Extremal Statistics on Non-Crossing Configurations*

Michael Drmota† Anna de Mier‡ Marc Noy§

Abstract

We analyze extremal statistics in non-crossing configurations on the n vertices of a convex polygon. We prove that the maximum degree and the largest component are of logarithmic order, and that, suitably scaled, they converge to a well-defined constant. We also prove that the diameter is of order $\sqrt{n}$. The proofs are based on singularity analysis, an application of the first and second moment method, and on the analysis of iterated functions.

1 Introduction and preliminaries

Let $p_1, \ldots, p_n$ be the vertices of a convex polygon in the plane, labelled counterclockwise. A *non-crossing graph* (or configuration) is a graph on these vertices such that when the edges are drawn as straight lines the only intersections occur at vertices. The *root* of a graph is vertex p_1. We call the edge p_1p_n (if present) the root edge.

From now on, all graphs are assumed to be non-crossing graphs. A triangulation is a graph with the maximum number of edges and it is characterized by the fact that all internal faces are triangles. A dissection is a graph containing all the boundary edges $p_1p_2, p_2p_3, \ldots, p_np_1$; a single edge p_1p_2 is also considered a dissection (see Figure 1). From a graph theoretic point of view, dissections are the same as 2-connected graphs. The root region of a dissection is the internal region adjacent to the root edge.

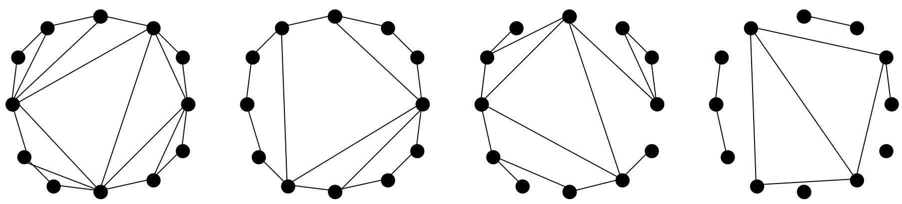

Figure 1: From left to right: a triangulation, a 2-connected graph (dissection), a connected graph, and an arbitrary graph.

The enumerative theory of non-crossing configurations is an old subject, going back to Euler; see, for instance, Comtet's book [3] for an account of classical results. Flajolet and Noy [10] reexamined these problems using the tools from analytic combinatorics [12] in a unified way. They showed that for all natural classes under consideration the number of non-crossing graphs with n vertices is asymptotically of the form

$$cn^{-3/2}\gamma^n,$$

for some positive constants c and γ. In addition, many basic parameters obey a Gaussian limit law with linear expectation and variance. These include: number of edges, number of components,

*Research partially supported by Projects MTM2011-24097 and DGR2009-SGR1040.
†Technische Universität Wien, Vienna. michael.drmota@tuwien.ac.at
‡Universitat Politècnica de Catalunya, Barcelona. anna.de.mier@upc.edu
§Universitat Politècnica de Catalunya, Barcelona. marc.noy@upc.edu

number of leaves in trees and number of blocks in partitions. The proofs in [10] are based on perturbation of singularities and extensions of the Central Limit Theorem.

In this paper we take a step further and analyze more complex parameters, specially *extremal* parameters. Some results have been obtained previously for triangulations [5, 13] and trees [4, 15], but here we aim at a systematic treatment of the subject, covering the most important extremal parameters and proving limit laws whenever possible. Our main results are summarized as follows.

- For graphs, connected graphs and 2-connected graphs, the degree of the root vertex converges to a discrete law. More precisely, if p_k is the probability that the root has degree k, then

$$\sum p_k w^k = \frac{A(w)}{(1-qw)^2},$$

  where $A(w)$ is a polynomial of degree two and q is a quadratic irrational with $0 < q < 1$. It follows that the tail of the distribution satisfies, for a suitable constant $c > 0$,

$$p_k \sim ckq^k, \qquad \text{as } k \to \infty.$$

- For graphs, connected graphs and 2-connected graphs, the maximum degree Δ_n is of logarithmic order. More precisely, for each class under consideration there exists a well-defined constant $c > 0$ such that

$$\frac{\Delta_n}{\log n} \to c \qquad \text{in probability.}$$

- The largest connected component M_n in graphs is of logarithmic order: there exists a well-defined constant $c > 0$ such that

$$\frac{M_n}{\log n} \to c \qquad \text{in probability.}$$

- For triangulations, connected graphs and 2-connected graphs, the diameter D_n is of order $\sqrt{n}$. For each class under consideration, there exist constants $0 < c_1 < c_2$ such that

$$c_1\sqrt{n} < \mathbb{E}D_n < c_2\sqrt{n}.$$

These results reflect the tree-like nature of non-crossing configurations. In particular, the diameter is of order $\sqrt{n}$, like the height of plane trees. The expected maximum degree in triangulations was shown to be asymptotically $\log n/\log 2$ in [5] (much more precise results were obtained in [13]). To our knowledge, the diameter of random configurations has not been studied before, even in the basic case of triangulations.

In the rest of this section we collect several technical preliminaries needed in the paper. In Section 2 we analyze the degree of the root vertex. Sections 3 and 4 are devoted to the maximum degree and the size of the largest component, respectively, and are based on the first and second moment method. Finally, in Section 5 we analyze the diameter, making use of iterated functions.

1.1 Generating functions

We denote by $G(z)$ and $C(z)$ the generating functions for arbitrary and connected graphs, respectively, counted by the number of vertices. Furthermore, let $B(z)$ be the generating function for 2-connected graphs, where z marks the number of vertices minus one. We have the following relations for the generating functions. The first one reflects the decomposition of a dissection as a sequence of dissections attached to the root region, as in [10]:

$$B(z) = z + \frac{B(z)^2}{1 - B(z)}. \tag{1}$$

Hence we obtain

$$B(z) = \frac{1 + z - \sqrt{1 - 6z + z^2}}{4} \tag{2}$$

which has a square-root singularity at $z = 3 - 2\sqrt{2}$. The next equation encodes the decomposition of a connected graph into 2-connected components (blocks):

$$C(z) = \frac{z}{1 - B(C(z)^2/z)}. \tag{3}$$

Indeed, a connected graph consists of a root and a sequence of blocks containing the root, in which each vertex is substituted by a pair of connected graphs (to the left and to the right) with one vertex identified. Eliminating B we obtain

$$C(z)^3 + C(z)^2 - 3zC(z) + 2z^2 = 0, \tag{4}$$

in accordance with [10]. The function $C(z)$ has a dominant singularity at $z = \sqrt{3}/18$, also of square-root type. The decomposition of an arbitrary graph into connected components gives, as in [10], the equation

$$G(z) = 1 + C(zG(z)). \tag{5}$$

It follows that $G(z)$ satisfies

$$G^2 + (2z^2 - 3z - 2)G + 3z + 1 = 0,$$

which leads to

$$G(z) = \frac{2 + 3z - 2z^2 - z\sqrt{1 - 12z + 4z^2}}{2}. \tag{6}$$

This function is singular at $z = 3/2 - \sqrt{2}$. Let us remark that

$$G(z) = 1 + z - 2z^2 + 2zB(2z).$$

We summarize the above discussion in the following table:

