Extremal Statistics on Non-Crossing Configurations[†]

Anna de Mier and Marc Noy

Departament de Matemàtica Aplicada II, Universitat Politècnica de Catalunya, Barcelona, Spain.

Abstract. We obtain several properties of extremal statistics in non-crossing configurations with n vertices. We prove that the maximum degree and the largest component are of logarithmic order, and 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.

Résumé. On obtient des propriétés de paramètres extrémales dans les configurations sans croisement avec n sommets. On démontre que le degré maximal et la composante plus large sont d'ordre logarithmique, et le diamètre est d'ordre $\sqrt{n}$. Les preuves utilisent l'analyse de singularités, une application de la méthode du premier et second moment, et l'analyse de fonctions itérées.

Keywords: Non-crossing configuration; Extremal parameters; Maximum degree; Diameter.

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 $\{p_1, p_2\}$ (if present) the root edge. The root region is the internal region adjacent to 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 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 ia also considered a dissection. From a graph theoretical point of view, dissections are the same as 2-connected graphs.

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 [9] reexamined these problems using the tools from analytic combinatorics [11] 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, number
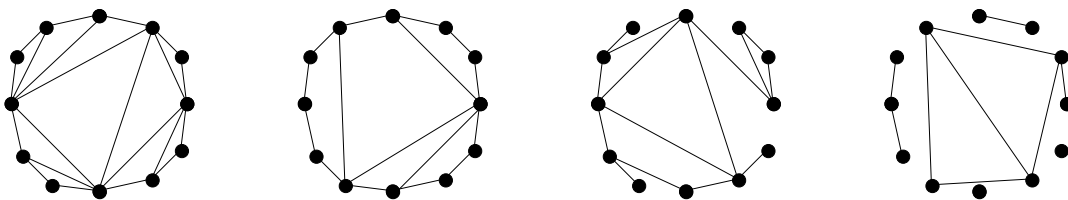
Fig. 1: From left to right: a triangulation, a 2-connected graph (dissection), a connected graph, and an arbitrary graph.

of leaves in trees and number of blocks in partitions. The proofs in [9] 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, 12] and trees [4, 13], 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 some $c > 0$,

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

- 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 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 $c > 0$ such that

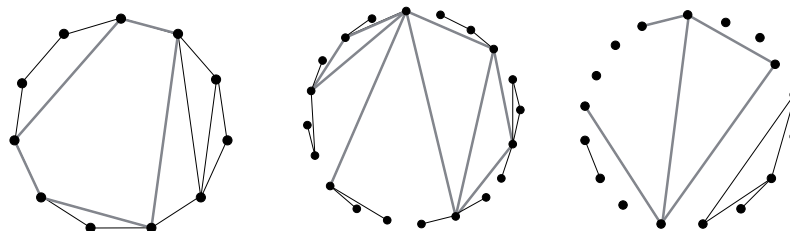$$\frac{M_n}{\log n} \to c \qquad \text{in probability.}$$

The same result holds for the largest 2-connected component in connected graphs.

- 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

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

Fig. 2: Left: a dissection consisting of a root face (in gray) with an arbitrary dissection glued at each non-root edge. Center: the root of the connected graph belongs to three blocks (in gray); to any other vertex of these blocks there are two connected graphs attached, possibly reduced to a single vertex. Right: between every two consecutive vertices of the root component (in gray) there is an arbitrary graph, possibly empty. In all three cases the root is the vertex on top.

To our knowledge, the diameter of random configurations has not been studied before, even in the basic case of triangulations. 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,

   In the rest of this section we collect several 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 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. Proofs are only sketched due to limitations of space.

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. They reflect the decomposition of a dissection as a sequence of dissections attached to the root region, of a connected graph into 2-connected components, and of an arbitrary graph into connected components [9] (see Figure 2).

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

$$G(z) \quad = 1 + C(zG(z)).$$

Hence we obtain

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

which has a square-root singularity at $z = 3 - 2\sqrt{2}$. Furthermore, $C = C(z)$ is the solution to

$$C^3 + C^2 - 3zC + 2z^2 = 0.$$

and has a dominant singularity at $z = \sqrt{3}/18$. Finally it follows that $G(z)$ satisfies the equation

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

which leads to

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

Functions $B(z)$ and $G(z)$ are clearly of square-root type; function $C(z)$ has also a local expansion in $\sqrt{1 - z18/\sqrt{3}}$.

