On the diameter of random planar graphs

Guillaume Chapuy
Simon Fraser University
Vancouver

Éric Fusy
École Polytechnique
Palaiseau

Omer Giménez and Marc Noy
Universitat Politècnica de Catalunya
Barcelona

Abstract

We show that the diameter $D(G_n)$ of a random labelled connected planar graph with n vertices is asymptotically almost surely of order $n^{1/4}$, in the sense that there exists a constant $c > 0$ such that

$$P(D(G_n) \in (n^{1/4-\epsilon}, n^{1/4+\epsilon})) \geq 1 - \exp(-n^{c\epsilon})$$

for ϵ small enough and n large enough ($n \geq n_0(\epsilon)$). We prove similar statements for rooted 2-connected and 3-connected maps and planar graphs.

1 Introduction

The diameter of random maps has attracted a lot of attention since the pioneering work by Chassaing and Schaeffer [4] on the radius $r(Q_n)$ of random quadrangulations with n vertices, where they show that $r(Q_n)$ rescaled by $n^{1/4}$ converges as $n \to \infty$ to an explicit (continuous) distribution related to the Brownian snake. This suggests that random maps of size n are to be rescaled by $n^{1/4}$ in order to converge; precise definitions of the convergence can be found in [12, 7], and the (spherical) topology of the limit is studied in [8, 14]; some general statements about the limiting profile and radius are obtained in [11, 13]. At the combinatorial level, the two-point function of random quadrangulations has surprisingly a simple exact expression, a beautiful result found in [3] that allows one to derive easily the limit distibution (rescaled by $n^{1/4}$) of the distance between two randomly chosen vertices in a random quadrangulation. In contrast, little is known about the profile of random *unembedded* connected planar graphs, even if it is strongly believed that the results should be similar as in the embedded case.

We have not been able to show a limit distribution for the profile (or radius, diameter) of a random connected planar graph rescaled by $n^{1/4}$; instead we have obtained large deviation results on the diameter that strongly support the belief that $n^{1/4}$ is the right scaling order. We say that a property A, defined for all values n of a parameter, holds asymptotically almost surely if

$$P(A) \to 1, \qquad \text{as } n \to \infty.$$

In this case we write a.a.s. In this paper we need a certain rate of convergence of the probabilities. Suppose property A depends on a real number $\epsilon > 0$ (usually very small). Then we say that A holds a.a.s. with exponential rate if there is a constant $c > 0$, such that for every ϵ small enough there exist an integer $n_0(\epsilon)$ so that

$$P(\text{not } A) \leq e^{-n^{c\epsilon}} \qquad \text{for all } n \geq n_0(\epsilon). \tag{1}$$

The diameter of a graph (or map) G is denoted by $D(G)$. The main results proved in this paper are the following.

Theorem 1 *The diameter of a random connected labelled planar graph with n vertices is, a.a.s. with exponential rate, in the interval*

$$(n^{1/4-\epsilon}, n^{1/4+\epsilon}).$$

Theorem 2 *Let $1 < \mu < 3$. The diameter of a random connected labelled planar graph with n vertices and $\lfloor \mu n \rfloor$ edges is in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$ a.a.s. with exponential rate.*

This contrasts with so-called "subcritical" graph families, such as trees, outerplanar graphs, series-parallel graphs, where the diameter is in the interval $(n^{1/2-\epsilon}, n^{1/2+\epsilon})$ a.a.s. with exponential rate. (see the remark just before the bibliography).

The basis of our proof is the result for planar maps mentioned above. Then we prove the result for 2-connected maps using the fact that a random map has a large 2-connected component a.a.s. A similar argument allows us to extend the result to 3-connected maps, which proves it also for 3-connected planar graphs, because they have a unique embedding in the sphere. We then reverse the previous arguments and go first to 2-connected and then connected planar graphs, but this is not straightforward. One difficulty is that the largest 3-connected component of a random 2-connected graph does not have the typical ratio between number of edges and number of vertices, and this is why we must study maps with a given ratio between edges and vertices. In addition, we must show that there is a 3-connected component of size $n^{1-\epsilon}$ a.a.s. with exponential rate, and similarly for blocks. Finally, we must show that the height of the tree associated to the decomposition of a 2-connected graph into 3-connected components is at most n^{ϵ}, and similarly for the tree of the decomposition of a connected graph into blocks.

