $f_n = [z^n]f(z)$ satisfies

$$f_n \bowtie \left(\frac{1}{R}\right)^n.$$

For functions with nonnegative coefficients, including all combinatorial generating functions, one can also adopt

$$R := \sup\{r \geq 0 \mid f \text{ is analytic at all points of } 0 \leq z < r\}.$$

In the sequel, most theorems are stated for functions $f(z)$ that are singular at $z = 1$, a condition that entails no loss of generality, as the location of the dominant singularities always induces a multiplicative exponential factor for coefficients. If $f(z)$ is singular at $z = \omega$, then $g(z) \equiv f(z \cdot \omega)$ satisfies, by the scaling rule of Taylor expansions,

$$[z^n]f(z) = \omega^{-n}[z^n]f(z \cdot \omega) = \omega^{-n}[z^n]g(z),$$

and $g(z)$ itself is singular on the unit circle, but not inside the disc.

Besides the exponential growth rate, we can also get a more precise estimate. From a series expansion of the generating function at its dominant singularity – the so called *singular expansion* – we can derive the sub-exponential part of the growth rate (see the "transfer theorem", Corollary 2.5 below).

Before we can formally state this, we need one more definition.

DEFINITION 2.4 (Definition VI.1. of [**FS06**]). Given two numbers ϕ, R with $R > 1$ and $0 < \phi < \frac{\pi}{2}$, the open domain $\Delta(\phi, R)$ is defined as

$$\Delta(\phi, R) = \{z \mid |z| < R,\ z \neq 1,\ |Arg(z-1)| > \phi\}.$$

A domain is a Δ-domain if it is a $\Delta(\phi, R)$ for some R and ϕ. A function is Δ-analytic if it is analytic in some Δ-domain.

Analyticity of a function $f(z)$ in a Δ-domain is the basic condition that allows a *transfer* of error terms in an asymptotic expansion of $f(z)$ to error terms of its coefficients. This "transfer" is formalised in the next corollary.

COROLLARY 2.5 (Corollary VI.1 of [**FS06**]). *Assume f is Δ-analytic and, as $z \to 1$, $z \in \Delta$,*

$$f(z) \sim (1-z)^{-\alpha}$$

with $\alpha \notin \mathbb{Z}_{\leq 0}$. Then the coefficients of f satisfy

$$[z^n]f(z) \sim \frac{n^{\alpha-1}}{\Gamma(\alpha)},$$

where $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}e^{-t}dt$ is the well-known gamma function.

As mentioned before, one can show that under suitable analytic conditions such a "transfer" also holds for error terms; for example one can show that if $f(z) = \mathcal{O}((1-z)^{-\alpha})$ then it holds that $f_n = \mathcal{O}(n^{\alpha-1})$. For further details see Chapter VI (especially Section VI.4) of [**FS06**].

Equipped with this, we can describe the overall approach which works roughly as follows. First, we have to find a *unique* decomposition of our graph class which we can translate into relations of the corresponding generating functions. The symbolic method developed in [**FS06**] facilitates this translation by providing a set of rules for standard combinatorial constructions as for instance taking a set of objects, rooting or pointing an object, taking a sequence of objects, etc. As long as we can describe our decomposition within this set of construction rules, we can systematically derive relations of the corresponding generating functions. See Table 1 for a basic set of translation rules.

The next step is to perform a singularity analysis, that is, we first have to find the dominant singularity and then to derive a singular expansion for our generating functions. The later step is usually a tedious but routine calculation which can be carried out using a computer algebra system such as MAPLE. For more complicated graph classes such as planar graphs [**GN05b**] or $K_{3,3}$-minor-free graphs (see Chapter 5), an edge-rooting in the decomposition often yields that we first have to solve an integral to make this step possible. As a heuristic for solving such an integral, we can often use the functional inverse of a part of the integrand for an appropriate substitution. Sometimes, more involved techniques are necessary, see Section

Combinatorial Construction	Exponential Generating Functions
Disjoint Union $\mathcal{C} = \mathcal{A} + \mathcal{B}$	$C(z) = A(z) + B(z)$
Product $\mathcal{C} = \mathcal{A} \star \mathcal{B}$	$C(z) = A(z) \cdot B(z)$
Sequence $\mathcal{C} = \mathrm{Seq}(\mathcal{A})$	$C(z) = \frac{1}{1-A(z)}$
Set $\mathcal{C} = \mathrm{Set}(\mathcal{A})$	$C(z) = \exp(A(z))$
Cycle $\mathcal{C} = \mathrm{Cyc}(\mathcal{A})$	$C(z) = \log\left(\frac{1}{1-A(z)}\right)$
Pointing $\mathcal{C} = \Theta\mathcal{A}$	$C(z) = A^{\bullet}(z) := z\frac{\partial}{\partial z}A(z)$
Substitution $\mathcal{C} = \mathcal{A} \circ \mathcal{B}$	$C(z) = A(B(z))$

TABLE 1. Translation rules: Combinatorial constructions for labelled objects and operations for the corresponding exponential generating functions.

5.3. Moreover, note, that a "typical" decomposition for a graph class frequently yields a recursive description. The following theorem allows to obtain a singular expansion for such a generating function. For further details see also the discussion before Theorem VI.6 in [**FS06**] and its proof.

THEOREM 2.6 (Theorem VI.6 of [**FS06**]). *Let $y(z)$ be implicitly defined by $y(z) = z \cdot \phi(y(z))$. Let ϕ be a nonlinear function satisfying the following conditions:*

(1) The function $\phi(u)$ is nonlinear, analytic at 0, where it has nonnegative coefficients, and satisfies $\phi(0) \neq 0$, i.e.

$$\phi(0) \neq 0, \quad [u^n]\phi(u) \geq 0, \quad \phi(u) \not\equiv \phi_0 + \phi_1 u.$$

(2) Within the open disc of convergence of ϕ at 0, $|z| < R$, where R denotes the radius of convergence, there exists a (necessarily unique) positive solution to the characteristic equation

$$\exists \tau,\ 0 < \tau < R, \quad \phi(\tau) - \tau\phi'(\tau) = 0.$$

Let $y(z)$ be the solution of $y = u \cdot \phi(y)$ satisfying $y(0) = 0$. Then, the quantity $\rho = \tau/\phi(\tau)$ is the radius of convergence of $y(z)$ at 0 (with τ the root of the characteristic equation), and the singular expansion of $y(z)$ near ρ is of the form

$$y(z) = \tau - d_1\sqrt{1 - \frac{z}{\rho}} + \sum_{j\geq 2}(-1)^j d_j \left(1 - \frac{z}{\rho}\right)^{\frac{j}{2}},$$

where $d_1 := \sqrt{\frac{2\phi(\tau)}{\phi''(\tau)}}$ and the d_j, $j \geq 2$, being some computable constants.

From the singular expansion we get asymptotic estimates for the coefficients by the transfer theorems (see Corollary 2.5 above). Figure 1 gives an overview of the steps needed to carry out a singularity analysis.

If we can introduce a variable into the generating functions which marks a parameter we are interested in (e.g., a marker for edges in a graph), we may again obtain a singular expansion of the bivariate generating function; this may eventually allow us to describe the behaviour of the parameter. The following theorem will be useful in the subsequent chapters. If the

Given a function $f(z)$ analytic at 0, whose coefficients are to be asymptotically analysed.

(1) Locate the (single) dominant singularity ρ of $f(z)$.

(2) Check continuation, i.e., establish that $f(z)$ is analytic in some domain of the form $\rho\Delta_0$, for a Δ-domain Δ_0, where $\rho\Delta_0$ is the image of Δ_0 by the mapping $z \mapsto \rho z$.

(3) Analyse the function $f(z)$ as $z \to \rho$ in the domain $\rho\Delta_0$ and determine a singular expansion of the form

$$f(z) \underset{z\to\rho}{=} f_1\left(\frac{z}{\rho}\right) + \mathcal{O}\left(f_2\left(\frac{z}{\rho}\right)\right),$$

where $f_2(z) \ll f_1(z)$ and $f_1(z)$ and $f_2(z)$ belong to the standard scale of functions $S = \{(1-z)^{-\alpha}\lambda(z)^{\beta}\}$, where $\lambda(z) := z^{-1}\log(1-z)^{-1}$ (see Chapter VI.4 of [**FS06**] for details).

(4) Use the transfer theorems to translate the main term and the error term to obtain that

$$[z^n]f(z) \underset{n\to\infty}{=} \rho^{-n}[z^n]f_1(z) + \mathcal{O}\left(\rho^{-n}[z^n]f_2(z)\right),$$

where for the corresponding exponent α in $f_1(z)$ and $f_2(z)$ it hast to hold that $\alpha \notin \mathbb{Z}_{\leq 0}$.

FIGURE 1. Summary of the process of singularity analysis for a single dominant singularity.

technical conditions (i) through (iii) are fulfilled, it gives us that the parameter marked by the variable u is normally distributed with given mean and variance.

THEOREM 2.7 (Theorem IX.10 (Algebraic singularity schema) of [**FS06**]). *Let $F(z,u)$ be a bivariate function that is bivariate analytic at $(z,u) = (0,0)$ and has nonnegative coefficients there. Assume the following conditions:*

(i) *Algebraic perturbation: there exist three functions A, B, C, analytic in a domain $\mathcal{D} = \{|z| \leq r\} \times \{|u-1| < \varepsilon\}$, for some $r > 0$ and $\varepsilon > 0$, such that for $\alpha \notin \mathbb{Z}_{\leq 0}$ the following representation holds,*

$$F(z,u) = A(z,u) + B(z,u)C(z,u)^{-\alpha}, \tag{2.3}$$

that $\rho < r$ is the unique (simple) root in $|z| \leq r$ of the equation $C(z,1) = 0$, and that $B(\rho,1) \neq 0$.

(ii) *Non-degeneracy: one has $\frac{\partial}{\partial z}C(\rho,1) \cdot \frac{\partial}{\partial u}C(\rho,1) \neq 0$, ensuring the existence of a nonconstant $\rho(u)$ analytic at $u = 1$, such that $C(\rho(u),u) = 0$ and $\rho(1) = \rho$.*

(iii) *Variability: one has*

$$v\left(\frac{\rho}{\rho(1)}\right) = -\frac{\frac{\partial^2}{\partial u^2}\rho(1)}{\rho(1)} - \frac{\frac{\partial}{\partial u}\rho(1)}{\rho(1)} + \left(\frac{\frac{\partial}{\partial u}\rho(1)}{\rho(1)}\right)^2 \neq 0 \tag{2.4}$$

Then the random variable with probability generating function

$$p_n(u) = \frac{[z^n]F(z,u)}{[z^n]F(z,1)}$$

converges in distribution to a Gaussian variable with a speed of convergence that is $O(n^{-\frac{1}{2}})$. *The mean* μ_n *and standard deviation* σ_n *converge asymptotically to* μn *and* $\sqrt{\sigma n}$, *where* $\mu = -\frac{\frac{\partial}{\partial u}\rho(1)}{\rho(1)}$ *and* σ *is given by Equation (2.4).*

2.3. Boltzmann Samplers

The framework of Boltzmann samplers was introduced and developed by Duchon, Flajolet, Louchard, and Schaeffer in [**DFLS04**] and only recently extended and applied by Fusy to derive an efficient sampler for planar graphs in [**Fus05**]. It constitutes a general and efficient framework to obtain a random generator for sampling objects from a combinatorial class uniformly at random. It is an attractive alternative to the recursive method introduced by Nijenhuis and Wilf [**NW78**], as it usually delivers faster samplers in theory and especially in practice.

Here, we briefly present the main ideas of the Boltzmann sampler framework in the case of labelled objects and exponential generating functions, following [**DFLS04**] and [**Fus05**]. For a detailed description of the topic we refer to [**DFLS04**] and [**Fus05**]. We will later use this framework to derive efficient sampling routines for two constrained graph classes, see Chapter 4. Moreover, we will introduce a new method in Section 4.7 to prove that "typical" instances of a constrained graph class have specific properties, which is based on this framework.

Note, that subsequently we will use notation, which was introduced in Section 2.1, as the framework described here relies on methods of the analytic combinatorics approach.

In practice, one might often prefer *approximate-size samplers*, which generate objects within a given range of sizes, to *exact-size samplers*, as long as the running time is better in the former case. The Boltzmann sampler framework provides exactly this behaviour: for many combinatorial classes it allows us to derive an algorithm whose running time is expected *linear* in the size of the output object for approximate-size sampling and expected *quadratic* for exact-size sampling.

Instead of fixing a particular output size for the random generation (as in the case of the recursive method), objects are drawn under a probability distribution spread over the whole class. The probability of an object to be drawn is essentially proportional to the exponential of its size n (e.g., number of nodes of a graph). More precisely, let $\mathcal{C}$ denote the class of labelled objects, in which we are interested, and let $C(x)$ denote its exponential generating function. In the Boltzmann model of parameter x, we assign to any object $\gamma \in \mathcal{C}$ of size $|\gamma|$ the probability

$$\Pr[\gamma] = \frac{1}{C(x)} \cdot \frac{x^{|\gamma|}}{|\gamma|!}, \tag{2.5}$$

if the above expression is well-defined – which is the case if x is within the radius of convergence of $C(x)$. A Boltzmann sampler $\Gamma C(x)$ for class $\mathcal{C}$ is a process that produces objects from $\mathcal{C}$ according to the corresponding Boltzmann model, i.e., according to the Boltzmann distribution (2.5) with parameter x. The parameter x can be tuned in such a way that the sampler prefers objects of the desired output size. One can prove that the random output

Construction	Boltzmann sampler
Empty atom $\mathbf{1}$	return $\mathbf{1}$
Unit atom / singleton $\mathcal{Z}$	return $\mathcal{Z}$
Disjoint union $\mathcal{C} = \mathcal{A} + \mathcal{B}$	$\Gamma C(x,y)$: if Bern $\left(\frac{A(x,y)}{A(x,y)+B(x,y)}\right)$ return $\Gamma A(x,y)$ else return $\Gamma B(x,y)$
Labelled product $\mathcal{C} = \mathcal{A} \star \mathcal{B}$	$\Gamma C(x,y)$: $\gamma \leftarrow (\Gamma A(x,y), \Gamma B(x,y))$ DistributeLabels(γ) return γ
Sequence $\mathcal{C} = Seq_{\geq d}(\mathcal{A})$	$\Gamma C(x,y)$: $k \leftarrow \mathrm{Geom}_{\geq d}(\Gamma A(x,y))$ $\gamma \leftarrow (\Gamma A(x,y), \ldots, \Gamma A(x,y))$ {k independent calls} DistributeLabels(γ) return γ
Set $\mathcal{C} = Set_{\geq d}(\mathcal{A})$	$\Gamma C(x,y)$: $k \leftarrow \mathrm{Po}_{\geq d}(\Gamma A(x,y))$ $\gamma \leftarrow (\Gamma A(x,y), \ldots, \Gamma A(x,y))$ {k independent calls} DistributeLabels(γ) return γ

x-substitution $\mathcal{C} = \mathcal{A} \circ_x \mathcal{B}$	$\Gamma C(x,y)$: $\gamma \leftarrow \Gamma A(B(x,y), y)$
	for each labelled atom $v \in \gamma$ do
	replace v by $\gamma_v \leftarrow \Gamma B(x,y)$
	end for
	DistributeLabels(γ)
	return γ

y-substitution $\mathcal{C} = \mathcal{A} \circ_y \mathcal{B}$	$\Gamma C(x,y)$: $\gamma \leftarrow \Gamma A(x, B(x,y))$
	for each labelled atom $e \in \gamma$ do
	replace e by $\gamma_e \leftarrow \Gamma B(x,y)$
	end for
	DistributeLabels(γ)
	return γ

TABLE 2. Translation rules: Combinatorial constructions to according Boltzmann sampler. Bern(p) denotes a Bernoulli generator which returns true with probability p and false otherwise. $\mathrm{Po}_{\geq d}(\lambda)$ and $\mathrm{Geom}_{\geq d}(\lambda)$ denote a Poisson and a geometric generator which return a value $x \geq d$ according to a Poisson and a geometric law respectively.

size N of an object satisfies $\mathbb{E}[N]_x = xC'(x)/C(x)$ – see Proposition 1 of [**DFLS04**]. Below we also present a sampler which does not need any tuning of the parameter x.

Duchon et al. [**DFLS04**] provide a set of general procedures, which translate combinatorial constructions like disjoint union, product, set, etc. into Boltzmann samplers – this is very similar to the symbolic method provided in [**FS06**] for generating functions (see also Section 2.2 and Table 1). Fusy [**Fus05**] extended this set of rules by further constructions, as for instance unrooting a rooted object. An overview of these rules is given in Table 2. For correctness proofs of these rules we refer to [**DFLS04**] and [**Fus05**].

Now, to obtain a realisation of a Boltzmann sampler, we need the following three missing ingredients.

Firstly, by definition, a Boltzmann sampler requires as input the value of the control parameter x that defines the Boltzmann model of use. Moreover, as generating functions may depend on each other in a specification, we may also need the finite collection of values of these generating functions evaluated at x to provide them as input to Bernoulli, geometric, Poisson, and other generators – see also Table 2. We can pre-compute these values *once* by a computer algebra system or a multiprecision package; note, that the power series expansions can be computed in low polynomial complexity.

Secondly, it remains to specify generators for the basic probabilistic laws, for instance a Poisson $\mathrm{Po}(\lambda)$, Bernoulli $\mathrm{Bern}(p)$, and geometric $\mathrm{Geom}(\lambda)$ generator. Under the assumption that we have a random generator that draws a random variable uniformly distributed over the real interval $(0, 1)$ at unit cost, we can easily implement such generators (for details see [**DFLS04**]). In the model with unit cost for the four elementary real-arithmetic operations, all these generators have linear running time in the outcome of the drawn variable.

In practice we may implement a Boltzmann sampler by truncating real numbers to some fixed precision, for example standard floating point numbers. Then, the resulting samplers operate in time that is linear in the size of the produced object, although they may deviate (slightly) from uniformity in a small number of cases. One approach to correct this lack of uniformity is to use an adaptive precision strategy.

Thirdly, we have to provide a routine that distributes labels $\{1, \ldots, n\}$ onto the atoms of the object. One can show that it suffices to do so at the end of the sampling process, i.e., the samplers are only sampling the *shape* of the object (see [**FZvC94**] (Section 3)). As sampling a permutation uniformly at random can be done in linear time (see [**Knu81**] (page 145)), this step will not affect the overall running time.

The efficiency of a Boltzmann sampler for a class $\mathcal{C}$ depends highly on the type of the singular expansion of its generating function $C(x)$. More precisely, in order to obtain an expected linear running time sampler, either the exponent $-\alpha$ of the singular expansion of C has to be negative *or*, in the special case that $-\alpha = \frac{1}{2}$, we can use a so-called *singular ceiled rejection sampler*. This sampler simply discards objects during sampling, as soon as they become too large.

We can formalise the conditions on the generating function as follows.

Definition 2.8 (Definition 2 of [**DFLS04**]). A function $f(z)$ analytic at 0 and with finite radius of analyticity $\rho > 0$ is said to be Δ-*singular* if it satisfies the following two conditions:

(i) The function admits ρ as its only singularity on $|z| = \rho$ and it is continuable in a domain

$$\Delta(\phi, r) = \{z \mid |z| < r,\ z \neq \rho,\ Arg(z - \rho) \notin (-\phi, \phi)\},$$

for some $r > \rho$ and some $0 < \phi < \frac{\pi}{2}$.

(ii) For z tending to ρ in the Δ-domain, $f(z)$ satisfies a singular expansion of the form

$$f(z) \underset{z \to \rho}{\sim} P(z) + c_0 \left(1 - \frac{z}{\rho}\right)^{-\alpha} + o\left(\left(1 - \frac{z}{\rho}\right)^{-\alpha}\right), \quad \alpha \in \mathbb{R} \setminus \{0, -1, -2, \ldots\},$$

where $P(z)$ is a polynomial. The quantity $-\alpha$ is called the singular exponent of $f(z)$.

We assemble the basic *rejection sampler* denoted by $\mu C(x; n, \varepsilon)$ from a Boltzmann generator $\Gamma C(x)$ for the class $\mathcal{C}$ as follows. For any x with $0 < x < \rho_C$, where ρ_C denotes the dominant singularity of the exponential generating function $C(x)$ for $\mathcal{C}$, for a target size n and $\varepsilon \geq 0$ a relative tolerance:

$\mu C(x; n, \varepsilon)$ { Returns an object of $\mathcal{C}$ in $I_{n,\varepsilon} := [(1-\varepsilon)n, (1+\varepsilon)n]$ }
repeat
  $\gamma := \Gamma C(x)$
until $|\gamma| \in I_{n,\varepsilon}$
return γ

The rejection sampler μC simply tries to sample an object of satisfactory size by repeatedly calling the sampler $\Gamma C(x)$. If we set $\varepsilon = 0$, we obtain an exact-size sampler. One can prove that this sampler indeed is quite efficient for a well-behaved generating function (see Definition 2.8).

THEOREM 2.9 (Theorem 5 of [**DFLS04**]). *Let $\mathcal{C}$ be a combinatorial class such that its generating function is Δ-singular with an exponent $-\alpha < 0$. Then the rejection sampler $\mu C(x_n; n, \varepsilon)$ corresponding to a fixed tolerance $\varepsilon > 0$ succeeds in a number of trials whose expected value is asymptotic to the constant*

$$\frac{1}{\xi_\alpha(\varepsilon)}, \text{ where } \xi_\alpha(\varepsilon) = \frac{\alpha}{\Gamma(\alpha)} \int_{-\varepsilon}^{\varepsilon} (1+s)^{\alpha-1} e^{-\alpha(1+x)} ds.$$

Recursive structures tend to conform to a universal complex-analytic pattern corresponding to a square-root singularity, that is, a singular exponent $-\alpha = 1/2$. As mentioned before, this specific behaviour may be exploited, resulting in a singular sampler, that is we set the input parameter x to the dominant singularity $x := \rho_C$ of $C(x)$. This yields heavy tails in the size distribution and in particular the expected size of an object is infinite. Thus we cannot use a straight-rejection sampler as before. But we can use an early interrupt strategy by modifying all sampling routines of a Boltzmann sampler as follows. Consider a general Boltzmann sampler $\Gamma C(x)$ and let m be an upper bound imposed on the size of the required objects. We build a modification $\Gamma C_{<m}(x)$ of $\Gamma C(x)$. We maintain a global counter c which counts the number of atoms produced at any given time during a partial execution of sampling by $\Gamma C(x)$. The counter is incremented each time an atom is created as long as $c \leq m$. However, as soon as c exceeds m, execution is interrupted and the special symbol $\perp$ is returned. Then rejection can be piled on top of this sampler. The corresponding routine looks as follows:

$\nu C(x; n, \varepsilon)$ { Ceiled rejection sampler }
repeat
  $\gamma := \Gamma C_{<m}(x; n(1+\varepsilon))$
until $(\gamma \neq \perp) \wedge (|\gamma| \geq n(1-\varepsilon))$
return γ

The ceiling technique optimises any Boltzmann sampler for any value of x; if we set the parameter to the singular value, i.e., $x = \rho$, we obtain an efficient sampling algorithm for

recursive classes. Before we can formally state this in a theorem, we need to introduce some more notation.

We say that a function $A(z)$ is *aperiodic* if $d := \gcd\{n \mid A_n \neq 0\}$ satisfies $d = 1$.

A recursive class $\mathcal{C}$ is defined as the component $\mathcal{C} = \mathcal{F}_1$ of a system of mutually dependent equations,

$$\{\mathcal{F}_1 = \Psi_1(\mathcal{Z}; \mathcal{F}_1, \ldots, \mathcal{F}_m), \ldots, \mathcal{F}_m = \Psi_m(\mathcal{Z}; \mathcal{F}_1, \ldots, \mathcal{F}_m)\}$$

where the Ψ's are *any* functional term involving any of the basic constructors previously defined (see Table 2 without substitution). The system is said to be *irreducible* if the dependency graph between the $\mathcal{F}_j$ is strongly connected (every nonterminal $\mathcal{F}_j$ depends on any other $\mathcal{F}_k$).

THEOREM 2.10 (Theorem 8 of [**DFLS04**]). *Let $\mathcal{C}$ be a combinatorial class given by a recursive specification that is irreducible and aperiodic. Then the singular ceiled rejection sampler $\nu C(\rho; n, \varepsilon)$, corresponding to a fixed tolerance $\varepsilon > 0$ succeeds in a number of trials whose expected value grows like $n^{\frac{1}{2}}/\zeta(\varepsilon)$ for a positive constant $\zeta(\varepsilon)$. Moreover, the cumulated size T_n of the generated and rejected objects during the call of $\nu C(\rho; n, \varepsilon)$ satisfies as $n \to \infty$*

$$\mathbb{E}[T_n] \sim \frac{n}{\varepsilon}\left((1-\varepsilon)^{\frac{1}{2}} + (1+\varepsilon)^{\frac{1}{2}}\right)$$

with its variance, $\sigma^2 = \mathbb{E}\left[T_n^2\right] - \mathbb{E}[T_n]^2$, being

$$\sigma^2 \sim \mathbb{E}[T_n]^2 + \frac{n^2}{\varepsilon}\left(\frac{1}{3}(1-\varepsilon)^{\frac{3}{2}} + (1+\varepsilon)^{\frac{3}{2}}\right).$$

Under these conditions, approximate-size sampling and exact-size sampling are of average-case complexity respectively $O(n)$ and $O(n^2)$.

2.4. Properties via Boltzmann Samplers

In Sections 2.1 and 2.2 we briefly introduced two common approaches from the literature for investigating constrained graph classes. In this section, we sketch the main ideas behind a new approach which is based on the Boltzmann sampler framework of Duchon et al. [**DFLS04**] described in Section 2.3. We will illustrate how this method works in detail in Chapter 4, where we prove that random cactus and random block graphs have certain properties with high probability.

Let us briefly answer the question "Why do we need yet another approach?". While both previously introduced methods, the combinatorial and the analytic combinatorics approach, have their strengths, clearly both of them have also some disadvantages.

The combinatorial approach is by its nature problem-tailored and thus one might have to adapt it already for slightly different graph classes. Moreover, it sometimes only yields rough upper and/or lower bounds. Finally, it seems to be quite difficult to obtain estimates of certain parameters as for example the expected number of edges.

The main problem with the analytic combinatorics approach is that it seems to be inherently difficult to investigate properties, which cannot be directly addressed with (a finite number of) parameters of the generating functions, such as the maximum node degree, the maximum size of a biconnected block, or the degree sequence.

With the new approach we want to address these problems. Our central idea is to analyse the execution of an algorithm, which *samples* uniformly at random graphs from the graph class in question. If we know what properties the "typical" output of such an algorithm has, then a random member of the considered graph class will also have those properties w.h.p. On the other hand, from a technical point of view, the analysis may become much easier if we reconceive, as just described.

More precisely, our approach is based on the analysis of the behaviour of *Boltzmann sampler* algorithms (see Section 2.3), and is an extension to a purely generating function based approach.

Roughly speaking, we proceed as follows: instead of investigating directly the properties of the graph class, we consider the execution of its corresponding Boltzmann sampler. We examine how the *shape* of a sampled object evolves over a run of the sampler and how this affects the related property. From this knowledge, we can eventually deduce the probability that a sampled output object has the property in interest.

Consider the sampling process (for technical reasons we will have to modify the sampler slightly) as drawing random variables from different probability distributions. The outcome of the variables determines the shape of the output object and as long as the property we are interested in depends on the shape (which is true for many interesting properties of a graph), we can define events which have to hold if an object has a certain property. If we can prove that all these events hold with high probability, we are done. To achieve this, we usually have to prove tail estimates for the corresponding probability distributions. For further details, see Chapter 4.

CHAPTER 3

Planar Graphs with given Average Degree

In this chapter we consider planar graphs with given average degree. More precisely, we are interested in properties of a random planar graph $R_{n,q}$ which is drawn uniformly at random from the class $\mathcal{P}(n, \lfloor qn \rfloor)$ of simple labelled planar graphs with n nodes and $\lfloor qn \rfloor$ edges, where $1 < q < 3$.

The proof techniques exploited in this section are purely combinatorial; we follow the approach of McDiarmid, Steger, and Welsh [**MSW05, MSW06**] (see also Section 2.1). A key tool of the method is to show that $(|\mathcal{P}(n, \lfloor qn \rfloor)|/n!)^{1/n}$ tends to a limit $\gamma(q)$ as n tends to infinity. In the subsequent we also study the behaviour of this function in more detail.

Note that some of the results were strengthened by Giménez and Noy [**GN05b, GN05a**] or can be improved using their results. We point out if one of our statements below also relies on their work.

3.1. Previous and Related Work

Planar graphs are well-known and well-studied combinatorial objects in graph theory. Roughly speaking, a graph is planar if it can be drawn in the plane in such a way that no two edges cross. A *random planar graph* R_n is a simple labelled planar graph that is drawn uniformly at random from the set $\mathcal{P}(n)$ of all simple planar graphs on the node set $\{1, \ldots, n\}$.

The random planar graph was first investigated in [**DVW96**] by Denise, Vasconcellos, and Welsh. They were mainly interested in the following two questions: how can one quickly generate such a random planar graph R_n, and what does this random planar graph look like?

For the first question there are some satisfactory answers. The first to present an algorithm that generates a random planar graph (exactly uniformly at random) in (expected) polynomial time were Bodirsky, Gröpl, and Kang [**BGK03**], who showed that one can generate a random planar graph on n nodes in expected time $O(n^3)$ per generation (see [**FPS05**] for an explanation) and $O(n^7(\log n)^2(\log\log n))$ preprocessing time. The algorithm is based on the decomposition of planar graphs into 1-, 2-, and 3-connected components and a generation algorithm for random 3-connected planar graphs. Recently, Fusy [**Fus05**] presented an algorithm which runs in expected quadratic time for exact size sampling and in expected linear time for approximate size sampling, thus dramatically improving on the running time. This new algorithm is based on the principle of Boltzmann samplers (see also Sections 2 and 4).

Denise, Vasconcellos, and Welsh [**DVW96**] gave a remarkably simple construction of a Markov chain such that the stationary distribution is uniform over all planar subgraphs of a given graph G. They defined a Markov chain with state space all planar subgraphs of G (with the same set of nodes as G) and the following transitions: choose an unordered pair

$\{i, j\}$ of distinct nodes uniformly at random. If the edge $e = \{i, j\}$ is present in the current state then delete it, if not then add it whenever planarity is preserved and $\{i, j\}$ is an edge of G, otherwise stay in the current state. With G as the complete graph K_n, this gives what appears to be a fairly effective way of generating a random planar graph; but it turns out to be difficult to give a bound on the mixing time – a problem which is still unsolved. Nevertheless using this Markov chain they could give experimental probabilities for various properties of the random planar graph; for instance the (corrected) experiments suggested that the expected number of edges of a random planar graph on n nodes is about $2.2n$, which later turned out to be a good estimate (see [**GN05b**] and [**Noy07**]).

There has also been some progress towards answers for the second question about properties of random planar graphs.

The key step in investigating the behaviour of the random planar graph R_n is to estimate the number of planar graphs. For the set $\mathcal{P}(n)$ of all labelled planar graphs on n nodes, McDiarmid, Steger, and Welsh [**MSW05**] showed that

$$\left(\frac{|\mathcal{P}(n)|}{n!}\right)^{\frac{1}{n}} \to \gamma \quad \text{as } n \to \infty,$$

where γ is the *planar graph growth constant*. Note that using the same approach they generalised their results to a much wider class of graphs in [**MSW06**].

Giménez and Noy [**GN04**] improve on earlier estimates [**DVW96, OPT03, BGH03**] and show using generating functions and singularity analysis that γ satisfies $\gamma \approx 27.2269$ to four decimal places. Even more recently they give an explicit analytic expression for γ, and show that

$$|\mathcal{P}(n)| \sim g \cdot n^{-\frac{7}{2}} \gamma^n n! \tag{3.1}$$

where the constant g has an explicit analytic expression and is about $4.97 \cdot 10^{-6}$ [**GN05b**]. A corresponding expression for the number of 2-connected planar graphs was given in [**BGW02**]. This was a major step towards establishing (3.1), see [**Noy07**] for an excellent discussion.

To deduce certain properties of the random planar graph McDiarmid, Steger, and Welsh [**MSW05**] only needed that there exists such a *planar graph growth constant* and not its exact value nor the sub-exponential behaviour. In particular, they were able to show, among other results, that a random planar graph R_n with high probability (w.h.p., that is, with probability tending to 1 as n tends to infinity) contains linearly many nodes of each given degree, has linearly many faces of each given size in any embedding, and contains linearly many node disjoint copies of any given fixed connected planar graph. Additionally, and perhaps most surprisingly, they showed that the probability that R_n is connected is bounded away from zero and from one by non-zero constants.

In [**GN05b**] (see also [**Noy07**]) it is shown that the number of edges $|E(R_n)|$ is asymptotically normally distributed, with mean $\sim \kappa n$ and variance $\sim \lambda n$, where the constants κ and λ have explicit analytic expressions and $\kappa \approx 2.213$ and $\lambda \approx 0.4303$. In particular this means that the expected number of edges of a random planar graph on n nodes is approximately $2.213n$, which was an open problem for quite some time. Thus the average degree in R_n is about 4.416 w.h.p. Furthermore, Giménez and Noy show in [**GN05b**] additional limit laws for the

random planar graph, for instance that the number of 2-connected components in a random connected planar graph and the number of appearances of a fixed connected planar graph in a random planar graph are asymptotically normally distributed. Furthermore, using their results one can strengthen some of the theorems in [**MSW05**] and the next sections; we point out where we rely on their work in the subsequent.

Most of these improvements are possible due to rather precise asymptotic expressions for $|\mathcal{P}(n, \lfloor qn \rfloor)|$ and connected planar graphs with given average degree $|\mathcal{P}_c(n, \lfloor qn \rfloor)|$, for $q \in (1, 3)$, which Giménez and Noy [**GN05a**] obtained using analytic methods, and we may write as:

$$|\mathcal{P}(n, \lfloor qn \rfloor)| \sim \alpha(q)\ n^{-4}\gamma(q)^n n! \tag{3.2}$$

and similarly

$$|\mathcal{P}_c(n, \lfloor qn \rfloor)| \sim \alpha_c(q)\ n^{-4}\gamma(q)^n n!, \tag{3.3}$$

where $\alpha(q)$ and $\alpha_c(q)$ are constants.

3.2. Results

In the following sections we are interested in the class $\mathcal{P}(n, m)$ of (simple) labelled planar graphs on n nodes with m edges, and in particular in $\mathcal{P}(n, \lfloor qn \rfloor)$, where the average degree is about $2q$. Subsequently, we show that for all $1 < q < 3$ the random planar graph $R_{n,q}$, which is drawn uniformly at random from the set $\mathcal{P}(n, \lfloor qn \rfloor)$, has properties similar to those of a random planar graph R_n. In particular we show the following results:

THEOREM 3.1. *Let $1 \leq q < 3$, let k be a positive integer, and for a graph G let $d_k(G)$ denote the number of nodes with degree equal to k. Then there exists a constant $\alpha_k = \alpha_k(q) > 0$ such that*

$$\Pr[d_k(R_{n,q}) < \alpha_k n] = e^{-\Omega(n)}.$$

THEOREM 3.2. *Let $1 < q < 3$, let $k \geq 3$ be an integer, and for a planar graph G let $f_k(G)$ denote the number of faces of size k* minimised *over all plane embeddings of G. Then there exists a constant $\beta_k = \beta_k(q) > 0$ such that*

$$\Pr[f_k(R_{n,q}) < \beta_k n] = e^{-\Omega(n)}.$$

COROLLARY 3.3. *Let $1 \leq q < 3$ and let H be a fixed connected planar graph, where H must be a tree if $q = 1$, and for a graph G let $f_H^*(G)$ denote the maximum number of pairwise node disjoint appearances of H contained in G. Then there exists a constant $\alpha = \alpha(H, q) > 0$ such that*

$$\Pr[f_H^*(R_{n,q}) < \alpha n] = e^{-\Omega(n)}.$$

Note that this last result implies in particular, that for every $\varepsilon > 0$ and any sufficiently large n, a random planar graph $R_{n,1+\varepsilon}$ with $\lfloor (1+\varepsilon)n \rfloor$ edges contains already a K_4 and is thus 4-chromatic. Moreover the probability of containing a K_4 is sufficiently large, that the algorithm that tests whether a K_4 is present, and if this is not the case finds an optimal colouring, can be made to run in expected polynomial time. Therefore one can find the chromatic number of a random planar graph with n nodes and $\lfloor qn \rfloor$, $1 < q < 3$, edges in expected polynomial time.

Only recently there was progress towards answering the question whether $R_{n,q}$ contains also a K_4 for $0 \le q < 1$. If $0 \le q < 1$ and H is a given connected planar graph with more than one cycle, then with high probability $R_{n,q}$ has no subgraph isomorphic to H. This was observed recently by Chris Dowden [**Dow06**]. Hence, in particular, $R_{n,q}$ is very unlikely to have a subgraph K_4 for $q < 1$.

Equations (3.2) and (3.3) from [**GN05a**] make it possible to strengthen a result of [**GMSW05**] about the connectivity of $R_{n,q}$.

THEOREM 3.4. *For each $0 \le q \le 3$, there exists a constant $\beta = \beta(q) \ge 0$ such that*

$$\Pr[R_{n,q} \text{ is connected}] \to \beta(q) = \begin{cases} 0 & \text{for } 0 \le q \le 1 \\ \alpha_c(q)/\alpha(q) & \text{for } 1 < q < 3 \\ 1 & \text{for } q = 3, \end{cases}$$

as $n \to \infty$, where the constants $\alpha(q)$ and $\alpha_c(q)$ are as in Equations (3.2) and (3.3) respectively.

In addition to the structural properties above, we prove lower bounds of the form $(1 + o(1)) \ln n / \ln\ln n$ on the maximum node degree and maximum face size in any plane embedding. Note, that McDiarmid and Reed have recently shown that the maximum degree is $\Theta(\log n)$ w.h.p. [**MR**].

One of the main tools for the proofs of the theorems stated above is a bound on the number of planar graphs with a given number of nodes. We show that for all $0 \le q \le 3$, there exists $\gamma(q) \ge 0$ such that

$$\left(\frac{|\mathcal{P}(n, \lfloor qn \rfloor)|}{n!}\right)^{\frac{1}{n}} \to \gamma(q) \quad \text{as} \quad n \to \infty. \tag{3.4}$$

(For $q = 3$, we interpret $\mathcal{P}(n, \lfloor qn \rfloor)$ as the set $\mathcal{P}(n, 3n-6)$ of triangulations.)

This result holds also if we replace $|\mathcal{P}(n, \lfloor qn \rfloor)|$ by *connected* planar graphs with given average degree $|\mathcal{P}_c(n, \lfloor qn \rfloor)|$ (with the same limiting value $\gamma(q)$). Let us first consider properties of the function $\gamma(q)$, and then look more closely at the limiting result. Recall that κ is the parameter for the mean of the number of edges of the random planar graph R_n, and γ is the planar graph growth constant.

THEOREM 3.5. *The function $\gamma(q)$ on $[0, 3]$ satisfies*

(i) $\gamma(q) = 0$ *for* $0 \le q < 1$, $\gamma(1) = e$, $\gamma(\kappa) = \gamma$, *and* $\gamma(3) = 256/27$.

(ii) $\gamma(q)$ *is continuous and log-concave on* $[1, 3]$, *and it is strictly increasing on* $[1, \kappa]$ *and strictly decreasing on* $[\kappa, 3]$.

