FPSAC 2013 Paris, France DMTCS proc. **AS**, 2013, 599–610

Asymptotic properties of some minor-closed classes of graphs

Mireille Bousquet-Mélou[1†] and Kerstin Weller[2]

[1] *CNRS, LaBRI, UMR 5800, Université de Bordeaux, 351 cours de la Libération, 33405 Talence Cedex, France*
[2] *Department of Statistics, University of Oxford, 1 South Parks Road, Oxford, OX1 3TG, United Kingdom*

Abstract. Let $\mathcal{A}$ be a minor-closed class of labelled graphs, and let $\mathcal{G}_n$ be a random graph sampled uniformly from the set of n-vertex graphs of $\mathcal{A}$. When n is large, what is the probability that $\mathcal{G}_n$ is connected? How many components does it have? How large is its biggest component? Thanks to the work of McDiarmid and his collaborators, these questions are now solved when all excluded minors are 2-connected.

Using exact enumeration, we study a collection of classes $\mathcal{A}$ excluding non-2-connected minors, and show that their asymptotic behaviour is sometimes rather different from the 2-connected case. This behaviour largely depends on the nature of the dominant singularity of the generating function $C(z)$ that counts connected graphs of $\mathcal{A}$. We classify our examples accordingly, thus taking a first step towards a classification of minor-closed classes of graphs. Furthermore, we investigate a parameter that has not received any attention in this context yet: the size of the root component. This follows non-gaussian limit laws (beta and gamma), and clearly deserves a systematic investigation.

Keywords: Labelled graphs – Excluded minors – Enumeration – Asymptotic properties

1 Introduction

We consider simple graphs on the vertex set $\{1, \ldots, n\}$. A set of graphs is a *class* if it is closed under isomorphisms. A class of graphs $\mathcal{A}$ is *minor-closed* if any minor of a graph of $\mathcal{A}$ is in $\mathcal{A}$. To each such class one can associate a set $\mathcal{E}$ of *excluded minors*: an (unlabelled) graph is excluded if its labelled versions do not belong to $\mathcal{A}$, but the labelled versions of each of its proper minors belong to $\mathcal{A}$. A remarkable result of Robertson and Seymour states that $\mathcal{E}$ is always finite [19].

For a minor-closed class $\mathcal{A}$, we study the asymptotic properties of a random graph $\mathcal{G}_n$ taken uniformly in $\mathcal{A}_n$, the set of graphs of $\mathcal{A}$ having n vertices: what is the probability p_n that $\mathcal{G}_n$ is connected? More generally, what is the number N_n of connected components? What is the size S_n of the *root component*, that is, the component containing the vertex 1? Or the size L_n of the largest component?

Thanks to the work of McDiarmid and his collaborators, a lot is known if all excluded graphs are 2-connected: then p_n converges to constant larger than $1/\sqrt{e}$, N_n converges in law to a Poisson distribution, $n - S_n$ and $n - L_n$ converge in law to the same discrete distribution. Details are given in Section 3. If some

[†]Both authors were partially supported by a CNRS-Oxford collaboration scheme, 2012

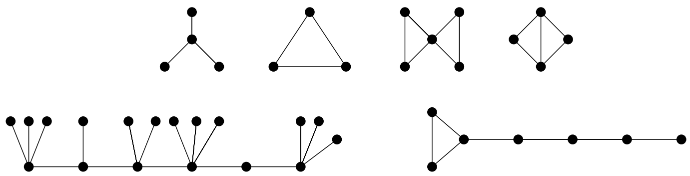
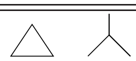
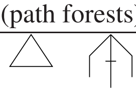
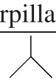
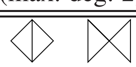
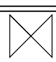


Fig. 1: Top: the 3-star, the triangle K_3, the bowtie and the diamond. Bottom: A caterpillar and the 4-spoon.

excluded minors are *not* 2-connected, the properties of $\mathcal{G}_n$ may be rather different (imagine we exclude the one edge graph...). This paper takes a preliminary step towards a classification of the possible behaviours by presenting an organized catalogue of examples. We refer to [7] for more examples, complete proofs and Boltzmann samplers for our classes of graphs.

For each class $\mathcal{A}$ that we study, we first determine the generating functions $C(z)$ and $A(z)$ that count connected and general graphs of $\mathcal{A}$, respectively. The minors that we exclude are always connected, which implies that $\mathcal{A}$ is *decomposable*: a graph belongs to $\mathcal{A}$ if and only if all its connected components belong to $\mathcal{A}$. In particular, $A(z) = \exp(C(z))$. Our exact and asymptotic results make extensive use of the