For lack of space, proofs are omitted in this extended abstract.

2 Quadrangulations and maps

We recall here the definitions of maps. A *planar map* (shortly called a map here) is a connected unlabelled graph embedded in the plane up to isotopic deformation. Loops and multiple edges are allowed. A *rooted map* is a map where an edge incident to the outer face is marked so as to have the outer face on its left; the *root-vertex* is the origin of the root. A *quadrangulation* is a map where all faces have degree 4.

We recall Schaeffer's bijection (itself an adaptation of an earlier bijection by Cori and Vauquelin [5]) between labelled trees and quadrangulations. A *rooted plane tree* is a rooted map with a unique face. A *labelled tree* is a rooted plane tree with a integer label $\ell(v) \in \mathbb{Z}$ on each vertex v so that the labels of the extremities of each edge $e = (v, v')$ satisfy $|\ell(v) - \ell(v')| \leq 1$, and such that the root vertex has label 0. A useful observation is that labelled trees are in bijection with rooted plane trees where a subset of the edges is oriented arbitrarily (for the onto mapping, one orients an edge with labels $(i, i+1)$ toward the vertex with label $i+1$ and one leaves an edge of type (i, i) unoriented). Thus the number of labelled trees with n edges is $3^n C_n$ with $C_n := (2n)!/n!/(n+1)!$ the nth Catalan number. A *signed* labelled tree is a pair (τ, σ) where τ is a labelled tree and σ is an element of $\{-1, +1\}$.

Theorem 3 (Schaeffer [15], Chassaing, Schaeffer [4]) *Signed labelled trees with n vertices are in bijection with rooted quadrangulations with n vertices and a secondary pointed vertex v_0. Each vertex v of a labelled tree corresponds to a non-pointed vertex ($\neq v_0$) in the associated quadrangulation Q, and $\ell(v) - \ell_{min} + 1$ gives the distance from v to v_0 in Q, where l_{min} is the minimum label in the tree.*

From this bijection, it is easy to show large deviation results for the diameter of a quadrangulation (the basic idea, originating in [4], is that the typical depth k of a vertex in the tree is $n^{1/2}$, and the typical discrepancy of the labels along a branch is $k^{1/2} = n^{1/4}$). The main result we use, from [6], is the property that (under general conditions) the height of a random tree of size n from a given family has diameter in $(n^{1/2-\epsilon}, n^{1/2+\epsilon})$ a.a.s. with exponential rate.

Lemma 4 (Flajolet et al. Theorem 3.1 in [6]) *Let $\mathcal{T}$ be a family of rooted trees endowed with a weight-function $w(.)$ so that the corresponding weighted series $y(z)$ is admissible (in a precise analytic sense not defined here).*

Let ξ be a height-parameter and let T_n be taken at random in $\mathcal{T}_n$ under the weighted distribution in size n. Then $\xi(T_n) \in (n^{1/2-\epsilon}, n^{1/2+\epsilon})$ a.a.s. with exponential rate.

Proposition 5 *The diameter of a random rooted quadrangulation with n vertices is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$.*

We also need a weighted version of the previous theorem. Recall that a rooted quadrangulation Q has a unique bicoloration of its vertices in black and white such that the origin of the root is black and each edge connects a black with a white vertex. Call it the *canonical bicoloration* of Q. Given $x > 0$, a rooted quadrangulation with v black vertices is weighted with parameter x if we assign to it weight x^v. The next theorem generalizes Proposition 5 to the weighted case. The analytical part of the proof is a little more delicate since the system specifying weighted labelled trees is two-lines, and has to be transformed to a one-line equation in order to apply Lemma 4.