Function	Equation	Singularity
$G(z)$	$G^2 + (2z^2 - 3z - 2)G + 3z + 1 = 0$	$\frac{3}{2} - \sqrt{2}$
$C(z)$	$C^3 + C^2 - 3zC + 2z^2 = 0$	$\frac{\sqrt{3}}{18}$
$B(z)$	$2B^2 - (1 + z)B + z = 0$	$3 - 2\sqrt{2}$

1.2 Singularity analysis

In what follows we will make use of power series of the *square-root type*. They are power series $y(z)$ with a square-root singularity at $z_0 > 0$, that is, $y(z)$ admits a local representation of the form

$$y(z) = g(z) - h(z)\sqrt{1 - z/z_0}, \tag{7}$$

for $|z - z_0| < \varepsilon$, for some $\varepsilon > 0$ and $|\arg(z - z_0)| > 0$, where $g(z)$ and $h(z)$ are analytic and non-zero at z_0. Moreover, $y(z)$ can be analytically continued to the region

$$D(z_0, \varepsilon) = \{z \in \mathbb{C} : |z| < z_0 + \varepsilon\} \setminus [z_0, \infty). \tag{8}$$

We denote $[z^n]y(z)$ the n-th coefficient in $y(z)$. The Transfer Theorem of Flajolet and Odlyzko (see [12]), implies the following estimate:

$$[z^n]y(z) \sim \frac{h(z_0)}{2\sqrt{\pi}} n^{-3/2} z_0^{-n}.$$

For the analysis of the root degree we need a particular application of this estimate, whose proof can be be found in [8].

Lemma 1.1. *Let $f(z) = \sum_{n\ge 0} a_n z^n$ denote the generating function of a sequence a_n of non-negative real numbers and assume that $f(z)$ has exactly one dominating square-root singularity at $z = \rho$ of the form*

$$f(z) = g(z) - h(z)\sqrt{1 - z/\rho},$$

where $g(z)$ and $h(z)$ are analytic at $z = \rho$ and $f(z)$ has an analytic continuation to the region $D(z_0, \varepsilon)$ for some $\varepsilon > 0$. Further, let $H(z,t)$ denote a function that is analytic for $|z| < \rho + \varepsilon$ and $|t| < f(\rho) + \varepsilon$, and such that $H_t(\rho, f(\rho)) \neq 0$, where H_t stands for the derivative with respect to t. Then the function

$$f_H(z) = H(z, f(z))$$

has a power series expansion $f_H(z) = \sum_{n\ge 0} b_n z^n$ and the coefficients b_n satisfy

$$\lim_{n\to\infty} \frac{b_n}{a_n} = H_t(\rho, f(\rho)). \tag{9}$$

In our applications, $H(z,t)$ depends also on an additional variable w, that can be treated as a parameter.

1.3 First and second moment method

In order to obtain results for the maximum statistics of the root degree we follow the methods of [9]. They are based on the so-called *first and second moment method* [1].

Lemma 1.2. *Let X be a discrete random variable on non-negative integers with finite first moment. Then*

$$\mathbb{P}\{X > 0\} \le \min\{1, \mathbb{E}\, X\}.$$

Furthermore, if X is a non-negative random variable which is not identically zero and has finite second moment then

$$\mathbb{P}\{X > 0\} \ge \frac{(\mathbb{E}\, X)^2}{\mathbb{E}\,(X^2)}.$$

We apply this principle for the random variable $Y_{n,k}$ that counts the number of vertices of degree greater than k in a random graph with n vertices. This variable is closely related to the maximum degree Δ_n by

$$Y_{n,k} > 0 \quad \Longleftrightarrow \quad \Delta_n > k.$$

One of our aims is to get bounds for the expected maximum degree $\mathbb{E}\,\Delta_n$. Due to the relation

$$\mathbb{E}\,\Delta_n = \sum_{k\ge 0} \mathbb{P}\{\Delta_n > k\} = \sum_{k\ge 0} \mathbb{P}\{Y_{n,k} > 0\}$$

we are led to estimate the probabilities $\mathbb{P}\{Y_{n,k} > 0\}$, which can be done via the first and second moment methods by estimating the first two moments

$$\mathbb{E}\, Y_{n,k} \quad \text{and} \quad \mathbb{E}\, Y_{n,k}^2.$$

Actually, we work with the probabilities $d_{n,k}$ that a random vertex in a graph of size n has degree k. They are related to the first moment by

$$\mathbb{E}\, Y_{n,k} = n \sum_{\ell > k} d_{n,\ell}. \tag{10}$$

Similarly we deal with probabilities $d_{n,k,\ell}$ that two different randomly selected vertices have degrees k and ℓ. Here we have

$$\mathbb{E}\, Y_{n,k}^2 = n \sum_{\ell > k} d_{n,\ell} + n(n-1) \sum_{\ell_1, \ell_2 > k} d_{n,\ell_1,\ell_2}.$$

The following technical lemma subsumes some results from [9] and it is stated according to our needs in this paper. The full proof can be found in Theorem 1.1 and Lemmas 3.1 and 3.2 from [9]. It guarantees that, if the generating functions associated to the degree of the root and the degree of a secondary vertex have a certain local expansion of square-root type, then *automatically* the maximum degree is $c \log n$ for a well-defined constant c. For our functions these conditions are not difficult to check.

Lemma 1.3 (Master Theorem). *Let $f(z,w) = \sum_{n,k} f_{n,k} z^n w^k$ be a double generating function and $g(z,w,t) = \sum_{n,k,\ell} g_{n,k,\ell} z^n w^k t^\ell$ be a triple generating function of non-negative numbers $f_{n,k}$ and $g_{n,k,\ell}$ such that the probabilities $d_{n,k}$ and $d_{n,k,\ell}$ that a random vertex in a graph of size n has degree k, and that two different randomly selected vertices have degrees k and ℓ, respectively, are given by*

$$d_{n,k} = \frac{f_{n,k}}{f_n} \quad \text{and} \quad d_{n,k,\ell} = \frac{g_{n,k,\ell}}{g_n},$$

where $f_n = \sum_k f_{n,k}$ and $g_n = \sum_{k,\ell} g_{n,k,\ell}$. Suppose that $f(z,w)$ can be represented as

$$f(z,w) = \frac{G(z,Z,w)}{1-y(z)w}, \tag{11}$$

where $Z = \sqrt{1-z/z_0}$, $y(z)$ is a power series with non-negative coefficients of square-root type,

$$y(z) = g(z) - h(z)\sqrt{1-z/z_0},$$

where $0 < g(z_0) < 1$ and the function $G(z,v,w)$ is analytic in the region

$$D' = \{(z,v,w) \in \mathbb{C}^3 : |z| < z_0 + \eta, |v| < \eta, |w| < 1/g(z_0) + \eta\}$$

for some $\eta > 0$ and satisfies $G(z_0, 0, 1/g(z_0)) \neq 0$.

Furthermore suppose that $g(z,w,t)$ can be represented as

$$g(z,w,t) = \frac{H(z,Z,w,t)}{Z\,(1-y(z)w)^2(1-y(z)t)^2}, \tag{12}$$

where the function $H(z,v,w,t)$ is non-zero and analytic at $(z,0,w,t) = (z_0, 0, 1/g(z_0), 1/g(z_0))$.