Excluded minors	$C(\rho)$	Sing. of $C(z)$	$\lim p_n$	number N_n of comp.	root comp. S_n	largest comp. L_n	Refs. and methods
2-connected	$< \infty$	?	$\geq e^{-\frac{1}{2}}$ < 1	$O(1)$ Poisson	$n - S_n$ $\to$ disc.	$n - L_n$ $\to$ disc.	[1, 16] [17]
a spoon, but no tree	$< \infty$	$(1-ze)^{3/2}$	> 0 $\leq e^{-\frac{1}{2}}$	id.	id.	id.	Sec. 8 sing. an.
 (path forests)	∞	simple pole	0	$\sqrt{n}$ gaussian	$\sqrt{n}$ xe^{-x}	$\sqrt{n}\log n$ Gumbel	Sec. 5 saddle
 (caterpillar for.)	∞	id.	0	id.	id.	?	Sec. 5 saddle
 (max. deg. 2)	∞	id. $(+\log)$	0	id.	id.	?	Sec. 5 saddle
	∞	log $(+\sqrt{\ })$	0	$\log n$ gaussian	n $\frac{1}{4}(1-x)^{-\frac{3}{4}}$	$\mathrm{PD}^{(1)}(1/4)$	Sec. 6 sing. an.
	∞	$1/\sqrt{\ }$	0	$n^{1/3}$ gaussian	$n^{2/3}$ $2\sqrt{x/\pi}e^{-x}$	?	Sec. 7 saddle

Tab. 1: Summary of the results: for each quantity N_n, S_n and L_n, we give the limit law (or its density) and an estimate of the expected value when it diverges (up to a constant). The symbol $\mathrm{PD}^{(1)}(1/4)$ stands for the first component of a Poisson-Dirichlet distribution of parameter $1/4$.

techniques of Flajolet and Sedgewick's book [11]: symbolic combinatorics, singularity analysis, saddle point method, and their application to the derivation of limit laws. We recall a few basic principles in Section 2, but then we only sketch the proofs, at best. We also need and prove two general results of independent interest related to the saddle point method (Theorems 3 and 4).

Our results are summarized in Table 1. A first dichotomy emerges: when $C(z)$ is finite at its radius of convergence ρ, the properties of $\mathcal{A}$ are qualitatively the same as in the 2-connected case (for which $C(\rho)$ is known to converge), except that the limit of p_n can be arbitrarily small (Section 8). When $C(\rho)$ diverges, a whole variety of behaviours can be observed, depending on the nature of the singularity of $C(z)$ at ρ (Sections 5 to 7): the probability p_n tends to 0 at various speeds; the number N_n of components goes to infinity at various speeds (but is invariably gaussian after normalization); the size S_n of the root component follows, after normalization, non-gaussian limit laws (gamma or beta). We only study the size L_n of the largest component in a few cases. Much remains to be done in this direction.

Let us conclude with a few words on the size of the root component. It appears that this parameter, which can be defined for any exponential family of objects, has not been studied systematically yet, and follows interesting (*i.e.*, non-gaussian!) continuous limit laws, after normalization. We are currently working on such a systematic study, in the spirit of what Bell *et al*. [4] and Gourdon [13] did for the number of components and the largest component, respectively. This project is also reminiscent of the study of the 2-connected root component in a planar map [3], which also leads to non-gaussian limit laws.

2 Generatingfunctionology for graphs

Let $\mathcal{E}$ be a finite set of (unlabelled) *connected* graphs that forms an antichain with respect to the minor order. Let $\mathcal{A}$ be the set of labelled graphs that do not contain any element of $\mathcal{E}$ as a minor. By $\mathcal{A}_n$ we denote the subset of $\mathcal{A}$ formed of graphs having n vertices and by a_n the cardinality of $\mathcal{A}_n$. The associated exponential generating function (g.f.) is $A(z) = \sum_{n\geq 0} a_n z^n/n!$. We use similar notation (c_n and $C(z)$) for the subset $\mathcal{C}$ of $\mathcal{A}$ consisting of (non-empty) connected graphs. Since the excluded minors are connected, $\mathcal{A}$ is decomposable, and $A(z) = \exp(C(z))$. Several refinements of this series are of interest, for instance the g.f. that keeps track of the number of (connected) components as well:

$$A(z,u) = \sum_{G\in\mathcal{A}} u^{c(G)} \frac{z^{|G|}}{|G|!} = \exp(uC(z)),$$

where $|G|$ is the *size* of G (the number of its vertices) and $c(G)$ the number of its components. We denote by N_n the number of components in a (uniform) random graph $\mathcal{G}_n$ of $\mathcal{A}_n$. Clearly,

$$\mathbb{P}(N_n = i) = \frac{[z^n]C(z)^i}{i![z^n]A(z)} \qquad \text{and} \qquad \mathbb{E}(N_n(N_n-1)\cdots(N_n-i+1)) = \frac{[z^n]C(z)^i A(z)}{[z^n]A(z)}. \qquad (1)$$

Several general results provide a limit law for N_n if $C(z)$ satisfies certain conditions: for instance the results of Bell *et al*. [4] that require $C(z)$ to converge at its radius; or the *exp-log schema* of [11, Prop. IX.14, p. 670], which requires $C(z)$ to have a logarithmic singularity. We use them when applicable, and prove a new result of this type, which applies when $C(z)$ diverges with an algebraic singularity (Theorem 4).