1.2 First and second moment method

In order to obtain results for the maximum statistics of the root degree we follow the methods of [8]. They are based on the so-called *first and second moment methods*.

Lemma 1.1 *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 $> k$ in a random graph with n vertices. This random 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

$$\mathbb{E}\,Y_{n,k} = n \sum_{\ell > k} d_{n,\ell}. \tag{1}$$

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

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

In what follows we will make use of *power series of the square-root type*, that is a power series $y(x)$ with a square root singularity at $x_0 > 0$, that is, $y(x)$ admits a local representation of the form

$$y(x) = g(x) - h(x)\sqrt{1 - x/x_0} \tag{2}$$

for $|x - x_0| < \varepsilon$ for some $\varepsilon > 0$ and $|\arg(x - x_0)| > 0$, where $g(x)$ and $h(x)$ are analytic and non-zero at x_0. Moreover, $y(x)$ can be analytically continued to the region

$$D = D(x_0, \varepsilon) = \{x \in \mathbb{C} : |x| < x_0 + \varepsilon\} \setminus [x_0, \infty). \tag{3}$$

The following technical lemma subsumes some results from [8] and is adapted for our needs (the proof is very similar). 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 the square-root type, then automatically the maximum degree is $c \log n$ for a well-defined constant c. For our functions these conditions are relatively easy to check.

Lemma 1.2 *Let $f(x, w) = \sum_{n,k} f_{n,k} x^n w^k$ be a double generating and $g(x, w, t) = \sum_{n,k,\ell} g_{n,k,\ell} x^n w^k t^\ell$ 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*

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

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

$$f(x, w) = \frac{G(x, X, w)}{1 - y(x)w}, \tag{4}$$

where $X = \sqrt{1 - x/x_0}$, $y(x)$ is a power series with non-negative coefficients of square-root type,

$$y(x) = g(x) - h(x)\sqrt{1 - x/x_0},$$

where $0 < g(x_0) < 1$ and the function $G(x, v, w)$ is analytic in the region

$$D' = \{(x, v, w) \in \mathbb{C}^3 : |x| < x_0 + \eta, |v| < \eta, |w| < 1/g(x_0) + \eta\}$$

for some $\eta > 0$ and satisfies $G(x_0, 0, 1/g(x_0)) \neq 0$.
Furthermore suppose that $g(x, w, t)$ can be represented as

$$g(x, w, t) = \frac{H(x, X, w, t)}{X (1 - y(x)w)^2 (1 - y(x)t)^2}, \tag{5}$$

where the function $H(x, v, w, t)$ is non-zero and analytic at $(x, 0, w, t) = (x_0, 0, 1/g(x_0), 1/g(x_0))$.
Let Δ_n denote the maximum degree of a random graph in this class of size n. Then we have

$$\frac{\Delta_n}{\log n} \to \frac{1}{\log g(x_0)^{-1}} \qquad in \ probability \tag{6}$$

$$\mathbb{E}\, \Delta_n \sim \frac{1}{\log g(x_0)^{-1}} \log n \qquad (n \to \infty). \tag{7}$$

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

$$d_{n,k} \sim ckg(x_0)^k \quad \text{and} \quad d_{n,k,\ell} \sim d_{n,k}d_{n,\ell} \sim c^2 k\ell g(x_0)^{k+\ell}$$

uniformly for $k \leq C \log n$ and $\ell \leq C \log n$ (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.

The key conditions in the lemma are equations (4) and (5) 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 $y(x)$. In (4) we find a linear factor in w and in (5) we find a quadratic factor both in w and in t. In Section 3 we point out how these conditions are satisfied in our case.

1.3   Iterated functions

The results on the diameter are based on the following lemma on iterated functions. In the next statement, the notation $A(z) \preceq B(z)$ means that the coefficients of $A(z) = \sum_n a_n z^n$ and $B(z) = \sum_n b_n z^n$ satisfy $a_n \leq b_n$.