Let Δ_n denote the maximum degree of a random graph in this class of size n. Then we have

$$\frac{\Delta_n}{\log n} \to \frac{1}{\log g(z_0)^{-1}} \qquad \textit{in probability} \tag{13}$$

$$\mathbb{E}\,\Delta_n \sim \frac{1}{\log g(z_0)^{-1}} \log n \qquad (n \to \infty). \tag{14}$$

We do not go into the details of the proof. We just mention that the main intermediate step is to prove that

$$d_{n,k} \sim ckg(z_0)^k \quad \text{and} \quad d_{n,k,\ell} \sim d_{n,k} d_{n,\ell} \sim c^2 k\ell g(z_0)^{k+\ell},$$

uniformly for $k \le C \log n$ and $\ell \le C \log n$ (for a certain constant $c > 0$ and an arbitrary constant $C > 0$). With the help of these asymptotic relations one gets asymptotic expansions for $\mathbb{E}\,Y_{n,k}$ and $\mathbb{E}\,Y_{n,k}^2$ that can be used to estimate the probabilities $\mathbb{P}\{\Delta_n > k\}$ from below and above and which lead to the final result. We direct the reader to [9] for a detailed discussion of the Master Theorem and its proof, based on the estimation of Cauchy integrals.

The key conditions in the lemma are equations (11) and (12) for the shape of the generating functions marking the degree of one and two vertices, respectively, and the square-root type of the univariate function $y(z)$. In (11) we find a linear factor in w in the denominator, and in (12) we find a quadratic factor both in w and in t, and a factor $\sqrt{1-z/z_0}$. In Section 3 we point out how these conditions are satisfied in our case.

1.4 Iterated functions

The results on the diameter are based on the following lemma on iterated functions. Such a lemma was first studied in the analysis of the height of random trees, as in [11]. A basic example is given by the class $\mathcal{T}$ of plane trees (rooted trees in which the children of a node are ordered from left to right). Let $\mathcal{T}^{[k]}$ the class of plane trees with height at most k. In the terminology of the symbolic method [12], we have the decomposition

$$\mathcal{T}^{[k+1]} = \{\rho\} \times Seq(\mathcal{T}^{[k]}),$$

where Seq denotes the sequence construction and ρ represent the root of the tree. This is because the subtrees attached to the children of the root of a tree in $\mathcal{T}^{[k+1]}$ have height at most k. This translates into an equation for the associated generating functions:

$$T^{[k+1]}(z) = \frac{z}{1 - T^{[k]}(z)}.$$

This is precisely the kind of equations that are covered in the lemma. Notice also that the generating function of trees with height exactly k is equal to $T^{[k]} - T^{[k-1]}$.

Lemma 1.4. *Suppose that $F(z,t)$ is an analytic function at $(z,t) = (0,0)$ such that the equation*

$$T(z) = F(z, T(z)) \tag{15}$$

has a solution $T(z)$ that is analytic at $z = 0$ and has non-negative Taylor coefficients. Suppose that $T(z)$ has a square-root singularity at $z = z_0$ and can be continued to a region of the form (8), such that $F_t(z_0, t_0) = 1$, $F_z(z_0, t_0) \neq 0$, and $F_{tt}(z_0, t_0) \neq 0$, where $t_0 = T(z_0)$.

Let $T_0(z)$ be a power series with $0 \leq_c T_0(z) \leq_c T(z)$[1] such that $T_0(z)$ is analytic at $z = z_0$, and let $T_k(z)$, $k \geq 1$ be iteratively defined by

$$T_k(z) = F(z, T_{k-1}(z)). \tag{16}$$

Assume that $T_{k-1}(z) \leq_c T_k(z) \leq_c T(z)$.

Let H_n be an integer valued random variable that is defined by

$$\mathbb{P}\{H_n \leq k\} = \frac{[z^n]\, T_k(z)}{[z^n]\, T(z)}$$

for those n with $[z^n]\, T(z) > 0$. Then

$$\mathbb{E}\, H_n \sim \sqrt{\frac{2\pi}{z_0 F_z(z_0, t_0) F_{tt}(z_0, t_0)}}\; n^{1/2}.$$

In the previous statement, H_n is a kind of generalized height parameter. Recursion (16) is analogous to our previous equation for trees of height at most k. The quotient $[z^n]T_k(z)/[z^n]T(z)$ is the proportion of elements of size n and 'height' at most k among the total number of elements of size n. The series $T(z)$ is in a sense a 'limit' of the $T_k(z)$, and is defined by the fixed point equation (15). The lemma is a direct extension of the results in [11]; see also Theorems 4.8 and 4.59 in [6] for the proof techniques.

The previous lemma can be more precise, in the sense that $H_n/\sqrt{n}$ converges to the maximum of a Brownian excursion of duration one (a Brownian motion conditioned to be positive and to take the value 0 at time 1). However, the estimate on the expectation is sufficient for our needs.

[1]The notation $A(z) \leq_c B(z)$ means that the coefficients of $A(z) = \sum_n a_n z^n$ and $B(z) = \sum_n b_n z^n$ satisfy $a_n \leq b_n$.

2 Degree of the root

In this section we determine the asymptotic distribution of the degree of the root vertex in graphs, and also in connected and 2-connected graphs. The probability generating function is in all cases a rational function with a quadratic factor in the denominator. This implies that the probability that the root has degree k is asymptotically, for large k, of the form

$$ckq^k,$$

where $q < 1$ is a constant that depends on the class of graphs under consideration.

Let $B(z, w)$, $C(z, w)$ and $G(z, w)$ be the corresponding generating functions, where w counts the degree of the root vertex p_1. A simple adaptation of the basic equations for $B(z), C(z)$ and $G(z)$ gives

$$B(z, w) = wz + \frac{wB(z, w)B(z)}{1 - B(z)}$$

$$C(z, w) = \frac{z}{1 - B(C(z)^2/z, w)} \tag{17}$$

$$G(z, w) = 1 + C(zG(z), w). \tag{18}$$

In the first equation, the term $B(z, w)$ on the right singles out the only dissection that contributes to the degree of the root vertex. In the second equation, the sum of the degrees of the blocks containing the root is added. And in the last equation, only the component containing the root contributes to its degree.

In particular, we have

$$B(z, w) = \frac{zw(1 - B(z))}{1 - (1 + w)B(z)} = \frac{wz}{1 - \frac{w}{2}(1 - z - \sqrt{1 - 6z + z^2})}. \tag{19}$$

Let d_k^B be the limiting probability that the root vertex in a 2-connected graph has degree k, that is,

$$d_k^B = \lim_{n\to\infty} \frac{[z^n][w^k]B(z, w)}{[z^n]B(z)},$$

and define d_k^C and d_k^G analogously.

Next we find the probability generating functions for the d_k^B, d_k^C and d_k^G, and show that they indeed exist. The following result can be found also in [8] and [2].

Theorem 2.1. *The limiting distribution of the root degree in 2-connected graphs is given by*

$$p_B(w) = \sum_{k\geq 1} d_k^B w^k = \frac{2(3 - 2\sqrt{2})w^2}{(1 - (\sqrt{2} - 1)w)^2}.$$

Proof. We use Lemma 1.1 with $f(z) = B(z)$ and $H(z, t) = zw(1 - t)/(1 - t - wt)$. We know that $B(z)$ has a unique dominating square-root singularity at $\rho = 3 - 2\sqrt{2}$. By substituting into Equation (2) we find that $B(\rho) = (2 - \sqrt{2})/2$. The result follows then by differentiating $H(z, t)$ with respect to t and evaluating at $(\rho, B(\rho))$. □

Notice that $p_B(1) = 1$, so that it is indeed a probability distribution. Now we prove a similar result for connected and arbitrary graphs.