We also study the size $r(G)$ of the *root component* (the component containing the vertex 1). Let

$$\bar{A}(z,v) = \sum_{G\in\mathcal{A}, G\neq\emptyset} v^{r(G)-1} \frac{z^{|G|-1}}{(|G|-1)!} = C'(zv)A(z). \qquad (2)$$

The choice of $|G|-1$ instead of $|G|$ slightly simplifies some calculations. Note that $\bar{A}(z,1) = A'(z)$. Denoting by S_n the size of the root component in $\mathcal{G}_n$, we have

$$\mathbb{P}(S_n = k) = \frac{c_k a_{n-k}\binom{n-1}{k-1}}{a_n} \qquad \text{and} \qquad \mathbb{E}((S_n-1)\cdots(S_n-i)) = \frac{[z^{n-i-1}]C^{(i+1)}(z)A(z)}{n[z^n]A(z)}. \quad (3)$$

Surprisingly, this parameter has not been studied before. Our examples lead to non-gaussian limit laws (gamma or beta, cf. Propositions 7 or 12). In fact, the form (2) of the generating function shows that this parameter is bound to give rise to interesting limit laws, as both the location and nature of the singularity change as v moves to $1-\varepsilon$ to $1+\varepsilon$. Using the terminology of [11, Sec. IX.11], a *phase transition* occurs. We are currently working on a systematic study of this parameter for exponential objects.

3 Classes defined by 2-connected excluded minors

We assume here that the class $\mathcal{A}$ excludes at least one minor, and that all excluded minors are 2-connected. This includes the class of forests, series-parallel graphs, planar graphs, and many more. Many results are known in this case. The general picture is that $\mathcal{A}$ shares many properties with the class of forests.

Proposition 1 *The series $C(z)$ and $A(z) = e^{C(z)}$ are finite at their (positive) radius of convergence ρ. Moreover, the sequence $(a_n/n!)_n$ is* smooth*, meaning that na_{n-1}/a_n tends to ρ as n grows.*

The probability that $\mathcal{G}_n$ is connected tends to $1/A(\rho)$, which is clearly in $(0,1)$. In fact, this limit is also larger than or equal to $1/\sqrt{e}$. This value is reached when $\mathcal{A}$ is the class of forests.

The fact that $\rho > 0$ is due to Norine *et al*. [18], and holds for any proper minor-closed class. The next results are due to McDiarmid [16]. The fact that $1/A(\rho) \geq 1/\sqrt{e}$, or equivalently, that $C(\rho) \leq 1/2$, was proved independently in [1] and [14].

For forests, all results are well-know (see for instance [11, p. 132]). We have $C(z) = T(z) - T(z)^2/2$, where $T(z) = ze^{T(z)}$ counts rooted trees. The series T, C and A have radius $\rho = 1/e$, and $A(\rho) = \sqrt{e}$.

The nature of the singularity of $C(z)$ at ρ depends on the class: $(1-z/\rho)^{3/2}$ for forests (and more generally, for *subcritical* minor-closed classes [8]), but $(1-z/\rho)^{5/2}$ for planar graphs. We refer to [12] for a more detailed discussion that applies to classes that exclude 3-connected minors.

Proposition 2 *The random variable $N_n - 1$ converges in law to a Poisson distribution of parameter $C(\rho)$:*

$$\mathbb{P}(N_n = i+1) \to \frac{C(\rho)^i}{i!e^{C(\rho)}}.$$

The random variables $n - L_n$ and $n - S_n$ both converge to a discrete limit distribution X given by

$$\mathbb{P}(X = k) = \frac{1}{A(\rho)} \frac{a_k \rho^k}{k!}. \quad (4)$$

Proof: The first two results (on N_n and L_n) are due to McDiarmid [16, Cor. 1.6]. The third one is in fact equivalent to the second (the root component is, with high probability, the largest one), but we give here an independent proof, as we will recycle its ingredients later for some classes of graphs that avoid non-2-connected minors. Let $k \geq 0$ be fixed. By (3),

$$\mathbb{P}(S_n = n-k) = \frac{c_{n-k} a_k \binom{n-1}{k}}{a_n} = \frac{a_k}{k!} \frac{c_{n-k}}{a_{n-k}} \frac{(n-1)! a_{n-k}}{(n-k-1)! a_n}.$$

By Proposition 1, the term c_{n-k}/a_{n-k}, which is the probability that $\mathcal{G}_{n-k}$ is connected, converges to $1/A(\rho)$. Moreover, the sequence $a_n/n!$ is smooth, so that $\frac{(n-1)!a_{n-k}}{(n-k-1)!a_n}$ converges to ρ^k. The result follows. □

4 General tools: Hayman admissibility and extensions

We consider in Sections 5 to 7 minor-closed classes of graphs such that $C(z)$ diverges at its radius of convergence ρ. This often results in $A(z)$ diverging rapidly at ρ, and leads us to estimate a_n using the saddle point method, or rather, a black box that applies to *Hayman-admissible* (or H-admissible) series: see [11, Thm. VIII.4, p. 565]. These series have useful closure properties [ibid., p. 568]. Here is one that we did not find in the literature.

Theorem 3 *Let $A(z) = F(z)G(z)$ where $F(z)$ and $G(z)$ are power series with real coefficients and radii of convergence $0 < \rho_F < \rho_G \leq \infty$. Assume that $F(z)$ has non-negative coefficients and is H-admissible, and that $G(\rho_F) > 0$. Then $A(z)$ is H-admissible.*

We will also need a *uniform* version of Hayman-admissibility for series of the form $e^{uC(z)}$.