(iii) $\gamma(q)$ *is computable, and analytic on* $(1, 3)$.

This result is based on Theorem 3.6 below together with the following three important contributions from [**GN05b, GN05a**]: (a) The expected number of edges in R_n is $\sim \kappa n$. (b) From the discussion in [**GN05a**] following Theorem 3, the function $\gamma(q)$ is continuous on the right at 1 (we do not have a combinatorial proof of this). (c) Part (iii).

Now let us look more closely at the limiting result (3.4), and give two directions in which it can be strengthened, one allowing more freedom in the number of edges and one being far

more precise. First, if $1 \leq q \leq 3$ and $m = m(n)$ satisfies $n \leq m \leq 3n-6$ and $m/n \to q$ as $n \to \infty$, then

$$\left(\frac{|\mathcal{P}(n,m)|}{n!}\right)^{\frac{1}{n}} \to \gamma(q). \tag{3.5}$$

This result holds also if we replace $\mathcal{P}$ by connected planar graphs $\mathcal{P}_c$. These results follow from the proof of Lemma 3.15 below, using also the fact that $\gamma(q)$ is continuous on the right at 1.

As $\gamma(q) = 0$ for $q < 1$ and $\gamma(1) = e$, we know that $\gamma(q)$ is discontinuous at 1 from the left. We can 'explain' this discontinuity as we approach 1 from below, by changing scale appropriately. Working at a scale of $n/\ln n$ rather than n for edge numbers, the discontinuity on the left at 1 disappears, see Theorem 3.17 below. The main feature now left open about $\gamma(q)$ is whether the slope stays finite as q approaches 1 from above, and approaches 3 from below. We see in Section 3.4 that this is not the case. More precisely, since the function $\lambda(q) = \ln \gamma(q)$ is concave and finite on $[1,3]$, its left and right derivatives exist in $(1,3)$ and are finite and non-increasing: we show that they tend to ∞ as $q \downarrow 1$ and to $-\infty$ as $q \uparrow 3$.

Finally, we show the existence of a growth constant for connected unlabelled planar graphs with given average degree, its relation to the labelled planar graph growth constant $\gamma(q)$, and that $R_{n,q}$ has w.h.p. an exponential number of automorphisms.

3.3. The number of planar graphs with n nodes and $\lfloor qn \rfloor$ edges

We first introduce some notation some of which was already used in the previous sections. Let $\mathcal{P}(n)$ denote the set of all simple labelled planar graphs on the node set $\{1,\ldots,n\}$. Let $\mathcal{P}(n,m)$ denote the set of all graphs in $\mathcal{P}(n)$ with exactly m edges and let $\mathcal{P}_c(n)$ and $\mathcal{P}_c(n,m)$ denote the sets of connected graphs in $\mathcal{P}(n)$ and $\mathcal{P}(n,m)$ respectively.

We are interested here in $|\mathcal{P}(n,m)|$ and in $|\mathcal{P}_c(n,m)|$. Clearly, $|\mathcal{P}_c(n,m)| \leq |\mathcal{P}(n,m)|$. Let us use $t(n)$ to denote the number of triangulations $|\mathcal{P}(n,3n-6)|$. We are now able to state the main theorem of this section.

THEOREM 3.6. *For each $0 \leq q \leq 3$, there is a finite constant $\gamma(q) \geq 0$ such that, as $n \to \infty$, if $0 \leq q < 3$ then both $(|\mathcal{P}_c(n,\lfloor qn \rfloor)|/n!)^{1/n}$ and $(|\mathcal{P}(n,\lfloor qn \rfloor)|/n!)^{1/n}$ tend to $\gamma(q)$, and $(t(n)/n!)^{1/n}$ tends to $\gamma(3)$.*

The function $\gamma(q)$ satisfies

(i) $\gamma(q) = 0$ for $0 \leq q < 1$, $\gamma(1) = e$, $\gamma(3) = 256/27$, and $\gamma(q) > 0$ for $1 \leq q \leq 3$.
(ii) $\gamma(q)$ is log-concave on $[1,3]$ and continuous on $(1,3]$.

Let us note here part of Theorem 3.6 above, namely that $\gamma(q) = 0$ for $0 \leq q < 1$: this follows from the fact that the number of graphs on n nodes and m edges is at most $\binom{\binom{n}{2}}{m} \leq \left(\frac{en^2}{2m}\right)^m$ and so by Stirling's formula

$$\left(\frac{|\mathcal{P}(n,\lfloor qn \rfloor)|}{n!}\right)^{\frac{1}{n}} \leq \frac{\left(\left(\frac{en^2}{2qn}\right)^{qn}\right)^{\frac{1}{n}}}{(n^n e^{-n})^{\frac{1}{n}}} = n^{q-1}\frac{e^{q+1}}{(2q)^q} \to 0$$

as $n \to \infty$. Thus in Theorem 3.6 the interest in $\gamma(q)$ is for $q \in [1,3]$.

To prove Theorem 3.6 we consider the following graph class. Let $\mathcal{C}(n,m)$ denote the set of planar graphs G on n nodes with $m' = m - \kappa(G) + 1$ edges, where $\kappa(G)$ denotes the number of components of G, and let $|\mathcal{C}(n,m)|$ denote the number of graphs in this class. Observe that if $m \leq n-2$ then $|\mathcal{C}(n,m)| = 0$. Also $\mathcal{C}(n,n-1)$ is the set of all forests on $1,\ldots,n$, and similarly $\mathcal{C}(n,n)$ is the set of all graphs on $1,\ldots,n$ with exactly one cycle. The class $\mathcal{C}(n,m)$ has the advantage over $\mathcal{P}(n,m)$ that one remains in the class when adding edges between components or deleting bridges.

Let $|\mathcal{C}_c(n,m)|$ denote the number of connected graphs in $\mathcal{C}(n,m)$. Observe that $|\mathcal{C}_c(n,m)| = |\mathcal{P}_c(n,m)|$, since both sides count connected planar graphs with n nodes and m edges.

Using a result from [**MSW05**] we will first prove in Section 3.3.1 that $|\mathcal{C}_c(n,\lfloor qn \rfloor)| \geq \frac{1}{e}|\mathcal{C}(n,\lfloor qn \rfloor)|$. With this result we will prove the analogue of the first part of Theorem 3.6 for the graph class $\mathcal{C}(n,\lfloor qn \rfloor)$. In Section 3.3.2 we will then derive that the first part of Theorem 3.6 holds for $\mathcal{P}(n,\lfloor qn \rfloor)$. Proofs of the remaining parts of Theorem 3.6, will be given in Section 3.4.1 and 3.4.2.

3.3.1. The class $\mathcal{C}(n,\lfloor qn \rfloor)$. We call a non-empty class $\mathcal{C}$ of finite graphs *weakly addable* if for each graph G in $\mathcal{C}$, any graph that is obtained from G by adding an edge joining two nodes in distinct components is also in $\mathcal{C}$. We will use the following theorem of [**MSW05**]:

THEOREM 3.7. [**MSW05**, Theorem 2.2] *Let $\mathcal{C}$ be any non-empty finite set of graphs, and let $\mathcal{C}$ be weakly addable. Let R denote a graph sampled uniformly at random from $\mathcal{C}$. Then the random number $\kappa(R)$ of components of R is stochastically dominated by $1+X$ where X has the Poisson distribution with mean 1. In particular,* $\Pr[R \text{ connected}] \geq 1/e$ *and* $\mathbb{E}[\kappa(R)] \leq 2$.

Observe that for $n-1 \leq m \leq 3n-6$, $\mathcal{C}(n,m)$ is weakly addable, and thus the next lemma is an immediate consequence of Theorem 3.7.

LEMMA 3.8. *Let n be a positive integer and let $n-1 \leq m \leq 3n-6$. Let G denote a graph sampled uniformly at random from $\mathcal{C}(n,m)$. Then the random number $\kappa(G)$ of components of G is stochastically dominated by $1+X$ where X has the Poisson distribution with mean 1. In particular,* $\Pr[G \text{ connected }] \geq \frac{1}{e}$ *and therefore* $|\mathcal{C}_c(n,m)| \geq \frac{1}{e}|\mathcal{C}(n,m)|$.

We will use Lemma 3.8 to prove a result analogous to the first part of Theorem 3.6 for the class $\mathcal{C}(n,\lfloor qn \rfloor)$.

LEMMA 3.9. *For each $1 \leq q \leq 3$, there exists a constant $\gamma(q)$ with $0 < \gamma(q) \leq \gamma$ such that as $n \to \infty$, for each $q < 3$*

$$\left(\frac{|\mathcal{C}(n,\lfloor qn \rfloor)|}{n!}\right)^{\frac{1}{n}} \to \gamma(q),$$

and

$$\left(\frac{t(n)}{n!}\right)^{\frac{1}{n}} \to \gamma(3).$$

PROOF. That $t(n)/n!$ tends to $\gamma(3) = 256/27$ follows from [**Tut62**] where it is shown that the number $t_u(n)$ of rooted (unlabelled) triangulations satisfies that $(t_u(n))^{1/n}$ tends to $256/27$ as

n tends to infinity. The result follows since we can label the nodes in $n!$ ways, and a labelled triangulation ($n \geq 4$) can be rooted in $4(3n-6)$ ways, since we have $3n-6$ possibilities for choosing the root edge, two for orienting the edge, and two for choosing the root face.

For $1 \leq q < 3$, we use Fekete's Lemma which states that if $g : \mathbb{N} \to \mathbb{R}^+$ is a function such that $g(i+j) \geq g(i) \cdot g(j)$ for all $i, j \in \mathbb{N}$, and $c = \sup_n g(n)^{\frac{1}{n}} < \infty$, then $g(n)^{\frac{1}{n}} \to c$ as $n \to \infty$. The proof in [**LW92**] of Fekete's lemma is easily modified to apply also for functions g with $g(i) = 0$ for finitely many i.

Suppose first that $1 \leq q < 3$. We set

$$\tilde{g}(n,q) = \frac{|\mathcal{C}(n, \lfloor qn \rfloor)|}{4e^2 n^2 n!} .$$

Since $\mathcal{C}(n, \lfloor qn \rfloor) \subset \mathcal{P}(n)$ we know that $c = \sup_n \tilde{g}(n,q)^{\frac{1}{n}} \leq \gamma$, so it remains to show that $\tilde{g}(i+j,q) \geq \tilde{g}(i,q) \cdot \tilde{g}(j,q)$.

Let n_0 be the smallest integer such that $|\mathcal{C}(n_0, \lfloor qn_0 \rfloor)| \neq 0$. Note that there exists a *connected* planar graph on n nodes and $\lfloor qn \rfloor$ edges for all $n \geq n_0$. Further, as $\lfloor qn \rfloor \geq n$, we can delete an edge of such a graph and obtain a connected graph. Assume w.l.o.g. that $i \leq j$. If $i < n_0$ then $|\mathcal{C}(i, \lfloor qi \rfloor)| = 0$ and therefore $\tilde{g}(i,q) = 0$ and the inequality holds. Otherwise $n_0 \leq i \leq j$ and we construct graphs on $i+j$ nodes with $\lfloor q(i+j) \rfloor$ edges and two components in the following way: first, we select i nodes out of $i+j$. Then we choose a connected graph on i nodes with $\lfloor qi \rfloor$ edges and a second connected graph on j nodes with $\lfloor qj \rfloor$ edges. The resulting graph G has $\lfloor qi \rfloor + \lfloor qj \rfloor$ edges and two components. We may have to delete an edge in order to construct a graph in $\mathcal{C}(i+j, \lfloor q(i+j) \rfloor)$, since any graph in $\mathcal{C}(i+j, \lfloor q(i+j) \rfloor)$ with two components needs to have $\lfloor q(i+j) \rfloor - 1$ edges and

$$\lfloor qi \rfloor + \lfloor qj \rfloor - 1 \leq \lfloor q(i+j) \rfloor - 1 \leq \lfloor qi \rfloor + \lfloor qj \rfloor .$$

If we have to delete an edge, we delete it in the component with i nodes and $\lfloor qi \rfloor$ edges. To avoid over-counting we have to divide by 2 as in the case $i = j$ we constructed every graph twice. Moreover we have to divide by $i^2/2$ as we may over-count $\binom{i}{2} \leq i^2/2$ graphs due to the deletion of one edge. Thus,

$$\begin{aligned} & |\mathcal{C}(i+j, \lfloor q(i+j) \rfloor)| \\ \geq \; & \frac{\binom{i+j}{i} |\mathcal{C}_c(i, \lfloor qi \rfloor)| \; |\mathcal{C}_c(j, \lfloor qj \rfloor)|}{i^2} \\ \geq \; & \frac{(i+j)!}{i!j!} \frac{1}{i^2} \frac{1}{e^2} |\mathcal{C}(i, \lfloor qi \rfloor)| \; |\mathcal{C}(j, \lfloor qj \rfloor)| . \end{aligned} \tag{3.6}$$

This implies

$$\begin{aligned} \tilde{g}(i+j,q) &= \frac{|\mathcal{C}(i+j, \lfloor q(i+j) \rfloor)|}{4e^2 (i+j)^2 (i+j)!} \\ &\geq \frac{1}{4e^2} \frac{1}{i!j!} \frac{1}{(i+j)^2} \frac{1}{e^2} \frac{1}{i^2} |\mathcal{C}(i, \lfloor qi \rfloor)| \; |\mathcal{C}(j, \lfloor qj \rfloor)| \\ &= \frac{|\mathcal{C}(i, \lfloor qi \rfloor)|}{4e^2 i^2 i!} \frac{|\mathcal{C}(j, \lfloor qj \rfloor)|}{4e^2 j^2 j!} \frac{j^2 4}{(i+j)^2} \\ &\geq \tilde{g}(i,q) \tilde{g}(j,q) \end{aligned}$$

as $\frac{4j^2}{(i+j)^2} \geq \frac{4j^2}{(2j)^2} = 1$. Since $(4e^2n^2)^{\frac{1}{n}} \to 1$ as $n \to \infty$, $\lim_{n\to\infty}(\tilde{g}(n,q))^{\frac{1}{n}} = \lim_{n\to\infty}(|\mathcal{C}(n,\lfloor qn\rfloor)|/n!)^{\frac{1}{n}} \leq \gamma$ exists. □

Let us check now that $\gamma(1) = e$. Since $\mathcal{C}(n, n-1)$ is the set of forests on $1,\ldots,n$ we have $|\mathcal{C}(n,n-1)| \sim e^{\frac{1}{2}}n^{n-2}$, see [**Rén59**]. But it is easy to check that $|\mathcal{C}(n,n)|/|\mathcal{C}(n,n-1)|$ is $\Omega(n)$ and $O(n^2)$, and so $\gamma(1) = e$ as required.

3.3.2. The class $\mathcal{P}(n,\lfloor qn\rfloor)$. Now let us prove the first part of Theorem 3.6. We will do so by showing that, for each $1 \leq q < 3$, we have

$$\gamma(q) \leq \liminf(|\mathcal{P}(n,\lfloor qn\rfloor)|/n!)^{\frac{1}{n}} \quad \text{and} \quad \limsup(|\mathcal{P}(n,\lfloor qn\rfloor)|/n!)^{\frac{1}{n}} \leq \gamma(q),$$

where $\gamma(q)$ is as in Lemma 3.9. (We already know about $q = 3$.)

LEMMA 3.10. *For each* $1 \leq q < 3$

$$\gamma(q) \leq \liminf_{n\to\infty}\left(\frac{|\mathcal{P}(n,\lfloor qn\rfloor)|)}{n!}\right)^{\frac{1}{n}}.$$

PROOF. As observed earlier, we have $|\mathcal{C}_c(n,\lfloor qn\rfloor)| = |\mathcal{P}_c(n,\lfloor qn\rfloor)| \leq |\mathcal{P}(n,\lfloor qn\rfloor)|$. Also, by Lemma 3.8, $\frac{1}{e}|\mathcal{C}(n,\lfloor qn\rfloor)| \leq |\mathcal{C}_c(n,\lfloor qn\rfloor)|$ and hence $\gamma(q) \leq \liminf(|\mathcal{P}(n,\lfloor qn\rfloor)|/n!)^{\frac{1}{n}}$. □

To prove the corresponding result for lim sup, we first need some helpful estimates that relate the number of planar graphs to the number of components in these graphs.

Recall that $\mathcal{P}(n)$ denotes the set of all labelled simple planar graphs on the nodes $1,\ldots,n$. Let $p(n)$ denote $|\mathcal{P}(n)|$. Let $p(n,\leq t)$ denote the number of graphs in $\mathcal{P}(n)$ having at most t components, let $p(n,\geq t)$ denote the number of graphs in $\mathcal{P}(n)$ with at least t components, and let $p(n,t)$ denote the number of graphs in $\mathcal{P}(n)$ with exactly t components.

For $\mathcal{C}(n,\lfloor qn\rfloor)$ and $\mathcal{P}(n,\lfloor qn\rfloor)$ we introduce a similar notation, so for example $|\mathcal{C}(n,\lfloor qn\rfloor,\leq t)|$ denotes the number of graphs in the class $\mathcal{C}(n,\lfloor qn\rfloor)$ with at most t components, and $|\mathcal{P}(n,\lfloor qn\rfloor,\geq t)|$ denotes the number of graphs in $\mathcal{P}(n,\lfloor qn\rfloor)$ having at least t components. Observe that $|\mathcal{P}(n,\lfloor qn\rfloor,k)| = |\mathcal{C}(n,\lfloor qn+k-1\rfloor,k)|$, as both sides count planar graphs with k components and $\lfloor qn\rfloor$ edges.

LEMMA 3.11. *For all integer* $k \geq 1$ *we have*

$$p(n,\geq k+1) \leq \frac{1}{k!}\cdot p(n)$$

and

$$|\mathcal{C}(n,\lfloor qn\rfloor+k,\geq k+1)| \leq \frac{1}{k!}\cdot|\mathcal{C}(n,\lfloor qn\rfloor+k)|.$$

PROOF. We know from Theorem 3.7 that

$$p(n,\geq k+1) \leq \sum_{i\geq k}\frac{e^{-1}}{i!}p(n) \leq \frac{1}{k!}p(n),$$

as

$$\sum_{i\geq k}\frac{e^{-1}}{i!} = \frac{e^{-1}}{k!}\sum_{i\geq k}\frac{k!}{i!} \leq \frac{e^{-1}}{k!}\sum_{i\geq k}\frac{1}{(i-k)!} = \frac{1}{k!}.$$

By Lemma 3.8, we have $|\mathcal{C}(n,\lfloor qn\rfloor,\geq k+1)| \leq \sum_{i\geq k}\frac{e^{-1}}{i!}|\mathcal{C}(n,\lfloor qn\rfloor)|$. Now we can proceed as above. □

We will later prove a more general version of the next lemma, see Lemma 3.20. Nevertheless, for ease of exposition we will give here a short proof for the special case needed here.

LEMMA 3.12. *Let* $1 \le q < 3$. *Let* $c > \ln \frac{\gamma}{\gamma(q)}$, *and let* $k = k(n) = \lceil cn/\ln n \rceil$. *Then*

$$\Pr[R_{n,q} \text{ has at least } k+1 \text{ components}] = e^{-\Omega(n)}.$$

PROOF. By Lemma 3.11,

$$|\mathcal{P}(n, \lfloor qn \rfloor, \ge k+1)| \le p(n, \ge k+1) \le \frac{1}{k!} \cdot p(n).$$

Hence

$$\left(\frac{|\mathcal{P}(n, \lfloor qn \rfloor, \ge k+1)|}{|\mathcal{P}(n, \lfloor qn \rfloor)|} \right)^{\frac{1}{n}} \le \left(\frac{1}{k!} \cdot \frac{p(n)}{|\mathcal{P}(n, \lfloor qn \rfloor)|} \right)^{\frac{1}{n}}.$$

But as $n \to \infty$, $(k!)^{\frac{1}{n}} \to e^c$, by [**MSW05**] $(p(n)/n!)^{\frac{1}{n}} \to \gamma$, and by Lemma 3.10 $\liminf (|\mathcal{P}(n, \lfloor qn \rfloor)|/n!)^{\frac{1}{n}} \ge \gamma(q)$. Hence

$$\limsup \left(\frac{|\mathcal{P}(n, \lfloor qn \rfloor, \ge k+1)|}{|\mathcal{P}(n, \lfloor qn \rfloor)|} \right)^{\frac{1}{n}} \le e^{-c} \frac{\gamma}{\gamma(q)} < 1,$$

and the lemma follows. □

LEMMA 3.13. *Let* $1 \le q < 3$. *Then*

$$\limsup \left(\frac{|\mathcal{P}(n, \lfloor qn \rfloor)|}{n!} \right)^{\frac{1}{n}} \le \gamma(q).$$

PROOF. Let $k_0 = k_0(n) = \lfloor n/\sqrt{\ln n} \rfloor$. As we know from Lemma 3.12 that $|\mathcal{P}(n, \lfloor qn \rfloor, \ge k_0 + 1)| = o(|\mathcal{P}(n, \lfloor qn \rfloor)|)$, we certainly have: $|\mathcal{P}(n, \lfloor qn \rfloor)| \le 2 \cdot |\mathcal{P}(n, \lfloor qn \rfloor, \le k_0)|$ for n sufficiently large. Therefore it suffices to show that

$$\limsup \left(\frac{|\mathcal{P}(n, \lfloor qn \rfloor, \le k_0)|}{n!} \right)^{\frac{1}{n}} \le \gamma(q).$$

By Lemma 3.8 and since we get each connected graph on n nodes with $\lfloor qn \rfloor + k$ edges by adding k edges to an appropriate connected graph with $\lfloor qn \rfloor$ edges (at most $\binom{\binom{n}{2}}{k}$ possibilities), we have:

$$\begin{aligned} |\mathcal{C}(n, \lfloor qn + k \rfloor)| &\le e\, |\mathcal{C}_c(n, \lfloor qn + k \rfloor)| \\ &\le e\, |\mathcal{C}_c(n, \lfloor qn \rfloor)| \cdot \binom{\binom{n}{2}}{k} \\ &\le e\, |\mathcal{C}(n, \lfloor qn \rfloor)| \cdot \left(\frac{en^2}{k} \right)^k. \end{aligned}$$

Next, $|\mathcal{C}(n, \lfloor qn+k \rfloor, k+1)| \le |\mathcal{C}(n, \lfloor qn+k \rfloor, \ge k+1)| \le \frac{1}{k!} |\mathcal{C}(n, \lfloor qn+k \rfloor)| \le \frac{e^k}{k^k} |\mathcal{C}(n, \lfloor qn+k \rfloor)|$ by Lemma 3.11. Summing up we get

$$|\mathcal{C}(n, \lfloor qn + k \rfloor, k+1)| \le e \left(\frac{en}{k} \right)^{2k} |\mathcal{C}(n, \lfloor qn \rfloor)|. \tag{3.7}$$

We conclude that for n sufficiently large

$$\begin{aligned}|\mathcal{P}(n,\lfloor qn\rfloor,\leq k_0)| &= \sum_{k=0}^{k_0-1}|\mathcal{C}(n,\lfloor qn+k\rfloor,k+1)|\\ &\overset{(3.7)}{\leq} e\sum_{k=1}^{k_0}\left(\frac{en}{k}\right)^{2k}|\mathcal{C}(n,\lfloor qn\rfloor)|\\ &\leq ek_0\left(\frac{en}{k_0}\right)^{2k_0}|\mathcal{C}(n,\lfloor qn\rfloor)|\\ &\leq e^{\frac{2n\ln\ln n}{\sqrt{\ln n}}}|\mathcal{C}(n,\lfloor qn\rfloor)|\end{aligned}$$

as $(\frac{en}{k+1})^{k+1}\geq(\frac{en}{k})^k$ for $k<n$. So

$$\left(\frac{|\mathcal{P}(n,\lfloor qn\rfloor,\leq k_0)|}{n!}\right)^{\frac{1}{n}}\leq e^{\frac{2\ln\ln n}{\sqrt{\ln n}}}\left(\frac{|\mathcal{C}(n,\lfloor qn\rfloor)|}{n!}\right)^{\frac{1}{n}}$$

for n sufficiently large, and since $e^{\frac{2\ln\ln n}{\sqrt{\ln n}}}\to 1$ as $n\to\infty$, we are done. □

The first part of Theorem 3.6 follows immediately from Lemmas 3.9, 3.10 and 3.13.

3.4. The function $\gamma(q)$

In this section we will investigate the behaviour of the function $\gamma(q)$ in more detail.

3.4.1. Continuity of $\gamma(q)$. Here, we will show that $\gamma(q)$ is log-concave and therefore continuous, thus proving more of Theorem 3.6.

The following inequality is useful in several places. We have

$$|\mathcal{P}(n,m+1)|\cdot(m+1)\geq|\mathcal{P}(n,m)|\cdot(3n-6-m),\tag{3.8}$$

as we have at least $(3n-6-m)$ possibilities to add an edge to a graph in $\mathcal{P}(n,m)$ and construct each graph in $\mathcal{P}(n,m+1)$ at most $(m+1)$ times. It follows that if $0\leq m_1<m_2\leq 3n-6$ then

$$\begin{aligned}|\mathcal{P}(n,m_1)| &\leq \frac{m_2!}{m_1!(m_2-m_1)!}|\mathcal{P}(n,m_2)|\\ &\leq \left(\frac{em_2}{m_2-m_1}\right)^{m_2-m_1}|\mathcal{P}(n,m_2)|.\end{aligned}$$

Thus if also $0<\delta<3$ and $m_2-m_1\leq\delta n$ then since $m_2\leq 3n$,

$$|\mathcal{P}(n,m_1)|\leq\left(\frac{3e}{\delta}\right)^{\delta n}|\mathcal{P}(n,m_2)|.\tag{3.9}$$

Lemma 3.14. *The function $\gamma(q)$ is log-concave on $[1,3]$ and continuous on the interval $(1,3]$.*

Proof. Let $1\leq x\leq z<3$, and let $y=x/2+z/2$. We proceed as in the proof of Lemma 3.9 (see Equation (3.6)), and obtain for even n,

$$\begin{aligned}\frac{|\mathcal{C}(n,\lfloor yn\rfloor)|}{n!} &\geq \frac{1}{n!}\frac{1}{n^2}\binom{n}{n/2}|\mathcal{C}_c(n/2,\lfloor xn/2\rfloor)|\,|\mathcal{C}_c(n/2,\lfloor zn/2\rfloor)|\\ &= \frac{1}{n^2}\frac{|\mathcal{C}_c(\frac{n}{2},\lfloor\frac{xn}{2}\rfloor)|}{\frac{n}{2}!}\frac{|\mathcal{C}_c(\frac{n}{2},\lfloor\frac{zn}{2}\rfloor)|}{\frac{n}{2}!}.\end{aligned}$$

But

$$\left(\frac{1}{n^2}\frac{|\mathcal{C}_c(\frac{n}{2},\lfloor\frac{xn}{2}\rfloor)|}{\frac{n}{2}!}\frac{|\mathcal{C}_c(\frac{n}{2},\lfloor\frac{zn}{2}\rfloor)|}{\frac{n}{2}!}\right)^{\frac{1}{n}} \to \gamma(x)^{\frac{1}{2}}\cdot\gamma(z)^{\frac{1}{2}}$$

as $n\to\infty$ and so

$$\gamma(y)\geq\gamma(x)^{\frac{1}{2}}\cdot\gamma(z)^{\frac{1}{2}}.$$

Thus

$$\log\gamma(y)\geq\frac{1}{2}\log\gamma(x)+\frac{1}{2}\log\gamma(z),$$

so $\gamma(y)$ is log-concave on $[1,3)$.

The above result may be extended to $[1,3]$. Let $1\leq x<3$ and let $y=x/2+3/2$. Let n be sufficiently large that $\frac{xn}{2}\leq\frac{3n}{2}-13$, and let n be even. We may form a planar graph G on $V=\{1,\ldots,n\}$ with $\lfloor yn\rfloor$ edges, as follows. First partition V into two sets of size $\frac{n}{2}$; then put a planar graph G_1 with $\lfloor\frac{xn}{2}\rfloor$ edges on the first set and a triangulation on the second set; and finally add six edges to G_1, maintaining planarity. The number of times a given graph G is constructed is at most $\binom{3n}{6}$. Hence

$$\begin{aligned}\frac{|\mathcal{P}(n,\lfloor yn\rfloor)|}{n!} &\geq \frac{1}{n!}\binom{n}{\frac{n}{2}}\cdot|\mathcal{P}\left(\frac{n}{2},\lfloor\frac{xn}{2}\rfloor\right)|\cdot t\left(\frac{n}{2}\right)\cdot(3n)^{-6}\\ &= \frac{|\mathcal{P}(\frac{n}{2},\lfloor\frac{xn}{2}\rfloor)|}{\frac{n}{2}!}\cdot\frac{t(\frac{n}{2})}{\frac{n}{2}!}\cdot(3n)^{-6}.\end{aligned}$$

But

$$\left(\frac{|\mathcal{P}(\frac{n}{2},\lfloor\frac{xn}{2}\rfloor)|}{\frac{n}{2}!}\cdot\frac{t(\frac{n}{2})}{\frac{n}{2}!}\cdot(3n)^{-6}\right)^{\frac{1}{n}}\to\gamma(x)^{\frac{1}{2}}\gamma(3)^{\frac{1}{2}}$$

as $n\to\infty$ and so

$$\log\gamma(y)\geq\frac{1}{2}\log\gamma(x)+\frac{1}{2}\log\gamma(3).$$

Since $\gamma(y)$ is log-concave on $[1,3]$, it follows that it is continuous on $(1,3)$. Also, $\gamma(q)$ tends to a limit at least $\gamma(3)$ as q increases to 3. But the Inequality (3.9) above shows that this limit is also at most $\gamma(3)$. Thus $\gamma(y)$ is continuous on $(1,3]$ as required. □

3.4.2. Uniform convergence to $\gamma(q)$.

LEMMA 3.15. *Let $1<a<3$ and let $\eta>0$. Then there exists n_0 such that for all $n\geq n_0$ and all m with $an\leq m\leq 3n-6$ we have*

$$\left|\left(\frac{|\mathcal{P}(n,m)|}{n!}\right)^{1/n}-\gamma\left(\frac{m}{n}\right)\right|<\eta.$$

PROOF. We may assume that $\eta<1$. We have seen that $\gamma(q)$ is continuous on $(1,3]$. Hence $\gamma(q)$ is uniformly continuous on $[a,3]$. Let $\delta_0>0$ be such that if $x,y\in[a,3]$ and $|x-y|<\delta_0$ then $|\gamma(x)-\gamma(y)|<\eta/5$.

Let $\eta'=\frac{\eta}{5\gamma}$. By (3.9) there exists an n_1 and $\delta_1>0$ such that if $n\geq n_1$, $0\leq m_1<m_2\leq 3n-6$ and $m_2-m_1\leq\delta_1 n+1$ then

$$|\mathcal{P}(n,m_1)|\leq(1+\eta')^n|\mathcal{P}(n,m_2)|.$$

Let $\delta = \min\{\delta_0, \delta_1\}$. Choose a finite list of points $a = q_0 < q_1 < \cdots < q_k = 3$ such that $|q_i - q_{i+1}| < \delta$. Note that for each $x \in [a, 3]$, if $x \in [q_i, q_{i+1}]$, then

$$|\mathcal{P}(n, \lfloor xn \rfloor)| \geq (1 - \eta')^n \, |\mathcal{P}(n, \lfloor q_i n \rfloor)|,$$

and

$$|\mathcal{P}(n, \lfloor xn \rfloor)| \leq (1 + \eta')^n \, |\mathcal{P}(n, \lfloor q_{i+1} n \rfloor)|.$$

Here we interpret $|\mathcal{P}(n, \lfloor q_{i+1} n \rfloor)|$ as $t(n)$ if $q_{i+1} = 3$. Let $n_0 \geq n_1$ be such that for each $n \geq n_0$,

$$\left| \left(\frac{|\mathcal{P}(n, \lfloor q_i n \rfloor)|}{n!} \right)^{1/n} - \gamma(q_i) \right| \leq \frac{\eta}{5}$$

for each $i = 0, \ldots, k$.

Let $n \geq n_0$, and let $an \leq m \leq 3n - 6$. Let i be such that $q_i n \leq m \leq q_{i+1} n$. Then

$$\begin{aligned} \left(\frac{|\mathcal{P}(n, m)|}{n!} \right)^{1/n} &\geq (1 - \eta') \left(\frac{|\mathcal{P}(n, \lfloor q_i n \rfloor)|}{n!} \right)^{1/n} \\ &\geq (1 - \eta')(\gamma(q_i) - \eta/5) \\ &> \gamma(m/n) - \eta, \end{aligned}$$

and

$$\begin{aligned} \left(\frac{|\mathcal{P}(n, m)|}{n!} \right)^{1/n} &\leq (1 + \eta') \left(\frac{|\mathcal{P}(n, \lfloor q_{i+1} n \rfloor)|}{n!} \right)^{1/n} \\ &\leq (1 + \eta')(\gamma(q_{i+1}) + \eta/5) \\ &< \gamma(m/n) + \eta. \end{aligned}$$

This completes the proof of Lemma 3.15. □

3.4.3. The Slope of $\gamma(q)$.

THEOREM 3.16. *The slope of $\gamma(q)$ tends to ∞ as $q \downarrow 1$ and to $-\infty$ as $q \uparrow 3$.*

PROOF. First we consider the case $q \downarrow 1$. For a planar graph G, let $\mathrm{add}(G)$ be the number of edges e that can be added to G such that $G + e$ stays planar; and let $\mathrm{add}(n, m)$ be the minimum value of $\mathrm{add}(G)$ over all graphs $G \in \mathcal{P}(n, m)$. Observe that

$$|\mathcal{P}(n, m)| \cdot \mathrm{add}(n, m) \leq |\mathcal{P}(n, m + 1)| \cdot (m + 1),$$

as we have at least $\mathrm{add}(n, m)$ possibilities to add an edge to a graph in $\mathcal{P}(n, m)$ and construct each graph in $\mathcal{P}(n, m + 1)$ at most $m + 1$ times. Thus

$$|\mathcal{P}(n, m)| \leq |\mathcal{P}(n, m + 1)| \cdot \frac{m + 1}{\mathrm{add}(n, m)}$$

and

$$\begin{aligned} |\mathcal{P}(n, m)| &\leq |\mathcal{P}(n, m + k)| \cdot \frac{(m + k) \cdot \ldots \cdot (m + 1)}{\mathrm{add}(n, m + k - 1) \cdot \ldots \cdot \mathrm{add}(n, m)} \\ &\leq |\mathcal{P}(n, m + k)| \cdot \frac{(m + k)^k}{(\mathrm{add}(n, m + k))^k}, \end{aligned} \tag{3.10}$$

as $\mathrm{add}(n, m)$ is monotone decreasing in m.

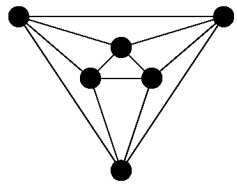

FIGURE 1. Octahedron, embedded in the plane.

It follows from Theorem 1.2 in [**GM04**] that, for any constant $K > 0$ there exists $\delta > 0$ (with $\delta < 1$) and n_0 such that $\mathrm{add}(n, m) \geq 2Kn$ for all $n \geq n_0$ and all $0 \leq m \leq (1+\delta)n$. Thus if $0 < s < t < \delta$ then

$$\begin{aligned} |\mathcal{P}(n, \lfloor (1+s)n \rfloor)| &\overset{(3.10)}{\leq} \left(\frac{(1+\delta)n}{2Kn}\right)^{\lfloor tn \rfloor - \lfloor sn \rfloor} |\mathcal{P}(n, \lfloor (1+t)n \rfloor)| \\ &\leq K^{-(t-s)n+1} |\mathcal{P}(n, \lfloor (1+t)n \rfloor)|, \end{aligned}$$

as $\lfloor tn \rfloor - \lfloor sn \rfloor \geq tn - 1 - sn$ and $(1+\delta)/2 < 1$. Hence using (3.4), dividing by $n!$, taking the n-th root, and taking the limit yields

$$\gamma(1+s) \leq K^{-(t-s)} \cdot \gamma(1+t).$$

Thus $\lambda(q) = \ln \gamma(q)$ satisfies

$$\frac{\lambda(1+t) - \lambda(1+s)}{t-s} \geq \ln K.$$

Hence the left derivative of $\lambda(q)$ at $1+t$ and the right derivative of $\lambda(q)$ at $1+s$ are at least $\ln K$. Thus the slope of $\lambda(q)$ tends to ∞ as $q \downarrow 1$, as required.

Now, consider the case $q \uparrow 3$. Let H be the octahedron (see Figure 1), embedded in the plane. In [**RW88**] it was shown that there exist n_0 and $\eta > 0$ such that for all $n \geq n_0$, at least half the triangulations T_n on n nodes contain at least ηn pairwise node disjoint induced copies of H. Let $0 < \varepsilon < \eta$. (Think of $\varepsilon \ll \eta$.) For each such T_n choose $\lfloor \varepsilon n \rfloor$ pairwise node disjoint induced copies of H and delete the lexicographically first edge on the cycle bounding the 'inner face' (in the unique embedding). This yields at least

$$\binom{\lceil \eta n \rceil}{\lfloor \varepsilon n \rfloor} \geq \left(\frac{\eta}{\varepsilon}\right)^{\varepsilon n - 1}$$

graphs in $\mathcal{P}(n, 3n - 6 - \lfloor \varepsilon n \rfloor)$, as $\binom{m}{k} \geq \left(\frac{m}{k}\right)^k$. Further, each such graph is 3-connected since H less an edge is 3-connected, and so by Whitney's theorem [**Whi32**] each has a unique embedding on the sphere. It follows that the generated graphs are all distinct. Thus

$$|\mathcal{P}(n, 3n - 6 - \lfloor \varepsilon n \rfloor)| \geq \left(\frac{\eta}{\varepsilon}\right)^{\varepsilon n - 1} \cdot \frac{1}{2} |\mathcal{P}(n, 3n-6)|$$

and so dividing by $n!$, taking the n-th root, and taking logarithms yields

$$\frac{1}{n} \ln \left(\frac{|\mathcal{P}(n, 3n - 6 - \lfloor \varepsilon n \rfloor)|}{n!} \right) \geq \varepsilon \ln \left(\frac{\eta}{\varepsilon}\right) + \frac{1}{n} \ln \left(\frac{|\mathcal{P}(n, 3n-6)|}{n!} \right) + \underbrace{\frac{1}{n} \ln \frac{\varepsilon}{2\eta}}_{o(1)}.$$

Hence, taking the limit and by Equation (3.5) we obtain

$$\lambda(3 - \varepsilon) \geq \varepsilon \ln \left(\frac{\eta}{\varepsilon}\right) + \lambda(3).$$

Thus

$$\frac{\lambda(3-\varepsilon)-\lambda(3)}{\varepsilon} \geq \ln\left(\frac{\eta}{\varepsilon}\right) = \ln\eta + \ln\frac{1}{\varepsilon}.$$

But this holds for each $0 < \varepsilon < \eta$. As the right hand side tends to ∞ as $\varepsilon \downarrow 0$, this completes the proof. □

3.4.4. Discontinuity at $\gamma(1)$. Recall from Theorem 3.5 that $\gamma(q) = 0$ for $0 \leq q < 1$ and $\gamma(1) = e$. The next result shows that this discontinuity vanishes as we approach 1 from below if the scale is changed appropriately.