Theorem 2.2. *The limiting distributions of the root degree in connected and arbitrary graphs are given, respectively, by*

$$p_C(w) = \sum_{k\geq 1} d_k^C w^k = \frac{(1 - \frac{1}{\sqrt{3}})^2}{2} \frac{w(w + 1 + \sqrt{3})}{(1 - (1 - \frac{1}{\sqrt{3}})w)^2},$$

$$p_G(w) = \sum_{k\geq 1} d_k^G w^k = \frac{(\sqrt{2} - 1)^2}{2} \frac{(1 + w)^2}{(1 - (\sqrt{2} - 1)w)^2}.$$

Proof. We use again Lemma 1.1, with $f(z) = C(z)$ and

$$H(z,t) = \frac{z}{1 - B(t^2/z, w)}.$$

The dominant singularity of $C(z)$ is $\rho = \sqrt{3}/18$ and it is of the square-root type. We need to find the derivative $H_t(z,t)$ and evaluate it at the point $(\rho, C(\rho))$. Substituting $z = \rho$ into Equation (4) and noting that the value of $C(\rho)$ must be positive we conclude that $C(\rho) = (\sqrt{3} - 1)/6$.

We next find the equation satisfied by $H(z,t)$ eliminating from Equations (19) and (2).

$$\begin{aligned}&\left(z + 2\,t^2w^2 - t^2w - zw\right) H(z,t)^2 + \\ &\left(2\,z^2w - 2\,z^2 + t^4w^2 - 3\,t^2w^2z\right) H(z,t) - z^3w + z^3 + t^2wz^2 + t^2w^2z^2 = 0.\end{aligned}$$

By differentiating with respect to t we find a linear equation for $H_t(z,t)$ which leads to the expression in the statement of the theorem upon substituting $t = C(\rho)$, $z = \rho$ and $H = C(\rho, w)$. To find this last value, we simply substitute $z = \rho$ in the equation above and of the two solutions we pick the one that evaluated at $w = 1$ gives $C(\rho)$, which is $(\sqrt{3}-1)(w-2-\sqrt{3})/(6(2\,w-3-\sqrt{3}))$.

The result for $p_G(w)$ can be proved in a similar way, but we prefer a more combinatorial proof, that explains in addition why we find the same value $\sqrt{2} - 1$ in the denominator of both $p_G(w)$ and $p_B(w)$.

As observed in [10], an arbitrary graph can be obtained from a polygon dissection by removing a subset of the edges of the polygon. If we put $b_n = [z^{n-1}]B(z)$ and $g_n = [z^n]G(z)$, we have that $g_n = b_n 2^n$. Let $b_{n,k} = [z^{n-1}][w^k]B(z,w)$ and $g_{n,k} = [z^n][w^k]G(z,w)$. We have the following relationship

$$g_{n,k} = b_{n,k}2^{n-2} + 2b_{n,k+1}2^{n-2} + b_{n,k+2}2^{n-2},$$

which reflects the different possibilities for obtaining a graph with root degree equal to k from a dissection of root degree k, $k+1$ or $k+2$. By dividing both sides of this equation by $g_n = b_n 2^n$ we obtain

$$4\frac{g_{n,k}}{g_n} = \frac{b_{n,k}}{b_n} + 2\frac{b_{n,k+1}}{b_n} + \frac{b_{n,k+2}}{b_n}.$$

This allows us in particular to express the limiting probability distribution for arbitrary graphs in terms of that for 2-connected graphs, giving the expression in the statement. □

Notice again that $p_C(1) = 1$ and $p_G(1) = 1$. In the next table we give approximate values for the different probabilities involved.

k	0	1	2	3	4	5	6
2-connected	0	0	0.3431	0.2843	0.1767	0.09755	0.05051
connected	0	0.2440	0.2956	0.2063	0.1216	0.06591	0.03340
arbitrary	0.08579	0.2426	0.2721	0.1838	0.1056	0.05592	0.02821

Table 1: Probability that the degree of the root is k in 2-connected, connected and arbitrary graphs.

From the previous explicit expressions it is immediate to obtain the tail of the distribution. In all cases it is of the form ckq^k, for suitable c and q.

Corollary 2.3. *We have the following estimates, as $k \to \infty$:*

$$\begin{aligned}p_k^B &\sim 2k(\sqrt{2}-1)^k \\ p_k^C &\sim \left(\tfrac{1}{2} + \tfrac{1}{\sqrt{3}}\right) k \left(1 - \tfrac{1}{\sqrt{3}}\right)^k \\ p_k^G &\sim k(\sqrt{2}-1)^k\end{aligned}$$

3 Maximum Degree

In this section we show that the maximum degree is of order $\log n$ and, suitably scaled, converges to a well-defined constant.

Theorem 3.1. *The maximum degree Δ_n for 2-connected, connected and arbitrary non-crossing graphs satisfies*

$$\frac{\Delta_n}{\log n} \to c \qquad \textit{in probability,}$$

where $c = 1/\log(q^{-1})$ and $q = \sqrt{2} - 1$ for 2-connected and arbitrary graphs, and $q = 1 - 1/\sqrt{3}$ for connected graphs. In all cases we also have

$$\mathbb{E}\,\Delta_n \sim c \log n \qquad \textit{as } n \to \infty.$$

Proof. We must show that the associated generating functions satisfy the conditions imposed on $f(z,w)$ and on $g(z,w,t)$ in Lemma 1.3. We treat first the case of a single root, and then that of a root plus a secondary vertex.

We rewrite (19) as

$$B(z,w) = \frac{wz}{1 - y_B(z)w},$$

where $y_B(z) = (1 - z - \sqrt{1 - 6z + z^2})/2$.

Now we consider $C(z,w)$. From the first equality in (19), the expression for $C(z,w)$ in (17) becomes

$$C(z,w) = \frac{z(1 - (1+w)B(C^2/z)}{1 - (1 + w + wC^2/z)B(C^2/z) - wC^2/z},$$

where we have set $C = C(z)$ for brevity. Eliminating from equation (3), we arrive at

$$C(z,w) = \frac{z + wz - wC(z)}{1 - y_C(z)w},$$

where $y_C(z) = (C(z)^2 + C(z) - z)/z$.

For $G(z,w)$, from (18), the expression for $C(z,w)$ just found and (5) we have

$$G(z,w) = 1 + \frac{zG(z) + wzG(z) - wC(zG(z))}{1 - y_c(zG(z))w} = 1 + \frac{w + (z + wz - w)G(z)}{1 - y_G(z)w},$$

where $y_G(z) = (G(z) - z - 1)/z$.

Next we consider the generating functions $B(z,w,t)$, $C(z,w,t)$, and $G(z,w,t)$. Recall that, in addition to the degree of the root p_1, the degree of a secondary vertex p_j (with $j \neq 1$) is marked by the variable t.