Theorem 6 *Let $0 < a < b$. The diameter of a random quadrangulation weighted by x is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$, uniformly over $x \in [a, b]$.*

We recall the classical bijection between rooted quadrangulations with n faces (and thus $n+2$ vertices) and rooted maps with n edges. Starting from Q endowed with its canonical bicoloration, add in each face a new edge connecting the two (diagonally opposed) black vertices. Return the rooted map M formed by the newly added edges and the black vertices, rooted at the edge corresponding to the root-face of Q, and with same root-vertex as Q. Conversely, to obtain Q from M, add a new white vertex v_f inside each face f of M (even the outer face) and add new edges from v_f to every corner around f; then delete all edges from M, and take as root-edge of Q the one corresponding to the incidence root-vertex/outer-face in M. Clearly, under this bijection, vertices of a map correspond to black vertices of the associated quadrangulation, and faces correspond to white vertices.

Map families are here weighted at their vertices, i.e., for a given parameter $x > 0$, a map with v vertices has weight x^v.

Theorem 7 *Let $0 < a < b$. The diameter of a random rooted map with n edges and weight x at the vertices is in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$, uniformly over $x \in [a, b]$.*

Here it is convenient to include the empty map in the families $\mathcal{M} = \cup_n \mathcal{M}_n$ of rooted maps and $\mathcal{C} = \cup_n \mathcal{C}_n$ of rooted 2-connected maps. As described by Tutte in [16], a rooted map M is obtained by taking a rooted 2-connected map C, called the *core* of M, and then inserting in each corner i of C an arbitrary rooted map M_i. The maps M_i are called the *pieces* of M. Denoting by $M(x, z)$ ($C(x, z)$, resp.) the series of rooted connected (2-connected, resp.) maps according to non-root vertices and edges, this decomposition yields

$$M(x, z) = C(x, H(x, z)), \quad \text{where } H(x, z) = zM(x, z)^2, \tag{2}$$

since a core with k edges has $2k$ corners where to insert rooted maps.

An important property of the composition scheme is to preserve the uniform distribution, as well as the (vertex-)weighted distribution. Precisely, let M be a rooted map with n edges and weight x at the vertices. Let C be the core of M, call k its size, and let $M_1, \ldots, M_{2k}$ be the pieces of M, call $n_1, \ldots, n_{2k}$ their sizes. Then, conditioned to have size k, C is a random rooted 2-connected map with k edges and weight x at vertices, and conditioned to have size n_i the ith piece M_i is a random rooted map with n_i edges and weight x at vertices.

Lemma 8 *Let $0 < a < b$, and let $x \in [a, b]$. Let $\rho^{(x)}$ be the radius of convergence of $z \mapsto M(x, z)$. Following [1], define*

$$\alpha^{(x)} = \frac{H(x, \rho^{(x)})}{\rho^{(x)} H_z(x, \rho^{(x)})}.$$

Let $n \geq 0$, and let M be a random rooted map with n edges and weight x at vertices. Let $X_n = |C|$ be the size of the core of M, and let $M_1, \ldots, M_{2|C|}$ be the pieces of M. Then

$$P\big(X_n = \lfloor \alpha^{(x)} n \rfloor,\ \max(|M_i|) \leq n^{3/4}\big) = \Theta(n^{-2/3})$$

uniformly over $x \in [a, b]$.

In [1] the authors derive the limit distribution of X_n and they show that $P(X_n = \lfloor \alpha^{(x)} n \rfloor) = \Theta(n^{-2/3})$. So Lemma 8 says that the asymptotic order of $P(X_n = \lfloor \alpha^{(x)} n \rfloor)$ is the same under the additional condition that all pieces are of size at most $n^{3/4}$ (one could actually ask $n^{2/3+\delta}$ for any $\delta > 0$). A closely related result proved in [9] is that, for any fixed $\delta > 0$, there is a.a.s. no piece of size larger than $n^{2/3+\delta}$ provided the core has size larger than $n^{2/3+\delta}$.