THEOREM 3.17. *Let $\beta \geq 0$ be a constant. If $m = m(n) = n - (\beta + o(1))(n/\ln n)$ then*

$$\left(\frac{|\mathcal{P}(n,m)|}{n!}\right)^{\frac{1}{n}} \to e^{1-\beta} \quad \text{as } n \to \infty.$$

We may now see that the result (3.5) concerning convergence for $q \in [1,3]$ can be extended to all of $[0,3]$, as long as we insist that if $q = 1$ then either $m \geq n$ for all sufficiently large n, or $(1 - m/n)\ln n \to 0$ or $\to \infty$ as $n \to \infty$. This follows immediately from Theorem 3.17, as the arguments given in the proof of Theorem 3.16 imply that for $0 \leq m < n$ we have $|\mathcal{P}(n,m)| \leq |\mathcal{P}(n,m+1)|$.

Theorem 3.17 is a special case of the following rather general theorem. To state the theorem we need the following definitions. A set $\mathcal{A}$ of labelled graphs, closed under isomorphism, is called *small* if there exists a constant C such that the set $\mathcal{A}(n)$ of graphs in $\mathcal{A}$ on the node set $\{1,\ldots,n\}$ has size at most $C^n n!$ for sufficiently large n. Recall that the class $\mathcal{A}$ is called *weakly addable* if for each graph G in $\mathcal{A}$, if u and v are nodes in distinct components of G, then the graph obtained from G by adding an edge joining u and v is also in $\mathcal{A}$. (See [**MSW06**] for the definition of *addable*.) Note that the class $\mathcal{P}$ of planar graphs is small and weakly addable, and, as seen before, the class $\mathcal{P}(n,m)$ is small but is definitely not weakly addable (which makes it harder to deal with).

THEOREM 3.18. *Let $\mathcal{A}$ be a small weakly addable set of graphs containing the set $\mathcal{F}$ of forests, and let $\mathcal{A}(n,m)$ denote the set of graphs in $\mathcal{A}$ on the node set $\{1,\ldots,n\}$ with exactly m edges. Let $\beta \geq 0$ be a constant. If $m = m(n) = n - (\beta + o(1))(n/\ln n)$ then*

$$\left(\frac{|\mathcal{A}(n,m)|}{n!}\right)^{\frac{1}{n}} \to e^{1-\beta} \quad \text{as } n \to \infty.$$

Note that $\mathcal{A}$ can be very different from $\mathcal{P}$, with a different growth constant.

In order to prove Theorem 3.18, and thus Theorem 3.17, in the next lemma we count forests with m edges, and general graphs with m edges and few components. After that we check that usually a graph in $\mathcal{A}(n,m)$ will not have too many components, so that we can use the last result. Let $\mathcal{G}(n,m)$ denote the set of all graphs on the node set $\{1,\ldots,n\}$ with exactly m edges, let $\mathcal{F}(n)$ denote the set of forests on node set $\{1,\ldots,n\}$, and let $\mathcal{F}(n,m)$ denote the subset of these forests which have exactly m edges. Let $\mathcal{F}_c(n)$ denote the set of connected forests (i.e. trees) on n nodes. Recall that for a graph G we denote by $\kappa(G)$ the number of components of G.

LEMMA 3.19. *Let $m = m(n)$ and $k = k(n)$ be integers such that $1 \leq n - m < k = o(n)$. Then*

$$|\mathcal{F}(n,m)| = e^{m \ln n + o(n)}$$

and

$$|\{G \in \mathcal{G}(n,m) : \kappa(G) \leq k\}| = e^{m \ln n + o(n)}.$$

PROOF. If $G \in \mathcal{G}(n,m)$ and $\kappa(G) \leq k$ then G contains a forest with $n-k$ edges, and we may obtain G by adding $m - n + k$ edges to this forest. Thus

$$|\{G \in \mathcal{G}(n,m) : \kappa(G) \leq k\}| \leq |\mathcal{F}(n,n-k)| \cdot \binom{\binom{n}{2}}{m-n+k}.$$

Also, by Theorem 3.7 (Theorem 2.2 of [**MSW05**]) we have $|\mathcal{F}(n)| \leq e|\mathcal{F}_c(n)|$ and $|\mathcal{F}(n,n-k)| \leq |\mathcal{F}(n)|/(k-1)!$ (see also Lemma 3.11 above). Observe that by Cayley's famous formula for the number of trees on n nodes, we have $|\mathcal{F}_c(n)| = n^{n-2}$, hence

$$|\mathcal{F}(n,n-k)| \leq \frac{en^{n-2}}{(k-1)!}.$$

Moreover, observe that if $t = t(n) = o(n)$ then

$$t \ln n - t \ln t = n \cdot \frac{\ln(n/t)}{n/t} = o(n). \tag{3.11}$$

Now, as $k = o(n)$ we have

$$\begin{aligned}
\frac{en^{n-2}}{(k-1)!} &\leq \frac{en^{n-2}}{\left(\frac{k-1}{e}\right)^{k-1}} = \frac{e^k n^{n-2}}{(k-1)^{k-1}} \\
&= \exp((n-2)\ln n + k - (k-1)\ln(k-1)) \\
&\leq \exp(n \ln n - k \ln k + o(n)) \\
&= \exp((n-k)\ln n + k \ln n - k \ln k + o(n)) \\
&= \exp((n-k)\ln n + o(n))
\end{aligned}$$

by Equation (3.11). Thus

$$\begin{aligned}
|\{G \in \mathcal{G}(n,m) : \kappa(G) \leq k\}| &\leq |\mathcal{F}(n,n-k)| \cdot \binom{\binom{n}{2}}{m-n+k} \\
&\leq \frac{en^{n-2}}{(k-1)!} \cdot \left(\frac{en^2}{2(m-n+k)}\right)^{m-n+k} \\
&\leq e^{(n-k)\ln n + o(n)} \cdot e^{(k-n+m)\left(\ln \frac{n}{k-n+m} + \ln n + \ln \frac{e}{2}\right)}.
\end{aligned}$$

Note that $m - n + k = k - (n-m) = o(n)$ and hence by Equation (3.11)

$$(k-n+m)\ln\frac{n}{k-n+m} = o(n).$$

Hence

$$\begin{aligned}
|\{G \in \mathcal{G}(n,m) : \kappa(G) \leq k\}| &\leq \exp\left((n-k)\ln n + (k-n+m)\ln n + o(n)\right) \\
&= \exp\left(m \ln n + o(n)\right).
\end{aligned}$$

But since $1 \leq n - m < k$, and by Cayley's formula

$$\begin{aligned}
|\{G \in \mathcal{G}(n,m) : \kappa(G) \leq k\}| &\geq |\mathcal{F}(n,m)| \geq |\mathcal{F}(m+1,m)| = (m+1)^{m-1} \\
&= \exp(m \ln n + o(n)),
\end{aligned}$$

where the last equality holds as $m \sim n$. This completes the proof of both inequalities in the lemma. □

Next we check that, with conditions as in Theorem 3.18, usually there will not be too many components, so that we can use the last result.

LEMMA 3.20. *Let $\mathcal{A}$ be a small weakly addable set of graphs containing the set $\mathcal{F}$ of forests. Let the constant γ be such that $|\mathcal{A}(n)| = O(\gamma^n n!)$. Let $\beta > 0$ be fixed, and let $m = m(n) = n - (\beta + o(1))(n/\ln n)$. Let the constant $c > 0$ satisfy $c > \beta + \ln\gamma - 1$. Then, for $R_{n,m}$ uniformly distributed over $\mathcal{A}(n,m)$,*

$$\Pr[\kappa(R_{n,m}) > cn/\ln n] = e^{-\Omega(n)}. \tag{3.12}$$

PROOF. Let $k = k(n) = \lceil cn/\ln n \rceil$. Then

$$|\{G \in \mathcal{A}(n,m) : \kappa(G) > k\}| \le |\{G \in \mathcal{A}(n) : \kappa(G) > k\}| \le \frac{|\mathcal{A}(n)|}{k!}$$

by Theorem 3.7 (Theorem 2.2 of [**MSW05**]). Assuming $m \le n-1$,

$$|\mathcal{A}(n,m)| \ge |\mathcal{F}(n,m)| \ge |\mathcal{F}(m+1,m)| = (m+1)^{m-1}$$

as we saw above. Hence, since $k\ln k \ge \frac{cn}{\ln n}\ln\frac{cn}{\ln n} = \frac{cn}{\ln n}(\ln c + \ln n - \ln\ln n) = cn + o(n)$, $|\mathcal{A}(n)| \le a\gamma^n n!$ for some constant a, and $(m+1)^{m-1} = \exp(m\ln n + o(n))$,

$$\begin{aligned}
\Pr[\kappa(R_{n,m}) > k] &\le \frac{|\{G \in \mathcal{A}(n) : \kappa(G) > k\}|}{|\mathcal{A}(n,m)|} \\
&\le \frac{|\mathcal{A}(n)|}{k!(m+1)^{m-1}} \\
&\le \frac{a\gamma^n n!}{\left(\frac{k}{e}\right)^k e^{m\ln n + o(n)}} \\
&= \exp\left(n\ln n - n + n\ln\gamma - k\ln k - (n - \beta n/\ln n)\ln n + o(n)\right) \\
&= \exp(n(-1 + \ln\gamma - c + \beta + o(1))) \\
&= e^{-\Omega(n)}
\end{aligned}$$

by our choice of c. □

Now we are able to prove Theorem 3.18.

PROOF OF THEOREM 3.18. Let c be as in Lemma 3.20. Then for n sufficiently large

$$\Pr[\kappa(R_{n,m}) > cn/\ln n] \le \frac{1}{2},$$

and so

$$\begin{aligned}
|\mathcal{A}(n,m)| &\le 2|\{G \in \mathcal{A}(n,m) : \kappa(G) \le cn/\ln n\}| \\
&\le 2|\{G \in \mathcal{G}(n,m) : \kappa(G) \le cn/\ln n\}| \\
&= e^{m\ln n + o(n)}
\end{aligned}$$

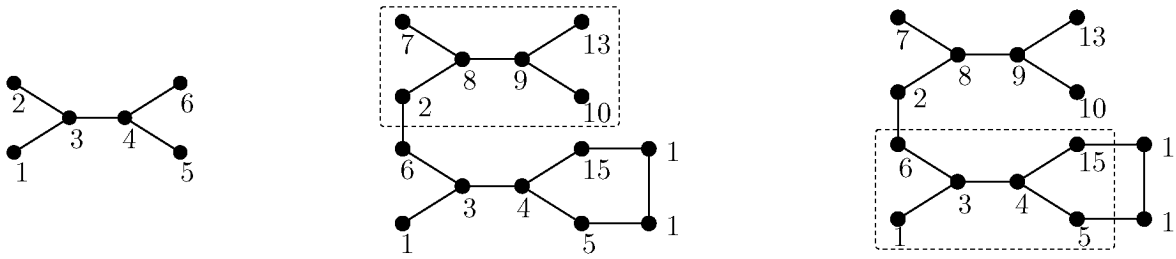

FIGURE 2. The graph H, an appearance of H, and a non-appearance of H.

by Lemma 3.19. Now, since $|\mathcal{A}(n,m)| \geq |\mathcal{F}(n,m)|$, we may use Lemma 3.19 again to get $|\mathcal{A}(n,m)| = e^{m \ln n + o(n)}$. Finally, for m as given,

$$\left(\frac{e^{m \ln n + o(n)}}{n!}\right)^{\frac{1}{n}} = \exp\left(\frac{1}{n}(m \ln n - n \ln n + n + o(n))\right) = e^{1-\beta+o(1)},$$

which completes the proof. □

3.5. Degrees, Faces, Subgraphs and Connectivity

To prove structural properties of the random planar graph with n nodes and $\lfloor qn \rfloor$ edges, for instance degree, face, and subgraph properties, we start by proving a more general result.

First we define what it means for a graph H to *appear* in a graph G, following [**MSW05**]. Let H be a graph on the node set $\{1, \ldots, h\}$, and let G be a graph on the node set $\{1, \ldots, n\}$ where $n > h$. Let $W \subset V(G)$ with $|W| = h$, and let the *root* r_W denote the least element in W. We say that H *appears* at W in G if (a) the increasing bijection from $\{1, \ldots, h\}$ to W gives an isomorphism between H and the induced subgraph $G[W]$ of G; and (b) there is exactly one edge in G between W and the rest of G, and this edge is incident with the root r_W. See Figure 2 for an example. The marked subgraph in the figure on the right is a non-appearance of H; in fact it conflicts with both conditions (a) and (b) that have to be met by an appearance: the increasing bijection from $\{1, \ldots, h\}$ to W is not an isomorphism as required, and there is more than one edge between W and the rest of G.

Let $f_H(G)$ be the number of appearances of H in G, that is the number of sets $W \subseteq V(G)$ such that H appears at W in G.

THEOREM 3.21. *Let $1 \leq q < 3$ and let H be a fixed connected planar graph on the node set $\{1, \ldots, h\}$, where H is a tree if $q = 1$. Then there exists $\alpha = \alpha(H, q) > 0$ such that*

$$\Pr[f_H(R_{n,q}) \leq \alpha n] = e^{-\Omega(n)}. \tag{3.13}$$

Before giving a proof for this theorem, let us note that it implies Corollary 3.3 – a similar statement but for pairwise node disjoint copies of a subgraph in $R_{n,q}$. Given a graph G, let $\mathcal{W}$ be the collection of all sets $W \subseteq V(G)$ such that H appears at W in G: then each set $W \in \mathcal{W}$ meets at most $h - 1$ other sets in $\mathcal{W}$, as the root of W has to be contained in each appearance that meets W and the root of H can be connected to at most $h - 1$ cut edges. Thus there is a set of at least $|\mathcal{W}|/h$ pairwise disjoint sets $W \in \mathcal{W}$.

Note, that if $0 \leq q < 1$ and H is a given connected planar graph with more than one cycle, then with high probability $R_{n,q}$ has no subgraph isomorphic to H. This was observed recently by Chris Dowden [**Dow06**].

However, the last theorem implies that, if we fix any q with $1 < q < 3$, then for large n, a random planar graph $R_{n,q}$ is very likely to contain a K_4 and thus be 4-chromatic. Moreover the probability of containing a K_4 is sufficiently large, that one can find an optimal colouring in expected polynomial time in the following way, as noted in [**MSW05**]. We may first test in linear time (see [**PY81**]) whether a K_4 is present. If a K_4 is present, we apply the quadratic time algorithm to four-colour planar graphs which follows from the proof of the four-colour-theorem [**RSST97**]. If no K_4 is found, which happens with probability $e^{-\Omega(n)}$, we apply the $\sqrt{n}$-separator theorem to colour the graph optimally in subexponential time $O(c^{\sqrt{n}})$ [**LT79**]. Therefore one can find the chromatic number of a planar graph with n nodes and $\lfloor qn \rfloor$ edges in quadratic expected time.

Proof of Theorem 3.21. We shall follow the lines of the proof in [**MSW05**] of Theorem 4.1 there, except that now we need to control the numbers of edges in the graphs we construct. (In [**BGR92**] a similar approach, together with generating functions, can be found for proving an analogous result for submaps of maps.)

We shall choose (small) positive α (see below for the precise value) and prove that, for n sufficiently large,

$$\Pr[f_H(R_{n,q}) \leq \alpha n] \leq e^{-\alpha n}. \tag{3.14}$$

For each $\varepsilon > 0$, there is a positive integer $n_0 = n_0(\varepsilon)$ such that for each $n \geq n_0$ we have by Theorem 3.6

$$(1-\varepsilon)^n n! \gamma(q)^n \leq |\mathcal{P}(n, \lfloor qn \rfloor)| \leq (1+\varepsilon)^n n! \gamma(q)^n. \tag{3.15}$$

Assume that Equation (3.14) does not hold for some chosen ε and some $n \geq n_0$. We intend to show that then for some suitable $\delta > 0$ and $\tilde{n} \sim (1+\delta)n$ we have

$$|\mathcal{P}(\tilde{n}, \lfloor q\tilde{n} \rfloor)| > (1+\varepsilon)^{\tilde{n}} \cdot \tilde{n}! \cdot \gamma(q)^{\tilde{n}},$$

contradicting (3.15).

Let us first sketch out roughly how we will achieve this contradiction, by constructing many graphs G' in $\mathcal{P}(\tilde{n}, \lfloor q\tilde{n} \rfloor)$ on the node set $\{1, \ldots, \tilde{n}\}$: later we shall return to fill in some more details. First we choose a subset of $\tilde{n} - n \sim \delta n$ *special* nodes ($\binom{\tilde{n}}{n}$ choices) and a graph $G \in \mathcal{P}(n, \lfloor qn \rfloor)$ on the remaining n nodes that satisfies $f_H(G) \leq \alpha n$. By assumption there are at least

$$e^{-\alpha n} \cdot |\mathcal{P}(n, \lfloor qn \rfloor)| \geq e^{-\alpha n}(1-\varepsilon)^n \gamma(q)^n n!$$

such graphs G.

Next we consider the special nodes. We can construct about $\frac{\delta n}{h}$ disjoint copies of H with these nodes. Then we connect these copies to G, each by a single edge, to build the graph G'. Now a counting argument as in [**MSW05**] shows that we have built more than $(1+\varepsilon)^{\tilde{n}}\, \tilde{n}!\, \gamma(q)^{\tilde{n}}$ distinct planar graphs with $\tilde{n}$ nodes and $\lfloor q\tilde{n} \rfloor$ edges, thus yielding the desired contradiction. The key point in the counting argument is that we can show that each graph G' is not constructed too often, since G had few appearances of H.

However, we have to be careful about the number of edges in G'; as H can be an arbitrary connected planar graph, the construction will not in general lead to $\lfloor q\tilde{n} \rfloor$ edges in G'. In this case, instead of directly dealing with H, we build connected planar graphs H_1 and H_2, which contain H as an induced subgraph and which yield the correct number of edges in G' when we connect them to G. If we can show that such graphs H_1 and H_2 must always exist then we can proceed as in the proof of [**MSW05**], replacing H by H_1 and H_2.

Let H have m_H edges. If $(m_H+1)/h = q$ then we may just use H. For if G has n nodes and $\lfloor qn \rfloor$ edges, and we add a disjoint copy of H together with one edge between H and G, the resulting graph has $n+h$ nodes and $\lfloor qn \rfloor + m_H + 1 = \lfloor q(n+h) \rfloor$ edges. Note that this case must hold if $q = 1$ and H is a tree. If $(m_H+1)/h \neq q$ we form H_1 and H_2 as follows.

Let x be the least integer at least 3 such that

$$\frac{x + m_H + 2}{x+h} \leq q \quad \text{and} \quad \frac{3x - 7 + m_H + 2}{x+h} > q.$$

Then let y be the greatest integer such that $\frac{y+m_H+2}{x+h} \leq q$. Since $1 < q < 3$ these constants x and y must exist. Observe that $x \leq y \leq 3x-7$, and if we set $q_1 = \frac{y+m_H+2}{x+h}$ and $q_2 = \frac{y+m_H+3}{x+h}$ then $q_1 \leq q < q_2$.

To build the graph H_1, we start with a copy of H on the nodes $x+1, \ldots, x+h$. Then we take a connected planar graph H_1' on the nodes $1, \ldots, x$ with y edges, and connect this graph to H by adding an edge between nodes 1 and $x+1$. Observe that H_1 has $h_1 = h + x$ nodes and $m_{H_1} = m_H + y + 1$ edges. Since $y \leq 3x - 7$ we can form a planar graph H_2' by adding an edge to H_1'. Let H_2 be formed as above but with H_2' instead of H_1'.

Now that we have seen how to construct the graphs H_1 and H_2, let us return to describe the proof in more detail. The constants α, β and ε may be taken exactly as in [**MSW05**] except with h replaced by h_1, i.e. let $\beta = e^2\gamma(q)^{h_1}(h_1+2)h_1!$ and $\alpha = \alpha(H,q) < 1/(c(H,q)\gamma(q)^{h_1}(h_1+2)h_1!)$ for a constant $c(H,q) > e^2$; this implies that $\alpha\beta < 1$ and we may therefore write $(\alpha\beta)^\alpha = 1 - 3\varepsilon$, where $0 < \varepsilon < \frac{1}{3}$. Note that

$$\frac{1-\varepsilon}{(1-3\varepsilon)(1+\varepsilon)^2} > 1. \tag{3.16}$$

Similarly we take $\delta = \alpha h_1$. Let $k = \lfloor \alpha n \rfloor$, and let $\tilde{n} = n + kh_1$.

We construct graphs G' in $\mathcal{P}(\tilde{n}, \lfloor q\tilde{n} \rfloor)$ on the node set $\{1, \ldots, \tilde{n}\}$, as follows. Let k_1 be the integer such that

$$k_1 = \lfloor qn \rfloor - \lfloor \tilde{n}q \rfloor + k(m_H + y + 3).$$

Note that $0 \leq k_1 \leq k$ by the choice of y as

$$k(m_H + y + 3) = \frac{\tilde{n} - n}{h_1}(m_H + y + 3) = (\tilde{n} - n)q_2 > (\tilde{n} - n)q$$

and the statement follows by the integrality of the terms in question. First we choose a subset of $kh_1 \sim \delta n$ nodes as special nodes and a graph $G \in \mathcal{P}(n, \lfloor qn \rfloor)$ on the remaining n nodes that satisfies $f_H(G) \leq \alpha n$. Next we partition the kh_1 special nodes into k blocks B of size h_1. Let r_B be the node with the smallest label over all nodes of the block B. Consider the blocks ordered by increasing value of the root node: divide the blocks into two sets of sizes k_1 and $k - k_1$, the first k_1 blocks in the order and the rest. On each of the first k_1 blocks B, we put a copy of H_1 such that the increasing bijection from $\{1, \ldots, h_1\}$ to B is an isomorphism

between H_1 and this copy. We do the same for the second set of $k - k_1$ blocks, except with H_2 instead of H_1. To connect each block B to G, we choose a non-special node v_B in G and add an edge between r_B and v_B.

This completes the construction of the graph G'. The number of edges in G' is $\lfloor q\tilde{n} \rfloor$ as

$$k_1(m_H + y + 2) + (k - k_1)(m_H + y + 3) = \lfloor q\tilde{n} \rfloor - \lfloor qn \rfloor.$$

Thus indeed G' is in the set $\mathcal{P}(\tilde{n}, \lfloor q\tilde{n} \rfloor)$, as required.

To complete the proof by a counting argument, we proceed as follows (we shall write x instead of $\lfloor x \rfloor$ or $\lceil x \rceil$ in the subsequent to avoid cluttering formulae). For each choice of special nodes, and each $G \in \mathcal{P}(n, \lfloor qn \rfloor)$ on the remaining n nodes, we construct

$$\binom{\delta n}{h_1 \cdots h_1} \cdot \frac{1}{(\alpha n)!} \cdot n^{\alpha n} = \frac{(\delta n)! n^{\alpha n}}{(h_1!)^{\alpha n}(\alpha n)!} \geq \frac{(\delta n)!}{(h_1! \alpha)^{\alpha n}}$$

graphs G'.

For considering how often the same graph G' is constructed in this way, we introduce the following notation: we call an oriented edge $e = uv$ *good*, if

- it is a bridge in G',
- the component C of $G' - e$ containing u has h_1 nodes,
- u is the root of C,
- and the increasing bijection of $\{1, \ldots, h_1\}$ to the node set of C is an isomorphism between C and H_1 or H_2.

Observe that each added oriented edge $r_B v_B$ for connecting a copy of H_1 or H_2 to G is *good*. Indeed there is exactly one *good* oriented edge for each appearance of H_1 or H_2 in G'. We shall see that G' contains at most $(h_1 + 2)\alpha n$ good oriented edges and so the number of times that G' can be constructed is at most $\binom{(h_1+2)\alpha n}{\alpha n} \leq ((h_1 + 2)e)^{\alpha n}$.

We may obtain a bound on the number of *good* edges in G' as follows: (a) By the construction of G' we added exactly αn oriented *good* edges $r_B v_B$. (b) By assumption G contains at most αn appearances of H and hence also at most αn *good* oriented edges $e = uv \in E(G)$. (c) If H_1 or H_2 contains a cut-edge, some 'extra' *good* edges may arise in G' when connecting H_1 (or H_2) to G (see Figure 3): Consider a block B, and let $\tilde{B}$ denote the connected graph formed from the induced subgraph $G'[B]$, which is isomorphic to H_1 (or H_2), together with the node v_B and the edge $r_B v_B$. Each *good* oriented edge must be a cut-edge in such a graph $\tilde{B}$ oriented away from v_B, and in each graph $\tilde{B}$ there are at most h_1 cut-edges. Thus there are at most $h_1 \alpha n$ additional *good* oriented edges.

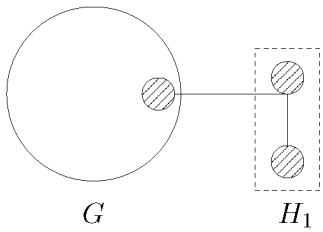

FIGURE 3. If a cut-edge in H_1 exists, some 'extra' *good* edges may arise if H_1 is connected to G.

Putting the above results together we obtain

$$\begin{aligned}
|\mathcal{P}((1+\delta)n, \lfloor q(1+\delta)n \rfloor)| &\geq \binom{(1+\delta)n}{\delta n} \cdot e^{-\alpha n}(1-\varepsilon)^n \gamma(q)^n n! \cdot \frac{(\delta n)!}{(h_1!\alpha)^{\alpha n}} \cdot ((h_1+2)e)^{-\alpha n} \\
&= [(1+\delta)n]! \cdot \gamma(q)^{(1+\delta)n} \cdot (1-\varepsilon)^n \cdot \left(e^2 \gamma(q)^{h_1}(h_1+2)h_1!\alpha\right)^{-\alpha n} \\
&\overset{(3.15)}{\geq} |\mathcal{P}((1+\delta)n, \lfloor q(1+\delta)n \rfloor)| \ (1+\varepsilon)^{-(1+\delta)n} \cdot (1-\varepsilon)^n \cdot (1-3\varepsilon)^{-n} \\
&\geq |\mathcal{P}((1+\delta)n, \lfloor q(1+\delta)n \rfloor)| \left(\frac{1-\varepsilon}{(1-3\varepsilon)(1+\varepsilon)^2}\right)^n \\
&\overset{(3.16)}{>} |\mathcal{P}((1+\delta)n, \lfloor q(1+\delta)n \rfloor)|,
\end{aligned}$$

yielding the desired contradiction. □

By choosing appropriate graphs H and applying Theorem 3.21, we can deduce Theorems 3.1 and 3.2, arguing as in [**MSW05**]. For Theorem 3.1 we let H be a star on the nodes $1, \ldots, k+1$ with centre at node $k+1$. For Theorem 3.2 we let H be a k-cycle on the nodes $1, \ldots, k$.

We now turn to copies in $R_{n,q}$ of a *plane* graph H, that is of a graph H embedded in the plane. What does it mean for H to 'appear' in $R_{n,q}$? Let H and G be two plane graphs. Let us say that H *appears* in G at the node set $W \subseteq V(G)$, if (a) the underlying graph of H appears at W in the underlying graph of G, (b) there is a continuous deformation of the plane taking H to the induced plane subgraph $G[W]$ of G, and (c) no node of $V(G) \setminus W$ is contained in an interior face of $G[W]$. By arguing as in the case of Theorem 3.2 (see also [**MSW05**]) we may obtain the following proposition.

PROPOSITION 3.22. *Let H be a connected plane graph. Let $f_H(G)$ denote the function which counts for a planar graph G the maximum over all embeddings of G of the number of pairwise node disjoint appearances of H. Let $1 < q < 3$. Then there exists a constant $\alpha = \alpha(H, q) > 0$ such that*

$$\Pr[f_H(R_{n,q}) < \alpha n] = e^{-\Omega(n)}.$$

If H is 3-connected then the claim remains true if $f_H(G)$ is defined by minimising *over all embeddings of G.*

Next we prove lower bounds on the maximum degree and maximum face size in a plane embedding. Note, McDiarmid and Reed have recently shown that the maximum degree is $\Theta(\log n)$ w.h.p.

THEOREM 3.23. *Let $1 < q < 3$. Then with high probability, in $R_{n,q}$ there is a set of at least $(1 + o(1)) \ln n / \ln \ln n$ pendant nodes with a common neighbour which lies on a triangle.*

PROOF SKETCH. The proof is essentially the same as the one of Corollary 5.3 in [**MSW06**], thus we only give a short proof sketch here.

For each $G \in \mathcal{P}(n, \lfloor qn \rfloor)$ we consider the set W of nodes of degree 1 which are adjacent to a node in a triangle. By Theorem 3.21 we know that there exists $\alpha > 0$ such that $R_{n,q}$ contains at least αn such nodes with probability tending to 1 as $n \to \infty$.

Now, let H denote the graph induced on the nodes not in W. Then the maximum number of edges in $R_{n,q}$ between a node in H and the nodes in W has the same distribution as the maximum bin load when we throw $|W|$ balls independently at random into n' bins, where n' denotes the number of nodes in triangles in H. But a well-known result concerning balls and bins shows that if we throw at least αn balls into at most n bins, then the maximum bin load is $(1 + o(1)) \ln n / \ln \ln n$ with high probability and the result follows. □

We can extend the last theorem as follows.

THEOREM 3.24. *Let $1 < q < 3$, and let $\varepsilon > 0$. For each n let T_n be a tree with at most $(1 - \varepsilon) \ln n / \ln \ln n$ nodes. Then with high probability in $R_{n,q}$ there is an appearance of T_n.*

PROOF. Let $k = k(n) = \lfloor (1 - \varepsilon) \ln n / \ln \ln n \rfloor$. Suppose that the probability that T_n fails to appear in $R_{n,q}$ is at least $\delta > 0$ for infinitely many n. By Theorem 3.23 there is an n_0 such that for each $n \geq n_0$, with probability at least $1 - \delta/2$ there is a set of at least $k + 1$ pendant nodes in $R_{n,q}$ with a common neighbour v which lies on a triangle. Let $n_1 \geq n_0$ be such that $2/k < \delta$ for each $n \geq n_1$.

Let $n \geq n_1$ be such that the probability that T_n fails to appear in $R_{n,q}$ is at least δ. Then we shall show that $\delta \leq 2/k$, a contradiction which will complete the proof.

Let S be the set of graphs G in $\mathcal{P}(n, \lfloor qn \rfloor)$ such that there is a set of at least $k + 1$ pendant nodes with a common neighbour v_G which lies on a triangle, and the tree T_n fails to appear. Thus $|S| \geq \delta/2 \ |\mathcal{P}(n, \lfloor qn \rfloor)|$.

Let $G \in S$. Choose a set W of k of the pendant nodes adjacent to v_G (there are at least k ways to do this), and use them to form an appearance of T_n with root adjacent to v_G. Clearly we may recover the graph G from the new graph. Hence

$$|\mathcal{P}(n, \lfloor qn \rfloor)| \geq k|S| \geq k\delta/2 \ |\mathcal{P}(n, \lfloor qn \rfloor)|,$$

and so $\delta \leq 2/k$, as required. □

Clearly, in every embedding of $R_{n,q}$ one face has to contain the entire tree T_n, and thus we obtain:

COROLLARY 3.25. *Let $1 < q < 3$. With high probability every plane embedding of $R_{n,q}$ contains a face incident with at least $(1 + o(1)) \ln n / \ln \ln n$ distinct edges, and with facial walk of length at least $(2 + o(1)) \ln n / \ln \ln n$.*

Finally, let us consider the connectedness of $R_{n,q}$. Note first that $\Pr[R_{n,q} \text{ is connected}] = 0$ for $0 \leq q < 1$, as in this case $\lfloor qn \rfloor < n-1$ for n sufficiently large. For $q = 3$ we consider triangulations which are clearly connected. For the remaining interval $1 < q < 3$, we know from (3.2) and (3.3) – strengthening a result of [**GMSW05**] – that $\Pr[R_{n,q} \text{ is connected}]$ tends to the positive constant $\alpha_c(q)/\alpha(q)$ as $n \to \infty$.

The remaining case is when $q = 1$. Let us check that here the probability of being connected tends to 0 as $n \to \infty$. To see this, note first that $|\mathcal{F}(n, n-3)| \geq \binom{n}{2} \cdot |\mathcal{F}_c(n-2, n-3)| = \binom{n}{2} \cdot (n-2)^{n-4} = \Omega(n^{n-2})$. If we add any three edges to a forest the graph obtained must be planar, and each planar graph in $\mathcal{P}(n,n)$ can be obtained at most n^3 times in this way. Hence $|\mathcal{P}(n,n)| = \Omega(n^{n+1})$. But each graph in $\mathcal{P}_c(n,n)$ can be obtained by adding an edge to a tree, so $|\mathcal{P}_c(n,n)| \leq n^{n-2} \cdot \binom{n}{2} = O(n^n)$. Thus indeed $|\mathcal{P}_c(n,n)|/|\mathcal{P}(n,n)| \to 0$ as $n \to \infty$, as required. Summarising the above, we get Theorem 3.4.

3.6. Unlabelled Graphs and the Number of Automorphisms

Next we turn to the number of automorphisms, unlabelled planar graphs and to the relation between the growth constants of labelled and unlabelled planar graphs. First, we introduce some more notation. We let $\mathrm{aut}(G)$ denote the number of automorphisms of a graph G. Let $\mathcal{UP}(n)$ denote the set of *unlabelled* planar graphs on n nodes, let $\mathcal{UP}(n,m)$ denote the set of unlabelled planar graphs on n nodes with exactly m edges, and let $\mathcal{UP}_c(n,m)$ denote the set of all *connected* graphs in $\mathcal{UP}(n,m)$.

It is shown in [**DVW96**] that

$$|\mathcal{UP}(n)|^{1/n} \to \gamma_u \tag{3.17}$$

as $n \to \infty$, where γ_u is the *unlabelled planar graph growth constant.* (We do not divide by $n!$ for unlabelled structures.) Here we show that there is a growth constant $\gamma_u^c(q)$ such that $|\mathcal{UP}_c(n, \lfloor qn \rfloor)|^{1/n} \to \gamma_u^c(q)$ as $n \to \infty$. It remains open to show the existence of such a constant for $\mathcal{UP}(n, \lfloor qn \rfloor)$.

THEOREM 3.26. *For each $1 < q \leq 3$ there is a finite constant $\gamma_u^c(q) \geq 0$ such that*

$$|\mathcal{UP}_c(n, \lfloor qn \rfloor)|^{\frac{1}{n}} \to \gamma_u^c(q)$$

as $n \to \infty$.

The proof follows closely the approach of [**DVW96**] and uses some ideas of the proof of Theorem 3.6.

PROOF. For $q = 3$ we interpret $\mathcal{UP}_c(n, \lfloor qn \rfloor)$ as $\mathcal{U}P(n, 3n-6)$, i.e., the class of unlabelled triangulations on n nodes, and the result follows from [**Tut62**] (and the constant is 256/27). So, assume $q < 3$. To prove the theorem we again want to apply Fekete's Lemma. Recall that it states that if $g : \mathbb{N} \to \mathbb{R}^+$ is a function such that

$$g(i+j) \geq g(i) \cdot g(j) \tag{3.18}$$

for all $i, j \in \mathbb{N}$, and $c = \sup_n g(n)^{\frac{1}{n}} < \infty$, then $g(n)^{\frac{1}{n}} \to c$ as $n \to \infty$. As mentioned before, the proof in [**LW92**] of Fekete's lemma is easily modified to apply also for functions g with $g(i) = 0$ for finitely many i.

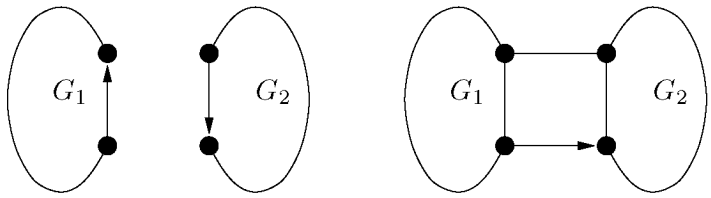

FIGURE 4. $G_1 \star G_2$. The directed edge connects the first with the second root.

Instead of arguing about $\mathcal{UP}_c(n, \lfloor qn \rfloor)$, we consider a new graph class of connected *birooted* unlabelled planar graphs $\mathcal{B}(n, \lfloor qn \rfloor)$ on n nodes with $\lfloor qn \rfloor$ edges, together with a binary operation on this class for merging two graphs. The class $\mathcal{B}(n, \lfloor qn \rfloor)$ is defined as follows: take a graph in $\mathcal{UP}_c(n, \lfloor qn \rfloor)$, choose an edge of the graph and call one end-node the first root r_1 and the other end-node the second root r_2. Note that $|\mathcal{B}(n, \lfloor qn \rfloor)| = 2\lfloor qn \rfloor \cdot |\mathcal{UP}_c(n, \lfloor qn \rfloor)|$. Moreover, by (3.17), $|\mathcal{UP}_c(n, \lfloor qn \rfloor)| \leq |\mathcal{UP}(n)| = (\gamma_u + o(1))^n$, and so $\sup_n |\mathcal{UP}_c(n, \lfloor qn \rfloor)|^{\frac{1}{n}} < \infty$.

We define the following binary operation $\star$ on graphs in $\mathcal{B}(n, \lfloor qn \rfloor)$. Given two graphs $G_1 \in \mathcal{B}(n_1, \lfloor qn_1 \rfloor)$ and $G_2 \in \mathcal{B}(n_2, \lfloor qn_2 \rfloor)$ we construct a new graph $G = G_1 \star G_2$ by creating an edge between the first root of G_1 and the second root of G_2 and another edge between the second root of G_1 and the first root of G_2. The first root of G_1 becomes the first root of G and the second root of G_2 becomes the second root of G; see Figure 4.

Observe that G has $\lfloor qn_1 \rfloor + \lfloor qn_2 \rfloor + 2$ edges. As $\lfloor qn_1 \rfloor + \lfloor qn_2 \rfloor \leq \lfloor q(n_1 + n_2) \rfloor \leq \lfloor qn_1 \rfloor + \lfloor qn_2 \rfloor + 1$, we have to delete 1 or 2 edges to obtain a graph in $\mathcal{B}(n_1 + n_2, \lfloor q(n_1 + n_2) \rfloor)$. We do so by deleting edges in G_1. If $n_1 \geq n_0 = \lceil 1/(q-1) \rceil$ then we can always do this without disconnecting G or deleting the root edge.

Observe that if $G = G_1 \star G_2 = G_1' \star G_2'$, then $G_1 = G_1'$ and $G_2 = G_2'$, as we find G_1 and G_2 by deleting the edge between r_1 and r_2 – the roots of G – and finding the middle edge of the unique path of length 3 connecting r_1 and r_2. The first root of G_2 and the second root of G_1 are then the endpoints of this edge. Thus

$$|\mathcal{B}(n_1 + n_2, \lfloor q(n_1 + n_2) \rfloor)| \geq \frac{|\mathcal{B}(n_1, \lfloor qn_1 \rfloor)| \, |\mathcal{B}(n_2, \lfloor qn_2 \rfloor)|}{n_1^4/8},$$

where we divide by $n_1^4/8 \geq \binom{\binom{n_1}{2}}{2}$ to avoid over-counting due to the deletion of 1 or 2 edges in G_1.

We set

$$\tilde{g}(n, q) = \begin{cases} \frac{|\mathcal{B}(n, \lfloor qn \rfloor)|}{2n^4} & \text{for } n \geq n_0 \\ 0 & \text{otherwise.} \end{cases}$$

Now we use the binary operator $\star$ as defined above to verify that $\tilde{g}(n, q)$ satisfies Inequality (3.18). Let $n = n_1 + n_2$. W.l.o.g. $n_1, n_2 \geq n_0$ as otherwise Inequality (3.18) clearly is fulfilled.