For 2-connected graphs we consider two different cases. The generating function $B_1(z,w,t)$ deals with the degree of p_2 and the generating function $B_2(z,w,t)$ with the general case p_j, $j \geq 3$. These functions were already computed in [9], as follows:

$$B_1(z,w,t) = zwt + \frac{z^2w^2t^2(1 + z(2A+1))}{(1 - zw(2A+1))(1 - zt(2A+1))}$$

$$B_2(z,w,t) = \frac{z^2w^2t^2(1 + z(2A+1))(P_1 + z(wt - w - t)P_2)}{(1 - z(4A+3))(1 - zw(2A+1))^2(1 - zt(2A+1))^2},$$

$$P_1 = 1 - z(4A+1), \quad P_2 = 1 - 2A + z(2A+1), \quad A = \frac{1 - 3z - \sqrt{1 - 6z + z^2}}{4z}.$$

By definition, $B(z,w,t) = B_1(z,w,t) + B_2(z,w,t)$. Notice that $z(2A+1) = (1 - z - \sqrt{1 - 6z + z^2})/2$ is precisely $y_B(z)$. The factor $1 - z(4A+3)$ in the denominator contributes precisely to the factor $Z = \sqrt{1 - z/z_0}$ (with $z_0 = 3 - 2\sqrt{2}$) in the denominator of (12). Hence the conditions of Lemma 1.3 are satisfied.

For connected graphs we also consider two different situations. We write $C(z,w,t) = C_1(z,w,t) + C_2(z,w,t)$, where the generating function $C_1(z,w,t)$ deals with the case when the secondary root is in one of the blocks that contain the root vertex. We claim

$$C_1(z,w,t) = \frac{z}{(1-B(C^2/z,w))^2} B(C^2/z,w,t) \frac{C(z,t)^2}{C^2}.$$

Indeed, the first factor corresponds to the blocks that do not contain the secondary root, the middle factor to the block that contains it, and the last factor corresponds to the two connected graphs attached to the secondary root. For $C_2(z,w,t)$ the secondary root lies in one of the connected graphs that are attached to the blocks that contain the root. We have

$$C_2(z,w,t) = \frac{z}{(1-B(C^2/z,w))^2} B_z(C^2/z,w) \frac{2C(z,t)C}{z},$$

where the middle factor identifies a block and a vertex v in that block, and the last factor picks a vertex in one of the connected graphs attached to v to become the secondary root.

Observe that the term $z(1-B(C^2/z))^2$ equals $C(z,w)^2/z$, hence

$$\frac{z}{(1-B(C^2/z,w))^2} = \frac{(z+wz-wC)^2}{z(1-y_C(z)w)^2}. \tag{20}$$

For the term $C(z,1,t)$, notice that if $C(z,t) = \sum c_{n,k} z^n t^k$, then $C(z,1,t) = \sum (n-1) c_{n,k} z^n t^k$, hence $C(z,1,t) = zC_z(z,t) - C(z,t)$. From (4) we obtain $C_z(z) = (-4z+3C)(3C^2+2C-3z)^{-1}$, which we can substitute in the derivative of $C(z,t)$ to find the following expression for $C(z,1,t)$:

$$C(z,1,t) = -\frac{tC\left((1-2t)C^2 + 4tzC - 2tz^2 - 2z^2\right)}{(3C^2+2C-3z)(1-ty_C(z))^2}. \tag{21}$$

The term $B_z(C^2/z,w)$ can be expressed in terms of C in a similar way, first using (1) to find $B_z(z,w)$ in terms of z, w and $B(z)$, and then using (3) to obtain

$$B_z(C^2/z,w) = \frac{wz^2(2z^2 - C^2 + 2w(C^2 - 2zC + z^2))}{(2z^2 - C^2)(z - w(C-z))^2}.$$

Finally, for the term $B(C^2/z,w,t)$ we need $A(C^2/z) = (z - 3C^2 - \sqrt{z^2 - 6zC^2 + C^4})/(4C^2)$. Eliminating from (4), one can check that the term inside the square root is a perfect square, namely, $z^2 - 6zC^2 + C^4 = (C^2 + 2C - 3z)^2$. It turns out that the correct square root is $-(C^2+2C-3z)$, therefore

$$A(C^2/z) = \frac{C - C^2 - z}{2C^2}.$$

For our purposes it is enough to write

$$B(C^2/z,w,t) = \frac{P(C,z,w,t)}{(C^2 - z(C^2+2C-2z))(C^2 - zw(C-z))^2(C^2 - zt(C-z))^2}$$

for an explicit polynomial $P(c,z,w,t)$.

It remains to check that $C(z,w,t)$ has the form given in (12). The terms $(1-wy_C(z))^2$ and $(1-ty_C(z))^2$ in the denominator appear clearly from (20) and (21). Among all other terms, one checks that the only one that contributes to a factor $Z = \sqrt{1-z/z_0}$ (now $z_0 = \sqrt{3}/18$) is the term $3C^2 + 2C - 3z$ from (21). This follows from the expansion of $C(z)$ in powers of Z, which is

$$C(z) = \frac{\sqrt{3}-1}{6} - \left(\frac{\sqrt{6}}{9} - \frac{\sqrt{2}}{6}\right) Z + O(Z^2).$$

Finally, for arbitrary graphs we have

$$G(z,w,t) = 1 + C(zG(z),w,t) + zC_z(zG(z),w)G_z(z,t).$$

The second summand corresponds to the case where the secondary vertex is in the same component as the root. The third summand corresponds to the opposite case, and the derivative C_z marks the component containing the secondary vertex; the term $G_z(z,t)$ marks the secondary vertex.

To find $G_z(z,t)$ we proceed as we did for $C_z(z,t)$ above and find

$$G_z(z,t) = \frac{Q(G(z),z,t)}{z^2(2G(z)+2z^2-3z-2)(1-ty_G(z))^2}$$

for a polynomial $Q(g,z,t)$. The substitution of $z = zG(z)$ in the already known terms $C(z,w,t)$ and $C_z(z,w)$ is straightforward since $C(zG(z)) = G(z)-1$. The terms $(1-wy_C(zG(z))^2$ and $(1-ty_C(zG(z))^2$ give rise to the required terms $(1-wy_G(z))^2$ and $(1-ty_G(z))^2$; it is routine to check that among all other terms, the ones that contribute a factor $\sqrt{1-z/z_0}$ in the denominator (with $z_0 = 3/2-\sqrt{2}$) are the term $2z^2-C(zG(z))^2 = 2z^2-(G(z)-1)^2$ coming from $C(zG(z),w,t)$ and the term $2G(z)+2z^2-3z-2$ from $G_z(z,t)$. $\square$

4 Size of the largest component

The size M_n of the largest component can be handled with the same tools as the maximum degree, but we need to refine the analysis. Let $X_{n,k}$ denote the number of components of size k in a random graph of size n and set

$$Y_{n,k} = \sum_{\ell>k} X_{n,\ell}$$

the number of components with more than k vertices. Then we have

$$Y_{n,k} > 0 \quad \Longleftrightarrow \quad M_n > k.$$

Hence, by applying the first and second moment method we can estimate the probabilities $\mathbb{P}\{M_n > k\}$ with the help of the first two moments $\mathbb{E}\,Y_{n,k}$ and $\mathbb{E}\,Y_{n,k}^2$.

By definition we have

$$X_{n,k} = \frac{1}{k}\sum_{i=1}^{n} \mathbf{1}_{[|\{\text{component of } p_i\}|=k]}$$

and consequently

$$\mathbb{E}\,X_{n,k} = \frac{n}{k}\, q_{n,k},$$

where $q_{n,k}$ denotes the probability that the root component has size k. By definition $q_{n,k}g_n$ is the number of graphs where the root component has size k; recall that g_n denotes the number of graphs of size n. Hence the the corresponding generating function is given by

$$G(z,u) = \sum_{n,k} q_{n,k}\, g_n z^n u^k = 1 + C(zuG(z)). \tag{22}$$

From this we obtain the following property.