Theorem 9 *For $0 < a < b$, the diameter of a random rooted 2-connected map with n edges and weight x at vertices is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$, uniformly over $x \in [a, b]$.*

In a similar way as when one goes from connected to 2-connected maps, there is a decomposition of 2-connected maps in terms of 3-connected components that allows to transfer the diameter concentration property from

2-connected to 3-connected maps. In this section it is convenient to exclude the loop-map from the family of 2-connected maps, so all 2-connected maps are loopless.

As shown by Tutte [16], a rooted 2-connected map C is either a series or parallel composition of 2-connected maps, or it is obtained from a rooted 3-connected map T where each non-root edge e is possibly substituted by a rooted 2-connected map C_e (identifying the extremities of e with the extremities of the root of C_e). In that case T is called the *3-connected core* of C and the components C_e are called the *pieces* of C. Call $C(x,z)$ ($\widehat{C}(x,z)$) the series counting rooted 2-connected maps (rooted 2-connected maps with a 3-connected core, resp.) according to vertices not incident to the root (variable x) and edges (variable z). Call $T(x,z)$ the series counting rooted 3-connected maps according to vertices not incident to the root (variable x) and edges (variable z). Then

$$\widehat{C}(x,z) = T(x, C(x,z)). \tag{3}$$

Accordingly, for a random rooted 2-connected map with n edges, weight x at vertices, and conditioned to have a 3-connected core T of size k, T is a random rooted 3-connected map with k edges and weight x at vertices; and each piece C_e conditioned to have a given size n_e is a random rooted 2-connected map with n_e edges and weight x at vertices.

Calling f_e the degree of the root face of C_e, we have

$$D(T) \leq D(C) \leq D(T) \cdot \max_e(f_e) + 2\max_e(D(C_e)). \tag{4}$$

The first inequality is trivial. The second one follows from the fact that a diametral path P in C starts in a piece, ends in a piece, and in between it passes by adjacent vertices $v_1, \ldots, v_k$ of H such that for $1 \leq i < k$, v_i and v_{i+1} are connected in H by an edge e and P travels in the piece C_e to reach v_{i+1} from v_i (since P is geodesic, its length in C_e is bounded by the distance from v_i to v_{i+1}, which is clearly bounded by f_e).

Theorem 10 *Let $0 < a < b$. The diameter of a random 3-connected map with n edges with weight x at the vertices is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$, uniformly over $x \in [a,b]$.*

3 Planar graphs

We need 3-connected graphs labelled at the edges (this is enough to avoid symmetries). The number of edges is now m, and n is reserved for the number of vertices. By Whitney's theorem 3-connected graphs have a unique embedding on the sphere (up to reflexion). Hence from the last theorem on 3-connected maps we obtain directly the following:

Theorem 11 *Let $0 < a < b$. The diameter of a random 3-connected planar graph with m edges with weight x at the vertices is, a.a.s. with exponential rate, in the interval $(m^{1/4-\epsilon}, m^{1/4+\epsilon})$.*

Before handling 2-connected planar graphs we treat the closely related family of (planar) *networks*. A *network* is a connected simple planar graph with two marked vertices called the poles, such that adding an edge between the poles, called the root-edge, makes the graph 2-connected. At first it is convenient to consider the networks as labelled at the edges.

Theorem 12 *Let $0 < a < b$. The diameter of a random network with m edges with weight x at the vertices is, a.a.s. with exponential rate, in the interval*

$$(m^{1/4-\epsilon}, m^{1/4+\epsilon}),$$

uniformly over $x \in [a, b]$.

Lemma 13 *Let $1 < a < b < 3$. For $N_{n,m}$ a network with n vertices and m labelled edges taken uniformly at random, $D(N_{n,m}) \in (n^{1/4-\epsilon}, n^{1/4+\epsilon})$ a.a.s. with exponential rate, uniformly over $m/n \in [a, b]$.*

An important remark is that networks with n vertices and m edges can be labelled either at vertices or at edges, and the uniform distribution in one case corresponds to the uniform distribution in the second case. Hence the result of Lemma 13 holds for random networks with n vertices and m edges and labelled at vertices.