Also, w.l.o.g $n_1 \leq n_2$. Then

$$\begin{aligned}\tilde{g}(n,q) &= \frac{|\mathcal{B}(n,\lfloor qn\rfloor)|}{2n^4} \\ &\geq \frac{|\mathcal{B}(n_1,\lfloor qn_1\rfloor)|}{n_1^4/8}\frac{|\mathcal{B}(n_2,\lfloor qn_2\rfloor)|}{2n_2^4}\frac{2n_2^4}{2n^4} \\ &\geq \tilde{g}(n_1,q)\,\tilde{g}(n_2,q),\end{aligned}$$

as $n_2 \geq \frac{n}{2}$.

It follows from Fekete's lemma that $\tilde{g}(n,q)$ tends to a limit as n tends to infinity. Hence $|\mathcal{B}(n,\lfloor qn\rfloor)|^{\frac{1}{n}}$, and so also $|\mathcal{UP}_c(n,\lfloor qn\rfloor)|^{\frac{1}{n}}$, tend to the same limit, which we denote by $\gamma_u^c(q)$; and the proof is complete. □

The next theorem shows that $R_{n,q}$ has with high probability an exponential number of automorphisms.

THEOREM 3.27. *Let $1 < q < 3$. There are constants $\alpha, \beta, \gamma > 0$ such that*

$$\Pr[2^{\alpha n} \leq \mathrm{aut}(R_{n,q}) \leq 2^{\beta n}] = 1 - o(2^{-\gamma n}).$$

PROOF. We mimic the proof of the corresponding result for R_n (Theorem 4.6 in [**MSW05**]).

Let H be the graph on the node set $\{1,2,3\}$ with the two edges $\{1,2\}$ and $\{2,3\}$. Then the number of automorphisms $\mathrm{aut}(G)$ is at least $2^{f_H^*(G)}$ where as before $f_H^*(G)$ is the maximum number of pairwise node disjoint appearances of H in G. Thus by Corollary 3.3, there are constants $\alpha, \gamma > 0$ such that

$$\Pr[\mathrm{aut}(R_{n,q}) < 2^{\alpha n}] = o(2^{-\gamma n}).$$

Now consider the upper bound on $\mathrm{aut}(R_{n,q})$. Let $\beta > 0$ satisfy $2^{\beta-\gamma} > \gamma_u/\gamma(q)$. The isomorphism class of a graph G in $\mathcal{P}(n,\lfloor qn\rfloor)$, i.e., the set of graphs in $\mathcal{P}(n,\lfloor qn\rfloor)$ isomorphic to G, has size $n!/\mathrm{aut}(G)$. Thus if $\mathrm{aut}(G) \geq 2^{\beta n}$ then the isomorphism class of G in $\mathcal{P}(n,\lfloor qn\rfloor)$ has size at most $n!/2^{\beta n}$. Hence

$$|\mathcal{UP}(n,\lfloor qn\rfloor)| \geq |\mathcal{P}(n,\lfloor qn\rfloor)|\Pr[\mathrm{aut}(R_{n,q}) \geq 2^{\beta n}]\cdot\frac{2^{\beta n}}{n!}$$

and so

$$\begin{aligned}\Pr[\mathrm{aut}(R_{n,q}) \geq 2^{\beta n}] &\leq \frac{|\mathcal{UP}(n,\lfloor qn\rfloor)|}{|\mathcal{P}(n,\lfloor qn\rfloor)|/n!}\cdot 2^{-\beta n} \\ &\leq \frac{|\mathcal{UP}(n)|}{|\mathcal{P}(n,\lfloor qn\rfloor)|/n!}\cdot 2^{-\beta n} \\ &\overset{(3.17)}{=} \left(\frac{\gamma_u}{\gamma(q)}\cdot 2^{-\beta} + o(1)\right)^n \\ &= o(2^{-\gamma n}),\end{aligned}$$

by choice of β. □

COROLLARY 3.28. *Let $1 < q < 3$. The labelled planar graph growth constant $\gamma(q)$ and the unlabelled connected planar graph growth constant $\gamma_u^c(q)$ satisfy $\gamma(q) < \gamma_u^c(q)$.*

PROOF. We follow the proof of Corollary 4.7 in [**MSW05**].

Again we use the observation that the isomorphism class of a graph G in $\mathcal{P}(n, \lfloor qn \rfloor)$ has size $n!/\mathrm{aut}(G)$. Thus for constants $\alpha > 0$ and $\gamma > 0$ as in Theorem 3.27 the number of graphs which are in isomorphism classes of size $> 2^{-\alpha n} n!$ is at most $2^{-\gamma n}|\mathcal{P}(n, \lfloor qn \rfloor)|$, which by (3.2) and (3.3) is at most $\frac{1}{2}|\mathcal{P}_c(n, \lfloor qn \rfloor)|$ for n sufficiently large. Hence, at least half of the graphs in $|\mathcal{P}_c(n, \lfloor qn \rfloor)|$ are in isomorphism classes of size at most $2^{-\alpha n} n!$ implying that

$$|\mathcal{UP}_c(n, \lfloor qn \rfloor)| \geq \frac{1}{2}|\mathcal{P}_c(n, \lfloor qn \rfloor)|/(2^{-\alpha n} n!),$$

that is

$$\frac{|\mathcal{P}_c(n, \lfloor qn \rfloor)|}{n!} \leq 2^{1-\alpha n}|\mathcal{UP}_c(n, \lfloor qn \rfloor)|.$$

It follows from (3.5) and (3.3) that $\gamma(q) \leq 2^{-\alpha}\gamma_u^c(q)$. □

It is known ([**BGW02**], proof of Theorem 3) that the probability that the number of automorphisms of a graph with n nodes exceeds a given value is non-decreasing as we move from labelled to unlabelled graphs. Hence we obtain the following corollary.

COROLLARY 3.29. *Let $1 < q < 3$ and let $U_{n,q}$ denote a graph sampled uniformly at random from the set of unlabelled simple planar graphs on n nodes with $\lfloor qn \rfloor$ edges. There is a constant $\delta > 0$ such that*

$$\Pr[\mathrm{aut}(U_{n,q}) \leq 2^{\delta n}] = o(2^{-\delta n}).$$

3.7. Conclusions

In [**DVW96**] the first results on the random planar graph were achieved. Since then the random planar graph has been better understood, but there are still many open questions. In particular, we know very little about more global structures in random planar graphs. For example, while we do know that the 3-core of a random planar graph is w.h.p. of linear size, we do not know bounds on the size of its components. If one restricts the number of edges, that is, if one considers $R_{n,q}$ instead of R_n then even more questions remain open.

CHAPTER 4

Cactus and Block Graphs

In this chapter we investigate cactus and block graphs. In contrast to Chapter 3, here we will use analytic combinatorics (see also Section 2.2). The main focus of this chapter, however, is on a newly developed method based on Boltzmann samplers (see also Sections 2.3 and 2.4) which we will use to show that a random cactus and a random block graph respectively has specific properties w.h.p.

4.1. Previous and Related Work

In the literature there are two common approaches, which can be used to prove that a random member of a constrained graph class has a specific property. Central to both is the *enumeration* of the objects in the class under consideration. The first approach is purely combinatorial. McDiarmid, Steger, and Welsh were the first to apply this method to obtain results about random planar graphs [**MSW05**], and later generalised their results to a wider class of graphs [**MSW06**]. Although the approach is quite powerful, it has some disadvantages: one might have to adapt it already for slightly different graph classes, see e.g., Chapter 3, and sometimes it only yields rough lower and/or upper bounds (see Section 4.7). Furthermore, it seems to be difficult to apply this method to estimate certain parameters, such as the expected number of edges.

The second approach is based on analytic combinatorics and uses generating functions [**FS06**]. It has recently led to immense progress in the enumeration and the understanding of properties of constrained graph classes, as for instance planar, series parallel, and outerplanar graphs [**BGKN05, GN05b**] (see also Chapter 5). However, it seems to be inherently difficult to investigate properties of such constrained graph classes, which cannot be directly addressed with (a finite number of) parameters of the generating functions, such as the maximum node degree, the maximum size of a biconnected block, or the degree sequence.

Here, we address this problem and propose a method, that is different from the approaches described above. Our central idea is to analyse the execution of an algorithm, which *samples* u.a.r. – uniformly at random – graphs from the graph class in question. If we know what properties the "typical" output of such an algorithm has, then a random member of the considered graph class will also have those properties. On the other hand, from a technical point of view, the analysis may become much easier if we reconceive, as just described.

More precisely, our approach is based on the analysis of the behaviour of *Boltzmann sampler* algorithms, and is an attractive alternative to the combinatorial method and an extension to a purely generating function based approach. The framework of Boltzmann samplers was

developed by Duchon, Flajolet, Louchard, and Schaeffer [**DFLS04**], and enabled the sophisticated development of very efficient, but surprisingly simple, algorithms – in practice and in theory – for sampling objects u.a.r. from such graph classes. As an additional advantage, these samplers can be derived systematically.

Let us briefly compare the combinatorial, the generating function, and our new approach on a simple example, namely Cayley trees. With the combinatorial approach we easily obtain the exact number $t_n := n^{n-2}$ of labelled trees on n nodes. The method relies on a bijection between trees on n nodes and the encoding of a tree as a string – the well-known Prüfer code. Moon [**Moo68**] used this to obtain precise asymptotic bounds for the maximum node degree in a random tree. Although it works quite well for trees, the problem-tailored approach yields difficulties (and often weaker results), as soon as the complexity of the graph class in consideration increases. If we use generating functions, we obtain by a singularity analysis an asymptotic estimate of the form $t_n \sim n! \left(\frac{e^n}{\sqrt{2\pi n^5}} + O(n^{-\frac{7}{2}})\right)$. This approach has the advantage that it works for more complex graph classes as well, and that it might be used to obtain results about the expected number of edges, for instance (of course, for graph classes different from trees). On the other hand, as soon as we are interested in parameters like the maximum degree, or the maximum size of a biconnected component, especially in more complex graph classes, it becomes much harder to introduce the right parameters in the generating functions and to analyse them. Contrary, our new method via Boltzmann samplers delivers the precise asymptotic behaviour of the maximum node degree, as well as the complete degree sequence for Cayley trees; moreover, it has the benefit that it generalises easily and mechanically to moderately complex graph classes, as we will demonstrate in this chapter. Our method yields precise estimates – that is the *precise* asymptotic behaviour – for the maximum node degree and maximum size of a biconnected block of a random graph within the class under consideration. We are confident that the method is applicable to other parameters and graph classes as well.

Here we will study and illustrate how our method works on two graph classes of moderate complexity, namely cactus and block graphs, which were first defined by Uhlenbeck [**Uhl**] and Riddell [**Rid51**], following a paper by Husimi [**Hus50**] on the cluster integrals in the theory of condensations. A *cactus graph* is a labelled connected graph, in which each edge is contained in at most one cycle. A *block graph* is a labelled connected graph, whose maximal biconnected blocks are cliques.

The problem of finding the number of cactus graphs was proposed by Uhlenbeck, who showed how this result can be used in the theory of condensation. Harary and Uhlenbeck [**HU53**] were the first to derive an exact counting formula – using generating functions – in terms of the number of nodes and the number of cycles n_i of length $i = 3, \ldots$. Later, first properties as for example the dissimilarity characteristic of these graphs were studied [**HN53**].

Nowadays, cactus and block graphs are well-known and well-studied combinatorial objects, and have applications in physics, computer science, and bioinformatics: cactus graphs are, for instance, used in combinatorial optimisation for encoding all minimum cuts of a graph [**Fle99**] in polynomial time which has further relevance for example in network failure detection [**KSS04**]. In physics they appear (in restricted form) often as an underlying model for

instance in the area of error-correcting codes [**VSK00**], p-spin interaction models [**RK92**], and polymer interactions [**SSCM04**].

When modelling real world scenarios on graphs, (weighted) block graphs come up quite naturally in many applications, especially if one wants to represent similarities or dissimilarities between groups of objects. One of these applications appears for instance in the area of bioinformatics, where they are used in the context of phylogenetic trees [**BHM02**].

As both graph classes appear quite often in practical applications, much effort was put into developing efficient algorithms specifically designed to work well on these graphs [**BMBS05, UU04, Won99, YC94, ZZ04**].

4.2. Results

We first derive for both graph classes in interest efficient samplers based on the Boltzmann sampler framework, which was recently introduced by Duchon et al. in [**DFLS04**]. They will be crucial for our further analysis.

THEOREM 4.1. *Let $n \in \mathbb{N}$ and $0 < \varepsilon < 1$. There exists a randomised algorithm with expected* quadratic *running time, which produces block graphs of size n uniformly at random. Furthermore, there exists a randomised algorithm with expected* linear *running time, which produces random block graphs with size in the interval $I_\varepsilon = \big(n(1-\varepsilon), n(1+\varepsilon)\big)$, such that the distribution of block graphs is uniform on each size $k \in I_\varepsilon$. The variance of the running time is in both cases quadratic in the expected running time. The same holds for cactus graphs.*

We also provide carefully optimised implementations of the above samplers which allow us to sample graphs of up to 10^7 nodes within a few seconds, see [**PW06**] and Section 4.6.2. We call a simple labelled graph whose connected components are cactus graphs a *desert graph*; and a simple labelled graph whose connected components are block graphs a *city graph*. Our main results on the structure of a random city and a random desert graph, i.e. a graph drawn uniformly at random from the set of all graphs on n nodes within the corresponding class, are summarised in the following theorems. Similar results also hold for random block and cactus graphs. In the subsequent, we mean by "$\doteq$" rounded to the last digit shown. First, we state the results which can be proved with our new method.

THEOREM 4.2. *Let D_n denote a random desert graph on n nodes. There exists a constant $K_0^\bullet \doteq 0.4563$, such that for every $\varepsilon > 0$ the following statements hold for almost all n:*

(i) The maximum node degree of D_n takes values in $\big((2-\varepsilon)\frac{\log n}{\log\log n}, (2+\varepsilon)\frac{\log n}{\log\log n}\big)$ w.h.p.
(ii) The largest cycle in D_n has size in $\big((1-\varepsilon)\log_{\frac{1}{K_0^\bullet}} n, (1+\varepsilon)\log_{\frac{1}{K_0^\bullet}} n\big)$ w.h.p.

THEOREM 4.3. *Let Y_n denote a random city graph on n nodes. For every $\varepsilon > 0$, the following statements hold for almost all n:*

(i) The maximum node degree of Y_n takes values in $\big((1-\varepsilon)\frac{\log n}{\log^{(3)} n}, (1+\varepsilon)\frac{\log n}{\log^{(3)} n}\big)$ w.h.p.
(ii) The largest clique in Y_n has size in $\big((1-\varepsilon)\frac{\log n}{\log\log n}, (1+\varepsilon)\frac{\log n}{\log\log n}\big)$ w.h.p.

Observe that in both cases, we obtain the *precise asymptotic behaviour* of the parameters in question. It seems that as soon as we consider more complex objects than trees, some of the involved constants become quite unpredictable and it might be very difficult – if not impossible – to obtain them by a purely combinatorial approach. Furthermore, note that both graph classes are examples, where the general lower bound on the maximum node degree of $(1+o(1))\frac{\log n}{\log\log n}$ for addable graph classes, given in [**MSW06**, Corollary 5.3], is not tight.

Roughly speaking, we proceed as follows to prove the above theorems: instead of investigating directly the properties of the graph class, we consider the execution of its corresponding Boltzmann sampler. We examine how the *shape* of a sampled object evolves over a run of the sampler and how this affects the related property. From this knowledge, we can eventually deduce the probability that a sampled output object has the property in interest.

Note that the chromatic number of a block graph equals its maximum clique size. Moreover, by the proof of Theorem 4.3 we also obtain that for all sufficiently large n the probability that a random block graph contains a clique of size at most $\alpha\frac{\log n}{\log\log n}$ is at least $1-n^{1-\alpha}$. This directly implies the following corollary.

Corollary 4.4. *Let $\chi(G)$ denote the chromatic number of a graph G. Let $\varepsilon>0$. Let B_n be a random block graph on n nodes. Then, as $n\to\infty$,*

$$\Pr\left[\chi(\mathrm{B}_n)\le(1+\varepsilon)\frac{\log n}{\log\log n}\right]\to 1.$$

As already mentioned, enumerating the objects in the graph class in question is important for understanding its structural properties. Here we exploit generating function techniques, which were only recently applied to obtain similar results for planar, series parallel, and outerplanar graphs [**BGKN05, GN05b**], and are well described in the forthcoming book "Analytic Combinatorics" by Flajolet and Sedgewick [**FS06**]. Altogether, we obtain a fairly complete picture of desert and city graphs. In the sequel, $\mathrm{Po}(\lambda)$ is the Poisson distribution with mean λ.

Theorem 4.5. *Let d_n denote the number of desert graphs on n nodes and let k_n denote the number of cactus graphs on n nodes. Then*

$$d_n\sim\alpha_d\,\rho_k^{-n}\,n^{-5/2}\,n!\qquad\text{and}\qquad k_n\sim\alpha_k\,\rho_k^{-n}\,n^{-5/2}\,n!,\tag{4.1}$$

where $\alpha_d\doteq 0.1605$, $\alpha_k\doteq 0.1201$, and $\rho_k^{-1}\doteq 4.1887$ are explicitly given constants.

Theorem 4.6. *Let y_n denote the number of city graphs on n nodes and let b_n denote the number of block graphs on n nodes. Then*

$$y_n\sim\alpha_y\,\rho_b^{-n}\,n^{-5/2}\,n!\qquad\text{and}\qquad b_n\sim\alpha_b\,\rho_b^{-n}\,n^{-5/2}\,n!,\tag{4.2}$$

where $\alpha_y\doteq 0.2515$, $\alpha_b\doteq 0.1807$ and $\rho_b^{-1}\doteq 3.7824$ are analytically computable constants.

Our next theorem is about the chromatic number of a random cactus graph. Observe that a cactus graph is either 2- or 3-colourable (if it is a tree or contains only even cycles and if it contains an odd cycle respectively). The next theorem states that the chromatic number of a random cactus graph is 2 with exponentially small probability.

THEOREM 4.7. *Let $\chi(G)$ denote the chromatic number of a graph G. Let K_n be a random cactus graph. Then, as $n \to \infty$,*

$$\Pr[\chi(\mathrm{K}_n) = 2] \sim \frac{\tilde{\alpha}_k \tilde{\rho}_k{}^{-n}}{\alpha_k \rho_k^{-n}} \to 0,$$

where $\alpha_k \doteq 0.1201$, $\tilde{\alpha}_k \doteq 0.1360$, $\rho_k^{-1} \doteq 4.1887$, and $\tilde{\rho}_k{}^{-1} \doteq 3.5505$ are explicitly given constants. Furthermore, $\Pr[\chi(\mathrm{K}_n) = 3] \to 1$ exponentially fast.

It is easily seen that the above theorem also holds for desert graphs. Next we turn to the expected number of components and then to the connectedness of a random desert and random cactus graph.

THEOREM 4.8. *Let D_n denote a random desert graph. Then the number of components of D_n is asymptotically distributed as $1+\mathrm{Po}(K_0)$, where $K_0 \doteq 0.2393$ is an explicitly given constant.*

This theorem directly implies the following corollary.

COROLLARY 4.9. *The expected number of components in a random desert graph is asymptotically equal to $1 + K_0 \doteq 1.2393$.*

THEOREM 4.10. *Let D_n be a random desert graph and let K_n be a random cactus graph. Then, as $n \to \infty$*

$$\begin{aligned} \Pr[\mathrm{D}_n \textit{ is connected}] &\to e^{-K_0} \doteq 0.7488, \\ \Pr[\mathrm{D}_n \textit{ is biconnected}] &\sim \alpha_d^{-1} \alpha_c \, n^{\frac{3}{2}} \, \rho_k^n \to 0, \\ \Pr[\mathrm{K}_n \textit{ is biconnected}] &\sim \alpha_k^{-1} \alpha_c \, n^{\frac{3}{2}} \, \rho_k^n \to 0, \end{aligned}$$

where K_0, α_c, α_k, and ρ_k are explicitly given constants.

The next parameter we study is the number of edges. Note that for each cycle in a cactus graph, we get an excess of one edge compared to a tree. The next theorem shows that the number of edges is asymptotically normally distributed with known parameters.

THEOREM 4.11. *The number of edges in a random desert graph is asymptotically normally distributed with mean μ_n and variance σ_n^2, which satisfy*

$$\mu_n \sim \kappa n \quad \textit{and} \quad \sigma_n^2 \sim \lambda n,$$

where $\kappa \doteq 1.1915$ and $\lambda \doteq 0.0627$ are explicitly given constants. The same is true for the number of edges in a random cactus graph.

Notice that the last theorem implies that the number of edges is sharply concentrated around its expected value. By the Polyhedral formula of Euler we obtain the following corollary.

COROLLARY 4.12. *Let C_n denote the number of cycles in a random desert graph. Then the mean value $\mu_n = \mathbb{E}[C_n]$ and variance $\sigma_n^2 = \mathrm{Var}[C_n]$ satisfy*

$$\mu_n \sim \kappa n \quad \textit{and} \quad \sigma_n^2 \sim \lambda n,$$

where $\kappa \doteq 0.1915$ and $\lambda \doteq 0.0627$ are explicitly given constants.

Analogously, we obtain the following results for random city and random block graphs.

THEOREM 4.13. *Let* Y_n *be a random city graph. Then the number of components of* Y_n *is asymptotically distributed as* $1 + \mathrm{Po}(B_0)$*, where* $B_0 \doteq 0.3304$ *is an analytically computable constant.*

COROLLARY 4.14. *The expected number of components in a random city graph is asymptotically equal to* $1 + B_0 \doteq 1.3304$.

THEOREM 4.15. *Let* Y_n *be a random city graph and let* B_n *be a random block graph. Then as* $n \to \infty$

$$\begin{aligned} \Pr[\mathrm{Y}_n \textit{ is connected}] &\to e^{-B_0} \doteq 0.7187, \\ \Pr[\mathrm{Y}_n \textit{ is biconnected}] &\to 0, \\ \Pr[\mathrm{B}_n \textit{ is biconnected}] &\to 0, \end{aligned}$$

where B_0 *is an analytically computable constant.*

As mentioned before, our main aim of examining properties of graph classes is to obtain the average-case performance of algorithms. In [**UU04**] Uehara and Uno provide algorithms for computing the longest path in a cactus graph and in a block graph. The running time of the former algorithm depends quadratically on the size of the largest cycle, whereas the running time of the later is $O(n + m)$ for a block graph on n nodes with m edges – see Theorem 2 and the proof of Theorem 4 in [**UU04**]. Now, note that by the proofs of Theorems 4.3 and 4.2, we also obtain that for all sufficiently large n the probability that a random block graph contains a clique of size at most $\alpha \frac{\log n}{\log\log n}$ is at least $1 - n^{1-\alpha}$ (as noted earlier) and similarly, that the probability that a random cactus graph contains a cycle of length at most $\alpha \log n$ is at least $1 - n^{1-\alpha}$, whenever $\alpha \geq 3$. From the upper bound on the maximum clique size we can deduce that a random block graph has at most $n \frac{\log n}{\log\log n}$ edges w.h.p. By these two observations we obtain the following corollary.

COROLLARY 4.16. *A longest path in a random cactus graph of size* n *can be computed in* $O(n(\log n)^2)$ *expected time and in a random block graph in* $O\left(\frac{n \log n}{\log\log n}\right)$ *expected time.*

4.3. Graph Classes and Notation

Before we proceed, let us introduce some notation, which we are going to use in the remainder of the chapter. Let $\mathcal{K}_n$ denote the set of all cactus graphs on n nodes and let $K(x) = \sum_n \frac{k_n}{n!} x^n$ denote the exponential generating function of cactus graphs, where $k_n := |\mathcal{K}_n|$. Moreover, let $K(x, y) = \sum_{n,m} \frac{k_{n,m}}{n!} x^n y^m$ denote the bivariate generating function of cactus graphs, where x marks nodes, y marks edges and $k_{n,m}$ is the number of cactus graphs on n nodes with exactly m edges. A graph whose connected components are cactus graphs is called a *desert* graph. Let $\mathcal{D}_n$ denote the set of all desert graphs on n nodes, and let $D(x) = \sum_n \frac{d_n}{n!} x^n$ denote the exponential generating function for those graphs, where $d_n := |\mathcal{D}_n|$. Similarly, let $D(x, y) = \sum_{n,m} \frac{d_{n,m}}{n!} x^n y^m$ denote the bivariate exponential generating function of desert graphs according to nodes and edges. Analogously, we denote by $C(x)$ and $C(x, y)$ the exponential generating function of a single edge and cycles of length at least 3. We proceed in

a similar fashion for block graphs. We denote by $B(x)$ the exponential generating function, which enumerates block graphs, and by $Y(x)$ the exponential generating function for *city* graphs, i.e., graphs whose connected components are block graphs. Finally, let $Q(x)$ denote the generating function for cliques of size at least 2.

Finally, we will often need a "pointing operator" to distinguish an atom of our combinatorial objects from all other atoms; for example we want to distinguish a node in our graph as a root node. On generating function level, this pointing operation means taking the derivative with respect to the according variable, and multiplying by it. To shorten notation, we define $G^\bullet(x) := x\frac{\partial}{\partial x}G(x)$, where G is an exponential generating function.

4.4. Decomposition and Generating Functions

A standard decomposition of a graph into connected and 2-connected blocks delivers us with the relations for our generating functions (see e.g. [**FS06**](p.95) and [**HP73**](p.10)).

LEMMA 4.17. *Let $D(x,y)$, $K(x,y)$, and $C(x,y)$ denote the bivariate exponential generating functions for desert graphs, cactus graphs and cycles of length at least 3 and a single edge. Then it holds*

$$D(x,y) = \exp\left(K(x,y)\right) \quad \text{and} \quad K^\bullet(x,y) = x\exp\left(\frac{\partial}{\partial x}C\left(K^\bullet(x,y),y\right)\right), \tag{4.3}$$

$$C(x,y) = \frac{1}{2}\log\left(\frac{1}{1-xy}\right) - \frac{1}{4}x^2y^2 + \frac{1}{2}x^2y - \frac{1}{2}xy. \tag{4.4}$$

Note that in (4.3) we write $\frac{\partial}{\partial x}C(K^\bullet(x,y),y)$ for $\frac{\partial}{\partial u}C(u,y)|_{u=K^\bullet(x,y)}$, i.e. we first take the derivative of $C(x,y)$ with respect to x and then evaluate this function at $(K^\bullet(x,y),y)$, but we will use the above simplified notation subsequently. Observe that setting $y = 1$ in the above lemma yields a relation for the corresponding univariate generating functions.

Similarly, we can derive the relations for city and block graphs.

LEMMA 4.18. *Let $Y(x)$, $B(x)$, and $Q(x)$ denote the exponential generating functions for city graphs, block graphs and cliques of size at least 2. Then it holds*

$$Y(x) = \exp\left(B(x)\right) \quad \text{and} \quad B^\bullet(x) = x\exp\left(\frac{\partial}{\partial x}Q\left(B^\bullet(x)\right)\right), \tag{4.5}$$

$$Q(x) = \exp(x) - 1 - x. \tag{4.6}$$

From these equations we can derive singular expansions of the involved generating functions (see Section 4.5). Furthermore, the above decomposition allows us to design the Boltzmann sampler for cactus and block graphs (see Section 4.6).

4.5. Singular Expansions and Asymptotic Estimates

In this section, we will derive singular expansions for the generating functions of desert, cactus, city, and block graphs. From these, we can then obtain asymptotic estimates for the number of graphs on n nodes in the corresponding class.

LEMMA 4.19. *Let $D(x,y)$ and $K(x,y)$ denote the bivariate exponential generating functions of desert and cactus graphs. Then, for all $y \in (0,1]$, $D(x,y)$ and $K(x,y)$ have the same singularity $\rho(y)$. Moreover, for $X = \left(1 - \frac{x}{\rho(y)}\right)$, they have singular expansions of the form*

$$K(x,y) = K_0(y) + K_1(y)X + K_2(y)X^{\frac{3}{2}} + O(X^2), \tag{4.7}$$

where $K_i(y)$, $i = 0,1,2$ are explicitly given functions of y and

$$D(x,y) = D_0(y) + D_1(y)X + D_2(y)X^{\frac{3}{2}} + O(X^2), \tag{4.8}$$

where $D_i(y)$, $i = 0,1,2$ are explicitly given functions of y.

PROOF. Let $y \in (0,1]$ be fixed. In order to obtain a singular expansion for $K(x,y)$ and $D(x,y)$, we need to locate the dominant singularity $\rho(y)$ of $D(x,y)$. Since $D(x,y) = \exp(K(x,y))$ and exp is an entire function, the functions $D(x,y)$ and $K(x,y)$ have the same singularities. Hence we will firstly concentrate on $K(x,y)$.

Notice that the singularities of $\frac{\partial}{\partial x}C(x,y)$ and $K^\bullet(x,y)$ are the same as those of $C(x,y)$ and $K(x,y)$ respectively. From (4.3) it follows that the functional inverse $\psi(u,y)$ of $K^\bullet(x,y)$ is

$$\psi(u,y) = u\exp\left(-\frac{\partial}{\partial u}C(u,y)\right), \tag{4.9}$$

as $\psi(K^\bullet(x,y),y) = x$.

The dominant singularity of $\psi(u,y)$ is the same as that of $C(x,y)$ which clearly is singular at $R(y) = \frac{1}{y}$. From Theorem 2.6 (Theorem VI.6 of [**FS06**]) it follows that we have to decide if $\frac{\partial}{\partial u}\psi(u,y)$ has a real zero $\tau(y) \in (0, R(y))$. If this is the case, then $\psi(u,y)$ ceases to be invertible at $\tau(y)$ and $K^\bullet(x,y)$ has a singularity at $\rho(y) = \psi(\tau(y),y)$. Solving $\frac{\partial}{\partial u}\psi(u,y) = 0$ yields a real zero at

$$\tau(y) = \frac{2}{3}\frac{-3y^2+y^3}{-2y^3+y^4} - \frac{6y^5-4y^6}{3(-2y^3+y^4)\beta(y)} + \frac{\beta(y)}{3(-2y^3+y^4)},$$

where

$$\beta(y) = \left(54y^7 - 45y^8 + 8y^9 + 3\sqrt{3}\sqrt{108y^{14} - 172y^{15} + 91y^{16} - 16y^{17}}\right)^{\frac{1}{3}}.$$

One can check that $0 < \tau(y) < R(y)$ for all $0 < y \leq 1$, so the singularity of $K^\bullet(x,y)$ is at $\rho = \rho(y) = \psi(\tau(y),y)$. In particular, for $y = 1$, we have a real zero at

$$\tau(1) = -\frac{(17+3\sqrt{33})^{\frac{1}{3}}}{3} + \frac{2}{3(17+3\sqrt{33})^{\frac{1}{3}}} + \frac{4}{3} \doteq 0.4563$$

and $K^\bullet(x,1) = K^\bullet(x)$ has a singularity at

$$\rho_k = \psi(\tau(1),1) = \left(-\frac{1}{3}\alpha + \frac{2}{3}\alpha^{-1} + \frac{4}{3}\right)e^{-\left(-\frac{2}{3}+\frac{2}{3}\alpha-\frac{4}{3}\alpha^{-1}\right)^{-1}-\frac{1}{6}+\frac{1}{6}\alpha-\frac{1}{3}\alpha^{-1}} \doteq 0.2387,$$

where $\alpha = (17+3\sqrt{33})^{\frac{1}{3}}$.

Let for the remainder of this proof $\psi'(u,y) = \frac{\partial}{\partial u}\psi(u,y)$ and define similarly higher derivatives. Now we are able to compute the asymptotic expansions. The Taylor expansion of $\psi(u,y)$ at $u = \tau(y)$ is

$$x = \psi(K^\bullet(x,y),y) = \rho(y) + \frac{\psi''(\tau(y),y)}{2!}(K^\bullet(x,y)-\tau(y))^2 + \frac{\psi'''(\tau(y),y)}{3!}(K^\bullet(x,y)-\tau(y))^3 + \ldots,$$

as $\psi'(u,y)$ vanishes at $\tau(y)$. Hence we have a locally quadratic dependency between $K^\bullet(x,y)$ and its functional inverse $\psi(u,y)$. We can derive a full formal expansion of $K^\bullet(x,y)$ in powers of $(1-x/\rho(y))$ by repeated substitution (see also the discussion before Theorem VI.6 and its proof in [**FS06**]) and finally we obtain

$$K^\bullet(x,y) = K_0^\bullet(y) + K_1^\bullet(y)\left(1-\frac{x}{\rho(y)}\right)^{\frac{1}{2}} + K_2^\bullet(y)\left(1-\frac{x}{\rho(y)}\right) + O\left(\left(1-\frac{x}{\rho(y)}\right)^{\frac{3}{2}}\right), \quad (4.10)$$

where

$$K_0^\bullet(y) = \tau(y),\ K_1^\bullet(y) = -\sqrt{\frac{2\rho(y)}{-\psi''(\tau(y),y)}},\ K_2^\bullet(y) = \frac{\rho(y)}{3}\frac{\psi'''(\tau(y),y)}{\psi''(\tau(y),y)^2}$$

are explicitly given constants. Now, we can derive a singular expansion for $K(x,y)$ by integrating $\frac{\partial}{\partial x}K(x,y) = \frac{K^\bullet(x,y)}{x}$. We get

$$K(x,y) = K_0(y) + K_1(y)X + K_2(y)X^{\frac{3}{2}} + K_3(y)X^2 + O\left(X^{\frac{5}{2}}\right), \quad (4.11)$$

where $X = \left(1-\frac{x}{\rho(y)}\right)$ and

$$K_1(y) = -K_0^\bullet(y),\ K_2(y) = -\frac{2}{3}K_1^\bullet(y),\ K_3(y) = -\frac{1}{2}(K_0^\bullet(y) + K_2^\bullet(y)).$$

$K_0(y)$ cannot be determined in this way, but we can compute it using integration by parts and substitution – the proof follows closely the lines of proof of Theorem 1 in [**GN05a**]. We change variables, $t = K^\bullet(s,y)$, thus $s = \psi(t,y) = t\exp\left(-\frac{\partial}{\partial t}C(t,y)\right)$ and obtain

$$K(x,y) = \int_0^x \frac{K^\bullet(s,y)}{s}ds = \int_0^{K^\bullet(x,y)} \frac{t}{\psi(t,y)}\left(\frac{\partial}{\partial t}\psi(t,y)\right)dt. \quad (4.12)$$

We use integration by parts and (4.12) becomes

$$\begin{aligned}\int_0^{K^\bullet(x,y)} \frac{t}{\psi(t,y)}\left(\frac{\partial}{\partial t}\psi(t,y)\right)dt &= (t\log\psi(t,y))\Big|_{t=0}^{K^\bullet(x,y)} - \int_0^{K^\bullet(x,y)} \log\psi(t,y)dt \\ &= K^\bullet(x,y)\log x - K^\bullet(x,y)\log K^\bullet(x,y) + K^\bullet(x,y) + C(K^\bullet(x,y),y).\end{aligned}$$

Hence we get

$$K_0(y) = K(\rho(y),y) = \tau(y)\log(\rho(y)) - \tau(y)\log\tau(y) + \tau(y) + C(\tau(y),y).$$

Finally, we can derive a singular expansion for $D(x,y) = \exp(K(x,y))$ by simply applying the exponential function to (4.11) and obtain

$$D(x,y) = e^{K_0(y)}\left(1 + K_1(y)X + K_2(y)X^{\frac{3}{2}}\right) + O\left(X^2\right). \quad (4.13)$$

Therefore, the coefficients $D_0(y), D_1(y)$ and $D_2(y)$ from (4.8) are given by

$$D_0(y) = e^{K_0(y)}, \quad D_1(y) = -e^{K_0(y)}K_0^\bullet(y) \quad \text{and} \quad D_2(y) = -\frac{2}{3}e^{K_0(y)}K_1^\bullet(y).$$

□

Lemma 4.20. *Let $Y(x)$ and $B(x)$ denote the exponential generating functions of city and block graphs. Then $Y(x)$ and $B(x)$ have the same singularity ρ_b. Moreover, for $X = \left(1-\frac{x}{\rho_b}\right)$, they have singular expansions of the form*

$$B(x) = B_0 + B_1X + B_2X^{\frac{3}{2}} + O(X^2), \quad (4.14)$$

where B_i, $i = 0, 1, 2$ are analytically computable constants and

$$Y(x) = Y_0 + Y_1 X + Y_2 X^{\frac{3}{2}} + O(X^2), \tag{4.15}$$

where Y_i, $i = 0, 1, 2$ are analytically computable constants.

PROOF. The proof is very similar to the one of Lemma 4.19. Therefore we omit some details to avoid repetition. In order to obtain a singular expansion for $Y(x)$ and $B(x)$, we first need to locate the dominant singularity ρ_b of $Y(x)$. Since $Y(x) = \exp(B(x))$ and exp is an entire function, $Y(x)$ and $B(x)$ have the same singularities. Hence we will concentrate on $B(x)$.

Notice that the singularities of $\frac{\partial}{\partial x}Q(x)$ and $B^\bullet(x)$ are the same as those of $Q(x)$ and $B(x)$ respectively. From (4.5) it follows that the functional inverse $\psi(u)$ of $B^\bullet(x)$ is

$$\psi(u) = u \exp\left(-\frac{\partial}{\partial x}Q(u)\right), \tag{4.16}$$

as $\psi(B^\bullet(x)) = x$. This time $Q(x)$ is an entire function and thus has a singularity at $R = \infty$. Again, we apply Theorem 2.6. One can check that solving $\frac{\partial}{\partial u}\psi(u) = 0$ yields a real zero at $\tau \in (0, R)$, where τ is the solution to the equation

$$1 - u\exp(u) = 0.$$

As $\psi(u)$ ceases to be invertible at τ, $B^\bullet(x)$ has a singularity at $\rho_b = \psi(\tau)$. Now, we mimic the proof of Lemma 4.19 to obtain a singular expansion for $B^\bullet(x)$ at ρ_b from the Taylor expansion of $\psi(u)$ at τ by repeated substitution and we obtain

$$B^\bullet(x) = B_0^\bullet + B_1^\bullet\left(1 - \frac{x}{\rho_b}\right)^{\frac{1}{2}} + B_2^\bullet\left(1 - \frac{x}{\rho_b}\right) + O\left(\left(1 - \frac{x}{\rho_b}\right)^{\frac{3}{2}}\right), \tag{4.17}$$

where

$$B_0^\bullet = \tau, \quad B_1^\bullet = -\sqrt{\frac{2\rho_b}{-\psi''(\tau)}} \quad \text{and} \quad B_2^\bullet = \frac{\rho_b}{3}\frac{\psi'''(\tau)}{\psi''(\tau)^2}$$

are analytically computable constants. Now we can derive the singular expansion (4.14) for $B(x)$ by integrating $\frac{\partial}{\partial x}B(x) = \frac{B^\bullet(x)}{x}$. Furthermore, the constant B_0 can be obtained by integration by parts and substitution analogously to the proof of Lemma 4.19. The coefficients we obtain are given by

$$B_0 = B(\rho_b) = \tau\log(\rho_b) - \tau\log\tau + \tau + Q(\tau), \quad B_1 = -B_0^\bullet \quad \text{and} \quad B_2 = -\frac{2}{3}B_1^\bullet.$$

Finally, we can derive the singular expansion (4.15) for $Y(x) = \exp(B(x))$ by simply applying the exponential function to (4.14) and we obtain the coefficients

$$Y_0 = e^{B_0}, \quad Y_1 = -e^{B_0}B_0^\bullet, \quad \text{and} \quad Y_2 = -\frac{2}{3}e^{B_0}B_1^\bullet.$$

□

Now, we can apply the transfer theorems of [**FS06**] (see also Section 2.2) to the singular expansions of Lemma 4.19 and we obtain asymptotic estimates for the number of desert and cactus graphs on n nodes.