Lemma 4.1. *There exist constants $c > 0$ and $\rho < 1$ such that uniformly for $k \le C\log n$ (where $C > 0$ is an arbitrary constant)*

$$\mathbb{E}\,Y_{n,k} \sim c\,n\rho^k k^{-3/2}. \tag{23}$$

Furthermore we have for every $\varepsilon > 0$

$$\mathbb{E}\,Y_{n,k} = O\left(n(\rho+\varepsilon)^k\right) \tag{24}$$

uniformly for all $n, k \ge 0$.

Proof. Set $r(z) = zG(z)/\rho_C$, where $\rho_C = \sqrt{3}/18$ denotes the singularity of $C(z)$. Then $r(z)$ has a square-root singularity at $z = \rho_G = 3/2 - \sqrt{2}$. Setting $Z = \sqrt{1 - z/\rho_G}$ we can, thus, represent $G(z,u)$ locally as

$$G(z,u) = g_1(z,u,Z) - h_1(z,u,Z)\sqrt{1 - ur(z)},$$

where g_1 and h_1 are analytic functions at $(z,u,Z) = (\rho_G, u_0, 0)$, $u_0 = \rho_C/(\rho_G G(\rho_G)) > 1$, and h_1 is non-zero at this point. Now we use the methods of [9] to obtain an asymptotic expansions of the form

$$[z^n u^k]\, G(z,u) \sim c_1 \rho_G^{-n} n^{-3/2} u_0^{-k} k^{-1/2}$$

for a certain constant $c_1 > 0$ (that is uniform for $k \le C \log n$, where $C > 0$ is an arbitrary constant). Of course this implies

$$\mathbb{E}\, X_{n,k} \sim c_2\, n \rho^k k^{-3/2}.$$

for some constant $c_2 > 0$ and with $\rho = 1/u_0$. Furthermore by using the inequality

$$[z^n u^k]\, G(z,u) \le (\rho + \varepsilon)^k [z^n] G(z, (\rho+\varepsilon)^{-1}),$$

we obtain for every $\varepsilon > 0$

$$\mathbb{E}\, X_{n,k} = O\left(n(\rho+\varepsilon)^k\right).$$

Clearly these two properties of $\mathbb{E}\, X_{n,k}$ imply (23) and (24). □

The estimates (23) and (24) imply an upper bound for $\mathbb{E}\, M_n$ of the form $\log n/\log(1/\rho)$. A corresponding lower bound follows by considering the second moment. Since

$$X_{n,k}^2 = \frac{1}{k^2} \sum_{i,j=1}^{n} \mathbf{1}_{[|\{\text{component of } p_i\}|=k]} \mathbf{1}_{[|\{\text{component of } p_j\}|=k]}$$

we obtain

$$\mathbb{E}\, X_{n,k}^2 = \frac{n}{k^2} \sum_{j=1}^{n} q_{n,k,k;j}$$

where $q_{n,k,k;j}$ denotes the probability that the component of the root p_1 as well as the component of p_j have size k. Similarly we have for $k \ne \ell$

$$\mathbb{E}\, X_{n,k} X_{n,\ell} = \frac{n}{k\ell} \sum_{j=1}^{n} q_{n,k,\ell;j}$$

where $q_{n,k,\ell;j}$ denotes the probability that the component of p_1 has size k and the component of p_j has size ℓ. Note that $q_{n,k,\ell;1} = 0$ if $k \ne \ell$ but for the sake of consistency we include this term in the formula.

If we know the behaviour of $\mathbb{E}\, X_{n,k} X_{n,\ell}$, we also obtain that of

$$\mathbb{E}\, Y_{n,k}^2 = \sum_{\ell_1,\ell_2 > k} \mathbb{E}\, X_{n,\ell_1} X_{n,\ell_2}.$$

In order to deal with these second moments we introduce another variable v that takes care of the size of a component of a vertex p_j, $1 \le j \le n$. The corresponding generating function is given by

$$G(z,u,v) = \sum_{n,k,\ell} \sum_{j=1}^{n} q_{n,k,\ell;j} g_n x^n u^k v^\ell$$

and satisfies the relation

$$G(z,u,v) = 1 + C'(uvzG(z))uvzG(z) + C'(uzG(z))uzG(z,1,v).$$

Consequently it is also given by

$$G(z,u,v) = 1 + C'(uvzG(z))uvzG(z) + C'(uzG(z))uz\frac{1 + C'(vzG(z))vzG(z)}{1 - zC'(zG(z))}.$$

This representation leads to the following property for the second moment.

Lemma 4.2. *We have uniformly for $k \le C\log n$ and $\ell \le C\log n$ (where $C > 0$ is an arbitrary constant)*

$$\mathbb{E}\, Y_{n,k}^2 \sim c^2\, n^2 \rho^{2k} k^{-3}(1 + o(1)), \tag{25}$$

where the constant c is the same as that of Lemma 4.1.

Proof. It is now an easy exercise to show that the asymptotic leading term of $G(z,u,v)$ can be represented as

$$\frac{C'(uzG(z))uzC'(vzG(z))vzG(z)}{1 - zC'(zG(z))} = \frac{\overline{H}(z,Z,u,v)}{Z\sqrt{1-ur(z)}\sqrt{1-vr(z)}},$$

where $\overline{H}(z,x,u,v)$ is non-zero and regular at $(z,x,u,v) = (z_0, 0, g(z_0), g(z_0))$.

Again we use the methods of [9] to obtain asymptotic expansions of the form

$$c'\, n^{-1/2}\rho_G^{-n}\rho^{k+\ell}(k\ell)^{-1/2}(1 + o(1))$$

for the coefficients. This leads to an asymptotic expansion for $\mathbb{E}\, X_{n,k}X_{n,\ell}$ of the form

$$\mathbb{E}\, X_{n,k}X_{n,\ell} = c'' n^2 \rho^{k+\ell}(k\ell)^{-3/2}(1 + o(1)).$$

Similarly to the calculations of Lemma 4.1 we also get a uniform upper bound of the form

$$\mathbb{E}\, X_{n,k}X_{n,\ell} = O\left(n^2(\rho+\varepsilon)^{k+\ell}\right)$$

(for every $\varepsilon > 0$).

Consequently the asymptotic expansion (25) for $\mathbb{E}\, Y_{n,k}^2$ follows, and it is easy to check the constant. □

With the help of Lemmas 4.1 and 4.2 we obtain corresponding lower bounds for $\mathbb{E}\, M_n$. To state our main result we use the exact value of ρ.

Theorem 4.3. *The size M_n of the largest component in non-crossing graphs satisfies*

$$\frac{M_n}{\log n} \to c \qquad \text{in probability,}$$

where $c = 1/\log(3\sqrt{3}(49\sqrt{2} - 69)/2)$. We also have

$$\mathbb{E}\, M_n \sim c\log n \qquad \text{as } n \to \infty.$$

A similar result holds for the size of the largest 2-connected component in connected graphs. We omit the details for the sake of conciseness.