Lemma 1.3 *Suppose that $F(z,t)$ is an analytic function at $(z,t) = (0,0)$ such that the equation $T(z) = F(z,T(z))$ 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 (3), 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 $0 \leq_c T_0(z) \leq_c T(z)$, such that $T_0(z)$ is analytic at $z = z_0$, and let $T_k(z)$, $k \geq 1$ be iteratively defined by

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

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

$$\mathbb{E}\, H_n \sim \sqrt{\frac{2\pi}{z_0 F_z(z_0,t_0) F_{tt}(z_0,t_0)}}\; n^{1/2} \qquad \text{and} \qquad \frac{H_n}{\sqrt{n}} \to \frac{2}{\sqrt{z_0 F_z(z_0,t_0) F_{tt}(z_0,t_0)}} \max_{0 \leq t \leq 1} e(t),$$

where $e(t)$ denotes the Brownian excursion of duration 1.

The lemma is a direct extension of [10]; see also [6, Th. 4.8, Th. 4.59] for the proof techniques.

2   Degree of the root

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

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

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

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

In particular, we have

$$B(z, w) = \frac{wz}{1 - \frac{w}{2}(1 - z - \sqrt{1 - 6z + z^2})}. \tag{8}$$

This means that $z = 3 - 2\sqrt{2}$ is a squareroot singularity for $|w| < \sqrt{2} + 1$. A similar phenomenon appears for $C(z, w)$ and $G(z, w)$. From these expressions we can obtain the limit distribution of the root degree. Let d_k^B be the limiting probability that the root in a 2-connected graph has degree k, that is,

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

Define analogously d_k^C and d_k^G for connected and arbitrary graphs. Using singularity analysis one obtains the degree distribution in all cases, encoded in the corresponding probability generating function.

Theorem 2.1 *The limiting distribution of the root degree in 2-connected, connected and arbitrary graphs are given by*

$$\sum_{k \geq 1} d_k^B w^k = \frac{2(3 - 2\sqrt{2})w^2}{(1 - (\sqrt{2} - 1)w)^2} = 0.34w^2 + 0.28w^3 + 0.18w^4 + 0.01w^5 + 0.05w^6 + \cdots.$$

$$\sum_{k \geq 1} d_k^C w^k = \frac{(1 - \frac{1}{\sqrt{3}})^2}{2} \frac{w(w + 1 + \sqrt{3})}{(1 - (1 - \frac{1}{\sqrt{3}})w)^2} = 0.24w + 0.29w^2 + 0.21w^3 + 0.12w^4 + 0.07w^5 + \cdots.$$

$$\sum_{k \geq 1} d_k^G w^k = \frac{(1 - \sqrt{2})^2}{2} \frac{(1 + w)^2}{(1 - (\sqrt{2} - 1)w)^2} = 0.09 + 0.24w + 0.27w^2 + 0.18w^3 + 0.11w^4 + \cdots.$$

The first item in the previous theorem was proved in [7], and independently in [1].

3 Maximum Degree

The functions $B(z, w)$, $C(z, w)$ and $G(z, w)$ are easily seen to satisfy the conditions imposed on $f(z, w)$ in Lemma 1.2. Next we consider the generating functions $B(z, w, t)$, $C(z, w, t)$, and $G(z, w, t)$, where we consider, in addition to the degree of the root p_1, the degree of another vertex p_j (with $j \neq 1$) 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, $j \geq 2$. In [8] these functions were already computed, as follows:

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

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

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

We set $B(z, w, t) = B_1(z, w, t) + B_2(z, w, t)$.

Now we check that the conditions of Lemma 1.2 are satisfied. Function $y(x)$ in the statement is our A which is certainly of square-root type. Function $B(z, w)$ derived in Equation (8) displays a factor $1 - Aw$ in the denominator, and $B(z, w, t)$ displays a quadratic factor, the role of $y(x)$ being played now by $C(z)^2 + C(z) - z$, which is of square-root type.

Similarly, the corresponding function $C(z, w, t)$ for connected can be expressed as

$$C(z, w, t) = \frac{P(C(z), w, t, z)}{(3C(z)^2 + 2C(z) - 3z)(z - w(C(z)^2 + C(z) - z))^2(z - t(C(z)^2 + C(z) - z))^2},$$

for an explicit polynomial $P(c, w, t, z)$. Since $B(z)$ and $C(z)$ are of square-root type it is immediate that the functions $B(z, w)$, $B(z, w, t)$ and $C(z, w)$, $C(z, w, t)$, respectively, satisfy the assumptions of Lemma 1.2. Finally the generating function $G(z, w, t)$ for all graphs satisfies

$$G(z, w, t) = C(zG(z), w, t) + \frac{\partial C(zG(z), w)}{\partial z}(G(z, t) + G(z, 1, t)),$$

Again it is easy to check that these functions satisfy the conditions in Lemma 1.1. Applying it we obtain the main result of this section.