Theorem 4 *Let $C(z)$ be a power series with non-negative coefficients and radius of convergence ρ. Assume $A(z) = e^{C(z)}$ has radius ρ and is H-admissible. Define*

$$b(r) = rC'(r) + r^2C''(r) \quad \textit{and} \quad V(r) = C(r) - \frac{(rC'(r))^2}{rC'(r) + r^2C''(r)}.$$

Assume that, as $r \to \rho$,

$$V(r) \to +\infty, \qquad \frac{C(r)}{V(r)^{3/2}} \to 0, \qquad b(r)^{1/\sqrt{V(r)}} = O(1). \tag{5}$$

Then $A(z,u) := e^{uC(z)}$ satisfies Conditions (1)–(6), (8) *and* (9) *of* [10, Def. 1]. *If N_n is a sequence of random variables such that*

$$\mathbb{P}(N_n = i) = \frac{[z^n]C(z)^i}{i![z^n]e^{C(z)}},$$

then the mean and variance of N_n satisfy:

$$\mathbb{E}(N_n) \sim C(\zeta_n), \qquad \mathbb{V}(N_n) \sim V(\zeta_n), \tag{6}$$

where $\zeta \equiv \zeta_n$ is the unique solution in $(0, \rho)$ of the saddle point equation $\zeta C'(\zeta) = n$. Moreover, the normalized random variable $\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\mathbb{V}(N_n)}}$ converges in law to a standard normal distribution.

Proof: We carefully check the eight conditions (the only that do not come for free are (2) and (3)). As explained in [10] just below Theorem 2, they give the estimates (6) of $\mathbb{E}(N_n)$ and $\mathbb{V}(N_n)$ and imply the existence of a gaussian limit law. □

We finish with a simple but useful result on products of series [5, Thm. 2].

Proposition 5 *Let $F(z) = \sum_n f_n z^n$ and $G(z) = \sum_n g_n z^n$ be power series with radii of convergence $0 \leq \rho_F < \rho_G \leq \infty$, respectively. Suppose that $G(\rho_F) \neq 0$ and f_{n-1}/f_n approaches a limit (which is necessarily ρ_F) as $n \to \infty$. Then $[z^n]F(z)G(z) \sim G(\rho_F)f_n$.*

5 Forests of paths or caterpillars: a simple pole

Let $\mathcal{A}$ be a decomposable class (for instance defined by excluding connected minors), with g.f. $A(z) = \exp(C(z))$. Assume that $C(z)$ has a unique dominant singularity ρ, which is an isolated simple pole:

$$C(z) = \frac{\alpha}{1 - z/\rho} + D(z), \quad \text{where} \quad D(\rho) = \beta \tag{7}$$

and D has radius of convergence strictly larger than ρ. Of course, we assume $\alpha > 0$.

Proposition 6 *If the above conditions hold, then, as $n \to \infty$,*

$$c_n \sim n!\, \alpha\rho^{-n} \quad \textit{and} \quad a_n \sim n!\, \frac{\alpha^{1/4} e^{\alpha/2+\beta}}{2\sqrt{\pi} n^{3/4}} \rho^{-n} e^{2\sqrt{\alpha n}}. \tag{8}$$

In particular, the probability that $\mathcal{G}_n$ is connected tends to 0 at speed $n^{3/4} e^{-2\sqrt{\alpha n}}$.

Proof: The asymptotic behaviour of c_n follows from [11, Thm. IV.10, p. 258]. For a_n, we first write

$$A(z) = F(z)G(z) \qquad \text{with} \qquad F(z) = \exp\left(\frac{\alpha}{1 - z/\rho}\right) \qquad \text{and} \qquad G(z) = e^{D(z)}, \tag{9}$$

where $G(z)$ has radius strictly larger than ρ. To estimate the coefficients of F, we apply the ready-to-use results of Macintyre and Wilson [15, Eqs. (10)–(14)], according to which, for $\alpha > 0$ and $\gamma \geq 0$,

$$[z^n] \frac{1}{(1-z)^\gamma} \exp\left(\frac{\alpha}{1-z}\right) \sim \frac{\alpha^{1/4} e^{\alpha/2}}{2\sqrt{\pi} n^{3/4}} \left(\frac{n}{\alpha}\right)^{\gamma/2} e^{2\sqrt{\alpha n}}. \tag{10}$$

This gives

$$f_n := [z^n]F(z) \sim \frac{\alpha^{1/4} e^{\alpha/2}}{2\sqrt{\pi} n^{3/4}} \rho^{-n} e^{2\sqrt{\alpha n}}.$$

In particular, f_{n-1}/f_n tends to ρ, so that we can apply Proposition 5 to (9) and conclude. □