5 Diameter

Let d_n be the maximum distance to the root vertex p_1. It is clear that the diameter D_n satisfies

$$d_n \le D_n \le 2d_n.$$

Hence, an estimate for d_n immediately gives bounds for D_n. In what follows, all the results are stated in terms of d_n, and the reader should keep in mind that they provide analogous bounds for the diameter D_n.

Before we discuss 2-connected graphs we consider the special case of triangulations of a convex polygon, which is interesting in itself. As it is well-known, the number of triangulations of a convex n-gon (where the root p_1 is not counted) is the Catalan number $\frac{1}{n}\binom{2n-2}{n-1}$. The associated generating function $T(z)$ satisfies the equation

$$T(z) = z + T(z)^2,$$

and has the explicit solution

$$T(z) = \frac{1-\sqrt{1-4z}}{2}.$$

Lemma 5.1. *Let $T_k(z)$ be the generating function of triangulations of a convex n-gon with $d_n \le k$. Then $T_0(z) = 0$ and*

$$T_k(z) = \frac{z}{1 - z - T_{k-1}(z)^2} \qquad (k \ge 1). \tag{26}$$

Proof. Let $T_{k,\ell}(z)$ be the generating function of triangulations of a convex n-gon, where for each vertex the distance to the root vertex p_1 is at most k or the distance to the vertex p_n is at most ℓ, where we require moreover that the edge (p_1, p_n) is not used in computing these distances. It is clear that $T_k(z) = T_{k,k-1}(z)$.

By considering the two triangulations adjacent to the root face we have the recurrence relations

$$T_{k,k}(z) = z + T_{k,k-1}(z)^2,$$
$$T_{k,k-1}(z) = z + T_{k,k-1}(z)T_{k-1,k-1}(z),$$

which lead directly to the recurrence

$$T_k(z) = \frac{z}{1 - z - T_{k-1}(z)^2}.$$

□

Theorem 5.2. *The expected value of d_n in triangulations is asymptotically given by*

$$\mathbb{E}\, d_n \sim \frac{2}{3}\sqrt{\pi n}.$$

Proof. We apply Lemma 1.4 with $F(z,t) = z/(1 - z - t^2)$ and the parameters

$$z_0 = \frac{1}{4}, t_0 = \frac{1}{2}, F_z(z_0, t_0) = 3, F_{tt}(z_0, t_0) = 6.$$

Clearly $1/4$ is the singularity of $T(z)$, and $t_0 = T(1/4) = 1/2$. The remaining relations are immediate. □

5.1 2-Connected graphs

For the analysis of 2-connected graphs we need a combinatorial decomposition different from the one we have used so far. Consider the root region of a 2-connected graph and let p_j be the vertex that follows p_1 when traversing the root region counterclockwise (see Figure 2). The vertices $p_1, \ldots, p_j$ induce a 2-connected graph, whereas the vertices $p_j, \ldots, p_n$ induced either a 2-connected graph or a graph that is 2-connected after the addition of the edge $p_j p_n$. This leads to the equation

$$B(z) = z + B(z)(2B(z) - z).$$

Lemma 5.3. *Let $B_k(z)$ be the generating function of 2-connected graphs with $d_n \le k$ Then $B_0(z) = 0$ and*

$$B_k(z) = \frac{z}{1 - z - 4B_{k-1}(z)^2 + 2zB_{k-1}(z)} \qquad (k \ge 1) \tag{27}$$

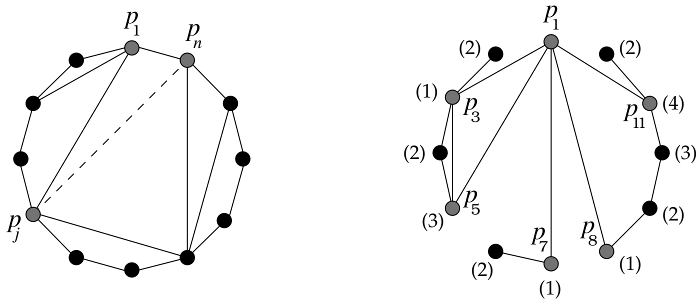

Figure 2: Decomposition of a 2-connected graph (left) and of a connected graph (right), where the numbers in parentheses indicate the value of $\overline{d}$ for each non-root vertex.

Proof. Let $B_{k,\ell}(z)$ be defined in the same way as in the proof of Lemma 5.1. We claim the recurrence relations

$$B_{k,k-1}(z) = z + B_{k,k-1}(z)(2B_{k-1,k-1}(z) - z)$$
$$B_{k-1,k-1}(z) = z + B_{k-1,k-2}(z)(2B_{k-1,k-2}(z) - z)$$

from which the statement of the lemma follows directly.

Both relations follow by looking at distances in the two 2-connected graphs that arise in the previous decomposition. □

Theorem 5.4. *The expected value of d_n in 2-connected graphs is given asymptotically by*

$$\mathbb{E}\, d_n \sim c\sqrt{\pi n}.$$

with $c = (3+\sqrt{2})2^{1/4}/7 \approx 0.7499$.

Proof. We apply Lemma 1.4 with $F(z,t) = z(1 - z - 4t^2 + 2zt)^{-1}$ and the parameters

$$z_0 = 3 - 2\sqrt{2}, t_0 = 1 - \frac{\sqrt{2}}{2}, F_z(z_0,t_0) = \frac{1+2\sqrt{2}}{2}, F_{tt}(z_0,t_0) = 8 + 2\sqrt{2}.$$

This is correct since z_0 is the singularity of $B(z)$ and $B(3-2\sqrt{2}) = 1 - \sqrt{2}/2$. □

5.2 Connected graphs

The diameter in connected graphs is also of order $\sqrt{n}$. As above we only discuss the largest distance d_n to the root vertex p_1, but instead of studying d_n directly, we look at two new parameters $\underline{d}_n$ and $\overline{d}_n$ that satisfy $\underline{d}_n \le d_n \le \overline{d}_n$.

For the lower bound, we consider the tree structure of the block decomposition of a connected graph. Let $\underline{d}_n$ be the maximum number of cut-points on a path to the root vertex (where the root vertex is never counted as a cut-point).

Lemma 5.5. *Let $C_k(z)$ be the generating function corresponding to those connected graphs with $\underline{d}_n \le k$. Then we have $C_0(z) = z/(1 - B(z))$ and*

$$C_k(z) = \frac{z}{1 - B(C_{k-1}(z)^2/z)}, \qquad (k \ge 1).$$

Proof. This is immediate from the block decomposition of connected graphs. □

Lemma 5.6. *The expected value of $\underline{d}_n$ is asymptotically given by*

$$\mathbb{E}\, \underline{d}_n \sim c\sqrt{\pi n}.$$

with $c = \sqrt{2}(1 - \sqrt{3}/3) \approx 0.5977$.