Proof of Theorem 4.5. Equation (4.1) follows by applying Corollary 2.5 (Corollary VI.1 of [**FS06**]) to the singular expansions of $D(x, y)$ and $K(x, y)$, which were derived in Lemma 4.19. We obtain

$$\alpha_d \sim D_2(1) \cdot \frac{3}{4\sqrt{\pi}} = e^{K_0(1)} \frac{1}{2\sqrt{\pi}} \sqrt{\frac{2\rho(1)}{-\psi''(\tau(1), 1)}} \doteq 0.1605$$

and similarly

$$\alpha_k \sim \frac{1}{2\sqrt{\pi}} \sqrt{\frac{2\rho(1)}{-\psi''(\tau(1), 1)}} \doteq 0.1201,$$

where $\psi(u, y)$ is given by (4.9) and $\psi''(u, y) = \frac{\partial^2}{\partial u^2}\psi(u, y)$. □

Theorem 4.6 can be proved in the same way but using the result of Lemma 4.20.

4.6. Sampling

In this section, we will use the knowledge about the generating functions derived so far to obtain efficient sampling procedures for the graph classes in interest. We will use Boltzmann samplers to derive these algorithms; see Section 2.3 for a brief introduction and [**DFLS04**] for a detailed description of the topic.

4.6.1. Boltzmann Samplers. Recall that the efficiency of a Boltzmann sampler for a combinatorial class $\mathcal{C}$ depends highly on the type of the singular expansion of its generating function $C(x)$, see Section 2.3. In order to obtain an expected linear running time sampler, either the exponent $-\alpha$ of the singular expansion of C has to be negative *or*, in the special case that $-\alpha = \frac{1}{2}$, we can use a so-called *singular ceiled rejection sampler*, which discards objects during sampling, as soon as they become too large.

Here, we will derive a ceiled rejection sampler for block graphs – the sampler for cactus graphs can be derived in the same way. The sampler has the general advantage of its kind, that is we don't have to tune it explicitly for a specific output size; moreover, in our case we have to pre-compute only two constants (see below). Furthermore, it turns out that the singular rejection sampler is very well suited for our further analysis.

Observe that the singular expansions for cactus graphs and block graphs do not have the right exponent $-\alpha = \frac{1}{2}$, see Lemma 4.19 and 4.20. But we can overcome this by *pointing* our structures – it then can easily be checked that the resulting generating function relations (see (4.3) and (4.5)) fulfil the preconditions of Theorem 2.10 (Theorem 8 of [**DFLS04**]). In the following, we give an exposition of our sampling algorithm for block graphs. We refer the reader to Section 2.3 for the transfer rules we use to compile our decomposition into an algorithm.

The singular rejection sampler repeatedly samples block graphs according to the Boltzmann distribution (2.5), until an object with size in $((1 - \varepsilon)n, (1 + \varepsilon)n)$ is sampled. It maintains a global variable *nodes*, which counts the number of nodes which were generated through recursive calls.

$\Gamma\nu B(n,\varepsilon)$: repeat
  $nodes \leftarrow 0$
  $\gamma \leftarrow \Gamma B^\bullet(n,\varepsilon)$
until $|\gamma| > (1-\varepsilon)n$
label the nodes of γ uniformly at random
return γ (and discard the root)

The sampler for rooted block graphs works as follows. Roughly speaking, a rooted block graph is a set of biconnected rooted block graphs, merged at their roots, in which every node is again replaced by a rooted block graph. This decomposition is the origin of the relations of the exponential generating functions for block graphs and cliques, see Lemma 4.18. The sampler works similarly: it starts with a single node, chooses according to a carefully chosen probability distribution a random number of cliques adjacent to it, and then replaces each generated node by a randomly drawn rooted block graph.

$\Gamma B^\bullet(n,\varepsilon)$: $\gamma \leftarrow$ a single root node r
$nodes \leftarrow nodes + 1$
$k \leftarrow \mathrm{Po}\left(\frac{\partial}{\partial x} Q(B^\bullet(\rho_b))\right)$
for $(j = 1 \dots k)$
  $\gamma' \leftarrow \Gamma Q^\bullet_{<(1+\varepsilon)n}(B^\bullet(\rho_b))$
  if (γ' equals $\perp$) return $\perp$
  $\gamma \leftarrow$ merge γ and γ' at their root nodes
for (each node $v \neq r$ of γ)
  $nodes \leftarrow nodes - 1$
  $\gamma_v \leftarrow \Gamma B^\bullet(n,\varepsilon)$
  if (γ_v equals $\perp$) return $\perp$
replace all nodes $v \neq r$ of γ with γ_v
return γ

Finally, the sampler $\Gamma Q^\bullet_{<m}$ for cliques chooses the size of the objects it outputs according to its probability in the Boltzmann model. It returns a default empty object ($\perp$), if the size of the generated object would have been too large. Before we can state it formally, let us define the distribution it samples from.

Definition 4.21. The clique distribution with parameter $x > 0$, denoted by $Cl(x)$ is defined by

$$\Pr[Cl(x) = k] := \begin{cases} \frac{x^k}{Q^\bullet(x)(k-1)!} & k \geq 2 \\ 0 & \text{otherwise,} \end{cases}$$

where $Q^\bullet(x) := x(e^x - 1)$ is the exponential generating function for rooted cliques of size ≥ 2.

With this, the sampler can be implemented as follows:

$\Gamma Q^{\bullet}_{<m}(x)$: $k \leftarrow \mathrm{Cl}\,(x)$
$nodes \leftarrow nodes + (k-1)$
if($nodes \geq m$)
generate $(k-1)-(nodes-m)$ nodes $\qquad(\star)$
return $\perp$
else return a rooted clique on k nodes

It can easily be checked that the sampler above generates cliques according to the Boltzmann distribution (2.5), if it does not early interrupt (returning $\perp$). Note that in the line marked by $(\star)$, the sampler returns $\perp$ *after* it has generated as many nodes as it is necessary to have generated precisely m nodes in the object sampled in the current execution of $\Gamma B^{\bullet}$. This is a detail convenient for the analysis which is useful in the proof of Theorem 4.28.

Observe, that although the decomposition tells us to use the derivative $\frac{\partial}{\partial x}Q(B^{\bullet}(\rho_b))$ for sampling cliques, we use a sampler $\Gamma Q^{\bullet}$ for *rooted* cliques, since this is more convenient. That these two variants are indeed equivalent, can be seen as follows: taking the derivative $\frac{\partial}{\partial x}Q(B^{\bullet}(\rho_b))$ means that we shift the series by one, i.e. we sample cliques of size $k-1$ and attach these to the node v for which we are currently sampling cliques, adding edges between all nodes of the clique and v. On the other hand, we can also sample a rooted clique of size k and merge the root node with the current node v. Note that the Boltzmann distribution is in both cases the same. For $k \geq 2$ we have according to (2.5) that the probability that we draw a clique of size k is in the first case $x^{k-1}/(Q'(x)(k-1)!) = x^{k-1}/((e^x-1)(k-1)!)$ and in the second case $x^k/(Q^{\bullet}(x)(k-1)!) = x^{k-1}/((e^x-1)(k-1)!)$ as we have k possibilities to root a graph on k nodes. Note that this observation holds in general.

This completes the specification of the sampler for block graphs. Now, by applying Theorem 2.10 (Theorem 8 of [**DFLS04**]), Theorem 4.1 follows immediately. In a completely analogous way we may obtain a sampler for cactus graphs. As a technical tool, we will later need the following lemma, which says that all possible output sizes are almost equally probable to occur, even if we constrain the number of atoms the sampler can generate.

LEMMA 4.22. *Let $0 < \varepsilon \leq \frac{1}{2}$. Whenever n is sufficiently large the following statement holds. For every $s \in \mathbb{N}$ and $t \in \mathbb{N}_0$ define the quantity*

$$p_{s,t} := \Pr\left[|\Gamma\nu B(n,\varepsilon)| = s \text{ and } \Gamma\nu B(n,\varepsilon) \text{ created precisely } t+s \text{ nodes}\right]. \tag{4.18}$$

For all $s = \alpha n$, where $\alpha \in (1-\varepsilon, 1+\varepsilon)$ it holds $p_{s,t} \sim \alpha^{-\frac{3}{2}} \cdot p_{n,t}$.

PROOF. Let $n_1, n_2 \in ((1-\varepsilon)n, (1+\varepsilon)n)$. Consider the execution of the sampler $\Gamma\nu B$, and observe that every sequence of generated and discarded objects resulting in an object O_1 of size n_1, can be transformed into a sequence which results in an object O_2 of size n_2. This can be done by keeping the random choices of the sampler which resulted in discarded objects, and by substituting the random choices in $\Gamma B^{\bullet}$ which generated O_1 with the choices which generate O_2. More formally, denote with $\mathcal{B}^{\bullet}_{\leq s}$ the set of rooted block graphs on at most s

nodes, and define for every $t \in \mathbb{N}_0$ the set of sequences of objects

$$S_t := \Big\{ (R_1, \ldots, R_\ell) \;\Big|\; 1 \le \ell \le t \;\text{ and }\; \sum_{i=1}^{\ell} |R_i| = t \;\text{ and } \\ \forall i \in \{1, \ldots, \ell\} : R_i \in \mathcal{B}^\bullet_{\le (1-\varepsilon)n} \cup \{\bot\} \Big\},$$

which form a *prefix* of a sequence that finally yields an object which is accepted by the sampler $\Gamma\nu B$. Note that here we assume that $|\bot| = \lceil (1+\varepsilon)n \rceil$, as $\Gamma B^\bullet$ interrupts its execution (and returns $\bot$) immediately after precisely $\lceil (1+\varepsilon)n \rceil$ nodes were generated. With this notation, the probability from (4.18) can be written as

$$p_{s,t} = \left(\sum_{\{R_1,\ldots,R_\ell\} \in S_t} \prod_{i=1}^{\ell} \Pr\left[\Gamma B^\bullet(n,\varepsilon) = R_i\right] \right) \cdot \Pr\left[|\Gamma B^\bullet(n,\varepsilon)| = s\right]. \tag{4.19}$$

It is easy to see that the first sum in this expression is the same for all $s \in ((1-\varepsilon)n, (1+\varepsilon)n)$, i.e. the probability that $\Gamma\nu B$ samples an object of size s differs only in the last term for different sizes. Hence, we only have to consider $\Pr[|\Gamma B^\bullet| = s]$.

Due to Theorem 4.6 the number of rooted block graphs $b^\bullet_N$ on N nodes is asymptotically $c' \rho_b^{-N} N^{-\frac{3}{2}} N!$, for an appropriately chosen constant c'. But this implies that the probability that $\Gamma B^\bullet$ outputs an object of size $(1-\varepsilon)n < N < (1+\varepsilon)n$ is asymptotically $c \cdot N^{-\frac{3}{2}}$, for a constant c. Indeed, by (2.5) we have

$$\Pr\left[|\Gamma B^\bullet(n,\varepsilon)| = N\right] \;=\; b^\bullet_N \cdot \frac{1}{B^\bullet(\rho_b)} \cdot \frac{\rho_b^N}{N!} \;\sim\; c' \rho_b^{-N} N^{-\frac{3}{2}} N! \cdot \frac{1}{B^\bullet(\rho_b)} \cdot \frac{\rho_b^N}{N!} \;\sim:\; c N^{-\frac{3}{2}}.$$

Combining this with (4.19) yields immediately the claim. □

Observe that we can derive an analogue lemma for cactus graphs in the same way.

4.6.2. Implementation and Statistics. Our implementation of the Boltzmann sampler for cactus graphs and for block graphs can be found at [**PW06**]. In this section we describe some details of it, which are highly relevant for the performance of our program.

We implemented the singular Boltzmann sampler for cactus graphs and for block graphs. Both of them turn out to be rather efficient. Note however, that we did not try to tune the code furthermore, even if further improvements – especially in the case of block graphs – are possible.

As follows from the previous section, the only computations we need to make in advance for the cactus graph sampler are to compute the values $K^\bullet(\rho_k)$ and $\frac{\partial}{\partial x} C(x)\big|_{x=K^\bullet(\rho_k)}$ needed in the sampler $\Gamma K^\bullet_{<m}(\rho_k)$ which reduces to calculating the constant $K^\bullet_0(1)$ from (4.10) and to evaluating the explicitly given generating function $\frac{\partial}{\partial x} C(x)$ at $x = K^\bullet_0(1)$ (and analogously for the block graph sampler). These computations can be accomplished up to arbitrary precision for example with MAPLE; in our implementation we used 80-bit floating point numbers[1].

If one wants to avoid the seldom cases, when the precision of numbers does not suffice to draw uniformly at random, one can also use a package for arbitrary precision numbers.

[1] IEEE 754 double-extended precision

A main task of the program is to build and merge graphs at given nodes. For an efficient implementation of this operation we use the *half-edge* data structure [**Wei85**], which allows us to implement the merging operation in constant time.

The procedures described in Section 4.6.1 are straightforward to implement in any programming language and using the *half-edge* data structure [**Wei85**] yields already a quite efficient sampler for up to 10^4 nodes. We propose here three major modifications which result in an enormous speed-up:

- we use our own memory management for the graph structure,
- we use an iterative instead of a recursive approach for substitution in the sampler and
- we use tables for the Poisson, cycle and clique distribution which have been computed in advance.

These modifications allow us to sample graphs with up to 10^7 nodes in only a few seconds.

The first modification is necessary, as the sampler rejects very often with too small objects if the target size is large. Thus the memory management overhead impacts overall running time, if we use the standard memory allocation functions. Therefore we propose two modifications: first, we allocate one block of memory at the beginning of our sampler, which is large enough to contain the whole graph if we hit the given upper bound. In case of cactus graphs space for $3n$ half-edges is enough. Since we don't have a linear bound on the number of edges in the case of block graphs, we would have to allocate space for $O(n \log n / \log \log n)$ half-edges. But this seems too pessimistic and in practice the sampler worked fine with space for $O(n)$ half-edges (see below for an explanation). Now, our second, more important modification is based on the observation that we can follow the recursive structure of the samplers given above to manage the reserved space like a *stack* to build the graph on. We use a variable as a kind of stack pointer which tells us where to write the next half-edge that is created. If the whole graph is rejected because the branching process dies out early or the object becomes too large, we just have to reset *one variable*, i.e., we delete the generated graph with a *single* instruction.

To avoid being limited by the stack depth and tremendously reducing the number of function calls, we propose the second modification: we use an iterative approach to replace nodes by rooted cactus graphs and block graphs respectively. For each node that is created in the sampling process we keep an "active" flag for one of the adjacent half-edges that tells us if this node has already been substituted by a rooted object. Then we just process our stack from bottom to top and substitute the next active node until no more active nodes can be found. Notice that we never have to *explicitly* carry out a substitution operation; instead we can attach newly generated blocks (i.e., cycles and cliques respectively) at the already existing node by modifying the pointers of the half-edge structures.

Our last performance improvement is to reduce the number of floating point operations by computing the tables for the distribution functions up to machine precision in advance.

An obvious possible improvement for the block graph sampler which we did not implement is to avoid building the cliques explicitly but to store only the nodes and the information

to which clique they belong. Then one just outputs the edges at the output stage of the program.

The described improvements made it possible to sample cactus and block graphs with up to $10,000,000$ nodes within about 30 seconds of running time on a Intel Xeon 2.2 Ghz machine with 1 GB of RAM running Debian Linux.

We gathered some statistics with our samplers which support the theoretical results that we will prove in Section 4.7. In the top of Figure 1 we plotted a histogram for the maximum degree of random block graphs and random cactus graphs taken from $5,000$ runs of our samplers with parameters $n = 10^6$ and $\varepsilon = 0.05$. Confirming our theoretical results, the maximum degree is slightly larger in random block graphs (green dashed) than in random cactus graphs (solid red).

A histogram for the maximum clique size in random block graphs and the maximum cycle size in random cactus graphs can be found in the bottom of Figure 1, where the data was again taken from $5,000$ runs of our samplers with the same parameters as above. Again, the gathered data supports our theoretical predictions of Theorems 4.2 and 4.3.

Finally, we counted the number of edges in random cactus graphs and random block graphs – the results are presented in Figure 2. We sampled graphs with sizes in $[100, 10^5]$ using a step size of 50 and $\varepsilon = 0.02$ where we have taken the average of 50 graphs for each target size. In the case of cactus graphs (top), the plot clearly has the expected structure due to Theorem 4.11. For block graphs (bottom) we get a somehow larger number of expected edges, but it still seems to be clearly linear in n, as it levels off at a value around $\approx 1.283n$. It remains an open problem to derive this observation theoretically – unfortunately we cannot prove how many edges we expect, as we cannot apply the theory developed in [**FS06**] to the bivariate generating function for cliques with respect to number of nodes and edges

$$\sum_{n\geq 2} \frac{x^n}{n!}\, y^{\binom{n}{2}}.$$

4.7. Structural Properties

4.7.1. Preliminaries. Next, we give all proofs and lemmas, which are additionally needed to prove our main result about the maximum degree and maximum size of a biconnected block in a random city and desert graph, see Theorems 4.2 and 4.3. The subsequent lemmas provide us for various probability distributions with estimates for their tails. We start with an easy observation, which will become helpful at several places. Recall Definition 4.21, and observe that for $k \geq 2$ and all $x > 0$ we have

$$\Pr\left[\mathrm{Cl}\,(x) = k\right] = \frac{x^k}{Q^{\bullet}(x)(k-1)!} \geq \frac{e^{-x}x^{k-1}}{(k-1)!} = \Pr[\mathrm{Po}\,(x) = k-1], \tag{4.20}$$

and similarly

$$\Pr\left[\mathrm{Cl}\,(x) = k\right] \leq \frac{e^x}{e^x - 1} \cdot \Pr[\mathrm{Po}\,(x) = k-1] =: c_x \cdot \Pr[\mathrm{Po}\,(x) = k-1]. \tag{4.21}$$

Before we proceed, let us state a well-known estimate for the tail of the Poisson distribution, which will be useful at several places in the proofs below.

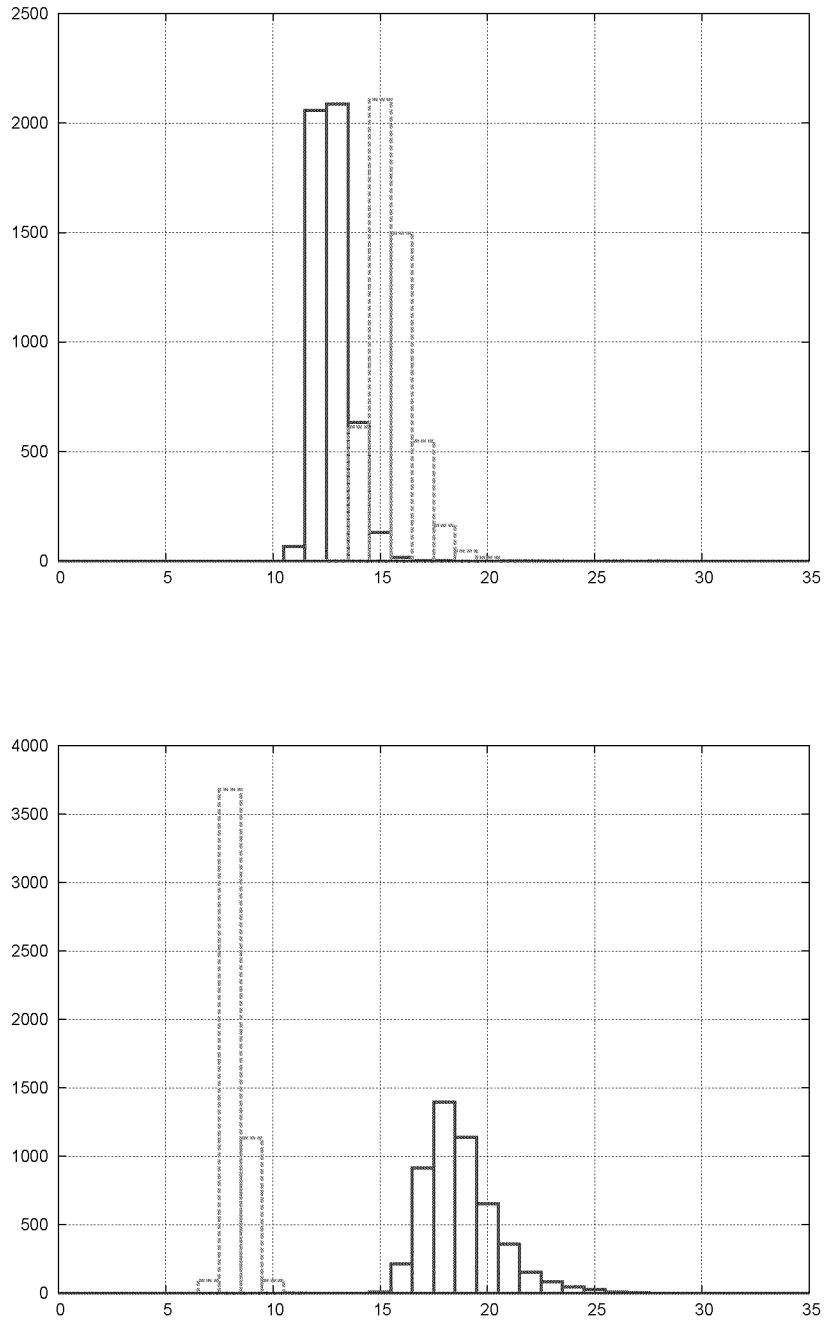

FIGURE 1. Maximum degree (top) and maximum cycle size, maximum clique size respectively (bottom) in a random cactus graph and a random block graph. Red solid lines for cactus graphs and green dashed lines for block graphs. The plot shows the result of $5,000$ runs of the Boltzmann samplers with parameters $n = 10^6$ and $\varepsilon = 0.05$. On the x-axis the maximum degree and maximum cycle length, maximum clique size respectively and on the y axis the amount of graphs.

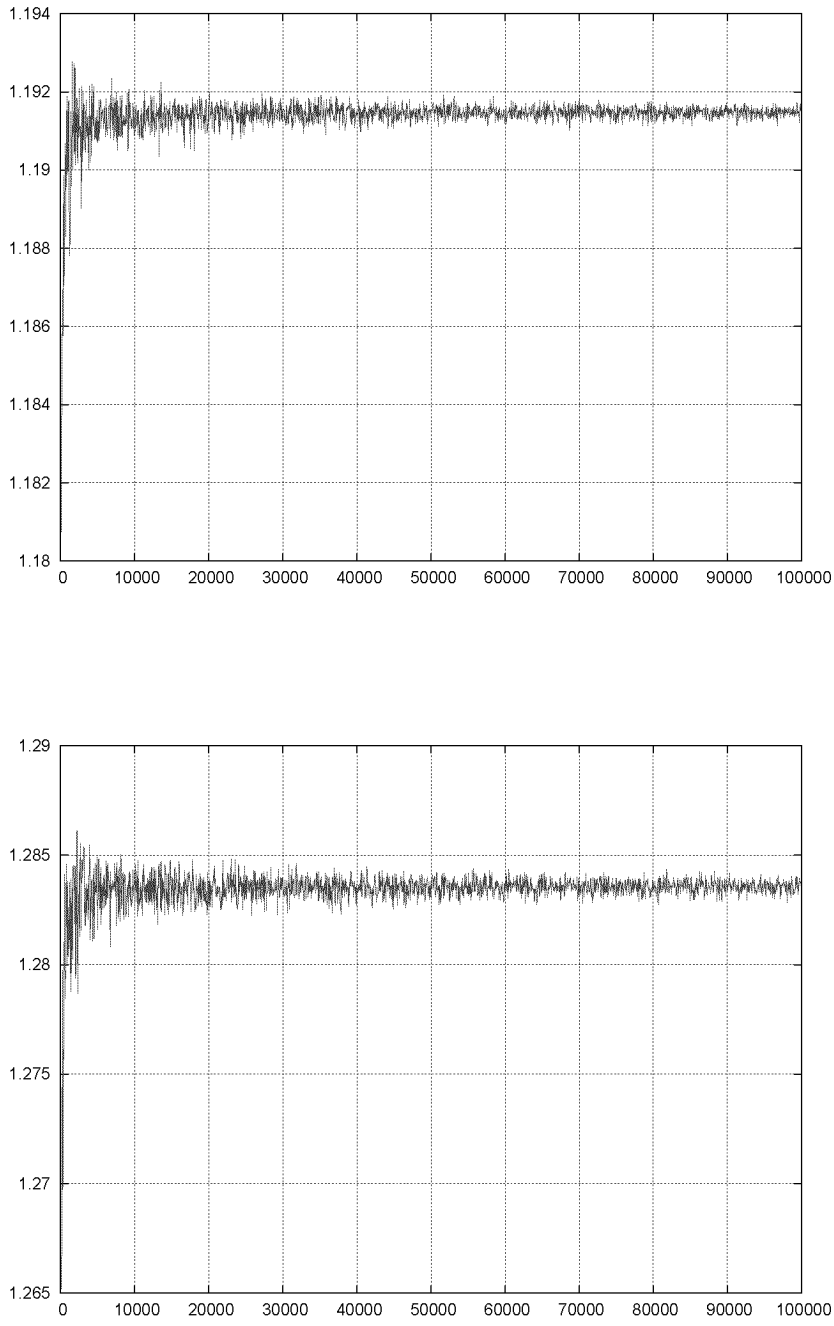

FIGURE 2. Average number of edges in a random cactus graph (top) and random block graph (bottom). The plot shows the average over 50 runs for $n = 100, \ldots, 10^5$ with a step size of 50 and $\varepsilon = 0.02$.

LEMMA 4.23. *Let $\alpha, \lambda > 0$ be constants. There exists $n_0 = n_0(\alpha, \lambda)$ such that for $n \geq n_0$*

$$\Pr\left[\mathrm{Po}(\lambda) \geq \alpha \frac{\log n}{\log\log n}\right] \sim \Pr\left[\mathrm{Po}(\lambda) = \left\lceil \alpha \frac{\log n}{\log\log n} \right\rceil\right] \sim n^{-\alpha+o(1)}.$$

PROOF. We omit $\lceil . \rceil$ to keep the calculations short. Let $u(n) := \alpha \frac{\log n}{\log \log n}$, and observe that

$$\Pr[\mathrm{Po}(\lambda) \geq u(n)] = \sum_{k \geq u(n)} e^{-\lambda} \frac{\lambda^k}{k!} = \Pr[\mathrm{Po}(\lambda) = u(n)] \cdot \sum_{k \geq 0} \frac{\lambda^k}{\prod_{i=1}^{k} (u(n)+i)} \sim \Pr[\mathrm{Po}(\lambda) = u(n)]$$

Now we show the second part of the statement. Using $x! = (1 + \Theta(x^{-1})) \cdot \sqrt{2\pi x} \left(\frac{x}{e}\right)^x$, we obtain

$$\Pr[\mathrm{Po}(\lambda) = u(n)] = e^{-\lambda} \frac{\lambda^{u(n)}}{(u(n))!} = (1+o(1)) e^{-\lambda} \left(\frac{\lambda e \log\log n}{\alpha \log n}\right)^{\alpha \frac{\log n}{\log\log n}} \left(\frac{\log\log n}{2\pi\alpha \log n}\right)^{1/2}.$$

Note that for every positive constant c we have $(c \log\log n)^{\frac{\log n}{\log\log n}} \cdot (\frac{\log\log n}{\log n})^{1/2} = n^{o(1)}$. For sufficiently large n we obtain

$$\Pr[\mathrm{Po}(\lambda) = u(n)] = \frac{e^{-\lambda}}{(2\pi\alpha)^{1/2}} \cdot (\log n)^{-\alpha \frac{\log n}{\log\log n}} \cdot n^{o(1)} = n^{-\alpha + o(1)}.$$

□

Before we proceed, let us first make a technical definition. We say that a random variable X is *sumClique*-distributed with parameters λ and μ, $X \sim \mathrm{SCl}(\lambda, \mu)$, if it is the sum of $\mathrm{Po}(\lambda)$ independent $\mathrm{Cl}(\mu) - 1$ distributed random variables. A crucial ingredient in our proof for the magnitude of the maximum degree in a random block graph will be the following lemma, which is a statement about the tail of a sumClique distributed random variable.

LEMMA 4.24. *Let $\lambda, \mu > 0$ and $0 < \varepsilon < \frac{1}{9}$ be constants. There is an $n_0 = n_0(\varepsilon) \geq 0$ such that whenever $n > n_0$ it holds*

$$\Pr\left[\mathrm{SCl}(\lambda,\mu) \geq (1-\varepsilon)\frac{\log n}{\log^{(3)}(n)}\right] \gg n^{-1+\frac{\varepsilon}{2}} \text{ and } \Pr\left[\mathrm{SCl}(\lambda,\mu) \geq (1+\varepsilon)\frac{\log n}{\log^{(3)}(n)}\right] \ll n^{-1-\frac{\varepsilon}{2}}.$$

PROOF. In the following we shall omit $\lceil . \rceil$ to keep the calculations concise, but it can easily be verified that our statements are also true in the general case. It is well-known that a sum of Poisson variables is distributed as a single Poisson variable with the sum of their parameters. Recall (4.20) and (4.21), and note that for all $t, s \in \mathbb{N}$, if $C_1, \ldots, C_t$ are independent clique-distributed variables with parameter μ, we obtain

$$\Pr[\mathrm{Po}(t\mu) = s] \;\leq\; \Pr\left[\sum_{i=1}^{t} C_i - 1 = s\right] \;\leq\; c_\mu^t \cdot \Pr[\mathrm{Po}(t\mu) = s] \tag{4.22}$$

Let us abbreviate $b(n) := \frac{\log n}{\log^{(3)} n}$ and $q(n) := \frac{\log n}{\log\log n}$. With (4.22) we obtain

$$\begin{aligned}\Pr[\mathrm{SCl}(\lambda,\mu) \geq (1-\varepsilon)b(n)] &\geq \Pr\left[\mathrm{Po}(\lambda) = \frac{\varepsilon}{9}q(n)\right] \cdot \Pr\left[\sum_{i=1}^{t} \mathrm{Cl}(\mu) - 1 \geq (1-\varepsilon)b(n) \mid t = \frac{\varepsilon}{9}q(n)\right] \\ &\geq \Pr\left[\mathrm{Po}(\lambda) = \frac{\varepsilon}{9}q(n)\right] \cdot \Pr\left[\mathrm{Po}\left(\frac{\varepsilon\mu}{9}q(n)\right) \geq (1-\varepsilon)b(n)\right].\end{aligned}$$

We estimate the terms on the right hand side of the above inequality one by one. The first probability is due to Lemma 4.23, if n is sufficiently large, $n^{-\frac{\varepsilon}{9}+o(1)}$. In the sequel we will show that for all constants $\alpha, \beta > 0$

$$\Pr[\mathrm{Po}(\alpha q(n)) \geq \beta b(n)] = n^{-\beta + o(1)}, \tag{4.23}$$

which immediately completes the proof of the first statement of the lemma. Indeed, with the definition of the Poisson distribution and the estimate $x! = (1+o(1))\left(\frac{x}{e}\right)^x \sqrt{2\pi x}$ we obtain

$$\begin{aligned}\Pr\left[\mathrm{Po}\left(\alpha q(n)\right) \geq \beta b(n)\right] &= \sum_{t \geq \beta b(n)} e^{-\alpha q(n)} \cdot \frac{(\alpha q(n))^t}{t!} \\ &= (1+o(1)) \cdot e^{-\alpha q(n)} \cdot \frac{(\alpha q(n))^{\beta b(n)}}{(\beta b(n))!} \\ &= n^{o(1)} \cdot \left(\frac{e\alpha}{\beta} \cdot \frac{\log^{(3)} n}{\log \log n}\right)^{\beta \frac{\log n}{\log^{(3)} n}} = n^{-\beta + o(1)}.\end{aligned}$$

Now we show the second statement of the lemma. Recall (4.22) and note that for all $t \geq \lambda_1 \geq \lambda_2 \geq 1$ we have $\Pr\left[\mathrm{Po}\left(\lambda_1\right) \geq t\right] \geq \Pr\left[\mathrm{Po}\left(\lambda_2\right) \geq t\right]$. We obtain with $q(n) = \frac{\log n}{\log \log n}$

$$\begin{aligned}\Pr\left[\mathrm{SCl}\left(\lambda, \mu\right) \geq (1+\varepsilon)b(n)\right] &= \sum_{t \geq 0} \Pr\left[\mathrm{Po}\left(\lambda\right) = t\right] \cdot \Pr\left[\sum_{i=1}^{t} \mathrm{Cl}\left(\mu\right) - 1 \geq (1+\varepsilon)b(n)\right] \\ &\leq \sum_{t=0}^{2q(n)} \Pr\left[\mathrm{Po}\left(\lambda\right) = t\right] \cdot \Pr\left[\mathrm{Po}\left(t\mu\right) \geq (1+\varepsilon)b(n)\right] \cdot c_\mu^t \\ &\qquad + \Pr\left[\mathrm{Po}\left(\lambda\right) \geq 2q(n)\right] \\ &\overset{\text{(Lem. 4.23)}}{\leq} n^{o(1)} \cdot \Pr\left[\mathrm{Po}\left(2q(n)\right) \geq (1+\varepsilon)b(n)\right] + n^{-2+o(1)} \\ &\overset{(4.23)}{\leq} n^{-1-\varepsilon+o(1)} + n^{-2+o(1)} \ll n^{-1-\frac{\varepsilon}{2}},\end{aligned}$$

as $c_\mu = \Theta(1)$. □

Similarly, to prove the estimates on the maximum clique size in a random block graph and the maximum node degree and maximum size of a cycle in a random cactus graph, we need the analogon of Lemma 4.24 for the tails of a clique and a cycle distributed random variable.

Lemma 4.25. *Let $\lambda, \varepsilon > 0$ be constants. There is an $n_0 = n_0(\varepsilon) \geq 0$ such that whenever $n > n_0$ it holds*

$$\Pr\left[\mathrm{Cl}\left(\lambda\right) \geq (1-\varepsilon)\frac{\log n}{\log \log n}\right] > n^{-1+\varepsilon+o(1)} \text{ and } \Pr\left[\mathrm{Cl}\left(\lambda\right) \geq (1+\varepsilon)\frac{\log n}{\log \log n}\right] < n^{-1-\varepsilon+o(1)}.$$

Proof. By (4.21) and Lemma 4.23 we have for $t = (1+\varepsilon)\frac{\log n}{\log \log n}$

$$\begin{aligned}\Pr\left[\mathrm{Cl}\left(\lambda\right) \geq t\right] &= \sum_{x=t}^{\infty} \Pr\left[\mathrm{Cl}\left(\lambda\right) = x\right] \leq \sum_{x=t}^{\infty} \frac{e^\lambda}{e^\lambda - 1} \Pr\left[\mathrm{Po}\left(\lambda\right) = x - 1\right] \\ &\leq \frac{e^\lambda}{e^\lambda - 1} \Pr\left[\mathrm{Po}\left(\lambda\right) \geq t - 1\right] \leq n^{-1-\varepsilon+o(1)}.\end{aligned}$$

Using (4.20) and Lemma 4.23 we obtain for $t = (1-\varepsilon)\frac{\log n}{\log \log n}$

$$\Pr\left[\mathrm{Cl}\left(\lambda\right) \geq t\right] \geq \Pr\left[\mathrm{Cl}\left(\lambda\right) = t\right] \geq \Pr\left[\mathrm{Po}\left(\lambda\right) = t - 1\right] \geq n^{-1+\varepsilon+o(1)}.$$

□

The cycle distribution needed by the sampler for (rooted) cactus graphs is defined as follows.

Definition 4.26. The cycle distribution denoted by $\mathrm{Cy}(x)$ is defined by

$$\Pr[\mathrm{Cy}(x) = k] = \begin{cases} \frac{x^2}{C^\bullet(x)} & k = 2 \\ \frac{x^k}{2C^\bullet(x)} & k > 2 \\ 0 & \text{otherwise,} \end{cases}$$

where $C^\bullet(x)$ is the exponential generating function for rooted cycles and a single rooted edge.

Lemma 4.27. *Let $\lambda, \varepsilon > 0$ be constants. There is an $n_0 = n_0(\varepsilon) \geq 0$ such that whenever $n > n_0$ it holds*

$$\Pr\left[\mathrm{Cy}(\lambda) \geq (1-\varepsilon)\log_{\frac{1}{\lambda}} n\right] \geq n^{-1+\varepsilon+o(1)} \text{ and } \Pr\left[\mathrm{Cy}(\lambda) \geq (1+\varepsilon)\log_{\frac{1}{\lambda}} n\right] \leq n^{-1-\varepsilon+o(1)}.$$

Proof. By the definition of the cycle distribution we have

$$\Pr\left[\mathrm{Cy}(\lambda) \geq (1+\varepsilon)\log_{\frac{1}{\lambda}} n\right] = \sum_{k=\lceil(1-\varepsilon)\log_{\frac{1}{\lambda}} n\rceil}^{\infty} \frac{\lambda^k}{2C^\bullet(\lambda)} = \frac{1}{2C^\bullet(\lambda)} \frac{\lambda^{\lceil(1+\varepsilon)\log_{\frac{1}{\lambda}} n\rceil}}{1-\lambda} = n^{-1-\varepsilon+o(1)}$$

and

$$\Pr\left[\mathrm{Cy}(\lambda) \geq (1-\varepsilon)\log_{\frac{1}{\lambda}} n\right] \geq \Pr\left[\mathrm{Cy}(\lambda) = \lceil(1-\varepsilon)\log_{\frac{1}{\lambda}} n\rceil\right] = \frac{\lambda^{\lceil(1-\varepsilon)\log_{\frac{1}{\lambda}} n\rceil}}{2C^\bullet(\lambda)} = n^{-1+\varepsilon+o(1)}.$$

□

4.7.2. Completing the Proofs. In this section we are going to demonstrate our new proof method in detail by estimating the maximum node degree in a random block graph. Additionally, we provide the proofs of the remaining theorems stated in Section 4.2.

Theorem 4.28. *Let $\varepsilon > 0$. For almost all n, it holds that the maximum node degree in a random block graph is w.h.p. in the interval $\mathcal{I}_{\varepsilon,n} := ((1-\varepsilon)d(n), (1+\varepsilon)d(n))$, where $d(n) := \frac{\log n}{\log^{(3)} n}$.*

Proof. The main idea in the following proof is to consider the process of object construction during a run of the singular Boltzmann sampler $\Gamma\nu B$ for block graphs, as it is described in Section 4.6. Recall that the output of $\Gamma\nu B$ is a random block graph according to the Boltzmann distribution (2.5), and here we want to make a statement about random block graphs B_n of a given size n. As described below, it turns out to be very convenient to study the properties of the *generated shapes* of $\Gamma\nu B$, instead of studying properties of B_n directly. As we shall see, these will translate to properties of random block graphs in a straightforward way.