It is proved in [2] that for a random network N_n with n vertices the ratio $r = \#edges/\#vertices$ is concentrated around a certain $\mu \approx 2.2$, implying that for $\delta > 0$ $P(r \notin [\mu - \delta, \mu + \delta])$ is exponentially small. Hence $D(N_n) \in (n^{1/4-\epsilon}, n^{1/4+\epsilon})$ a.a.s. with exponential rate. The same holds for the diameter of a random 2-connected planar graph B_n with n vertices (indeed 2-connected planar graphs are a subset of networks, the ratios of the cardinalities being of order n). We obtain:

Theorem 14 *The diameter of a random 2-connected planar graph with n vertices is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$.*

We prove here from Theorem 14 that a random connected planar graph with n vertices has diameter in $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$ a.a.s. with exponential rate. We use the well known decomposition of a connected planar graph C into 2-connected blocks such that the incidences of the blocks with the vertices form a tree. An important point is that if C is chosen uniformly at random among connected planar graphs with n vertices, then each block B of C is uniformly distributed when conditioned to have a given size. Formulated on pointed graphs, the block-decomposition ensures that a pointed planar graph is obtained as follows: take a collection of 2-connected pointed planar graphs, and merge their pointed vertices into a single vertex; then attach at each non-marked vertex v in these blocks a pointed connected planar graph C_v. Calling $C(z)$ ($B(z)$) the series counting pointed connected (2-connected, resp.) planar graphs, this yields the equation

$$F(z) = z\exp(B'(F(z))), \quad \text{where } F(z) = zC'(z). \tag{5}$$

Note that the inverse of $F(z)$ is the function $\phi(u) = u\exp(-g(u))$, where $g(u) := B'(u)$. Call ρ the radius of convergence of $C(z)$ and R the radius of convergence of $B(u)$.

Lemma 15 *A random connected planar graph with n vertices has a block of size at least $n^{1-\epsilon}$ a.a.s. with exponential rate.*

Lemma 15 directly implies that a random connected planar graph with n vertices has diameter at least $n^{1/4-\epsilon}$. Indeed it has a block of size $k \geq n^{1-\epsilon}$ a.a.s. with exponential rate and since the block is uniformly distributed in size k, it has diameter at least $k^{1/4-\epsilon}$ a.a.s. with exponential rate.

Let us now prove the upper bound, which relies on the following lemma:

Lemma 16 *The block-decomposition tree τ of a random connected planar graph with n vertices has diameter at most n^{ϵ} a.a.s. with exponential rate.*

Lemma 16 easily implies that the diameter of a random connected planar graph C with n vertices is at most $n^{1/4+\epsilon}$ a.a.s. with exponential rate. Indeed, calling τ the block-decomposition tree of C and B_i the blocks of C, one has

$$D(C) \leq D(\tau) \cdot \max_i D(B_i).$$

Lemma 16 ensures that $D(\tau) \leq n^{\epsilon}$ a.a.s. with exponential rate. Moreover Theorem 14 easily implies that a random 2-connected planar graph of size $k \leq n$ has diameter at most $n^{1/4+\epsilon}$ a.a.s. with exponential rate, whatever $k \leq n$ is (proof by splitting in two cases: $k \leq n^{1/4}$ and $n^{1/4} \leq k \leq n$). Hence, since each of the blocks has size at most n, $\max_i D(B_i) \leq n^{1/4+\epsilon}$ a.a.s. with exponential rate. Therefore we have completed the proof of Theorem 1.

Theorem 17 *The diameter of a random connected planar graph with n vertices is, a.a.s. with exponential rate, in the interval $(n^{1/4-\epsilon}, n^{1/4+\epsilon})$.*

Similarly one shows that a random planar graph with n vertices has a connected component of size at least $n^{1-\epsilon}$ a.a.s. with exponential rate, which yields Theorem 1.

To show Theorem 2, one needs to extend the statements of Theorem 14 and Lemmas 15, 16 to the case of a random graph of size n with weight $y > 0$ on each edge. Then, one uses the fact (proved in [10]) that for each $\mu \in (1, 3)$ there exists $y > 0$ such that a random planar graph with n edges and weight y on edges has probability $\Theta(n^{-1/2})$ to have $\lfloor \mu n \rfloor$ edges.