Proof. By combining the recurrence in Lemma 5.5 and Equation (2), we find the following explicit expression for $C_k(z)$ in terms of $C_{k-1}(z)$:

$$C_k(z) = \frac{3}{2}z - \frac{1}{2}C_{k-1}^2 - \frac{1}{2}\sqrt{z^2 - 6zC_{k-1}^2 + C_{k-1}^4}.$$

The result follows by applying Lemma 1.4 with

$$F(z,t) = 3z/2 - t^2/2 - \sqrt{z^2 - 6zt^2 + t^4}/2$$

and parameters

$$z_0 = \frac{\sqrt{3}}{18}, t_0 = \frac{1}{6}(\sqrt{3}-1), F_z(z_0,t_0) = \frac{3-\sqrt{3}}{2}, F_{tt}(z_0,t_0) = 9(5+3\sqrt{3}).$$

□

For the upper bound, we use an alternative decomposition of connected graphs [10]. Take the root of a connected graph and let $p_{i_1}, \ldots, p_{i_d}$ be its neighbours, with $i_1 < i_2 < \cdots < i_d$. The subgraph induced by $\{p_{i_j}, p_{i_j+1}, \ldots, p_{i_{j+1}}\}$ is either connected or it has exactly two connected components. Also, the subgraphs induced by $\{p_2, p_3, \ldots, p_{i_1}\}$ and $\{p_{i_d}, p_{i_d+1}, \ldots, p_n\}$ are connected. (See Figure 2.) This decomposition produces the equation

$$C(z) = z\left(1 + \frac{C(z)^2}{2z - C(z) - C(z)^2}\right). \tag{28}$$

Now define an application $\overline{d_G}$ from the set of vertices of a connected graph G to $\mathbb{N}$ recursively as follows. If x is the root-vertex, then $\overline{d_G}(x) = 0$. Otherwise, the vertex x belongs, according to the decomposition scheme, to at least one connected subgraph C which has either one or two vertices that are neighbours of the root. If x is adjacent to the root of G, then x belongs to two such subgraphs (which could be reduced to a single vertex); in this case, let C be the one that contains the vertex with smallest label. Define $\overline{d_G}(x) = \overline{d_C}(x) + 1$, where the root of C is taken to be the vertex with smallest label among those adjacent to the root of G. For instance, in the graph in Figure 2, the values of $\overline{d}$ are indicated for each vertex.

Let $\overline{d_n}$ be the maximum of $\overline{d}_G$ in a connected graph G with n vertices. Clearly $\overline{d_n}$ is an upper bound for d_n. The following lemma is immediate from the alternative decomposition of connected graphs.

Lemma 5.7. *Let $\overline{C}_k(z)$ be the generating function corresponding to those connected non-crossing graphs with $\overline{d}_n \leq k$. Then we have $\overline{C}_0(z) = z$ and*

$$\overline{C}_k(z) = z\left(1 + \frac{\overline{C}_{k-1}(z)^2}{2z - \overline{C}_{k-1}(z) - \overline{C}_{k-1}(z)^2}\right), \qquad (k \geq 1).$$

Lemma 5.8. *The expected value of $\overline{d}_n$ is asymptotically given by*

$$\mathbb{E}\,\overline{d}_n \sim c\sqrt{\pi n},$$

with $c = \sqrt{2}(1+\sqrt{3}/3)/2 \approx 1.1153$.

Proof. We apply Lemma 1.4 for $F(z,t) = z(1 + t^2/(2z - t - t^2))$ and parameters

$$z_0 = \frac{\sqrt{3}}{18}, t_0 = \frac{1}{6}(\sqrt{3}-1), F_z(z_0,t_0) = 9 - 5\sqrt{3}, F_{tt}(z_0,t_0) = 18(1+\sqrt{3}).$$

□

Corollary 5.9. *The expected value of d_n in connected graphs satisfies*

$$c_1\sqrt{n} \leq \mathbb{E}\,d_n \leq c_2\sqrt{n}$$

for suitable constants c_1 and c_2.

6 Concluding remarks

One of the motivations for this work was to solve an open problem in [14], namely to count *bipartite* non-crossing graphs. This can be solved by observing that a dissection is a bipartite graph if and only if all the regions have even size, and an arbitrary graph is bipartite if and only if all its components and blocks are bipartite.

Let $B_b(z)$, $C_b(z)$ and $G_b(z)$ be the corresponding generating functions for bipartite configurations. The relations between $G_b(z)$ and $C_b(z)$ and between $C_b(z)$ and $B_b(z)$, are given by the analogous of Equations (5) and (3). To find the equation for $B_b(z)$, we modify Equation (1) by allowing only cycles of even length in the decomposition of dissections. This gives rise to

$$B_b(z) = z + \frac{B_b(z)^3}{1 - B_b(z)^2}. \tag{29}$$

This determines everything and we obtain

$$\begin{aligned} zB_b(z) &= z^2 + z^4 + 4z^6 + 21z^8 + 126z^{10} + \cdots \\ C_b(z) &= z + z^2 + 3z^3 + 13z^4 + 65z^5 + \cdots \\ G_b(z) &= 1 + z + 2z^2 + 7z^3 + 34z^4 + 196z^5 + \cdots \end{aligned}$$

Asymptotics can be obtained too and we have

$$\begin{aligned} [z^n]B_b(z) &\sim c_1 n^{-3/2}\gamma_1^n, & \gamma_1 &\sim 2.9696, \\ [z^n]C_b(z) &\sim c_2 n^{-3/2}\gamma_2^n, & \gamma_2 &\sim 7.5289, \\ [z^n]G_b(z) &\sim c_3 n^{-3/2}\gamma_3^n, & \gamma_3 &\sim 8.9129. \end{aligned}$$

As a final remark, it is shown in [7] that for every $k \geq 1$ the random variable that counts the number of vertices of degree k in dissections is asymptotically normal with linear expectation and variance. With the same techniques, based on multivariate functional equations, it is possible to prove the same result for connected and arbitrary non-crossing graphs.

References

[1] N. Alon, J. H. Spencer, *The probabilistic method*, third edition, John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[2] N. Bernasconi, K. Panagiotou, A. Steger, On properties of random dissections and triangulations, *Combinatorica* 30 (2010), 627–54.

[3] L. Comtet, *Advanced combinatorics. The art of finite and infinite expansions*, revised and enlarged edition, D. Reidel Publishing Co., Dordrecht, 1974.

[4] E. Deutsch, M. Noy, Statistics on non-crossing trees, *Discrete Math.* 254 (2002), 75–87.

[5] L. Devroye, P. Flajolet, F. Hurtado, M. Noy, W. Steiger, Properties of random triangulations and trees, *Discrete Comput. Geom.* 22 (1999), 105–117.

[6] M. Drmota, *Random Trees*, Springer, Wien-New York, 2009.

[7] M. Drmota, O. Giménez, M. Noy, Vertices of given degree in series-parallel graphs, *Random Structures Algorithms* 36 (2010), 273–314.

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

[9] M. Drmota, O. Giménez, M. Noy, The maximum degree of series-parallel graphs, *Combin. Probab. Comput.* 20 (2011), 529–570.

[10] P. Flajolet, M. Noy, Analytic combinatorics of non-crossing configurations, *Discrete Math.* 204 (1999), 203–229.

[11] P. Flajolet, A. Odlyzko, The average height of binary trees and other simple trees. *J. Comput. System Sci.* 25 (1982), 171–213.

[12] P. Flajolet, R. Sedgewick, *Analytic combinatorics*, Cambridge U. Press, Cambridge, 2009.

[13] Z. Gao, N. C. Wormald, The distribution of the maximum vertex degree in random planar maps, *J. Combin. Theory Ser. A* 89 (2000), 201–230.

[14] C. Huemer, *Structural and Enumerative Problems for Plane Geometric Graphs*, Ph.D. Thesis, Universitat Politècnica de Catalunya, 2007.

[15] A. Panholzer, The height distribution of nodes in non-crossing trees, *Ars Combin.* 69 (2003), 19–32.