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

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

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

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

4   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 $\mathbb{E}\,Y_{n,k}^2$. However, in contrast to the maximum degree it is not enough to calculate the probability that the root component has size k, since the number of components is also a random variable. By definition we have

$$X_{n,k} = \sum_{c\ \text{component}} \mathbf{1}_{[|c|=k]}$$

Let $p_{n,m}$ denote probability that a graph with n vertices has m components, $p_{n,m,k}$ the probability that a graph with n vertices has m components and the root component has size k, and $p_{n,k|m}$ the probability that a graph with n vertices has a root component of size k conditioned on the event that the graph has m components. Clearly we have $p_{n,m,k} = p_{n,k|m}p_{n,m}$, and consequently

$$\mathbb{E} X_{n,k} = \sum_{m \geq 1} p_{n,m} m p_{n,k|m} = \sum_{m \geq 1} m p_{n,m,k}.$$

The quantity $p_{n,m,k}$ can be handled as follows. Let y mark the number of components and u the size of the component of the root p_1. Then we have

$$G(z, y, u) = \sum_{n,m,k} p_{n,m,k}\, g_n x^n y^m u^k = 1 + yC(zuG(z, y, 1)),$$

where g_n denotes the number of graphs of size n. Using this expression we prove:

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

$$\mathbb{E} Y_{n,k} \sim c n \rho^k k^{1/2}. \tag{9}$$

Furthermore we have for every $\varepsilon > 0$, uniformly for all $n, k \geq 0$,

$$\mathbb{E} Y_{n,k} = O\left(n(\rho + \varepsilon)^k\right). \tag{10}$$

The asymptotic relations (9) and (10) imply an upper bound for $\mathbb{E} M_n$ of the form $\sim \log n / \log(1/\rho)$. A corresponding lower bound follows by considering the second moment. Since

$$X_{n,k}^2 = \sum_{c_1, c_2 \text{ components}} \mathbf{1}_{[|c_1|=k]} \mathbf{1}_{[|c_2|=k]}$$

we obtain

$$\mathbb{E} X_{n,k}^2 = \mathbb{E} X_{n,k} + \sum_{m \geq 1} m^2 p_{n,m,k,k}^{(2)},$$

where $p_{n,m,k,k}^{(2)}$ denotes the probability that a graph with n vertices has m components, the root component has size k and a uniformly and randomly chosen component different from the root component has also size k. Similarly we have for $k \neq \ell$

$$\mathbb{E} X_{n,k} X_{n,\ell} = \sum_{m \geq 1} m^2 p_{n,m,k,\ell}^{(2)},$$

where $p_{n,m,k,\ell}^{(2)}$ denotes the probability that a graph with n vertices has m components, the root component has size k and a uniformly and randomly chosen component different from the root component has size ℓ.

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

component different from the root component. The corresponding generating function is given by

$$G(z, y, u, v) = \sum_{n,m,k,\ell} p^{(2)}_{n,m,k,\ell}\, g_{n,m}(m-1)x^n y^m u^k v^\ell$$

$$= yC'(uzG(z,y))\,(uzG(z,y,v) + uzG(z,y,1,v))\,,$$

where $g_{n,m}$ denotes the number of graphs with n vertices and m components. Using this we prove:

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

$$\mathbb{E}\, Y_{n,k}^2 \geq c^2\, n^2 \rho^{2k} k(1 + o(1)) \tag{11}$$

where the constant c is as in Lemma 4.1.

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 lack of space.

5   Diameter

The diameter of a connected graph is the maximum distance between any pair of vertices. In this section we show that the diameter D_n of connected and 2-connected non-crossing graphs is of order $\sqrt{n}$. To get equations for the *exact* value of the diameter is usually very hard (to our knowledge, only in the case of trees exact results are known [14, 2]). Instead we prove results on the parameter d_n, which is the maximum distance from a vertex to the root p_1. Since clearly we have

$$d_n \leq D_n \leq 2d_n,$$

it is enough to obtain bounds of the right order of magnitude for d_n.