Proposition 7 *Assume* (7) *holds.*
1. *The mean and variance of N_n satisfy:*

$$\mathbb{E}(N_n) \sim \sqrt{\alpha n}, \qquad \mathbb{V}(N_n) \sim \sqrt{\alpha n}/2, \tag{11}$$

and the random variable $\frac{N_n - \sqrt{\alpha n}}{(\alpha n/4)^{1/4}}$ converges in law to a standard normal distribution.
2. *For $i \geq 0$, the i^{th} moment of S_n satisfies, as $n \to \infty$,*

$$\mathbb{E}(S_n^i) \sim (i+1)!(n/\alpha)^{i/2}.$$

Consequently, the normalized variable $S_n/\sqrt{n/\alpha}$ converges in distribution to a gamma(2) *law, of density xe^{-x} on $[0, \infty)$. A local limit law also holds: for $x > 0$ and $k = \lfloor x\sqrt{n/\alpha} \rfloor$,*

$$\sqrt{n/\alpha}\, \mathbb{P}(S_n = k) \sim xe^{-x}.$$

Proof: 1. We apply Theorem 4. The H-admissibility of $A(z)$ follows from Theorem 3, using (9) and the H-admissibility of $\exp(\alpha/(1-z/\rho))$ (see [11, p. 562]). Conditions (5) are then readily checked, using

$$C(r) \sim \frac{\alpha}{1-z/\rho}, \quad b(r) \sim \frac{2\alpha}{(1-z/\rho)^3} \quad \text{and} \quad V(r) \sim \frac{\alpha}{2(1-z/\rho)}.$$

We thus conclude that the normalized version of N_n converges in law to a standard normal distribution. For (11), we use (6) with the saddle point estimate $\zeta_n = \rho - \rho\sqrt{\alpha/n} + O(1/n)$.

2. We apply (3), with

$$C^{(i+1)}(z) = \frac{\alpha(i+1)!}{\rho^{i+1}(1-z/\rho)^{i+2}} + D^{(i+1)}(z). \tag{12}$$

As in the proof of Proposition 6, we combine Proposition 5, (9) and (10) to obtain

$$\frac{a_n}{(n-1)!}\mathbb{E}((S_n-1)\cdots(S_n-i)) \quad = \quad \alpha(i+1)!\,\frac{\alpha^{1/4}e^{\alpha/2+\beta}}{2\sqrt{\pi}n^{3/4}}\left(\frac{n}{\alpha}\right)^{i/2+1}\rho^{-n}e^{2\sqrt{\alpha n}}.$$

Combined with (8), this gives the limiting i^{th} moment of S_n. Since these moments characterize the above gamma distribution, we conclude [11, Thm. C.2] that $S_n/\sqrt{n/\alpha}$ converges in law to this distribution.

For the local limit law, we simply combine the first part of (3) with (8). □

We now apply these results to two classes for which $C(z)$ has a simple pole: forests of paths, and forests of *caterpillars* (a caterpillar is a tree made of a simple path to which leaves are attached; see Figure 1).

Proposition 8 *The generating functions of paths and of caterpillars are respectively*

$$C_p(z) = \frac{z(2-z)}{2(1-z)} \qquad \textit{and} \qquad C_c(z) = \frac{z^2(e^z-1)^2}{2(1-ze^z)} + ze^z - \frac{z^2}{2}.$$

For both series, Condition (7) *is satisfied and Propositions* 6 and 7 *hold. For paths we have* $\rho = 1$, $\alpha = 1/2$ *and* $\beta = 0$. *For caterpillars,* $\rho \simeq 0.567$ *is the only real such that* $\rho e^\rho = 1$,

$$\alpha = \frac{(1-\rho)^2}{2(1+\rho)} \simeq 0.06 \quad \textit{and} \quad \beta = \frac{\rho\left(10+3\,\rho-4\,\rho^2-\rho^3\right)}{4\,(1+\rho)^2} \simeq 0.59.$$

For forests of paths, we have also studied the size L_n of the largest component.

Proposition 9 *In forests of paths, the size of the largest component converges to a Gumbel distribution:*

$$\mathbb{P}\left(\frac{L_n - \sqrt{n/2}\log n}{\sqrt{n/2}} < x\right) \to \exp\left(-\frac{e^{-x/2}}{\sqrt{2}}\right).$$

Proof: The generating function of paths of size less than k is $C^{[k]}(z) = z/2 + (z-z^k)/(2(1-z))$. We then use Cauchy's formula and a saddle point approach. □

Graphs with maximal degree 2: a simple pole plus a logarithm. Let $\mathcal{A}$ be the class of graphs avoiding the 3-star. The connected components of such graphs are paths and cycles. The series $C(z)$ has now, in addition to a simple pole, a logarithmic singularity at its radius. The logarithm changes the asymptotic behaviour of the numbers a_n, but the other results remain unaffected. The proofs are very similar to the above ones.

Proposition 10 *The generating function of connected graphs of $\mathcal{A}$ is*

$$C(z) = \frac{z(2-z+z^2)}{4(1-z)} + \frac{1}{2}\log\frac{1}{1-z}.$$

The generating function of graphs of $\mathcal{A}$ is $A(z) = e^{C(z)}$. We have, for large n,

$$c_n = \frac{n!}{2} + \frac{(n-1)!}{2} \qquad \textit{and} \qquad a_n \sim n!\,\frac{1}{2\sqrt{e\pi}n^{1/2}}e^{\sqrt{2n}}.$$

In particular, the probability that $\mathcal{G}_n$ is connected tends to 0 at speed $n^{1/2}e^{-\sqrt{2n}}$.