Before we proceed, let us modify slightly the sampling algorithm defined in Section 4.6. This is done for solely technical reasons, and will become clear later. The idea behind the singular rejection sampler $\Gamma\nu B$ is to repeatedly sample and reject rooted block graphs, until $\Gamma B^\bullet$ outputs an object of the desired size. $\Gamma B^\bullet$ proceeds as follows: for every generated node v, it calculates a random value p_v according to a Poisson law with parameter $\lambda := \frac{\partial}{\partial x} Q(x)\big|_{x=B^\bullet(\rho_b)}$, and then calculates p_v independent random values according to a clique law with parameter $\mu := B^\bullet(\rho_b)$. Then it generates rooted cliques with sizes given by the random values, joins them at their roots, and proceeds in an identical way for all newly created nodes, until the process dies out. The important modification in the sampling procedure below is that we

let the sampler $\Gamma\nu B$ make *precisely* $\lceil n \log n \rceil$ random choices, abort after that immediately its execution, and output the first object of size in $\mathcal{I} := (\frac{n}{2}, \frac{3n}{2})$, if there is any that was generated during a call of $\Gamma B^\bullet$:

$\widetilde{\Gamma}\nu B(n)$: for$(i = 1 \dots \lceil n \log n \rceil)$
$p_i \leftarrow$ random value according to $\mathrm{Po}(\lambda)$
$(C_1^{(i)}, \dots, C_{p_i}^{(i)}) \leftarrow p_i$ independent random values according to $\mathrm{Cl}(\mu)$
$d_i :\leftarrow \sum_{1 \le j \le p_i} (C_j^{(i)} - 1)$ $\qquad (\star\star)$
run the sampler $\Gamma\nu B(n, \frac{1}{2})$ with the above random values, and do not stop its execution if it generated an object with size in $(\frac{n}{2}, \frac{3n}{2})$, and abort it as soon as all $\lceil n \log n \rceil$ random values were exhausted
if(above execution of $\Gamma\nu B(n, \frac{1}{2})$ generated an object with size in $(\frac{n}{2}, \frac{3n}{2})$)
return the first such object (and label its nodes uniformly at random)
else return $\perp$

The values d_i calculated in the line marked with $(\star\star)$ are needed later in the analysis. Before we proceed, let us make some important observations. $\widetilde{\Gamma}\nu B$ behaves very similar to the sampler $\Gamma\nu B$ described in Section 4.6. The only difference is that if it generates an object of size in $\mathcal{I}$, it does not output it directly, but continues its execution until it has generated at least $n \log n$ nodes. On the other hand, if it does not succeed in sampling an object of the desired size, it returns a default object $\perp$. Moreover, observe that the sampler $\Gamma B^\bullet$, called by the algorithm above, if it built a graph on s nodes, then it used *exactly* s (consecutive) values $p_i, \dots, p_{i+s}$, and the corresponding random clique-distributed values to do so.

In order to prove the theorem, we will proceed in two steps. Let $\mathcal{B}'$ denote the set of all block graphs with maximum degree in $\bigcup_{s \ge 1} \mathcal{I}_{\varepsilon,s}$. First, we are going to show

$$\Pr\left[\widetilde{\Gamma}\nu B(n) \in \mathcal{B}'\right] = 1 - o(1), \tag{4.24}$$

i.e., with high probability, the sampler $\widetilde{\Gamma}\nu B(n)$ will output an object, which is not $\perp$, and which has the property that its maximum degree lies in the desired interval. Then, in the second part of the proof, we will show that the above statement indeed proves the theorem.

First we show (4.24). Observe that the sampler $\Gamma\nu B$, when called by $\widetilde{\Gamma}\nu B$, samples only the *shape* of a graph – the labels are distributed on the nodes at the end of the process. For a labelled graph G define $s(G)$ to be its shape, i.e., the corresponding unlabelled graph, and for a shape S, let $\ell(S)$ be the set of labelled graphs with shape S, and let $|S|$ denote the number of nodes in S.

Furthermore, set $s(\mathcal{B}') := \bigcup_{G \in \mathcal{B}'} s(G)$. Now, for a shape S with $|S| \in \mathcal{I}$, let $\Pr[S]$ be the probability that S is the first shape of size in $\mathcal{I}$ generated by the repeated execution of $\Gamma B^\bullet$, and let $\Pr[G \mid S]$ be the probability that the sampler outputs the labelled graph G, given that the shape S had been generated. Observe that $\Pr[G \mid S] = \frac{\mathrm{aut}(S)}{|S|!}$, if $S = s(G)$, and zero

otherwise, where $\mathrm{aut}(S)$ denotes the number of automorphisms of S. With this, we get

$$\begin{aligned}\Pr\left[\Gamma\nu B(n) \in \mathcal{B}'\right] &= \sum_{S \in s(\mathcal{B}')} \Pr[S] \cdot \sum_{G \in \ell(S)} \Pr[G \mid S] \\ &= \sum_{S \in s(\mathcal{B}')} \Pr[S] \cdot |\ell(S)| \cdot \frac{\mathrm{aut}(S)}{|S|!} = \sum_{S \in s(\mathcal{B}')} \Pr[S],\end{aligned} \tag{4.25}$$

where the last step follows from the fact $|\ell(S)| = \frac{|S|!}{\mathrm{aut}(S)}$. Hence, in order to show (4.24), it will suffice to prove that the shape generated by $\Gamma\nu B$, during its execution in $\widetilde{\Gamma}\nu B$, has with high probability the property that its maximum degree is in the desired interval.

Let us now focus on the evolution of the node degrees during the sampling process. The sampler repeatedly calls $\Gamma B^\bullet(n, \frac{1}{2})$, which starts with a single node v. This algorithm calculates a random value p_v according to a Poisson law with parameter λ, and then calculates p_v random values $C_1^{(v)}, \ldots, C_{p_v}^{(v)}$ according to a clique law with parameter μ. This determines the *size* and *structure* of the neighbourhood of v. Then it creates p_v cliques with sizes $C_1^{(v)}, \ldots, C_{p_v}^{(v)}$, and joins them together at their roots (and v). Finally, for each of the newly created nodes, the same procedure is repeated independently until the process dies out, or it is interrupted, because too many nodes were generated. Thus, the degree of a node is given by the outcome of a *sumClique*-distributed random variable with parameters λ and μ, plus the size of the clique, in which this node is contained, minus one. All in all, a *lower bound* for the degree of a node v generated during the sampling process is its *sumClique*-value $\mathrm{scl}(v)$, and an *upper bound* is the size of the clique it is contained in plus $\mathrm{scl}(v)$.

Let $\varepsilon' := \frac{\varepsilon}{2}$. With the above discussion in mind, consider the execution of $\widetilde{\Gamma}\nu B(n)$, which generates the random values p_i, $C_j^{(i)}$ and d_i (in the line marked with $(\star\star)$ in the exposition of the algorithm), and let us define the following four events:

($\mathcal{A}$) $\widetilde{\Gamma}\nu B(n) \neq \perp$.

($\mathcal{B}$) Every sequence of $\frac{n}{2}$ consecutive values d_i contains a value larger than $(1-\varepsilon')d(n)$, i.e.,

$$\mathcal{B} := \left\{(d_1, \ldots, d_{\lceil n \log n \rceil}) \;\middle|\; \forall 1 \leq i \leq \lceil n \log n \rceil - \frac{n}{2} : \exists i \leq j \leq i + \frac{n}{2} : d_j \geq (1-\varepsilon')d(n)\right\}.$$

($\mathcal{C}$) There is no $1 \leq i \leq \lceil n \log n \rceil$ such that $d_i \geq (1+\varepsilon')d(n)$.

($\mathcal{D}$) For all $1 \leq i \leq \lceil n \log n \rceil$ and $1 \leq j \leq p_i$ we have $C_j^{(i)} \leq 2\frac{\log n}{\log \log n}$.

The motivation behind the above events is that if they occur simultaneously, then the shape generated by $\widetilde{\Gamma}\nu B$ will have the property that its maximum node degree is in $\mathcal{I}_{\varepsilon,n}$. In fact, suppose that $|\widetilde{\Gamma}\nu B(n)| \neq \perp$, i.e. $\mathcal{A}$ holds. Then there is an index i_0 and a number $s \in (\frac{n}{2}, \frac{3n}{2})$, such that the sampler $\Gamma B^\bullet$, given the random choices $p_{i_0}, \ldots, p_{i_0+s}$ and the corresponding values $(C_1^{(i_0)}, \ldots, C_{p_{i_0}}^{(i_0)}), \ldots, (C_1^{(i_0+s)}, \ldots, C_{p_{i_0+s}}^{(i_0+s)})$, outputs a shape of size s. Now, suppose that additionally $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{D}$ occur. Then, due to $\mathcal{B}$, there is an index $i_0 \leq i' \leq i_0 + s$, such that $d_{i'} = \sum_{j=1}^{p_{i'}} (C_j^{(i')} - 1) \geq (1-\varepsilon')d(n)$, i.e. $\widetilde{\Gamma}\nu B(n)$ has a node of at least that degree. Furthermore, due to $\mathcal{C}$ and $\mathcal{D}$, there is for sufficiently large n no node with degree larger than

$$2\frac{\log n}{\log \log n} + (1+\varepsilon')d(n) = 2\frac{\log n}{\log \log n} + (1+\varepsilon')\frac{\log n}{\log^{(3)} n} < (1+\varepsilon)d(n).$$

Hence, with (4.25) we have

$$\Pr\left[\widetilde{\Gamma}\nu B(n) \in \mathcal{B}'\right] \geq \Pr\left[\mathcal{A} \text{ and } \mathcal{B} \text{ and } \mathcal{C} \text{ and } \mathcal{D}\right]. \tag{4.26}$$

In order to show (4.24), we have to prove that the probabilities for the events $\overline{\mathcal{A}}$, $\overline{\mathcal{B}}$, $\overline{\mathcal{C}}$, and $\overline{\mathcal{D}}$ are all $o(1)$.

Let us first calculate $\Pr\left[\overline{\mathcal{A}}\right]$. Recall that with Theorem 2.10 (Theorem 8 of [**DFLS04**]), we obtain that the expected number of nodes N_n generated by $\Gamma\nu B(n, \frac{1}{2})$, until it outputs an object, is $\Theta(n)$, and that the variance is $\Theta(n^2)$. Hence, using the fact that $\widetilde{\Gamma}\nu B$ generates at least $n \log n$ nodes, we get

$$\Pr\left[\widetilde{\Gamma}\nu B(n) = \perp\right] \leq \Pr\left[N_n \geq n \log n\right] \leq \frac{\operatorname{Var}\left[N_n\right]}{(n \log n - \mathbb{E}\left[N_n\right])^2} = \Theta\left(\frac{1}{\log^2 n}\right) = o(1). \tag{4.27}$$

Next, we consider the probability of the event $\overline{\mathcal{B}}$. Observe that the d_i's are independent sumClique-distributed random variables with parameters λ and μ. Due to Lemma 4.24, the probability that a $\operatorname{SCl}(\lambda, \mu)$ random variable is larger than $(1 - \varepsilon')d(n)$ is at least $n^{-1+\frac{\varepsilon'}{2}}$, which implies that the probability that $\frac{n}{2}$ independent $\operatorname{SCl}(\lambda, \mu)$ variables are all smaller than $(1 - \varepsilon')d(n)$ is at most $(1 - n^{-1+\frac{\varepsilon'}{2}})^{\frac{n}{2}}$. As the number of ways to choose $\frac{n}{2}$ consecutive indexes out of $\lceil n \log n \rceil$ indexes, i.e., the number of possible i's in the definition of $\mathcal{B}$, is at most $\lceil n \log n \rceil$, we obtain

$$\Pr\left[\overline{\mathcal{B}}\right] \leq \lceil n \log n \rceil \cdot (1 - n^{-1+\frac{\varepsilon'}{2}})^{\frac{n}{2}} \leq \lceil n \log n \rceil \cdot e^{-\frac{1}{2} n^{\frac{\varepsilon'}{2}}} = o(1). \tag{4.28}$$

Now we consider the probability of the event $\overline{\mathcal{C}}$. By Lemma 4.24, the probability that a $\operatorname{SCl}(\lambda, \mu)$ variable is larger than $(1 + \varepsilon')d(n)$, is at most $n^{-1-\frac{\varepsilon'}{2}}$. Hence, we obtain

$$\Pr\left[\overline{\mathcal{C}}\right] \leq \lceil n \log n \rceil \cdot n^{-1-\frac{\varepsilon'}{2}} = o(1). \tag{4.29}$$

Finally, we will estimate $\Pr\left[\overline{\mathcal{D}}\right]$. Due to Lemma 4.25, the probability that a $\operatorname{Cl}(\mu)$ random variable is larger than $2\frac{\log n}{\log \log n}$, is for sufficiently large n at most $n^{-2+o(1)}$. As the number of clique-distributed random variables is distributed as the sum of $\lceil n \log n \rceil$ independent $\operatorname{Po}(\lambda)$ variables, we obtain

$$\Pr\left[\overline{\mathcal{D}}\right] \leq \sum_{t \geq 0} \Pr\left[\operatorname{Po}\left(\lceil n \log n \rceil \lambda\right) = t\right] \cdot t \cdot n^{-2+o(1)} \leq \lambda \lceil n \log n \rceil \cdot n^{-2+o(1)} = o(1),$$

which proves with (4.26), (4.27), (4.28), and (4.29) the claim (4.24).

To complete the proof, we now show how (4.24) implies the theorem. In fact, let B_n be a random block graph on n nodes. Let $\delta > 0$ be arbitrarily small, and choose a sufficiently large n such that

$$\Pr\left[\widetilde{\Gamma}\nu B(n) \in \mathcal{B}'\right] > 1 - \delta \quad \text{and} \quad \sum_{\frac{n}{2} < s < \frac{3n}{2}} \Pr\left[|\widetilde{\Gamma}\nu B(n)| = s\right] > 1 - \delta.$$

Such an n always exists, due to (4.24) and (4.27). But then we have

$$\Pr\left[\widetilde{\Gamma}\nu B(n) \in \mathcal{B}'\right] < \delta + \sum_{\frac{n}{2} < s < \frac{3n}{2}} \Pr\left[\widetilde{\Gamma}\nu B(n) \in \mathcal{B}' \;\middle|\; |\widetilde{\Gamma}\nu B(n)| = s\right] \Pr\left[|\widetilde{\Gamma}\nu B(n)| = s\right],$$

which yields with our assumptions on δ

$$1-2\delta \;<\; \sum_{\frac{n}{2}<s<\frac{3n}{2}} \Pr\left[\mathsf{B}_s \in \mathcal{B}'\right] \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right]. \tag{4.30}$$

Before we proceed, we are going to show that for all $s = \alpha n$, where $\alpha \in (\frac{1}{2}, \frac{3}{2})$, the probabilities $\Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right]$ differ in at most a multiplicative factor. In fact, denote as in Lemma 4.22 with $p_{s,t}$ the probability that $\Gamma\nu B$ returns an object of size s, and having generated precisely $t+s$ nodes during its execution. Then, for all s as above

$$\Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right] = \sum_{t\le\lceil n\log n\rceil} p_{s,t} \overset{\text{(Lemma 4.22)}}{=} \alpha^{-\frac{3}{2}} \cdot \sum_{t\le\lceil n\log n\rceil} p_{n,t},$$

and the claim follows, as the last sum is independent of s.

But this yields with (4.30) that, say, for more than $(1-30\sqrt{\delta})n$ numbers $s_1, \ldots$ in the interval $(\frac{n}{2}, \frac{3n}{2})$, we have $\Pr\left[\mathsf{B}_{s_i} \in \mathcal{B}'\right] \ge 1-\sqrt{\delta}$, as otherwise the sum on the right-hand side of (4.30) would have been smaller than $1-2\delta$. Assume the contrary, i.e. let $i_1, \ldots, i_x$, $x \ge 30\sqrt{\delta}n$ with $\Pr\left[\mathsf{B}_{i_j} \in \mathcal{B}'\right] < 1-\sqrt{\delta}$ for $j = 1, \ldots, x$. Let $b := \sum_{j=1}^{x} \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = i_j\right]$. By (4.30)

$$\begin{aligned} 1-2\delta &< \sum_{\frac{n}{2}<s<\frac{3n}{2}} \Pr\left[\mathsf{B}_s \in \mathcal{B}'\right] \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right] \\ &< b(1-\sqrt{\delta}) + (1-b) \\ &< 1 - b\sqrt{\delta}, \end{aligned}$$

and thus $b < 2\sqrt{\delta}$.

Let $S := \sum_{t\le\lceil n\log n\rceil} p_{n,t}$. Then we have by the above observation

$$\left(\frac{3}{2}\right)^{-\frac{3}{2}} S \;\le\; \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right] \;\le\; \left(\frac{1}{2}\right)^{-\frac{3}{2}} S. \tag{4.31}$$

Moreover, since we know that $\sum_{\frac{n}{2}<s<\frac{3n}{2}} \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right] > 1-\delta$, we have

$$1-\delta < \sum_{\frac{n}{2}<s<\frac{3n}{2}} \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = s\right] \sim S \sum_{\frac{n}{2}<s<\frac{3n}{2}} \left(\frac{s}{n}\right)^{-\frac{3}{2}} < S\,n \left(\frac{1}{2}\right)^{-\frac{3}{2}},$$

and thus $S > \frac{1-\delta}{n2^{\frac{3}{2}}}$.

Summing up, we get

$$\left(\frac{3}{2}\right)^{-\frac{3}{2}} Sx \le \sum_{j=1}^{x} \Pr\left[\left|\widetilde{\Gamma}\nu B(n)\right| = i_j\right] \le 2\sqrt{\delta}$$

and hence $x \le \left(\frac{3}{2}\right)^{\frac{3}{2}} \frac{2\sqrt{\delta}n2^{\frac{3}{2}}}{1-\delta} < 30\sqrt{\delta}$ yielding the desired contradiction.

Hence, for every $\delta > 0$, in every interval of the form $(\frac{n}{2}, \frac{3n}{2})$, when n is sufficiently large, we have at most $30\sqrt{\delta}n$ numbers with $\Pr\left[\mathsf{B}_{s_i} \in \mathcal{B}'\right] \le 1-\sqrt{\delta}$. This completes the proof of the theorem. □

Let us briefly sketch why the result also holds for random city graphs. Let $d(n) = \frac{\log n}{\log^{(3)} n}$ and $\mathcal{I}_{\varepsilon,n}$ be defined as in Theorem 4.28 above. By Theorem 6.4 of [**MSW06**] we know that for any $\varepsilon > 0$ at least $1-\varepsilon$ of all city graphs have a giant component which contains all but at

most k nodes, where $k = k(\varepsilon)$. From the proof of Theorem 4.28 it follows that for all large enough n the probability that a random block graph has no node of degree in $\mathcal{I}_{\varepsilon,n}$ is at most $\frac{1}{\log^2 n}$. Now, a counting argument yields that we have at most

$$(1 - o(1))k\binom{n}{k} b_{n-k} \frac{1}{\log^2 n} 2^{k^2} \ll y_n$$

city graphs on n nodes with degree not in $\mathcal{I}_{\varepsilon,n}$, which is $o(y_n)$. Hence the theorem also holds for random city graphs, which proves the first part of Theorem 4.3. We can proceed similarly for the remaining parts of Theorem 4.2 and 4.3 – see the following proofs.

Now, we give the proofs for the other properties stated in Theorems 4.2 and 4.3. As they follow exactly the pattern of the proof of Theorem 4.28, we keep them short and highlight only the important differences. First, observe that we can easily adapt the sampler $\widetilde{\Gamma}\nu B$ defined in the proof of Theorem 4.28 for the case of cactus graphs: instead of $\Gamma\nu B$ we call the singular rejection sampler $\Gamma\nu K$ as a subroutine, which can be compiled systematically from the relations of the generating functions in a similar way as for block graphs, see Section 4.6. The sampler $\widetilde{\Gamma}\nu K$ draws $\lceil n \log n \rceil$ Poisson distributed random variables $p_1, \ldots, p_{\lceil n \log n \rceil}$, but with a different parameter $\lambda := \frac{\partial}{\partial x} C(x)\big|_{x=K^\bullet(\rho_k)}$. Moreover, we draw the $(C_1^{(i)}, \ldots, C_{p_i}^{(i)})$ independent random values according to the cycle distribution $\mathrm{Cy}(\mu)$, where $\mu := K^\bullet(\rho_k)$. The remainder of the proof is the same, except for the following: we have to adapt the definition of the d_i's (marked with $(\star\star)$ in the exposition of $\widetilde{\Gamma}\nu B$ on page 66) in case of the maximum node degree in a random cactus graph, and we have to redefine the events $\mathcal{B}, C, D$ accordingly. We always keep event $\mathcal{A}$. We will give these new definitions and changes in the following proofs.

Proof of Theorem 4.2, part (1). From the discussion before, it follows that we can follow the scheme of proof of Theorem 4.28. We define the sampler $\widetilde{\Gamma}\nu K$ by adapting the sampler $\widetilde{\Gamma}\nu B$ as described before. Let $d(n) := \frac{\log n}{\log \log n}$, let $\varepsilon > 0$ and define $\mathcal{I}_{\varepsilon,n} := ((2-\varepsilon)d(n), (2+\varepsilon)d(n))$.

Observe that the maximum node degree in a random cactus graph is determined similar as in a random block graph. First the number of components for a node v is drawn according to a Poisson distribution with an appropriate parameter λ (see discussion before). Then for each component a $\mathrm{Cy}(\mu)$ distributed random variable is drawn to determine the size of the corresponding cycle. Note, that every adjacent edge contributes 1 and every adjacent cycle 2 to the degree of node v. The following function will be convenient:

$$f(\mathrm{Cy}(\mu)) = \begin{cases} 1 & \text{if } \mathrm{Cy}(\mu) = 2 \\ 2 & \text{if } \mathrm{Cy}(\mu) \geq 3 \end{cases}$$

Now, define $d_i :\leftarrow \sum_{1 \leq j \leq p_i} f(C_j^{(i)})$ for $i = 1, \ldots, \lceil n \log n \rceil$. Hence, d_i is a lower bound on the node degree of node i and $d_i + 2$ is an upper bound as we assumed that node i is part of a cycle and thus has already degree 2. Recall that the d_i's are drawn independently.

Obverse that we can now follow the argumentation of the proof of Theorem 4.28. What is left, is to define new events $\mathcal{B}, C$ – it follows from the above discussion that we don't need an equivalent of event $\mathcal{D}$ – and show that they occur w.h.p. Let $\varepsilon' := \frac{\varepsilon}{2}$. We define the events as follows:

($\mathcal{B}$) Every sequence of $\frac{n}{2}$ consecutive values d_i contains a value larger than $(2-\varepsilon')d(n)$, i.e.,

$$\mathcal{B} := \left\{(d_1,\dots,d_{\lceil n\log n\rceil}) \;\middle|\; \forall 1 \le i \le \lceil n\log n\rceil - \frac{n}{2} : \exists i \le j \le i+\frac{n}{2} : d_j \ge (2-\varepsilon')d(n)\right\}.$$

($\mathcal{C}$) There is no $1 \le i \le \lceil n\log n\rceil$ such that $d_i \ge (2+\varepsilon')d(n)$.

First, we consider the probability of the event $\overline{\mathcal{B}}$. Let $\alpha := \Pr[\mathrm{Cy}(\mu)=2]$ and let $I_\mu := f(\mathrm{Cy}(\mu))$. We have by Lemma 4.23

$$\Pr\left[\sum_{i=1}^{\mathrm{Po}(\lambda)} I_\mu \ge (2-\varepsilon')d(n)\right] \ge \Pr\left[\mathrm{Po}(\lambda) = (1-\frac{\varepsilon'}{2})d(n)\right]\cdot\Pr\left[\sum_{i=1}^{(1-\frac{\varepsilon'}{2})d(n)} I_\mu \ge (2-\varepsilon')d(n)\right]$$
$$\ge n^{-1+\frac{\varepsilon'}{2}}\cdot(1-\alpha)^{(1-\frac{\varepsilon'}{2})d(n)} = n^{-1+\frac{\varepsilon'}{2}+o(1)}.$$

With this and as the number of ways to choose $\frac{n}{2}$ consecutive indexes out of $\lceil n\log n\rceil$ indexes, i.e., the number of possible i's in the definition of $\mathcal{B}$, is at most $\lceil n\log n\rceil$, we obtain

$$\Pr[\overline{\mathcal{B}}] \le \lceil n\log n\rceil\cdot(1-n^{-1+\frac{\varepsilon'}{2}})^{\frac{n}{2}} \le \lceil n\log n\rceil\cdot e^{-\frac{1}{2}n^{\frac{\varepsilon'}{2}}} = o(1). \tag{4.32}$$

Now we consider the probability of the event $\overline{\mathcal{C}}$. Observe that a sum of j $f(\mathrm{Cy}(\mu))$ distributed random variables can only be larger than some value t, if we sum over at least $\frac{t}{2}$ such values, as $f(\mathrm{Cy}(\mu)) \le 2$. Thus we have

$$\Pr\left[\sum_{i=1}^{\mathrm{Po}(\lambda)} f(\mathrm{Cy}(\mu)) \ge (2+\varepsilon')d(n)\right] \le \sum_{j\ge 0}\Pr[\mathrm{Po}(\lambda)=j]\Pr\left[\sum_{i=1}^{j} f(\mathrm{Cy}(\mu)) \ge (2+\varepsilon')d(n)\right]$$
$$\le \sum_{j\ge(1+\frac{\varepsilon'}{2})d(n)} \Pr[\mathrm{Po}(\lambda)=j] \le n^{-1-\frac{\varepsilon'}{2}},$$

where the last step follows by Lemma 4.23. Hence, we obtain

$$\Pr[\overline{\mathcal{C}}] \le \lceil n\log n\rceil\cdot n^{-1-\frac{\varepsilon'}{2}} = o(1). \tag{4.33}$$

This completes the proof. □

Proof of Theorem 4.2, part (ii). Again, we can follow the pattern of proof of Theorem 4.28. We define the sampler $\widetilde{\Gamma}\nu K$ by adapting the sampler $\widetilde{\Gamma}\nu B$ as described before. Let $c(n) := \log_{\frac{1}{\lambda}} n$, let $\varepsilon > 0$ and define $\mathcal{I}_{\varepsilon,n} := ((1-\varepsilon)c(n),(1+\varepsilon)c(n))$. Note that the maximum size of a cycle in a random cactus graph is determined by the random variables $C_j^{(i)}$ which are independently $\mathrm{Cy}(\mu)$ distributed.

Again, we can follow the argumentation of the proof of Theorem 4.28. What is left, is to define new events $\mathcal{B}$ and $\mathcal{C}$ and show that they occur w.h.p. Let $\varepsilon' := \frac{\varepsilon}{2}$. We define the events as follows:

($\mathcal{B}$) Every sequence of $\frac{n}{2}$ consecutive values $C_j^{(i)}$ contains a value larger than $(1-\varepsilon')c(n)$, i.e.,

$$\mathcal{B} := \left\{(C_1^{(1)},\dots,C_{p_1}^{(1)},\dots,C_1^{(\lceil n\log n\rceil)},\dots,C_{p_{\lceil n\log n\rceil}}^{(\lceil n\log n\rceil)}) \;\middle|\; \right.$$
$$\left. \forall 1 \le l \le \lceil n\log n\rceil - \frac{n}{2} : \exists l \le i \le l+\frac{n}{2} : \exists 1 \le j \le p_i : C_j^{(i)} \ge (1-\varepsilon')c(n)\right\}.$$

($\mathcal{C}$) There is no $1 \le i \le \lceil n\log n\rceil, 1 \le j \le p_i$ such that $C_j^{(i)} \ge (1+\varepsilon')c(n)$.

First, we consider the probability of the event $\overline{\mathcal{B}}$. Due to Lemma 4.27 we have

$$\Pr\left[\sum_{i=1}^{\mathrm{Po}(\lambda)} \mathrm{Cy}(\mu) \geq (1-\varepsilon')c(n)\right] \geq \Pr[\mathrm{Po}(\lambda)=1] \cdot \Pr\left[\mathrm{Cy}(\mu) \geq (1-\varepsilon')c(n)\right] \geq n^{-1+\varepsilon'+o(1)}.$$

With this and as the number of ways to choose $\frac{n}{2}$ consecutive indexes out of $\lceil n\log n\rceil$ indexes, i.e., the number of possible l's in the definition of $\mathcal{B}$, is at most $\lceil n\log n\rceil$, we obtain

$$\Pr[\overline{\mathcal{B}}] \leq \lceil n\log n\rceil \cdot (1-n^{-1+\varepsilon'+o(1)})^{\frac{n}{2}} \leq \lceil n\log n\rceil \cdot e^{-\frac{1}{2}n^{\varepsilon'+o(1)}} = o(1). \tag{4.34}$$

Now we consider the probability of the event $\overline{\mathcal{C}}$. Due to Lemma 4.27, the probability that a $\mathrm{Cy}(\mu)$ random variable is larger than $(1+\varepsilon')c(n)$, is for sufficiently large n at most $n^{-1-\varepsilon'+o(1)}$. As the number of cycle-distributed random variables is distributed as the sum of $\lceil n\log n\rceil$ independent $\mathrm{Po}(\lambda)$ variables, we obtain

$$\Pr[\overline{\mathcal{C}}] \leq \sum_{t\geq 0} \Pr[\mathrm{Po}(\lceil n\log n\rceil\lambda)=t]\cdot t\cdot n^{-1+\varepsilon'+o(1)} \leq \lambda\lceil n\log n\rceil\cdot n^{-1-\varepsilon'+o(1)} = o(1).$$

□

Proof of Theorem 4.3, part (ii). Again, we can follow the pattern of proof of Theorem 4.28. We can use the same sampling procedure $\widetilde{\Gamma}\nu B$. Let $c(n) := \frac{\log n}{\log\log n}$, let $\varepsilon > 0$ and define $\mathcal{I}_{\varepsilon,n} := ((1-\varepsilon)c(n),(1+\varepsilon)c(n))$. Note that the maximum size of a clique in a random block graph is determined by the random variables $C_j^{(i)}$ which are independently $\mathrm{Cl}(\mu)$ distributed.

We can follow the argumentation of the proof of Theorem 4.28, we only have to define new events $\mathcal{B}$ and $\mathcal{C}$ and show that they occur w.h.p. Let $\varepsilon' := \frac{\varepsilon}{2}$. We define the events as follows:

($\mathcal{B}$) Every sequence of $\frac{n}{2}$ consecutive values $C_j^{(i)}$ contains a value larger than $(1-\varepsilon')c(n)$, i.e.,

$$\mathcal{B} := \Big\{(C_1^{(1)},\ldots,C_{p_1}^{(1)},\ldots,C_1^{(\lceil n\log n\rceil)},\ldots,C_{p_{\lceil n\log n\rceil}}^{(\lceil n\log n\rceil)}) \;\Big|\; \forall 1\leq l\leq \lceil n\log n\rceil - \frac{n}{2} : \exists l\leq i\leq l+\frac{n}{2} : \exists 1\leq j\leq p_i : C_j^{(i)} \geq (1-\varepsilon')c(n)\Big\}.$$

($\mathcal{C}$) There is no $1\leq i\leq\lceil n\log n\rceil, 1\leq j\leq p_i$ such that $C_j^{(i)} \geq (1+\varepsilon')c(n)$.

First, we consider the probability of the event $\overline{\mathcal{B}}$. Due to Lemma 4.25 we have

$$\Pr\left[\sum_{i=1}^{\mathrm{Po}(\lambda)} \mathrm{Cl}(\mu) \geq (1-\varepsilon')c(n)\right] \geq \Pr[\mathrm{Po}(\lambda)=1] \cdot \Pr\left[\mathrm{Cl}(\mu) \geq (1-\varepsilon')c(n)\right] \geq n^{-1+\varepsilon'+o(1)}.$$

With this and as the number of ways to choose $\frac{n}{2}$ consecutive indexes out of $\lceil n\log n\rceil$ indexes, i.e., the number of possible l's in the definition of $\mathcal{B}$, is at most $\lceil n\log n\rceil$, we obtain

$$\Pr[\overline{\mathcal{B}}] \leq \lceil n\log n\rceil \cdot (1-n^{-1+\varepsilon'+o(1)})^{\frac{n}{2}} \leq \lceil n\log n\rceil \cdot e^{-\frac{1}{2}n^{\varepsilon'+o(1)}} = o(1). \tag{4.35}$$

Now we consider the probability of the event $\overline{\mathcal{C}}$. Due to Lemma 4.25, the probability that a $\mathrm{Cl}(\mu)$ random variable is larger than $(1+\varepsilon')c(n)$, is for sufficiently large n at most $n^{-1-\varepsilon'+o(1)}$.

As the number of clique-distributed random variables is distributed as the sum of $\lceil n \log n \rceil$ independent $\mathrm{Po}(\lambda)$ variables, we obtain

$$\Pr\left[\overline{\mathcal{C}}\right] \;\leq\; \sum_{t\geq 0} \Pr\left[\mathrm{Po}\left(\lceil n\log n\rceil \lambda\right) = t\right] \cdot t \cdot n^{-1+\varepsilon'+o(1)} \;\leq\; \lambda\lceil n\log n\rceil \cdot n^{-1-\varepsilon'+o(1)} \;=\; o(1).$$

□

In the next proofs we apply (more classical) generating function techniques to obtain the remaining structural results about random block and random cactus graphs as stated in Section 4.2.

Proof of Theorem 4.7. As a tree is bipartite, we can colour it using only 2 colours. Moreover we can colour any even cycle with 2 colours. Thus, we only need 3 colours, if a cactus graph contains an odd cycle as a biconnected component. In the same way as in Section 4.5 we can derive a singular expansion for the class $\tilde{\mathcal{K}}$ of cactus graphs which have only even cycles as biconnected components and hence are 2-colourable. Again, we can obtain asymptotic estimates for the coefficients $\tilde{k}_n$ of the generating function – i.e. the number of cactus graphs on n nodes which contain only even cycles – from the singular expansion by Corollary 2.5 (Corollary VI.1 of [**FS06**]). Then the probability for a random cactus graph for being 2-colourable is $\frac{\tilde{k}_n}{k_n}$. In the sequel we show that

$$\tilde{k}_n \sim \tilde{\alpha}_k \, n^{-\frac{5}{2}} \, \tilde{\rho}_k{}^{-n},$$

where $\tilde{\alpha}_k \doteq 0.1360$ and $\tilde{\rho}_k{}^{-1} \doteq 3.5505$. This concludes with Theorem 4.5 the proof.

Let $\tilde{C}(x)$ denote the exponential generating function for even cycles of length at least 3 and a single edge. With $C(x)$ the exponential generating function of all cycles of length at least 3 and a single edge as defined in Section 4.5, we have

$$\tilde{C}(x) = \frac{1}{2}\left(C(x) + C(-x)\right),$$

as $C(-x)$ cancels every odd term and the factor $\frac{1}{2}$ avoids that we count every even term twice.

Let $\tilde{K}(x)$ denote the exponential generating function for cactus graphs that contain only even cycles. Now set $y = 1$; again from the well-known decomposition of a graph into 2-connected components we obtain the relations

$$\tilde{K}^{\bullet}(x) = x \exp\left(\frac{\partial}{\partial x}\tilde{C}\left(\tilde{K}^{\bullet}(x)\right)\right) \text{ and } \tilde{C}(x) = \frac{1}{4}\log\left(\frac{1}{1-x}\right) + \frac{1}{4}\log\left(\frac{1}{1+x}\right) + \frac{1}{4}x^2. \quad (4.36)$$

Notice that the singularities of $\frac{\partial}{\partial x}\tilde{C}(x)$ and $\tilde{K}^{\bullet}(x)$ are the same as those of $\tilde{C}(x)$ and $\tilde{K}(x)$. From (4.36) it follows that the functional inverse $\tilde{\psi}(u)$ of $\tilde{K}^{\bullet}(x)$ is

$$\tilde{\psi}(u) = u \exp\left(-\frac{\partial}{\partial x}\tilde{C}(u)\right), \quad (4.37)$$

as $\tilde{\psi}(\tilde{K}^{\bullet}(x)) = x$. This time $\tilde{C}(x)$ has two singularities of modulus 1 at $R_1 = 1$ and $R_2 = -1$.

We use again the functional inverse of $\tilde{K}^{\bullet}(x)$ to obtain a singular expansion for $\tilde{K}^{\bullet}(x)$ and then by integration a singular expansion for $\tilde{K}(x)$. As before, we then apply Theorem 2.6

(Theorem VI.6 of [**FS06**]). One can check that solving $\frac{\partial}{\partial u}\tilde{\psi}(u) = 0$ has a real zero at $\tau \in (0, R_1)$, where τ is the solution to the equation

$$-2 + 4x^2 - 2x^4 + 2x - x^3 + x^5 = 0,$$

and $\tau \doteq 0.5476$.

Thus the singularity of $\tilde{K}^{\bullet}(x)$ is at $\tilde{\rho}_k = \tilde{\psi}(\tau)$. From the singular expansion we obtain the following asymptotic estimate for the number of cactus graphs with even cycles only

$$\tilde{c}_n \sim \tilde{\alpha}_k \, n^{-\frac{5}{2}} \, \tilde{\rho}_k{}^{-n},$$

where $\tilde{\alpha}_k \doteq 0.1360$ and $\tilde{\rho}_k{}^{-1} = \tilde{\psi}(\tau)^{-1} \doteq 3.5505$. This completes the proof. □

Proof of Theorem 4.8. The proof is the same as in [**GN05a**] except for the difference in the constants due to the different underlying graph class. For fixed k the generating function of desert graphs with exactly k connected components (i.e. cactus graphs) is $K(x)^k/k!$. Moreover, for fixed k we have

$$[x^n]K(x)^k \sim kK_0^{k-1}[x^n]K(x),$$

as kK_0^{k-1} is the leading term in the singular expansion of $K(x)$. Thus the probability that a random desert graph has exactly k components is asymptotically

$$\frac{[x^n]K(x)/k!}{[x^n]D(x)} \sim \frac{kK_0^{k-1}}{k!}e^{-K_0} = \frac{K_0^{k-1}}{(k-1)!}e^{-K_0},$$

which implies the theorem. □

We may prove the corresponding result for city and block graphs, i.e. Theorem 4.13, in an analogous way.

Proof of Theorem 4.10. By Theorem 4.5 we have as $n \to \infty$

$$\Pr[\mathrm{D}_n \text{ is connected}] \sim \frac{\alpha_k}{\alpha_d} = e^{-K_0} \doteq 0.7488.$$

Also

$$\Pr[\mathrm{D}_n \text{ is biconnected}] \sim \frac{\alpha_c \, n^{\frac{3}{2}} \, \rho_k^n}{\alpha_d} \to 0,$$

as $\rho_k < 1$ and $n \to \infty$. Finally,

$$\Pr[\mathrm{K}_n \text{ is biconnected}] \sim \frac{\alpha_c \, n^{\frac{3}{2}} \, \rho_k^n}{\alpha_k} \to 0,$$

as $\rho_k < 1$ and $n \to \infty$. □

We can prove Theorem 4.15 in the same way, using the estimates of Theorem 4.6.

Proof of Theorem 4.11. To prove the theorem, it suffices to derive a singular expansion for the bivariate generating function of desert graphs, where the second parameter marks the number of edges. Then we can apply Theorem 2.7 (Theorem IX.10 of [**FS06**]) to obtain the result.