We start with 2-connected graphs. Let d_n be the maximum distance to the root vertex p_1. Using the decomposition of 2-connected graphs according to the root region, one can prove the following.

Lemma 5.1 *Let $B_k(z)$ be the generating function of 2-connected non-crossing graphs with $d_n \leq k$. Then $B_0(z) = 0$ and*

$$B_k(z) = \frac{z}{1 - z - 4B_{k-1}(z)^2 + 2zB_{k-1}(z)} \qquad (k \geq 1) \tag{12}$$

Theorem 5.2 *The expected value of d_n in 2-connected graphs is asymptotically given by*

$$\mathbb{E}\, d_n \sim c\sqrt{\pi n}. \qquad \text{with } c = (3 + \sqrt{2})2^{1/4}/7 = 0.749916\ldots$$

The proof is an application of Lemma 1.3 with $F(z,t) = z(1 - z - 4t^2 + 2zt)^{-1}$ and

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

Next we turn to connected graphs. As above we discuss only 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. The following lemma follows directly from the decomposition of connected graphs.

Lemma 5.3 *Let $C_k(z)$ be the generating function corresponding to those connected non-crossing configurations with $\underline{d}_n \le k$. Then we have $C_0(z) = z/(1 - B(z))$ and*

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

Theorem 5.4 *The expected value of $\underline{d}_n$ is asymptotically given by*

$$\mathbb{E}\, \underline{d}_n \sim c\sqrt{\pi n} \qquad \text{with } c = \sqrt{2}(1 - \sqrt{3}/3) = 0.597716\ldots$$

The proof combines the recurrence in Lemma 5.3 and the equation for $B(z)$, from where 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 1.3 for $F(z,c) = 3z/2 - c^2/2 - \sqrt{z^2 - 6zc^2 + c^4}/2$ and

$$z_0 = \frac{\sqrt{3}}{18}, c_0 = \frac{1}{6}(\sqrt{3} - 1), F_z(z_0, c_0) = \frac{3 - \sqrt{3}}{2}, F_{cc}(z_0, c_0) = 9(5 + 3\sqrt{3}).$$

For the upper bound, we use an alternative decomposition of connected graphs [9]. Take the root of a connected graph and consider the orderer list $\{v_1, \ldots, v_d\}$ of its neighbours. The subgraph induced between two consecutive neighbours v_i and v_{i+1} is either a connected graph not reduced to a point or two disjoint connected graphs containing v_i and v_{i+1}, respectively. This decomposition produces the equation

$$C(z) = z\left(\frac{C(z)^2}{2z - C(z) - C(z)^2}\right). \tag{13}$$

Now to 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 may belong to two such subgraphs; in this case, pick C to be the one that contains the vertex with smallest label. Define $\overline{d}_G(x) = \overline{d}_C(x) + 1$. For instance, for each n there is only one graph with $\overline{d}(x) \le 1$ for all vertices x.

Then $\overline{d}_n$ is the maximum of $\overline{d}$ in a connected graph with n vertices. Clearly $\overline{d}_n$ is an upper bound for the diameter. The following lemma is immediate from the alternative decomposition of connected graphs.

Lemma 5.5 *Let $\overline{C}_k(z)$ be the generating function corresponding to those connected non-crossing configurations with $\overline{d}_n \leq k$. Then we have $\overline{C}_0(z) = z$ and*

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

Theorem 5.6 *The expected value of $\overline{d}_n$ is asymptotically given by*

$$\mathbb{E}\,\overline{d}_n \sim c\sqrt{\pi n}, \qquad \text{with } c = \sqrt{2}(1 + \sqrt{3}/3)/2 = 1.115355\dots$$

The proof follows from another application of Lemma 1.3.

We conclude by remarking that a similar result can be prove for triangulations of a polygon. If d_n is as before for triangulations, we have

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

References

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

[2] N. Broutin, P. Flajolet. The distribution of height and diameter in random non-plane binary trees. Random Structures Algorithms (to appear).

[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. Degree distribution in random planar graphs. *J. Combin. Theory Ser. A* 118 (2011), 2102–2130.

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

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

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

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

[12] 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.

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

[14] G. Szekeres. Distribution of labelled trees by diameter. Lecture Notes in Math. 1036 (1983), pp. 392-397.