The number of components and the size of the root component behave as in Proposition 7, with $\alpha = 1/2$.

6 Excluding the diamond and the bowtie: a logarithm dominates

Let $\mathcal{A}$ be the class of graphs avoiding the diamond and the bowtie (shown in Figure 1). The connected components are trees or unicyclic graphs, and have been counted a long time ago by Wright [20].

Proposition 11 *Let $T(z) = ze^{T(z)}$ be the g.f. of rooted trees. The g.f. of connected graphs of $\mathcal{A}$ is*

$$C(z) = \frac{T}{2} - \frac{3T^2}{4} + \frac{1}{2}\log\frac{1}{1-T}.$$

The generating function of graphs of $\mathcal{A}$ is $A(z) = e^{C(z)}$. As $n \to \infty$,

$$c_n \sim n!\frac{e^n}{4n} \quad \textit{and} \quad a_n \sim n!\,\frac{1}{(2e)^{1/4}\Gamma(1/4)}\frac{e^n}{n^{3/4}}. \tag{13}$$

In particular, the probability that $\mathcal{G}_n$ is connected tends to 0 at speed $n^{-1/4}$ as $n \to \infty$.

Proof: A connected graph of $\mathcal{A}$ is either an (unrooted) tree (with g.f. $T - T^2/2$), or consists of a cycle, in which each vertex is replaced by a rooted tree. The generating function of cycles is

$$Cyc(z) = \frac{1}{2}\sum_{n\geq 3}\frac{z^n}{n} = \frac{1}{2}\left(\log\frac{1}{1-z} - z - \frac{z^2}{2}\right), \tag{14}$$

and this gives the expression of $C(z)$. We then estimate c_n and a_n via *singularity analysis* [11, VI.4]. □

Proposition 12 1. *The mean and variance of N_n satisfy $\mathbb{E}(N_n) \sim \mathbb{V}(N_n) \sim \log n/4$, and the random variable $\frac{N_n - \log n/4}{\sqrt{\log n/4}}$ converges in law to a standard normal distribution.*

2. *For $i \geq 0$, the i^{th} moment of S_n satisfies, as $n \to \infty$,*

$$\mathbb{E}(S_n^i) \sim \frac{\Gamma(5/4)i!}{\Gamma(i+5/4)}n^i.$$

Consequently, the normalized variable S_n/n *converges in distribution to a beta law, of density* $(1-x)^{-3/4}/4$ *on* $[0,1]$. *A local limit law also holds: for* $x \in (0,1)$ *and* $k = \lfloor xn \rfloor$,

$$n\,\mathbb{P}(S_n = k) \sim \frac{1}{4}(1-x)^{-3/4}.$$

3. *The normalized variable* L_n/n *converges in law to the first component of a Poisson-Dirichlet distribution of parameter* $1/4$.

Proof: Once the singular expansion of $C(z)$ is obtained, the first result follows from [11, Prop. IX.14, p. 670]. To study the moments of S_n, we apply (3). Using $T(z) = ze^{T(z)}$, we obtain, for $i \geq 1$,

$$C^{(i+1)}(z) \sim \frac{i!}{4}\left(\frac{e}{1-ze}\right)^{i+1}.$$

The estimate of the i^{th} moment of S_n then follows again from singularity analysis. Since these moments characterize the above beta distribution, we conclude [11, Thm. C.2] that S_n/n converges in law to this distribution. For the local limit law, we start from (3), and use (13).

Finally, the third result follows from general results on *logarithmic structures* [2]. □

Remark. A subdominant term in $\sqrt{1-ze}$ occurs in the expansion of $C(z)$, but has no influence on the asymptotic results. They would be the same (with possibly different constants) for any $C(z)$ having a purely logarithmic singularity.

7 Excluding the bowtie: a singularity in $(1-z/\rho)^{-1/2}$

We now denote by $\mathcal{A}$ the class of graphs avoiding the bowtie (shown in Figure 1).

Proposition 13 *Let* $T(z) = ze^{T(z)}$ *be the g.f. of rooted trees. The g.f. of connected graphs in* $\mathcal{A}$ *is*

$$C(z) = \frac{T^2\left(1 - T + T^2\right)\mathrm{e}^T}{1-T} + \frac{1}{2}\log\left(\frac{1}{1-T}\right) + \frac{T\left(12 - 54\,T + 18\,T^2 - 5\,T^3 - T^4\right)}{24(1-T)}.$$

As $n \to \infty$,

$$c_n \sim n!\,\frac{e-5/4}{\sqrt{2\pi}}\frac{e^n}{\sqrt{n}} \quad \text{and} \quad a_n \sim n!\,\frac{(e-5/4)^{1/6}e^{19/8-11e/3}}{\sqrt{6\pi}}\frac{e^n}{n^{2/3}}\,e^{3(e-5/4)^{2/3}n^{1/3}/2}.$$

Proof: This is the most delicate enumeration result of the paper. We have

$$C(z) = T(z) - T(z)^2/2 + \bar{C}(T(z)),$$

where $\bar{C}(z)$ counts graphs with minimal degree 2 avoiding the bowtie. After studying the properties of these graphs, we conclude that they are either cycles, or K_4 with one edge possibly replaced by a chain of vertices of degree 2, or the graphs of Figure 2. Counting these various classes gives $\bar{C}(z)$, and thus $C(z)$. □

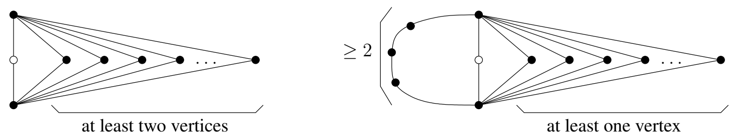


Fig. 2: Some graphs avoiding the bowtie. The white vertex is optional.