Let $D(x, y)$ denote the bivariate generating function for desert graphs. In Lemma 4.19 it was shown that $D(x, y)$ has the singular expansion

$$D(x, y) = D_0(y) + D_1(y)X + D_2(y)X^{\frac{3}{2}} + O(X^2),$$

where $X = \left(1 - \frac{x}{\rho(y)}\right)$ and the $D_i(y)$ and $\rho(y)$ are given explicitly. Now, we apply Theorem 2.7; for this, we set the bivariate functions A, B, C as follows:

$$A(x, y) = D_0(y) + D_1(y)X + O(X^2), \quad B(x, y) = K_2(y) \quad \text{and} \quad C(x, y) = \left(1 - \frac{x}{\rho(y)}\right),$$

and $\alpha = -\frac{3}{2}$. With these definitions, it is easily checked that the preconditions of Theorem 2.7 (Theorem IX.10 in [**FS06**]) are fulfilled. It follows that the number of edges in a random desert graph with n nodes is normally distributed with expectation and variance given by the expressions

$$\mu_n = -\frac{\frac{\partial}{\partial u}\rho(u)\Big|_{u=1}}{\rho(1)} \cdot n \quad \text{and} \quad \sigma_n^2 = \left(-\frac{\frac{\partial^2}{\partial u^2}\rho(u)\Big|_{u=1}}{\rho(1)} - \frac{\mu_n}{n} + \left(\frac{\mu_n}{n}\right)^2\right) \cdot n.$$

A similar calculation proves the same result for random cactus graphs; this proves the theorem. □

4.8. Conclusions

In this chapter we introduced a new method, that is based on the analysis of Boltzmann samplers and may be used to obtain precise estimates for the maximum degree and maximum size of a biconnected block of a "typical" member of a constrained graph class. We demonstrated the method on two moderately complex graph classes, namely cactus and block graphs. It would be interesting to apply the method to more complex graph classes, for instance outerplanar graphs, planar graphs or triangulations, as well as to study different parameters, especially the degree sequence.

CHAPTER 5

$K_{3,3}$-minor-free Graphs

In this chapter we are interested in the number of simple labelled $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs, where maximal means that adding any edge to such a graph yields a $K_{3,3}$-minor. We use analytic combinatorics (see also Section 2.2) to obtain asymptotic estimates for the number of $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs. We also derive limit laws for some parameters of a random $K_{3,3}$-minor-free graph.

5.1. Previous and Related Work

It is known that there exists a constant c, such that there are at most $c^n n!$ $K_{3,3}$-minor-free graphs on n nodes. This follows from a result of Norine, Seymour, Thomas, and Wollan [**NSTW06**] which states that every proper subclass of all graphs which is closed under isomorphism and taking minors has at most $c^n n!$ graphs on n nodes. Obviously, this gives also an upper bound on the number of maximal $K_{3,3}$-minor-free graphs as they are a proper subclass of $K_{3,3}$-minor-free graphs.

In [**MSW06**], McDiarmid, Steger and Welsh give conditions in which cases an upper bound of the form $c^n n!$ on the number of graphs $|\mathcal{C}_n|$ on n nodes in graph class $\mathcal{C}$ yields that $(|\mathcal{C}_n|/n!)^{\frac{1}{n}} \to c > 0$ as $n \to \infty$. These conditions are satisfied for $K_{3,3}$-minor-free graphs, but not in the case of maximal $K_{3,3}$-minor-free graphs as one condition requires that two disjoint copies of a graph of the class in question form again a graph of the class.

Thus we know that there exists a growth constant c for $K_{3,3}$-minor-free graphs, but not its exact value. For maximal $K_{3,3}$-minor-free graphs we only have an upper bound. Lower bounds on the number of graphs in our classes can be obtained by considering (maximal) planar graphs. Due to Kuratowski's theorem [**Kur30**] planar graphs are $K_{3,3}$- *and* K_5-minor-free. Hence, the class of planar graphs is contained in the class of $K_{3,3}$-minor-free graphs. The class of triangulations is contained in the class of maximal $K_{3,3}$-minor-free graphs due to Wagner's theorem [**Wag37**] (with the exception of all triangulations on 5 nodes). Hence, we can use the number of planar graphs and the number of triangulations as lower bounds. Determining the number (of graphs of sub-classes) of planar graphs attracted considerable attention [**BGW02, GN04, GN05b, BGKN05, BLMK**] in recent years. Giménez and Noy [**GN05b**] obtained precise asymptotic estimates for the number of planar graphs. The asymptotic number of triangulations was given by Tutte [**Tut62**] already in 1962. Investigating how much the number of planar graphs (triangulations) differs from (maximal) $K_{3,3}$-minor-free graphs was also the main motivation for our research. Here we will derive asymptotic estimates for the number of (maximal) $K_{3,3}$-minor-free graphs. For $K_{3,3}$-minor-free graphs we consider 3-connected, 2-connected, connected and not necessarily connected graphs. Investigating how much the number of planar graphs (triangulations)

differs from (maximal) $K_{3,3}$-minor-free graphs is a first important step in examining how "typical" instances of these graph classes will differ.

For 3-connected $K_{3,3}$-minor-free graphs, the change from planar graphs can be easily described: it follows from a theorem of Wagner [**Wag37**] that the set of 3-connected $K_{3,3}$-minor-free graphs consists of all 3-connected planar graphs and the complete graph on 5 nodes. Thus, on this connectivity level our graph class differs only in the existence of one additional graph from planar graphs. But as we will show, adding K_5 to the set of 3-connected graphs yields a slightly larger exponential growth rate for 2-connected, connected, and not necessarily connected $K_{3,3}$-minor-free graphs. It also slightly changes other parameters, for instance the expected number of edges in a random $K_{3,3}$-minor-free graph. For maximal $K_{3,3}$-minor-free graphs the growth rate also increases slightly compared to triangulations.

5.2. Results

We derive precise asymptotic estimates for the number of simple labelled $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs on n nodes, and we establish several limit laws for parameters in random $K_{3,3}$-minor-free graphs. More precisely, we show that the number g_n, c_n, and b_n of not necessarily connected, connected and 2-connected $K_{3,3}$-minor-free graphs on n nodes and the number m_n of maximal $K_{3,3}$-minor-free graphs on n nodes satisfy

$$\begin{aligned} g_n &\sim \alpha_g \, n^{-\frac{7}{2}} \, \rho_g^{-n} \, n!, \\ c_n &\sim \alpha_c \, n^{-\frac{7}{2}} \, \rho_c^{-n} \, n!, \\ b_n &\sim \alpha_b \, n^{-\frac{7}{2}} \, \rho_b^{-n} \, n!, \\ m_n &\sim \alpha_m \, n^{-\frac{7}{2}} \, \rho_m^{-n} \, n! \end{aligned}$$

where $\alpha_g \doteq 0.42643 \cdot 10^{-5}$, $\alpha_c \doteq 0.41076 \cdot 10^{-5}$, $\alpha_b \doteq 0.37074 \cdot 10^{-5}$, $\alpha_m \doteq 0.40553 \cdot 10^{-4}$, $\rho_c^{-1} = \rho_g^{-1} \doteq 27.22935$, $\rho_b^{-1} \doteq 26.18659$, and $\rho_m^{-1} \doteq 9.496119$ are analytically computable constants. Moreover, we derive limit laws for $K_{3,3}$-minor-free graphs, for instance we show that the number of edges is asymptotically normally distributed with mean κn and variance λn, where $\kappa \doteq 2.21338$ and $\lambda \doteq 0.43044$ are analytically computable constants. Thus the expected number of edges is only slightly larger than for planar graphs [**GN05b**].

To establish these results for $K_{3,3}$-minor-free graphs, we follow the approach taken for planar graphs [**BGW02, GN05b**]: we use a well-known decomposition along the connectivity structure of a graph, i.e. into connected, 2-connected and 3-connected components, and translate this decomposition into relations of our generating functions. This is possible as the decomposition for $K_{3,3}$-minor-free graphs which is due to Wagner [**Wag37**] fits well into this framework. Then we use singularity analysis to obtain asymptotic estimates and limit laws for several parameters from these equations.

For maximal $K_{3,3}$-minor-free graphs the situation is different, as the decomposition which is again due to Wagner has further constraints (it restricts which edges can be used to merge two graphs into a new one). The functional equations for the generating functions of *edge-rooted* maximal graphs are easy to obtain but in order to go to unrooted graphs, special integration techniques based on rational parametrisation of rational curves are needed. In this way we

can derive equations for the generating functions which involve the generating function for triangulations derived by Tutte [**Tut62**].

Overview and Notation

In the subsequent sections, we proceed as follows. We turn to maximal $K_{3,3}$-minor-free and $K_{3,3}$-minor-free graphs in Sections 5.3 and 5.4 respectively. In each of these sections, we will first derive relations for the generating functions based on a decomposition of the considered graph class and then apply singularity analysis to obtain asymptotic estimates for the number (and properties) of the graphs in these classes.

Throughout this chapter variable x marks nodes and variable y marks edges. Unless we specify the contrary, the graphs we consider are labelled and the corresponding generating functions are exponential. We often need to distinguish an atom of our combinatorial objects; for instance we want to mark a node in a graph as a root node. On generating function level this means taking the derivative with respect to the corresponding variable and multiplying the result by this variable. To simplify the formulas, we use the following notation. Let $G(x,y)$ and $G(x)$ be generating functions, then we abbreviate $G^\bullet(x,y) = x\frac{\partial}{\partial x}G(x,y)$ and $G^\bullet(x) = x\frac{\partial}{\partial x}G(x)$. Additionally, we use the following standard graph notation: let G be a graph, then we denote by $V(G)$ and $E(G)$ the node- and edge-set of graph G.

5.3. Maximal $K_{3,3}$-minor-free graphs

Already in the 1930s, Wagner [**Wag37**] described precisely the structure of maximal $K_{3,3}$-minor-free graphs. Roughly speaking his theorem states that a maximal graph not containing $K_{3,3}$ as a minor is formed by gluing planar triangulations and the exceptional graph K_5 along edges, in such a way that no edge glues two different triangulations. Before we state the theorem more precisely, we introduce the following notation (similar to [**Tho99**], see also Section 5.4.2).

Definition 5.1. Let G_1 and G_2 be graphs with disjoint node-sets, where each edge is either coloured blue or red. Let $e_1 = (a,b) \in E(G_1)$ and $e_2 = (c,d) \in E(G_2)$ be an edge of G_1 and G_2 respectively. For $i = 1,2$ let G_i' be obtained by deleting edge e_1 and colouring edge e_2 blue if e_1 and e_2 were both coloured blue and red otherwise. Let G be the graph obtained from the union of G_1' and G_2' by identifying nodes a and b by c and d respectively. Then we say that G is a *strict 2-sum* of G_1 and G_2. We say that a strict 2-sum is *proper* if edges e_1 and e_2 are neither red.

Theorem 5.2 (Wagner's theorem [**Wag37**]). *Let $\mathcal{T}$ denote the set of all labelled planar triangulations except for the triangulations on 5 nodes. Let each edge of a graph in $\mathcal{T}$ be coloured red. Let each edge of the complete graph K_5 be coloured blue. A graph is maximal $K_{3,3}$-minor-free if and only if it can be obtained from graphs in $\mathcal{T}$ and K_5 by proper, strict 2-sums.*

Let $\mathcal{A}$ be the family of maximal graphs not containing $K_{3,3}$ as a minor. Let $\mathcal{H}$ be the family of edge-rooted graphs in $\mathcal{A}$, where the root belongs to a triangulation, and let $\mathcal{F}$ be edge-rooted graphs in $\mathcal{A}$, where the root does not belong to a triangulation.

Let $T_0(x,y)$ be the generating function of edge-rooted planar triangulations (the root-edge is included) except for the triangulations on 5 nodes, and let $K_0(x,y)$ be the generating function of edge-rooted K_5 (the root-edge is not included). Let $A(x,y)$, $F(x,y)$, and $H(x,y)$ be the generating functions associated respectively to the families $\mathcal{A}$, $\mathcal{F}$, and $\mathcal{H}$. In all cases the two endpoints of the root edge do not bear labels, and the root edge is oriented; this amounts to multiplying the corresponding generating function by $2/x^2$. Notice that

$$K_0 = \frac{2}{x^2}\frac{\partial}{\partial y}\left(y^{10}\frac{x^5}{5!}\right) = y^9\frac{x^3}{6}.$$

Since edge-rooted graphs from $\mathcal{A}$ are the disjoint union of $\mathcal{H}$ and $\mathcal{F}$, we have

$$H(x,y) + F(x,y) = \frac{2}{x^2}y\frac{\partial A(x,y)}{\partial y}. \tag{5.1}$$

The fundamental equations that we need are the following:

$$H = T_0(x,F) \tag{5.2}$$

$$F = y\exp\left(K_0(x,H+F)\right) \tag{5.3}$$

The first equation means that a graph in $\mathcal{H}$ is obtained by substituting every edge in a planar triangulation by an edge-rooted graph whose root does not belong to a triangulation (because of the statement of Wagner's theorem). The second equation means that a graph in $\mathcal{F}$ is obtained by taking (an unordered) set of K_5's in which each edge is substituted by an edge-rooted graph either in $\mathcal{H}$ or in $\mathcal{F}$.

Eliminating H we get the equation

$$F = y\exp\left(K_0(x,F+T_0(x,F))\right). \tag{5.4}$$

Hence, for fixed x,

$$\psi(u) = u\exp\left(-K_0(x,u+T_0(x,u)\right) = u\exp\left(-\frac{x^3}{6}(u+T_0(x,u))^9\right) \tag{5.5}$$

is the functional inverse of $F(x,y)$.

In order to analyse F using Equation (5.3) we need to know the series $T_0(x,y)$ in detail. Let T_n be the number of (labelled) planar triangulations with n nodes. Since a triangulation has exactly $3n-6$ edges, we see that

$$T(x,y) = \sum T_n y^{3n-6}\frac{x^n}{n!}$$

is the generating function of triangulations. And since

$$T_0(x,y) = \frac{2}{x^2}y\frac{\partial T(x,y)}{\partial y} - \frac{3}{2}x^3y^9,$$

where we subtract the last polynomial to exclude the triangulation on 5 nodes, it is enough to study T.

Let now t_n be the number of rooted (unlabelled) triangulations with n nodes in the sense of Tutte and let $t(x) = \sum t_n x^n$ be the corresponding *ordinary* generating function. We know (see [**Tut62**]) that $t(x)$ is equal to

$$t = x^2\theta(1-2\theta)$$

where $\theta(x)$ is the algebraic function defined by

$$\theta(1-\theta)^3 = x.$$

It is known that the dominant singularity of θ is at $R = 27/256$ and $\theta(R) = 1/4$.

There is a direct relation between the numbers T_n and t_n. An unlabelled rooted triangulation can be labelled in $n!$ ways, and a labelled triangulation ($n \geq 4$) can be rooted in $4(3n-6)$ ways, since we have $3n-6$ possibilities for choosing the root edge, two for orienting the edge, and two for choosing the root face. Hence we have

$$t_n n! = 4(3n-6)T_n, \quad n \geq 4, \qquad t_3 = T_3 = 1.$$

The former identity implies easily the following equation connecting the exponential generating function $T(x,y)$ and the ordinary generating function $t(x)$:

$$y\frac{\partial T}{\partial y} = y^3\frac{x^3}{4} + \frac{t(xy^3)}{4y^6}.$$

Hence we have

$$T_0(x,y) = \frac{2}{x^2}y\frac{\partial T}{\partial y} = y^3\frac{x}{2} + \frac{t(xy^3)}{2x^2y^6} - \frac{3}{2}x^3y^9.$$

The last equation is crucial since it expresses T_0 in terms of a known algebraic function. It is convenient to rewrite it as

$$T_0(x,y) = y^3\frac{x}{2} + \frac{1}{2}L(x,y)(1-2L(x,y)) - \frac{3}{2}x^3y^9, \qquad \text{where } L(x,y) = \theta(xy^3). \tag{5.6}$$

The series $L(x,y)$ is defined through the algebraic equation

$$L(1-L)^3 - xy^3 = 0. \tag{5.7}$$

This defines a rational curve, i.e. a curve of genus zero, in the variables L and y (here x is taken as a parameter) and admits the rational (actually polynomial) parametrisation

$$L = \lambda(t) = -\frac{t^3}{x^2}, \qquad y = \xi(t) = -\frac{t^4 + x^2 t}{x^3}. \tag{5.8}$$

This is a key fact that we use later.

We have set up the preliminaries needed in order to analyse the series $A(x,y)$, which is the main goal of this section. From (5.1) it follows that

$$A(x,y) = \frac{x^2}{2}\int_0^y \frac{H(x,t)}{t}\,dt + \frac{x^2}{2}\int_0^y \frac{F(x,t)}{t}\,dt.$$

The following lemma expresses $A(x,y)$ directly in terms of H and F *without* integrals.

LEMMA 5.3. *The generating function $A(x,y)$ of maximal graphs not containing $K_{3,3}$ as a minor can be expressed as*

$$A(x,y) = \frac{-x^2}{60}\left(27(H+F)\log\left(\frac{F}{y}\right) + 10L + 20L^2 + 15\log(1-L) - 30F - 5xF^3 + 5x^3F^9\right), \tag{5.9}$$

where $L = L(x, F(x,y))$, $H = H(x,y)$ and $F = F(x,y)$ are defined through (5.7), (5.2) and (5.3).

PROOF. Integration by parts gives

$$A(x,y) = \frac{x^2}{2}\int_0^y \frac{H(x,t)+F(x,t)}{t}\,dt = \frac{x^2}{2}(H+F)\log(y) - \frac{x^2}{2}I \tag{5.10}$$

where

$$I = \int_0^y \log(t)\left(H'(x,t)+F'(x,t)\right)\,dt$$

and derivatives are with respect to the second variable. Because of (5.5), the change of variable $s = F(x,t)$ gives $t = \psi(s)$ and

$$\log(t) = \log(s) - \frac{x^3}{6}\left(s + T_0(x,s)\right)^9.$$

Since $H = T_0(x,F)$ we have $H' = T_0'(x,F)F'$ and so

$$\begin{aligned} I &= \int_0^F \left(\log(s) - \frac{x^3}{6}(s+T_0(x,s))^9\right)\left(1+T_0'(x,s)\right)\,ds \\ &= -\frac{x^3}{6}\frac{(F+T_0(x,F))^{10}}{10} + \int_0^F \log(s)\left(1+T_0'(x,s)\right)\,ds \\ &= -\frac{1}{10}(H+F)\log\left(\frac{F}{y}\right) + \int_0^F \log(s)\left(1+T_0'(x,s)\right)\,ds, \end{aligned}$$

where the last equality follows from Equation (5.3).

It remains to compute the last integral. From (5.6) it follows easily that

$$T_0' = \frac{3y^2x}{2}\left(1+\frac{1}{(1-L)^2}\right) - \frac{27}{2}x^3y^8. \tag{5.11}$$

Now we change variables according to (5.8) taking $s = \xi(t)$, so that $L = \lambda(t)$. Let ζ be the inverse function of ξ, so that $t = \zeta(s)$. Observe that $\zeta(s)$ satisfies

$$\zeta^4 + x^2\zeta + x^3 s = 0.$$

Then we have

$$\begin{aligned} &\int_0^F \log(s)\left(1+T_0'(x,s)\right)\,ds \qquad (5.12)\\ = &\int_0^{\zeta(F)} \log(\xi(t))\left(1+\frac{3\xi(t)^2x}{2}\left(1+\frac{1}{(1-\lambda(t))^2}\right)\right)\xi'(t)\,dt \\ &-\int_0^{\zeta(F)} \log(\xi(t))\frac{27}{2}x^3\left(\xi(t)\right)^8\xi'(t)\,dt. \end{aligned}$$

The second integral on the right hand side can easily be solved and after substituting the expressions for $\xi(t)$ and $\lambda(t)$, the integrand in the first integral is equal to

$$f(x,t) = -\frac{1}{2x^8}\left(4\,t^3+x^2\right)\left(2\,x^5+3\,t^8+6\,t^5x^2+6\,t^2x^4\right)\ln\left(-\frac{t^4+x^2t}{x^3}\right).$$

The function $f(x,t)$ can be integrated in elementary terms, resulting in

$$\begin{aligned} \int_0^{\zeta(F)} f(x,t)dt = &\left(-\frac{5\zeta^6}{2x^4} - \frac{\zeta^{12}}{2x^8} - \frac{\zeta^3}{x^2} - \frac{\zeta^4}{x^3} - \frac{\zeta}{x} - \frac{3\zeta^9}{2x^6}\right)\log\left(-\frac{\zeta^4+x^2\zeta}{x^3}\right) \\ &+ \frac{7\zeta^6}{6x^4} - \frac{\zeta^3}{6x^2} + \frac{\zeta}{x} + \frac{\zeta^4}{x^3} + \frac{\zeta^9}{2x^6} + \frac{\zeta^{12}}{6x^8} - \frac{1}{2}\log\left(1+\frac{\zeta^3}{x^2}\right), \end{aligned}$$

where $\zeta = \zeta(F)$. All terms with ζ are powers of either of the two forms

$$-\frac{\zeta^4 + x^2\zeta}{x^3} = \xi(\zeta(F)) = F, \qquad -\frac{\zeta^3}{x^2} = \lambda(\zeta(F)) = L(x, F),$$

so we can write the integral of $f(x,t)$ explicitly in terms of F and $L = L(x,F)$, and overall we get for the integral of Equation (5.12)

$$\left(-\frac{1}{2}L^4 + \frac{3}{2}L^3 - \frac{5}{2}L^2 + L + F - \frac{3}{2}x^3F^9\right)\log(F) + \frac{L^4}{6} - \frac{L^3}{2} + \frac{7L^2}{6} + \frac{L}{6} + \frac{\log(1-L)}{2} - F + \frac{x^3F^9}{6}.$$

We simplify this expression further using that, according to Equations (5.2), (5.6) and (5.7),

$$H = T_0(x,F) = \frac{1}{2}\left(xF^3 + L(1-2L) - 3x^3F^9\right) = \frac{1}{2}(-L^4 + 3L^3 - 5L^2 + 2L - 3L^3(1-L)^9). \tag{5.13}$$

Obtaining the final expression for $A(x,y)$ is just a matter of going back to Equation (5.10) and adding up all terms. □

Summarising, we have an explicit expression for A in terms of x, y, $H(x,y)$ and $F(x,y)$ which involves only elementary functions and the algebraic function $L(x,y)$. Moreover, note that $H(x,y)$ can be written in terms of $L(x,F)$ by Equation (5.13). Our goal is to carry out a full singularity analysis of the univariate generating function $A(x) = A(x,1)$. To this end we first perform singularity analysis on $F(x) = F(x,1)$.

Lemma 5.4. *The dominant singularity of $F(x)$ is the unique $\rho > 0$ such that $\rho F(\rho)^3 = 27/256$. The approximate value is $\rho \doteq 0.10530617$. The value $F(\rho) \doteq 1.0005143$ is the solution of*

$$t = \exp\left(\frac{27^3}{6\cdot 256^3}\left(1 + \left(\frac{59}{512} - \frac{3}{2}\left(\frac{27}{256}\right)^3\right)\frac{1}{t}\right)^9\right). \tag{5.14}$$

Proof. The function $F(x)$ satisfies

$$\Phi(x,F) = \exp\left(\frac{x^3}{6}\left(F + T_0(x,F)\right)^9\right) - F. \tag{5.15}$$

Thus the dominant singularity ρ of $F(x)$ may come from T_0 or from a branch point when solving (5.15). Assume that there is no such branch point. Then, since $L(x,y) = \theta(xy^3)$ and the dominant singularity of θ is at $27/256$, we have that $L(\rho, F(\rho)) = 1/4$ and $\rho F(\rho)^3 = 27/256$. Substituting on $\Phi(x,F) = 0$ we obtain Equation (5.14), where t stands for $F(\rho)$. The approximate value is $t \doteq 1.0005143$, which gives $\rho \doteq 0.10530617$, slightly smaller than $R = 27/256 = 0.10546875$.

We now prove that there is no branch point when solving Φ. If this were the case, then there would exist $\tilde{\rho} \le \rho$ such that $\partial_F\Phi(\tilde{\rho}, F(\tilde{\rho})) = 0$, where

$$\begin{aligned}\frac{\partial}{\partial F}\Phi(x,F(x)) =& \frac{3}{1024}(-3L^2 + 3L - 27(L(1-L)^3)^3 + 2F + 3xF^3)x^3 \\ &\cdot (2F + xF^3 + L - 2L^2 - 3(L(1-L)^3)^3)^8 - 1.\end{aligned} \tag{5.16}$$

follows by differentiating Equation (5.15), by using $\Phi(x,F(x)) = 0$ and Equations (5.7), (5.11), and (5.13).

Consider $\partial_F \Phi(x, F, L)$ as a function of three independent variables, where $x \geq 0$, $F \geq 1$ and $0 \leq L \leq 1/4$. It follows easily that it is an increasing function on any of them. Hence

$$0 = \partial_F \Phi(\tilde{\rho}, F(\tilde{\rho}), L(\tilde{\rho}, F(\tilde{\rho}))) \leq \partial_F \Phi(\rho, F(\tilde{\rho}), 1/4),$$

since, by assumption, $\tilde{\rho} \leq \rho$. On the other hand $\partial_F \Phi(\rho, t, 1/4) \doteq -0.9940$, so by comparing this with $\partial_F \Phi(\rho, F(\tilde{\rho}), 1/4)$ we deduce that $t < F(\tilde{\rho})$. Since $1 = F(0) < t$, by continuity of $F(x)$ there exists a value $x \in (0, \tilde{\rho})$ such that $F(x) = t$, and by the monotonicity of $\Phi(x, F)$ for fixed F there is a unique solution x to $\Phi(x, t) = 0$. This solution is precisely $x = \rho$, contradicting $\tilde{\rho} \leq \rho$.

□

Proposition 5.5. *Let ρ and t be as in Lemma 5.4. The singular expansion of $F(x)$ at ρ is*

$$F(x) = t + F_2 X^2 + F_3 X^3 + \mathcal{O}(X^4),$$

where $X = \sqrt{1 - x/\rho}$, and F_2 and F_3 are given by

$$F_2 = \frac{6t(16777216t + 9010867)\log(t)}{Q}, \qquad F_3 = \frac{t\log(t)M}{Q^{5/2}}$$

$$Q = 301989888t\log(t) - 33554432t + 127927431\log(t) - 3807575,$$

$$M = 6291456P^{\frac{1}{2}}(33554432t + 34268175\log(t) + 3807575),$$

$$P = -201326592t - 205609050\log(t) - 22845450.$$

Proof. To obtain the coefficients of the singularity expansion, including the fact that $F_1 = 0$, we apply indeterminate coefficients F_i, L_i of X^i to Equations (5.15) and

$$L(x)(1 - L(x))^3 - xF(x)^3 = 0,$$

where $X = \sqrt{1 - x/\rho}$, so that $x = \rho(1 - X^2)$. These calculations are tedious, but can be done with a computer algebra system such as Maple. □

Proposition 5.6. *Let ρ and t be as in Lemma 5.4. The dominant singularity of $A(x)$ is ρ, and its singular expansion at ρ is*

$$A(x) = A_0 + A_2 X^2 + A_4 X^4 + A_5 X^5 + \mathcal{O}(X^6),$$

where $X = \sqrt{1 - x/\rho}$ and A_0, A_2, A_4 and A_5 are given by

$$A_0 = \frac{243}{1310720}\frac{1}{t^6}\left(-27t\log(t) + 30t - \frac{102804525}{33554432}\log(t) - \frac{54165615}{16777216} - 15\log\left(\frac{3}{4}\right)\right)$$

$$A_2 = \frac{1}{Qt^6}\left(2519424\left((t + C)\log(t) - \frac{4}{3}t + \frac{2}{3}\log\left(\frac{3}{4}\right) + \frac{17858671}{150994944}\right)N\right)$$

$$A_4 = -\frac{1}{Q^3 t^6}\left(765887196662209204715 52N^2\left((t + C)K\log(t)^2 + \left(K\log\left(\frac{3}{4}\right) - 2t^2 - \frac{5858615}{8388608}t\right.\right.\right.$$
$$\left.\left.\left. + \frac{60561340172263}{1125899906842624}\right)\log(t) + \frac{2}{9}(t + C)\left(t - \frac{458789}{16777216} - \frac{1}{2}\log\left(\frac{3}{4}\right)\right)\right)\right)$$

$$A_5 = \frac{-7931762302491582015247220735}{5Q^{\frac{9}{2}}t^6}\left(P^{\frac{1}{2}}N^2\left(\frac{34268175}{33554432}\log(t) + t + C\right)^2\right),$$

where $C = 3807575/33554432$, $K = \frac{14214159}{33554432} + t$, $N = K\log(t) - C - \frac{1}{9}t$, *and* Q, M *and* P *are as in Proposition* 5.5.

PROOF. We just compute the singular expansion

$$A(x) = \sum_{k\geq 0} A_k X^k,$$

by plugging the expansions for $F(x)$ and $L(x)$ of Proposition 5.4 in (5.9).

□

THEOREM 5.7. *The number* A_n *of maximal graphs with* n *nodes not containing* $K_{3,3}$ *as a minor is asymptotically*

$$A_n \sim a \cdot n^{-7/2}\gamma^n n!,$$

where $\gamma = 1/\rho \doteq 9.496119$ *and* $a = -15A_5/8\sqrt{\pi} \doteq 0.40553 \cdot 10^{-4}$.

PROOF. This is a standard application of singularity analysis (see Corollary 2.5 (Corollary VI.1 of [**FS06**])) to the singular expansion of $A(x)$ obtained in the previous lemma. □

5.4. $K_{3,3}$-minor-free graphs

In this section, we derive the asymptotic number of $K_{3,3}$-minor-free graphs and properties of random $K_{3,3}$-minor-free graphs.

5.4.1. Notation. We introduce some more notation, which we will need in this section. Let $G(x,y)$, $C(x,y)$ and $B(x,y)$ denote the exponential generating functions of simple labelled *not necessarily connected*, *connected* and 2-*connected* $K_{3,3}$-minor-free graphs respectively, where x marks nodes and y marks edges. We will use the additional variable q to mark the number of K_5's used in the "construction process" of a $K_{3,3}$-minor-free graph (see below for a more precise explanation), but we won't give it explicitly in the argument list of our generating functions to simplify expressions.

Finally, for stating Lemma 5.10 and its proof, we will need the following notation from [**Wal82**]. We denote by a *brick* a 3-connected graph with at least 4 nodes, by a *(multi-)block* a 2-connected (multi-)graph with at least 2 nodes, by a *network* N a multigraph with two distinguished nodes – called its poles and labelled 0 and ∞ – such that the multigraph N^* obtained from N by adding an edge between the poles of N is 2-connected.

Furthermore, we define the following network types:

- *chain*: network consisting of 2 or more edges connected in series with the poles at its terminal nodes
- *bond*: network consisting of 2 or more edges connected in parallel with the poles at its terminal nodes
- *pseudo-brick*: a network N such that N^* is a brick

Let M be a multi-block with at least 2 edges and let $X = \{N_e, e \in E(M)\}$ be a set of networks, disjoint from each other and from M, each having at least one edge. Let $G = M(X)$ be the multi-block or network obtained from M by choosing an orientation (u,v) for each

edge $e \in E(M)$ and replacing $e = (u, v)$ by N_e identifying u with 0 and v with ∞. Then $G = M(X)$ is called a *superposition* with *core* M and components $N_e \in X$. A decomposition of a network or multi-block is a representation of it as a superposition.

A network N is called an *h-network*, a *p-network* or an *s-network* if it admits a decomposition whose core is a pseudo-brick, a bond or a chain, respectively.

5.4.2. Decomposition and Generating Functions. We want to apply singularity analysis to derive asymptotic estimates for the number of $K_{3,3}$-minor-free graphs. To achieve this, we first present a decomposition of this graph class which is due to Wagner [**Wag37**]. Then we will translate it into relations of our generating functions.

As seen in Theorem 5.2 above, Wagner [**Wag37**] characterised the class of maximal $K_{3,3}$-minor-free graphs. As a direct consequence we also obtain a decomposition for $K_{3,3}$-minor-free graphs. We will present here a more recent formulation of it, given by Thomas, Theorem 1.2 of [**Tho99**]. Roughly speaking the theorem states that a graph has no minor isomorphic to $K_{3,3}$ if and only if it can be obtained from a planar graph or K_5 by merging on an edge, a node, or taking disjoint components. To state the theorem more precisely, we need the following definition of [**Tho99**].

DEFINITION 5.8. Let G_1 and G_2 be graphs with disjoint node-sets, let $k \geq 0$ be an integer, and for $i = 1, 2$ let $X_i \subseteq V(G_i)$ be a set of pairwise adjacent nodes of size k. For $i = 1, 2$ let G_i' be obtained by deleting some (possibly none) edges with both ends in X_i. Let $f : X_1 \to X_2$ be a bijection, and let G be the graph obtained from the union of G_1' and G_2' by identifying x with $f(x)$ for all $x \in X_1$. In those circumstances we say that G is a *k-sum* of G_1 and G_2.

Now, we can state the theorem as a consequence of Wagner's theorem in the following way.

THEOREM 5.9 ([**Wag37**], see also Theorem 1.2 of [**Tho99**]). *A graph has no minor isomorphic to $K_{3,3}$ if and only if it can be obtained from planar graphs and K_5 by means of* 0-, 1-, *and* 2-*sums.*

Observe that for 2-connected $K_{3,3}$-minor-free graphs we only have to take 2-sums into account as 0- and 1-sums do not yield a 2-connected graph. This way the decomposition of Wagner fits perfectly well into a result of Walsh [**Wal82**] which delivers us – similar to planar graphs (see [**BGW02**]) – with the necessary relations for our generating functions.

The second ingredient for obtaining relations for our generating functions is to use a well-known decomposition of a graph along its connectivity-structure, i.e. into connected, 2-connected, and 3-connected blocks.

For ease of exposition, we give the following lemma (and a proof) which is a special case of Proposition 1.2 of [**Wal82**] (see also Proposition 5.19 in Section 5.4.6 for the original formulation by Walsh) and adopt also the notation of Walsh for the reader already familiar with it.

LEMMA 5.10. *Let $\mathcal{H}$ be a fixed finite set of* 3*-connected graphs. Let $\mathcal{C}$ be the class of all graphs which are $\mathcal{H}$-minor-free, i.e. they do not contain any minor which is isomorphic to a graph $H \in \mathcal{H}$. Let $F(x, y)$ and $B(x, y)$ denote the bivariate generating functions of* 3*-connected*

graphs with at least 4 nodes and 2-connected graphs with at least 2 nodes in $\mathcal{C}$. Similarly, let $K(x,y)$, $D(x,y)$ and $P(x,y)$ denote the generating functions for networks with non-adjacent poles, all networks with at least one edge and s-networks, respectively, where all of these networks are contained in $\mathcal{C}$. Then it holds

$$\frac{2}{x^2}\frac{\partial}{\partial D}F(x,D) = \log(K(x,y)) - P(x,y), \tag{5.17}$$

$$K(x,y) = \frac{2}{x^2}\frac{\partial}{\partial y}B(x,y), \tag{5.18}$$

$$D(x,y) = (1+y)K(x,y) - 1, \tag{5.19}$$

$$P(x,y) = xD(x,y)\left(D(x,y) - P(x,y)\right). \tag{5.20}$$

Proof. To prove the theorem, we count all networks with *non-adjacent* poles. From a theorem of Trakthenbrot [**Tra58**] (see Theorem 5.20 in Section 5.4.6) it follows that these networks are either h-, s-, or p-networks with non-adjacent poles. Thus we have

$$H(x,y) + P(x,y) + S(x,y) = K(x,y),$$

where $H(x,y)$ and $S(x,y)$ count h-networks and p-networks with non-adjacent poles.

Now, observe that $H(x,y) = \frac{2}{x^2}\frac{\partial}{\partial y}F(x,D(x,y))$, as we take a 3-connected graph choose an edge and its end nodes as the poles and plug-in any network for each edge. Now, to prove Equation (5.17), it is left to find an expression for $K(x,y) - S(x,y)$. Notice that $U(x,y) = K(x,y) - S(x,y)$ are exactly all non-p-networks with non-adjacent poles. But then we can obtain all networks with non-adjacent poles by taking a parallel union of those networks, i.e. $K(x,y) = \exp(U(x,y))$ and thus $U(x,y) = \log(K(x,y))$. Hence, altogether we obtain (5.17). Observe that by this construction, i.e. using a parallel union of non-p-networks with non-adjacent poles, we cannot introduce a 3-connected minor, since there is always a 2-cut between the networks in such a parallel union. So, if we want to drop the condition of being 3-connected for the excluded minors, we would have to argue why we do not build graphs that are no longer in the class $\mathcal{C}$.

Equations (5.18) - (5.20) can be derived as follows. If we take a 2-connected graph, choose one edge as a root edge, which is not included, and label its end nodes 0 and ∞ we clearly get all networks with non-adjacent poles. Thus (5.18) holds.

Furthermore, we get all networks with at least one edge by considering all networks with non-adjacent poles and adding an edge between the poles of each such network, and thus Equation (5.19) holds.

Finally, $E(x,y) = D(x,y) - P(x,y)$ counts all networks which are not s-networks. Thus the series union of such networks exhausts all s-networks by Trakhtenbrot's Theorem [**Tra58**] (Theorem 5.20). But then $P(x,y) = (1/x)(1/(1-xE(x,y)) - 1 - xE(x,y)) = xE(x,y)^2(1 - xE(x,y))^{-1}$, as the series network has at least length 2 and we multiply $E(x,y)$ by x and the whole term by $(1/x)$ since merging two poles yields an extra internal node. Substituting $E(x,y) = D(x,y) - P(x,y)$ yields Equation (5.20). Again, the construction of a series union cannot introduce a minor in our graph, as we have cut-nodes between the networks. For this construction 2-connectivity of the minors would be sufficient. □

As already mentioned in the proof of Lemma 5.10, it is crucial that we obtain by the parallel or series union of networks again a graph within our class. *Maximal* K_5-minor-free graphs are an example, where the construction yields graphs which are no longer in the original graph class; the series union of networks can yield a graph which is not maximal anymore, although the networks themselves are maximal K_5-minor-free.

Using Lemma 5.10 we get the following result for $K_{3,3}$-minor-free graphs.

LEMMA 5.11. *Let $G(x,y)$, $C(x,y)$ and $B(x,y)$ denote the bivariate exponential generating functions for not necessarily connected, connected, and 2-connected $K_{3,3}$-minor-free graphs. Then we have*

$$G(x,y) = \exp\left(C(x,y)\right) \quad \text{and} \quad C^\bullet(x,y) = x\exp\left(\frac{\partial}{\partial x}B\left(C^\bullet(x,y),y\right)\right). \tag{5.21}$$

Moreover, let $M(x,y)$ denote the bivariate generating function for 3-connected planar maps which fulfils

$$M(x,y) = x^2y^2\left(\frac{1}{1+xy} + \frac{1}{1+y} - 1 - \frac{(1+U)^2(1+V)^2}{(1+U+V)^3}\right), \tag{5.22}$$

where $U(x,y)$ and $V(x,y)$ are algebraic functions given by

$$U = xy(1+V)^2, \quad V = y(1+U)^2, \tag{5.23}$$

then we have

$$\frac{\partial}{\partial y}B(x,y) = \frac{x^2}{2}\left(\frac{1+D(x,y)}{1+y}\right), \tag{5.24}$$

where $D(x,y)$ is defined implicitly by $D(x,0) = 0$ and

$$\frac{M(x,D)}{2x^2D} + \frac{qx^3D^9}{6} - \log\left(\frac{1+D}{1+y}\right) + \frac{xD^2}{1+xD} = 0, \tag{5.25}$$

where q marks the monomial for K_5.