We conclude with a remark on so-called "subcritical" graph families, these are the families where the system

$$y = z \exp(B'(y)) =: F(z, y) \tag{6}$$

to specify pointed connected from pointed 2-connected graphs in the family is admissible, i.e., $F(z, y)$ is analytic at (ρ, τ) where ρ is the radius of convergence of $y = y(z)$ and $\tau = y(\rho)$.

Define the *block-distance* of a vertex v in a vertex-pointed connected graph G as the minimal number of blocks one can use to travel from the pointed vertex to v; and define the *block-height* of G as the maximum of the block-distance over all vertices of G. With the terminology of Lemma 4, one easily checks that the block-height is a height-parameter for the system (6). Hence by Lemma 4, the block-height h of a random pointed connected graph G with n vertices from a subcritical family is in $[n^{1/2-\epsilon}, n^{1/2+\epsilon}]$ a.a.s. with exponential rate. Clearly $D(G) \geq h - 1$ since the distance between two vertices is at least the block-distance minus 1. Hence $D(G) \geq n^{1/2-\epsilon}$ a.a.s. with exponential rate. For the upper bound, note that $D(G) \leq h \cdot \max_i(|B_i|)]$, where the B_i's are the blocks of G. Because of the subcritical condition one easily shows that $\max_i(|B_i|) \leq n^{\epsilon}$ a.a.s. with exponential rate. This implies that $D(G) \leq n^{1/2+\epsilon}$ a.a.s. with exponential rate.

References

[1] 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.

[2] E. Bender, Z. Gao, and N. Wormald. The number of labeled 2-connected planar graphs. *Electron. J. Combin.*, 9:1–13, 2002.

[3] J. Bouttier, P. Di Francesco, and E. Guitter. Geodesic distance in planar graphs. *Nucl. Phys.*, B663:535–567, 2003.

[4] P. Chassaing and G. Schaeffer. Random planar lattices and integrated superBrownian excursion. *Probab. Theory Related Fields*, 128(2):161–212, 2004.

[5] R. Cori and B. Vauquelin. Planar maps are well labeled trees. *Canad. J. Math.*, 33(5):1023–1042, 1981.

[6] P. Flajolet, Z. Gao, A. M. Odlyzko, and L. B. Richmond. The distribution of heights of binary trees and other simple trees. *Combin. Probab. Comput.*, 2(145-156), 1993.

[7] J. F. Le Gall. The topological structure of scaling limits of large planar maps. *Invent. Math.*, 169(3):621–670, 2007.

[8] J. F. Le Gall and F. Paulin. Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere. *Geom. Funct. Anal.*, 18:893–918, 2008.

[9] J. Gao and N. Wormald. The size of the largest components in random planar maps. *SIAM J. Discrete Math.*, 12(2):217–228, 1999.

[10] O. Giménez and M. Noy. Asymptotic enumeration and limit laws of planar graphs. *J. Amer. Math. Soc.*, 22:309–329, 2009.

[11] J. F. Marckert and G. Miermont. Invariance principles for random bipartite planar maps. *Ann. Probab.*, 35(5):1642–1705, 2007.

[12] J. F. Marckert and A. Mokkadem. Limit of normalized random quadrangulations: The brownian map. *Ann. Probab.*, 34(6):2144–2202, 2006.

[13] G. Miermont. An invariance principle for random planar maps. In *Fourth Colloquium in Mathematics and Computer Sciences CMCS'06, DMTCS Proceedings AG*, pages 39–58, 2006.

[14] G. Miermont. On the sphericity of scaling limits of random planar quadrangulations. *Elect. Comm. Probab.*, 13:248–257, 2008.

[15] G. Schaeffer. *Conjugaison d'arbres et cartes combinatoires aléatoires.* PhD thesis, Université Bordeaux I, 1998.

[16] W. T. Tutte. A census of planar maps. *Canad. J. Math.*, 15:249–271, 1963.