Proposition 14 1. *The mean and variance of N_n satisfy:*

$$\mathbb{E}(N_n) \sim (e - 5/4)^{2/3} n^{1/3}, \qquad \mathbb{V}(N_n) \sim \frac{2}{3}(e - 5/4)^{2/3} n^{1/3},$$

and the random variable $\frac{N_n - \mathbb{E}(N_n)}{\sqrt{\mathbb{V}(N_n)}}$ converges in law to a standard normal distribution.
2. *For $i \geq 0$, the i^{th} moment of S_n satisfies, as $n \to \infty$,*

$$\mathbb{E}(S_n^i) \sim \frac{\Gamma(i + 3/2)}{\Gamma(3/2)} \left(\frac{2n^{2/3}}{(e - 5/4)^{2/3}} \right)^i.$$

Consequently, the normalized variable $(e - 5/4)^{2/3} S_n / (2n^{2/3})$ converges in distribution to a gamma$(3/2)$ *law, of density $2\sqrt{x}e^{-x}/\sqrt{\pi}$ on $[0, \infty)$. A local limit law also holds: for $x > 0$ and $k = \lfloor x \frac{2n^{2/3}}{(e-5/4)^{2/3}} \rfloor$,*

$$\frac{2n^{2/3}}{(e - 5/4)^{2/3}} \mathbb{P}(S_n = k) \sim 2\sqrt{\frac{x}{\pi}} e^{-x}.$$

8 When trees dominate: $C(z)$ converges at ρ

Let $\mathcal{A}$ be a decomposable class of graphs (for instance, a class defined by excluding *connected* minors) with set of components $\mathcal{C}$. Assume that $\mathcal{C}$ contains all trees (counted by $T - T^2/2$), and that

$$C(z) = T(z) - T(z)^2/2 + D(z), \tag{15}$$

where $D(z)$ has radius strictly larger than $1/e$ (the radius of T). We say that $\mathcal{A}$ is *dominated by trees*. Some examples are presented below. In this case, the properties that hold for forests (Section 3) still hold, except that the limit of c_n/a_n is now smaller than $1/\sqrt{e}$.

Proposition 15 *Assume $\mathcal{A}$ is dominated by trees. As $n \to \infty$,*

$$c_n \sim n! \frac{e^n}{\sqrt{2\pi} n^{5/2}} \quad \text{and} \quad a_n \sim A(1/e) c_n.$$

In particular, the probability that $\mathcal{G}_n$ is connected tends to $1/A(1/e) = e^{-1/2 - D(1/e)}$.
The asymptotic behaviours of N_n, L_n and S_n are described by Proposition 2, with $\rho = 1/e$.

Proof: The asymptotic behaviours of c_n and a_n are obtained via singularity analysis. For N_n, we can either start from (1) and apply singularity analysis, or use directly [4, Thm. 2]. For S_n, the two ingredients used in the proof of Proposition 2, namely smoothness of $a_n/n!$ and convergence of c_n/a_n, still hold here. For L_n, we use the fact that the root component is with high probability the biggest one. □

We now give examples where trees dominate.

Proposition 16 *Let $k \geq 1$. Let $\mathcal{A}$ be a minor-closed class of graphs containing all trees, but not the k-spoon (shown in Figure 1). Then $\mathcal{A}$ is dominated by trees, and the results of Proposition 15 apply.*

Proof: We partition the set $\mathcal{C}$ of connected graphs of $\mathcal{A}$ into three subsets: the set of trees, counted by $T - T^2/2$, the set $\mathcal{C}_1$ of unicyclic graphs (counted by C_1), and finally the set $\mathcal{C}_2$ containing graphs with at least two cycles (counted by C_2). We prove that C_1 has radius strictly larger than $1/e$, and that C_2 is entire. □

Proposition 17 *Let T_k be the g.f. of rooted trees of height less than k. That is, $T_k = ze^{T_{k-1}}$ with $T_1 = z$. Let $\mathcal{A}^{(k)}$ be the class of graphs avoiding the diamond, the bowtie and the k-spoon. Then (15) holds with*

$$D(z) \equiv D^{(k)}(z) = \frac{1}{2}\left(\log \frac{1}{1 - T_k(z)} - T_k(z) - \frac{T_k(z)^2}{2}\right).$$

The class $\mathcal{A}^{(k)}$ is dominated by trees, and Proposition 15 applies. In particular, the probability that a random graph of $\mathcal{A}_n^{(k)}$ is connected tends to $1/A^{(k)}(1/e)$ as $n \to \infty$. This value tends to 0 as k increases.