PROOF. Equations (5.21) follow from a well-known decomposition of a graph into its connected components and 2-connected blocks (see e.g. [**FS06**](p.95) and [**HP73**](p.10)). Using Euler's Polyhedral formula, Equations (5.22) and (5.23) follow from [**MS68**], where Mullin and Schellenberg derived the generating function for rooted 3-connected n-node m-face planar maps.

Solving Equation (5.18) for $\frac{\partial}{\partial y}B(x,y)$ and substituting $K(x,y)$ from Equation (5.19) we obtain (5.24).

Finally, substituting $\frac{\partial}{\partial D}F(x,D) = M(x,D)/4D + \frac{\partial}{\partial D}x^5D^{10}/5!$ and (5.20) into (5.17), we obtain Equation (5.25). Observe that the substitution for $F(x,D)$ follows from Theorem 5.9, which states that 3-connected $K_{3,3}$-minor-free graphs consist of 3-connected planar graphs and K_5. Now, as $4y\frac{\partial}{\partial y}G_3(x,y) = M(x,y)$, where $G_3(x,y)$ counts 3-connected planar graphs, we substitute $M(x,D)/4D$ for the 3-connected planar graphs. For K_5 we add $\frac{\partial}{\partial D}x^5D^{10}/5!$. observe that we have to add only *one* monomial for K_5 as the case where the edge between the poles of the corresponding network is not present is already included in $M(x,D)$ since K_5 minus an edge is planar. □

5.4.3. Singular Expansions and Asymptotic Estimates. We use the relations of the generating functions obtained so far to derive singular expansions for the generating functions of the different connectivity levels. We start from 3-connected $K_{3,3}$-minor-free graphs and then go up the connectivity hierarchy level by level. Eventually, this will allow us to derive asymptotic estimates for the number of and properties of not necessarily connected $K_{3,3}$-minor-free graphs in the subsequent section.

We start with 3-connected $K_{3,3}$-minor-free graphs. We have to introduce only a slight modification into the formulas already known for planar graphs ([**BGW02, GN05b**]).

From Lemma 5.11 we can derive analogously to [**BGW02**] a singular expansion for $D(x, y)$. It will turn out that the singularity of $D(x, y)$ changes only slightly compared to the case of 2-connected planar graphs, but yields a larger exponential growth rate.

To simplify expressions, we will use the following notation. The equation $Y(t) = y$ has a unique solution in $t = t(y)$ in a suitable small neighbourhood of 1. Then we denote by $R(y) = \zeta(t(y))$. See Section 5.4.5 for expressions for $Y(t)$ and ζ.

LEMMA 5.12. *For fixed y in a small neighbourhood of 1, $R(y)$ is the unique dominant singularity of $D(x, y)$. Moreover, $D(x, y)$ has a branch-point at $R(y)$, and the singular expansion at $R(y)$ is of the form*

$$D(x, y) = D_0(y) + D_2(y)X^2 + D_3(y)X^3 + O(X^4)$$

where $X = \sqrt{1 - x/R(y)}$ and the $D_i(y)$, $i = 0, \ldots, 3$ are given in Section 5.4.5.

PROOF. We mimic the proof of Lemma 3 in [**BGW02**].

$D(x, y)$ is defined by Equation (5.25). As it changed only slightly and we did not introduce a new possible source for the singularities by adding the monomial for K_5, we only have to assure that the arguments of [**BGW02**] still hold.

As was shown in [**BR84**] the singularity ζ of $M(x, y)$ is related to y by Equations (5.23) and the equation $1 + U + V - 3UV = 0$ with $x = \zeta$. Now, if we set $U_0 = 1/(3t)$, we obtain $V = -(3t - 1)/(3(t - 1))$ from the last equation. Moreover, using (5.23) we get $\zeta = -((t - 1)^3(3t + 1))/(16t^3)$.

Replacing y by D and using Equations (5.23) and (5.25) we obtain $Y(t)$ and $D_0(y)$ as given in Section 5.4.5.

Now, replacing y by D in Equations (5.22) and (5.23) and using additionally (5.25) we obtain the following two equations in x, U and D

$$\begin{aligned} U &= xD(1 + D(1 + U)^2)^2 \\ 0 &= \frac{D}{2}\left(\frac{1}{1 + xD} + \frac{1}{1 + D} - 1 - \frac{(1 + U)^2\left(1 + D(1 + U)^2\right)^2}{(1 + U + D(1 + U)^2)^3}\right) + \frac{qx^3D^9}{6} \\ &\quad - \log\left(\frac{1 + D}{1 + y}\right) + \frac{xD^2}{1 + xD}. \end{aligned}$$

From these two equations we can see that U and D have asymptotic expansions of square root type, i.e. in $X = \sqrt{1 - x/\zeta}$, around the singularity ζ. Substituting $D = \sum D_i(y)X^i$ and

$U = \sum U_i(y)X^i$ into the two equations, and equating coefficients of powers of X we obtain the remaining $D_i(y)$ and $U_i(y)$ which are given in Section 5.4.5.

What is left, is to assure that $M(x, y)$ is still the source of the singularity of $D(x, y)$, i.e. we need to show that

(a) a branch point in solving (5.25) and
(b) $1 + \mathrm{xD} = 0$ and/or $\log((1 + D)/(1 + y))$ becomes unbounded

do not provide singularities in the disc $|x| \leq \zeta$. We follow again closely the proof of Lemma 3 in [**BGW02**].

Let us first consider source (a). Let $H(D, y)$ denote the left side of (5.25). Then

$$\frac{\partial}{\partial y}H = \frac{1}{1+y}$$

and

$$\frac{\partial}{\partial D}H = \frac{1}{2x^2}\frac{\partial}{\partial D}\left\{\frac{M(x,D)}{D}\right\} - \frac{1 - xD^2(2+xD)}{(1+D)(1+xD)^2} + \frac{3qx^3D^8}{2}$$

Since $\zeta > 0$, $D_0 = D(\zeta, Y(t)) > 0$ and the power series for D and M has nonnegative coefficients, we get by (5.22) and (5.23)

$$\begin{aligned}
\left|\frac{\partial}{\partial D}H(x,D)\right| &\geq \left|\frac{1-xD^2(2+xD)}{(1+D)(1+xD)^2}\right| - \left|\frac{\partial}{\partial D}\frac{M(x,D)}{2x^2D}\right| + \left|\frac{3qx^3D^8}{2}\right| \\
&\geq \left|\frac{1-\zeta D_0(y)^2(2+\zeta D_0(y))}{(1+D_0(y))(1+\zeta D_0(y))^2}\right| - \frac{\partial}{\partial D}\frac{M(x,D)}{2x^2D}\Big|_{x=\zeta, D=D_0(y)} + \frac{3q\zeta^3 D_0(y)^8}{2}\Big|_{q=1} \\
&\geq \frac{t^2(1-t)(400+1808t+2527t^2+1155t^3+237t^4+17t^5}{2(1+3t)^2(1+2t)^2(3+t)^2} + \frac{19683t^7(1-t)}{8192(3t+1)^5},
\end{aligned}$$

which is greater 0 when $|x| \leq \zeta$. Therefore x is not a singularity from source (a).

Finally, we consider source (b). Since $M(x, D)$ is well defined, it follows from (5.25) that the last two terms must both be unbounded. Hence, $1 + xD = 0$ and $1 + D = 0$. So $x = 1$ and $D = -1$ which contradicts the fact that $D(1, Y(t)) > 0$. □

Next we solve Equation (5.24) for $B(x, y)$ and obtain:

LEMMA 5.13. *Let $W(x, z) = z(1 + U(x, z))$. The generating function $B(x, y)$ of 2-connected $K_{3,3}$-minor-free graphs admits the following expression as a formal power series:*

$$B(x,y) = \beta(x, y, D(x,y), W(x, D(x,y))) + \frac{qx^5 D(x,y)^{10}}{120}, \tag{5.26}$$

where

$$\beta(x,y,z,w) = \frac{x^2}{2}\beta_1(x,y,z) - \frac{x}{4}\beta_2(x,z,w),$$

and

$$\begin{aligned}
\beta_1(x,y,z) &= \frac{z(6x-2+xz)}{4x} + (1+z)\log\left(\frac{1+y}{1+z}\right) - \frac{\log(1+z)}{2} + \frac{\log(1+xz)}{2x^2} \\
\beta_2(x,z,w) &= \frac{2(1+x)(1+w)(z+w^2)+3(w-z)}{2(1+w)^2} - \frac{1}{2x}\log(1+xz+xw+xw^2) \\
&+ \frac{1-4x}{2x}\log(1+w) + \frac{1-4x+2x^2}{4x}\log\left(\frac{1-x+wz-xw+xw^2}{(1-x)(z+w^2+1+w)}\right).
\end{aligned}$$

PROOF. The proof follows closely the lines of proof of Lemma 5 in [**GN05b**]. To avoid repetition we present only the calculations which are different to [**GN05b**] due to the different underlying graph class.

From Equation (5.24) we obtain

$$B(x,y) = \frac{x^2}{2}\int_0^y \frac{1+D(x,t)}{1+t}dt = \frac{x^2}{2}\log(1+y) + \frac{x^2}{2}\int_0^y \frac{D(x,t)}{1+t}dt.$$

Integrating by parts we get

$$\int_0^y \frac{D(x,t)}{1+t}dt = \log(1+t)D(x,y) - \int_0^y \log(1+t)\frac{\partial}{\partial t}D(x,t)dt.$$

Consider x as a fixed value. From Equation (5.25) it follows that

$$\phi(u) = -1 + (1+u)\exp\left(-\frac{M(x,u)}{2ux^2} - \frac{xu^2}{1+xu} - \frac{qx^3u^9}{6}\right)$$

is a functional inverse of $D(x,y)$ with respect to y, as $\phi(D(x,y)) = y$. Notice that compared to the case of planar graphs, we introduce here the only difference in the calculations since Equation (5.25) changed, but the rest stayed the same.

We substitute $s = D(x,y)$ and we have $t = \phi(s)$. Then

$$\int_0^y \log(1+t)\frac{\partial}{\partial t}D(x,t)dt = \int_0^{D(x,y)} \left(\log(1+s) - \frac{xs^2}{1+xs}\right)ds - \int_0^{D(x,y)} \frac{M(x,s)}{2x^2s}ds - \int_0^{D(x,y)} \frac{qx^3s^9}{6}ds,$$

where the last term appears due to the change of the graph class.

The first and the last integral have a simple primitive and we are left with an integral involving $M(x,y)$. Now, we can proceed as in the proof of Lemma 5 in [**GN05b**]: using Equations (5.22) and (5.23) and the definition of $W(x,z)$, we can solve this integral using substitution (and the functional inverse of $W(x,z)$ according to the second variable) in a similar way as the integral above. As the calculations for solving this integral do not differ from the case of planar graphs, we refer to [**GN05b**] for further details. □

We can use this lemma to obtain the singular expansion for $B(x,y)$.

LEMMA 5.14. *For fixed y in a small neighbourhood of 1, the dominant singularity of $B(x,y)$ is equal to $R(y)$. The singular expansion at $R(y)$ is of the form*

$$B(x,y) = B_0(y) + B_2(y)X^2 + B_4(y)X^4 + B_5(y)X^5 + O(X^6) \tag{5.27}$$

where $X = \sqrt{1 - x/R(y)}$, and the $B_i(y)$, $i = 0,\ldots,5$ are analytic functions in a neighbourhood of 1.

PROOF. From Equation (5.26) we can see that for y close to 1 the only singularities come from the singularities of $D(x,y)$; hence the first claim of the theorem follows.

The singular expansion for $B(x,y)$ can be obtained using Equation (5.26) and the singular expansion for $D(x,y)$. We substitute the singular expansion for $D(x,y)$, $U(x,D(x,y))$ in (5.26). Then we set $x = \zeta(t)(1-X^2)$ and $y = Y(t)$ and expand the resulting expression.

Now, collecting and simplifying the coefficients of the X^i for $i = 1, \ldots, 5$ is a tedious calculation, but can be done with a computer algebra system such as MAPLE. This yields the expressions for the $B_i(y)$ given in Section 5.4.5. □

For connected and not necessarily connected $K_{3,3}$-minor-free graphs, we can derive singular expansions by carrying out an analogous calculation as in the proof of Theorem 1 in [**GN05b**]. We only have to adapt for the different $D_i(y)$ and $B_i(y)$. One can easily check that the intermediate step of Claim 1 in [**GN05b**] still holds and the rest of the calculations stays the same. The coefficients of the expansions, which we obtain in this way, can be found in Section 5.4.5.

LEMMA 5.15. *For fixed y in a small neighbourhood of 1, the dominant singularity of $C(x, y)$ and $G(x, y)$ is equal to $R(y)$. The singular expansions at $R(y)$ are of the form*

$$C(x, y) = C_0(y) + C_2(y)X^2 + C_4(y)X^4 + C_5(y)X^5 + O(X^6) \tag{5.28}$$

and

$$G(x, y) = G_0(y) + G_2(y)X^2 + G_4(y)X^4 + G_5(y)X^5 + O(X^6) \tag{5.29}$$

where $X = \sqrt{1 - x/R(y)}$, and the $C_i(y)$ and $G_i(y)$, $i = 0, \ldots, 5$, are analytic functions in a neighbourhood of 1.

From Lemmas 5.14 and 5.15 we obtain the following asymptotic estimates using Corollary 2.5 (Corollary VI.1 of [**FS06**]).

THEOREM 5.16. *Let g_n, c_n, and b_n denote the number of not necessarily connected, connected and biconnected, respectively, $K_{3,3}$-minor-free graphs on n nodes. Then it holds*

$$g_n \sim \alpha_g \, n^{-\frac{7}{2}} \, \rho_g^{-n} \, n!, \tag{5.30}$$

$$c_n \sim \alpha_c \, n^{-\frac{7}{2}} \, \rho_c^{-n} \, n!, \tag{5.31}$$

$$b_n \sim \alpha_b \, n^{-\frac{7}{2}} \, \rho_b^{-n} \, n!, \tag{5.32}$$

where $\alpha_g \doteq 0.42643 \cdot 10^{-5}$, $\alpha_c \doteq 0.41076 \cdot 10^{-5}$, $\alpha_b \doteq 0.37074 \cdot 10^{-5}$, $\rho_c^{-1} = \rho_g^{-1} \doteq 27.22935$, and $\rho_b^{-1} \doteq 26.18659$ are analytically computable constants.

5.4.4. Structural Properties. If we consider a random $K_{3,3}$-minor-free graph, i.e. drawing a $K_{3,3}$-minor-free graph on n nodes uniformly at random from all such graphs on n nodes, we can derive the following properties using the algebraic singularity schema of Theorem 2.7 (Theorem IX.10 of [**FS06**]).

THEOREM 5.17. *The number of edges in a not necessarily connected and connected random $K_{3,3}$-minor-free graph is asymptotically normally distributed with mean μ_n and variance σ_n^2, which satisfy*

$$\mu_n \sim \kappa n \quad \text{and} \quad \sigma_n^2 \sim \lambda n,$$

where $\kappa \doteq 2.21338$ and $\lambda \doteq 0.43044$ are analytically computable constants.

Recall that we introduced the variable q in the equations of the generating functions above to mark the monomial for K_5. We can use this variable to obtain a limit law for the number of K_5 used in the construction process (following the above decomposition, see Theorem 5.9) of

a random $K_{3,3}$-minor-free graph. The next theorem shows that a linear number of K_5 is used, but the constant is very small; this is exactly what one would expect as the expected number of edges is only slightly larger than for planar graphs (see Theorem 5.17 and [**GN05b**]) – note that merging a K_5 in the construction process adds only 3 new nodes, but 9 new edges.

THEOREM 5.18. *Let* $\mathrm{C}(G)$ *denote the number of* K_5 *used in the construction of a random* $K_{3,3}$*-minor-free graph* G *according to Theorem* 5.9. *Then* $\mathrm{C}(G)$ *is asymptotically normally distributed with mean* μ_n *and variance* σ_n^2*, which satisfy*

$$\mu_n \sim \kappa n \quad \text{and} \quad \sigma_n^2 \sim \lambda n,$$

where $\kappa \doteq 0.92391 \cdot 10^{-4}$ *and* $\lambda \doteq 0.92440 \cdot 10^{-4}$ *are analytically computable constants. The same holds for a random connected* $K_{3,3}$*-minor-free graph.*

5.4.5. Expressions for the Coefficients. Here, we give the expressions for the coefficients of the singular expansions of $D(x,y)$, $U(x,y)$, $B(x,y)$, $C(x,y)$, and $G(x,y)$ as well as the expressions for the singularities. We use the same approach as in [**BGW02**] and parametrise the expressions in the complex variable t.

$$\begin{aligned}
h &= \frac{t^2}{8192(3t+1)^6(2t+1)(t+3)}\big(13122qt^9+45927qt^8-1658880t^7+19683qt^7\\
&\quad -12496896t^6-8847360t^5+6832128t^4+10399744t^3+4739072t^2+958464t+73728\big)\\
Y(t) &= -\frac{2t+1}{(3t+1)(t-1)}e^{-h}-1\\
\zeta &= -\frac{(t-1)^3(3t+1)}{16t^3}
\end{aligned}$$

$$\begin{aligned}
Q &= 78732t^9-1328940t^8-26889705t^7-153744066t^6-415828997t^5-522964992t^4\\
&\quad -342073344t^3-121237504t^2-22151168t-1638400\\
K &= 78732t^{11}+472392t^{10}-2668221t^9-816345t^8+92026557t^7+562023429t^6\\
&\quad +1040556032t^5+926367744t^4+455663616t^3+127336448t^2+19005440t+1179648
\end{aligned}$$

$$\begin{aligned}
U_0 &= \frac{1}{3t}\\
U_1 &= -\left(-\frac{2}{27}\frac{(3t+1)K}{t^3(t+1)Q}\right)^{\frac{1}{2}}\\
U_2 &= -\frac{(3t+1)^2}{54t^2(t+1)^2Q^2}\big(6198727824t^{20}+180231719760t^{19}+891036025560t^{18}\\
&\quad -129029367636 00t^{17}-197722642310 71t^{16}-182139652514 8269t^{15}\\
&\quad -138162723611 45022t^{14}-794239712173 7354t^{13}-324711461744 767867t^{12}\\
&\quad -931873748086 8966665t^{11}-188127580290 7541504t^{10}-271350292543 7276160t^9\\
&\quad -284365301063 3469952t^8-219073166103 7666304t^7-124651452495 0953984t^6\\
&\quad -521994799964 094464t^5-158674913803 108352t^4-34025656742 98880t^3\\
&\quad -48763217211 55584t^2-41894828992 1024t-16312285790208\big)
\end{aligned}$$

$$
\begin{aligned}
D_0 &= -\frac{3t^2}{(3t+1)(t-1)}\\
D_1 &= 0\\
D_2 &= -\frac{t(2t+1)^2}{(3t+1)(t-1)Q}\left(19683t^8+118098t^7-1592325t^6-10616832t^5-30670848t^4\right.\\
&\quad \left.+7602176t^3+24444928t^2+9830400t+1179648\right)\\
D_3 &= \frac{131072}{9Q^2}\left(\left(-\frac{(3t+1)K}{t^3(t+1)Q}\right)^{\frac{1}{2}}\sqrt{6}t^2(3t+1)(t+3)^2(2t+1)^2K\right)\\
P_1 &= 1549681956t^{24}-68328432252t^{23}-646991330895t^{22}+1383569088336t^{21}\\
&\quad -57934645367238t^{20}-1030641858893628t^{19}-5581315778170878t^{18}\\
&\quad -9690527546116164t^{17}+14823917538797880t^{16}+66591676440148968t^{15}\\
&\quad -6807229356797163t^{14}-121180156627243452t^{13}-38691868953118942t^{12}\\
&\quad +93938978979606528t^{11}+65141137737269248t^{10}-21686663626104832t^9\\
&\quad -36470289308778496t^8-96595012320010 24t^7+46686861426851 84t^6\\
&\quad +41198956963512 32t^5+13298026906910 72t^4+22334346677452 8t^3\\
&\quad +17853474406400t^2+207232172032t-40265318400\\
P_2 &= -472392t^{12}-2991816t^{11}+15064542t^{10}+10234512t^9-550526652t^8\\
&\quad -3556193688t^7-7367383050t^6-7639318528t^5-4586717184t^4-1675345920t^3\\
&\quad -368705536t^2-45088768t-2359296\\
B_0 &= \frac{1}{4t^6}\left(-\frac{9}{256}\left(t+\frac{1}{3}\right)^2(t-1)^6\ln\left(\frac{-2\,t-1}{3\,t^2-2\,t-1}\right)\right.\\
&\quad +\left(-\frac{3}{32}\,t^7-\frac{9}{512}\,t^8+\frac{7}{128}\,t^6+\frac{1}{32}\,t^3-\frac{15}{256}\,t^4-\frac{3}{16}\,t^5-\frac{1}{512}+\frac{3}{128}\,t^2\right)\ln\left(\frac{(3\,t+1)^2(t-1)^2}{(t+1)^4}\right)\\
&\quad +\left(t^3\ln\left(1+\frac{3}{16}\,\frac{(t-1)^2}{t}\right)+\frac{1}{2}\,t^3\ln\left(\frac{1}{16}\,\frac{(t+1)^2(3\,t+1)}{t^2}\right)\right.\\
&\quad \left.\left.-\frac{3}{8}\left(t^4-\frac{4}{3}\,t^3+2\,t^2-\frac{1}{3}\right)\ln\left(-(t-1)^{-1}\right)\right)t^3\right)\\
&\quad -\frac{(t-1)^2}{41943040t^4(3\,t+1)^5(t+3)}\left(19683\,t^{13}-131220\,t^{12}-183708\,t^{11}+360921744\,t^{10}\right.\\
&\quad \left.+2005423731\,t^9+3887177580\,t^8+5603033310\,t^7+4821770240\,t^6+2013921280\,t^5\right.\\
&\quad \left.+229048320\,t^4-97157120\,t^3-31436800\,t^2-2048000\,t+122880\right)\\
B_1 &= 0\\
B_2 &= \frac{9\left(t+\frac{1}{3}\right)(t-1)^3}{1024\,t^6}\left(2\left(t+\frac{1}{3}\right)(t-1)^3\ln\left(\frac{-2\,t-1}{3\,t^2-2\,t-1}\right)\right.\\
&\quad +\left(t-\frac{1}{3}\right)(t+1)^3\ln\left(\frac{(3\,t+1)^2(t-1)^2}{(t+1)^4}\right)+\frac{32}{3}\,t^3\ln\left(-(t-1)^{-1}\right)\Bigg)\\
&\quad +\frac{(t-1)^4}{8388608t^4(t+3)(3\,t+1)^5}\left(19683\,t^{11}-13122\,t^{10}-190269\,t^9+122862096\,t^8\right.\\
&\quad +626914188\,t^7+555393024\,t^6+28803072\,t^5-163438592\,t^4-81084416\,t^3\\
&\quad \left.-14852096\,t^2-720896\,t+49152\right)
\end{aligned}
$$

$$
\begin{aligned}
B_3 &= 0\\
B_4 &= -\frac{P_1}{8388608t^4(t+3)(3t+1)^5Q} - \frac{9(t+\frac{1}{3})^2(t-1)^6(-2\ln(t+1)+\ln(2t+1))}{1024t^6}\\
B_5 &= -\frac{\sqrt{\frac{3P_2}{t^3(t+1)Q}}P_2^2(t-1)^6}{2880(3t+1)^5(t+1)tQ^2},
\end{aligned}
$$

$$
\begin{aligned}
C_0 &= R + B_0 + B_2\\
C_1 &= 0\\
C_2 &= -R\\
C_3 &= 0\\
C_4 &= -\frac{R+\frac{R^2}{2B_4-R}}{2}\\
C_5 &= B_5\left(1-\frac{2B_4}{R}\right)^{-\frac{5}{2}}
\end{aligned}
$$

$$
\begin{aligned}
G_0 &= e^{C_0}\\
G_1 &= 0\\
G_2 &= e^{C_0}C_2\\
G_3 &= 0\\
G_4 &= e^{C_0}\left(C_4+\frac{C_2^2}{2}\right)\\
G_5 &= e^{C_0}C_5
\end{aligned}
$$

5.4.6. Network Theorems. For easier referencing we state here two theorems from literature, which we need at several places in this chapter. The first theorem is by Walsh [**Wal82**] and concerns – roughly speaking – the relation of the generating functions of 2-connected and 3-connected graphs of a graph class. The second theorem is by Trakhtenbrot [**Tra58**] and states that there exists a decomposition of a network, which is unique and involves only three different kind of network types. For an explanation of the notation used in both theorems, see Section 5.4.1.

PROPOSITION 5.19 (Proposition 1.2 of [**Wal82**]). *Let X be a set of bricks, X' be the set of pseudo-bricks N such that $N^* \in X$, X'' be the set of networks obtained by requiring the cores of h-networks to be taken from X', and Y be the set of blocks N^* such that $N \in X''$. Then*

the following equations are valid if $F(x,y)$ counts X and $B(x,y)$ counts Y.

$$\frac{2}{x^2}\frac{\partial}{\partial D}F(x,D) = \log(K(x,y)) - P(x,y) \tag{5.33}$$

$$K(x,y) = \frac{2}{x^2}\frac{\partial}{\partial y}B(x,y) \tag{5.34}$$

$$D(x,y) = (1+y)K(x,y) - 1 \tag{5.35}$$

$$P(x,y) = xD(x,y)\left(D(x,y) - P(x,y)\right) \tag{5.36}$$

THEOREM 5.20 (Trakhtenbrot [**Tra58**]). *Any network with at least 2 edges belongs to exactly one of the 3 classes: h-networks, p-networks, s-networks. An h-network has a unique decomposition and a p-network (resp. an s-network) can be uniquely decomposed into components which are not themselves p-networks (s-networks), where uniqueness is up to orientation of the edges of the core, and also up to their order if the core is a bond.*

5.5. Conclusions

In this chapter we obtained asymptotic estimates for the number (and properties) of $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs. While this is a first important step to better understand these graph classes, it would be interesting to investigate more parameters and to examine maximal $K_{3,3}$-minor-free graphs if we allow the number of edges to vary. Moreover, due to Kuratowski's theorem, it is natural to consider also the class of (maximal) K_5-minor-free graphs.

Bibliography

[AH77] K. Appel and W. Haken, *Every planar map is four colorable. I. Discharging*, Illinois Journal of Mathematics **21** (1977), no. 3, 429–490.

[AH89] ———, *Every planar map is four colorable*, Contemporary Mathematics, vol. 98, American Mathematical Society, Providence, RI, 1989, With the collaboration of J. Koch.

[AHK77] K. Appel, W. Haken, and J. Koch, *Every planar map is four colorable. II. Reducibility*, Illinois Journal of Mathematics **21** (1977), no. 3, 491–567.

[BGH03] N. Bonichon, C. Gavoille, and N. Hanusse, *An information-theoretic upper bound of planar graphs using triangulations*, Springer Lecture Notes in Computer Science (STACS 2003) **2608** (2003), 499–510.

[BGK03] M. Bodirsky, C. Gröpl, and M. Kang, *Generating labeled planar graphs uniformly at random*, Springer Lecture Notes in Computer Science (ICALP 2003) **2719** (2003), 1095–1107.

[BGKN05] M. Bodirsky, O. Giménez, M. Kang, and M. Noy, *On the number of series parallel and outerplanar graphs*, 2005 European Conference on Combinatorics, Graph Theory and Applications (EuroComb'05), DMTCS Proceedings, vol. AE, 2005, pp. 383–388.

[BGR92] E. A. Bender, Z. C. Gao, and L. B. Richmond, *Submaps of maps. I. General* 0-1 *laws*, Journal of Combinatorial Theory, Series B **55** (1992), no. 1, 104–117.

[BGW02] E. A. Bender, Z. C. Gao, and N. C. Wormald, *The number of labeled 2-connected planar graphs*, Electronic Journal of Combinatorics **9** **#R43** (2002).

[BHM02] H.-J. Bandelt, K. T. Huber, and V. Moulton, *Quasi-median graphs from sets of partitions*, Discrete Applied Mathematics **122** (2002), no. 1-3, 23–35.

[BLMK] M. Bodirsky, M. Löffler, C. McDiarmid, and M. Kang, *Random cubic planar graphs*, to appear.

[BMBS05] B. Ben-Moshe, B. K. Bhattacharya, and Q. Shi, *Efficient algorithms for the weighted 2-center problem in a cactus graph.*, International Symposium on Algorithms and Computation 2005, Lecture Notes in Computer Science, vol. 3827, 2005, pp. 693–703.

[Bol85] B. Bollobás, *Random graphs*, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1985.

[BR84] E. A. Bender and L. B. Richmond, *The asymptotic enumeration of rooted convex polyhedra*, Journal of Combinatorial Theory, Series B **36** (1984), no. 3, 276–283.

[Coo71] S. A. Cook, *The complexity of theorem-proving procedures*, STOC '71: Proceedings of the third annual ACM symposium on Theory of computing (New York, NY, USA), ACM Press, 1971, pp. 151–158.

[DFLS04] P. Duchon, P. Flajolet, G. Louchard, and G. Schaeffer, *Boltzmann samplers for the random generation of combinatorial structures*, Combinatorics, Probability and Computing **13** (2004), no. 4-5, 577–625.

[Dow06] C. Dowden, Manuscript, 2006.

[DVW96] A. Denise, M. Vasconcellos, and D. Welsh, *The random planar graph*, Congressus numerantium **113** (1996), 61–79.

[Fle99] L. Fleischer, *Building chain and cactus representations of all minimum cuts from Hao-Orlin in the same asymptotic run time*, Journal of Algorithms **33** (1999), no. 1, 51–72.

[FPS05] E. Fusy, D. Poulalhon, and G. Schaeffer, *Dissections and trees, with applications to optimal mesh encoding and to random sampling*, Proceedings of the 16th ACM-SIAM Symposium on Discrete Algorithms (SODA 2005) (2005), 690–699.

[FS06] P. Flajolet and R. Sedgewick, *Analytic Combinatorics*, Book in preparation, April 2006.

[Fus05] E. Fusy, *Quadratic exact-size and linear approximate-size random generation of planar graphs*, DMTCS Proceedings **AD** (2005), 125–138.

[FZvC94] P. Flajolet, P. Zimmermann, and B. van Cutsem, *A calculus for the random generation of labelled combinatorial structures*, Theoretical Computer Science **132** (1994), no. 1-2, 1–35.

[GGNW] S. Gerke, O. Giménez, M. Noy, and A. Weißl, *On the number of $K_{3,3}$-minor-free and maximal $K_{3,3}$-minor-free graphs*, Submitted for publication.

[GJ79] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*, W. H. Freeman & Co., New York, NY, USA, 1979.

[GM04] S. Gerke and C. McDiarmid, *On the number of edges in a random planar graph*, Combinatorics, Probability, and Computing **13** (2004), 165–183.

[GMSW05] S. Gerke, C. McDiarmid, A. Steger, and A. Weißl, *Random planar graphs with n nodes and a fixed number of edges*, SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms, 2005, pp. 999–1007.

[GMSW07] ———, *Random planar graphs with given average degree*, Combinatorics, Complexity and Chance, A tribute to Dominc Welsh, Oxford University Press, 2007, pp. 83–102.

[GN04] O. Giménez and M. Noy, *Estimating the growth constant of labelled planar graphs*, Proceedings of the 3rd Colloquium on Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities (2004), 133–139.

[GN05a] ———, *Asymptotic enumeration and limit laws of planar graphs*, arXiv:math.CO/0501269, 2005.

[GN05b] ———, *The number of planar graphs and properties of random planar graphs*, 2005 International Conference on Analysis of Algorithms, DMTCS Proceedings, vol. AD, 2005, pp. 147–156.

[HN53] F. Harary and R. Z. Norman, *The dissimilarity characteristic of Husimi trees*, Annals of Mathematics. Second Series **58** (1953), 134–141.

[HP73] F. Harary and E. M. Palmer, *Graphical enumeration*, Academic Press, New York, 1973.

[HU53] F. Harary and G. E. Uhlenbeck, *On the number of Husimi trees. I*, Proceedings of the National Academy of Sciences **39** (1953), 315–322.

[Hus50] K. Husimi, *Note on Mayers' theory of cluster integrals*, The Journal of Chemical Physics **18** (1950), 682–684.

[JŁR00] S. Janson, T. Łuczak, and A. Rucinski, *Random graphs*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience, New York, 2000.

[Kar72] R. M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), Plenum, New York, 1972, pp. 85–103.

[Knu81] D. E. Knuth, *The art of computer programming. Vol. 2, Seminumerical algorithms*, second ed., Addison-Wesley, 1981.

[KSS04] J. Kleinberg, M. Sandler, and A. Slivkins, *Network failure detection and graph connectivity*, SODA '04: Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms (Philadelphia, PA, USA), Society for Industrial and Applied Mathematics, 2004, pp. 76–85.

[Kur30] C. Kuratowski, *Sur le problème des courbes gauches en topologie*, Fundamenta Mathematicae **15** (1930), 217–283.

[LT79] R. J. Lipton and R. E. Tarjan, *A separator theorem for planar graphs*, SIAM Journal on Applied Mathematics **36** (1979), no. 2, 177–189.

[LW92] J. H. van Lint and R. M. Wilson, *A course in combinatorics*, Cambridge University Press, Cambridge, 1992.

[Moo68] J. W. Moon, *On the maximum degree in a random tree*, The Michigan Mathematical Journal **15** (1968), 429–432.

[MR] C. McDiarmid and B. Reed, Personal Communication.

[MR95] R. Motwani and P. Raghavan, *Randomized algorithms*, Cambridge University Press, Cambridge, 1995.

[MS68] R. C. Mullin and P. J. Schellenberg, *The enumeration of c-nets via quadrangulations*, Journal of Combinatorial Theory **4** (1968), 259–276.

[MSW05] C. McDiarmid, A. Steger, and D. Welsh, *Random planar graphs*, Journal of Combinatorial Theory, Series B **93** (2005), 187–205.

[MSW06] ———, *Random graphs from planar and other addable classes*, Topics in Discrete Mathematics (2006), 231–246.

[Noy07] M. Noy, *Random planar graphs and the number of planar graphs*, Combinatorics, Complexity and Chance, A tribute to Dominc Welsh, Oxford University Press, 2007, pp. 213–233.

[NSTW06] S. Norine, P. Seymour, R. Thomas, and P. Wollan, *Proper minor-closed families are small*, Journal of Combinatorial Theory, Series B **96** (2006), 754–757.

[NW78] A. Nijenhuis and H. S. Wilf, *Combinatorial algorithms*, second ed., Academic Press, New York, 1978.

[OPT03] D. Osthus, H. J. Prömel, and A. Taraz, *On random planar graphs, the number of planar graphs and their triangulations*, Journal of Combinatorial Theory, Series B **88** (2003), 119–134.

[PW] K. Panagiotou and A. Weißl, *Properties of graph classes via boltzmann samplers*, Submitted for publication.

[PW06] ———, *http://www.ti.inf.ethz.ch/as/people/weissl/sampler.zip*, 2006.

[PY81] C. Papadimitriou and M. Yannakakis, *The clique problem for planar graphs*, Information Processsing Letters **13** (1981), 131–133.

[Rén59] A. Rényi, *Some remarks on the theory of trees*, Magyar Tudomanyos Akademia Matematikai Kutato Intezetenek Kozlemenyei **4** (1959), 73–85.

[Rid51] R. J. Riddell, *Contributions to the theory of condensation, dissertation*, 1951.

[RK92] H. Rieger and T. R. Kirkpatrick, *Disordered p-spin interaction models on Husimi trees*, Physical Review B **45** (1992), 9772–9777.

[RSST97] N. Robertson, D. Sanders, P. D. Seymour, and R. Thomas, *The four-colour theorem*, Journal of Combinatorial Theory, Series B **70** (1997), 2–44.

[RW88] L. B. Richmond and N. C. Wormald, *Random triangulations of the plane*, European Journal of Combinatorics **9** (1988), 61–71.

[SSCM04] P. Serra, J. F. Stilck, W. L. Cavalcanti, and K. D. Machado, *Polymers with attractive interactions on the husimi lattice*, Journal of Physics A: Mathematical and General **37** (2004), no. 37, 8811–8821.

[SSW05] J. Schwartz, A. Steger, and A. Weißl, *Fast algorithms for weighted bipartite matching*, WEA (S. E. Nikoletseas, ed.), Lecture Notes in Computer Science, vol. 3503, Springer, 2005, pp. 476–487.

[Tho99] R. Thomas, *Recent excluded minor theorems for graphs*, Surveys in Combinatorics, Cambridge University Press, 1999, pp. 201–222.

[Tra58] B. A. Trakhtenbrot, *Towards a theory of non-repeating contact schemes*, Trudi Mathematicheskogo Instituta Academii Nauk SSSR **51** (1958), 226–269.

[Tut62] W. T. Tutte, *A census of planar triangulations*, Canadian Journal of Mathematics **14** (1962), 21–38.

[Uhl] G. E. Uhlenbeck, *Some basic problems of statistical mechanics*, Gibbs Lecture, AMS, 1950.

[UU04] R. Uehara and Y. Uno, *Efficient algorithms for the longest path problem*, Algorithms and computation, Lecture Notes in Computer Science, vol. 3341, Springer, 2004, pp. 871–883.

[VSK00] R. Vicente, D. Saad, and Y. Kabashima, *Error-correcting code on a cactus: a solvable model*, Europhysics Letters **51** (2000), 698–704.

[Wag37] K. Wagner, *Über eine Erweiterung des Satzes von Kuratowski*, Deutsche Mathematik **2** (1937), 280–285.

[Wal82] T. R. S. Walsh, *Counting labelled three-connected and homeomorphically irreducible two-connected graphs*, Journal of Combinatorial Theory, Series B **32** (1982), no. 1, 1–11.

[Wei85] K. Weiler, *Edge-based data structures for solid modeling in curved-surface environments*, IEEE Computer Graphics and Application **5** (1985), no. 1, 21–40.

[Whi32] H. Whitney, *Congruent Graphs and the Connectivity of Graphs*, American Journal of Mathematics **54** (1932), no. 1, 150–168.

[Wil84] H. S. Wilf, *Backtrack: an $O(1)$ expected time algorithm for the graph coloring problem*, Information Processing Letters **18** (1984), no. 3, 119–121.

[Wil90] ———, *Generating functionology*, Academic Press Professional, Inc., San Diego, CA, USA, 1990.

[Won99] P. K. Wong, *Optimal path cover problem on block graphs.*, Theoretical Computer Science **225** (1999), no. 1-2, 163–169.

[YC94] J. H. Yan and G. J. Chang, *The path-partition problem in block graphs*, Information Processing Letters **52** (1994), no. 6, 317–322.

[ZZ04] B. Zmazek and J. Zerovnik, *The obnoxious center problem on weighted cactus graphs.*, Discrete Applied Mathematics **136** (2004), no. 2-3, 377–386.

Curriculum Vitae

Andreas Christian Weißl

born on April 18, 1978 in Munich, Germany

1988 - 1997	Gymnasium bei St. Anna, Augsburg, Germany Degree: Abitur
1997 - 1998	Civil Service, Kolping Bildungswerk Augsburg, Germany
1998 - 2003	Studies of Computer Science at TU München, Germany Degree: Diplom-Informatiker (Univ.)
2003	Ph.D. student at TU München, Germany
since Oct. 2003	Ph.D. student at ETH Zurich, Switzerland