By specializing the proof of Proposition 13, we have also counted graphs avoiding the 2-spoon.

Proposition 18 *The g.f. of connected graphs avoiding the 2-spoon satisfies (15) with*

$$D(z) = \frac{1}{2}\left(\log \frac{1}{1 - ze^z} - ze^z - \frac{z^2 e^{2z}}{2}\right) + \frac{z^4}{4!} + z^2 e^{2z}\left(e^z - 1 - z - \frac{z^2}{4}\right).$$

9 Final comments

The nature of the dominant singularities of $C(z)$ is clearly a central parameter of the class, as the quantities N_n and S_n seem to depend largely on it (see Table 1). Is it possible to determine the nature of this singularity from the properties of the excluded minors? For instance, $C(\rho)$ is finite when all excluded minors are 2-connected, but Section 8 shows that this happens as well with non-2-connected minors. Which excluded minors give rise to a simple pole in $C(z)$? to a logarithmic singularity? to a singularity in $(1 - z/\rho)^{-1/2}$?

Other parameters. We have focussed on certain parameters that are well understood for 2-connected excluded minors. But others have been investigated in different contexts: the number of edges, the size of the largest 2-connected component, the distribution of vertex degrees [6, 8, 9, 12]. When specialized to the theory of minor-closed classes, these papers generally assume that all excluded minors are (at least) 2-connected. Including (some of) these parameters in our results may be the topic of future work.

Acknowledgements. KW would like to thank C. McDiarmid for many inspiring discussions as well as for constant support. We also thank J.-F. Marckert and B. Salvy for pointing out Refs. [2] and [15].

References

[1] L. Addario Berry, C. McDiarmid, and B. Reed. Connectivity for bridge-addable monotone graph classes. *Combin. Probab. Comput.*, 21(6):803–815, 2012. ArXiv:1110.0009.

[2] R. Arratia, A. D. Barbour, and S. Tavaré. *Logarithmic combinatorial structures: a probabilistic approach*. EMS Monographs in Mathematics. European Mathematical Society (EMS), Zürich, 2003.

[3] C. Banderier, P. Flajolet, G. Schaeffer, and M. Soria. Random maps, coalescing saddles, singularity analysis, and Airy phenomena. *Random Structures Algorithms*, 19(3-4):194–246, 2001.

[4] J. P. Bell, E. A. Bender, P. J. Cameron, and L. B. Richmond. Asymptotics for the probability of connectedness and the distribution of number of components. *Electron. J. Combin.*, 7:RP 33, 2000.

[5] E. A. Bender. Asymptotic methods in enumeration. *SIAM Rev.*, 16:485–515, 1974.

[6] N. Bernasconi, K. Panagiotou, and A. Steger. The degree sequence of random graphs from subcritical classes. *Combin. Probab. Comput.*, 18(5):647–681, 2009.

[7] M. Bousquet-Mélou and K. Weller. Asymptotic properties of some minor-closed classes of graphs. ArXiv:1303.3836, 2013.

[8] M. Drmota, É. Fusy, M. Kang, V. Kraus, and J. Rué. Asymptotic study of subcritical graph classes. *SIAM J. Discrete Math.*, 25(4):1615–1651, 2011.

[9] M. Drmota, O. Giménez, and M. Noy. Degree distribution in random planar graphs. *J. Combin. Theory Ser. A*, 118(7):2102–2130, 2011.

[10] M. Drmota, B. Gittenberger, and T. Klausner. Extended admissible functions and Gaussian limiting distributions. *Math. Comp.*, 74(252):1953–1966 (electronic), 2005.

[11] P. Flajolet and R. Sedgewick. *Analytic combinatorics*. Cambridge Univ. Press, Cambridge, 2009.

[12] O. Giménez, M. Noy, and J. Rué. Graph classes with given 3-connected components: Asymptotic enumeration and random graphs. *Random Structures Algorithms*, to appear.

[13] X. Gourdon. Largest component in random combinatorial structures. *Discrete Math.*, 180(1-3):185–209, 1998.

[14] M. Kang and K. Panagiotou. On the connectivity of random graphs from addable classes. *J. Combin. Theory Ser. B*, 103(2):306–312, 2013.

[15] A. J. Macintyre and R. Wilson. Operational methods and the coefficients of certain power series. *Math. Ann.*, 127:243–250, 1954.

[16] C. McDiarmid. Random graphs from a minor-closed class. *Combin. Probab. Comput.*, 18(4):583–599, 2009.

[17] C. McDiarmid, A. Steger, and D. J. A. Welsh. Random planar graphs. *J. Combin. Theory Ser. B*, 93(2):187–205, 2005.

[18] S. Norine, P. Seymour, R. Thomas, and P. Wollan. Proper minor-closed families are small. *J. Combin. Theory Ser. B*, 96(5):754–757, 2006.

[19] N. Robertson and P.D. Seymour. Graph minors I–XX. *J. Combin. Theory Ser. B*, 1983–2004.

[20] E. M. Wright. The number of connected sparsely edged graphs. *J. Graph Th.*, 1(4):317–330, 1977.