On the treewidth and related parameters of random geometric graphs*

Dieter Mitsche[1] and Guillem Perarnau[2]

1 Dept. of Mathematics, Ryerson University, 350 Victoria St., Toronto, ON, Canada. dmitsche@ryerson.ca

2 MA4 - Universitat Politècnica de Catalunya, C/ Jordi Girona 1-3, 08034 Barcelona, Spain. guillem.perarnau@ma4.upc.edu

Abstract

We give asymptotically exact values for the treewidth $\mathrm{tw}(G)$ of a random geometric graph $\mathcal{G}(n,r)$ in $[0,\sqrt{n}]^2$. More precisely, we show that there exists some $c_1 > 0$, such that for any constant $0 < r < c_1$, $\mathrm{tw}(G) = \Theta(\frac{\log n}{\log \log n})$, and also, there exists some $c_2 > c_1$, such that for any $r = r(n) \geq c_2$, $\mathrm{tw}(G) = \Theta(r\sqrt{n})$. Our proofs show that for the corresponding values of r the same asymptotic bounds also hold for the pathwidth and treedepth of a random geometric graph.

1998 ACM Subject Classification G.2.2 Graph theory

Keywords and phrases Random geometric graphs, treewidth, treedepth

Digital Object Identifier 10.4230/LIPIcs.STACS.2012.408

1 Introduction

Starting with the seminal paper of Gilbert [5], random geometric graphs have in recent decades received a lot of attention as a model for large communication networks such as sensor networks. Network agents are represented by the vertices of the graph, and direct connectivity is represented by edges. For applications of random geometric graphs, we refer to Chapter 3 of [7], and for a survey of many theoretical results, we refer to Penrose's monograph [12].

Given a set V of n vertices and a nonnegative real $r = r(n)$, a random geometric graph is defined as follows: each vertex is placed at some position of the square $S_n = [0,\sqrt{n}]^2$, chosen independently and uniformly at random. This choice of the square is only for convenience; by suitable scaling of r we could have chosen the square $[0,1]^2$ and the results were still valid.

Note that with probability 1 no two vertices choose the same position. We will identify each vertex with each position, that is, $u \in V$ refers also to the geometrical position of u in the square. Then we define $\mathcal{G}(n,r)$ as the random graph having V as the vertex set with $|V| = n$, and with an edge connecting each pair of vertices $u, v \in V$ at distance $d(u,v) \leq r$, where $d(\cdot,\cdot)$ denotes the Euclidean distance. In order to simplify calculations, we will use the well-known idea of Poissonization (see [12]): we assume that the vertices of $\mathcal{G}(n,r)$ are generated according to a Poisson point process of intensity 1 over the square $S_n = [0,\sqrt{n}]^2$.

* This research is partially supported by the Catalan Research Council under project 2009SGR1387. The first author is partially supported by the ICT Program of the European Union under contract number 215270 (FRONTS). The second author wants to thank the FPU grant from the Ministerio de Educación de España.

Conditioned under the fact that this Poisson point process generates exactly n vertices (which happens with probability $\Theta(1/\sqrt{n})$), this model and the standard model of random geometric graphs have the same uniform distribution of the n vertices, and we will use this equivalence from now on.

All our stated results are asymptotic as $n \to \infty$. We use the usual notation *a.a.s.* to denote *asymptotically almost surely*, i.e. with probability $1 - o(1)$. It is well known that the property of the existence of a giant component of order $\Theta(n)$ undergoes a sharp threshold in $\mathcal{G}(n,r)$ (see e.g. [6]), but the exact value r is not yet known. However, there exist two positive constants c^-, c^+ such that for $r \le c^-$, *a.a.s.* the largest component of $\mathcal{G}(n,r)$ is *a.a.s.* of order $O(\log n)$, whereas for $r \ge c^+$, a component of order $\Theta(n)$ is present (see [12]).

In this paper, we study the behaviour of two tree-like parameters, the treewidth and the treedepth, on random geometric graphs.

The treewidth of a graph measures the similarity between a tree and G. It was introduced by Robertson and Seymour in [17] inside their series of articles on graph minors. It has several applications in graph theory and algorithmics; one good example is Courcelle's Theorem [1].

For a graph $G = (V, E)$ on n vertices, we call (T, W) a *tree decomposition* of G, where W is a set of vertex subsets $W_1, \dots, W_s \subseteq V(G)$ and T is a forest with vertices in W, such that

1. $\bigcup W_i = V(G)$
2. For any $e = uv \in E(G)$ there exists a set W_i such that $u, v \in W_i$
3. For any $v \in V(G)$, the subgraph induced by the $W_i \ni v$ is connected as a subgraph of T.

The *width* of a tree-decomposition is $\mathrm{w}(T,W) = \max_i |W_i| - 1$, and the *treewidth* of a graph G can be defined as

$$\mathrm{tw}(G) = \min_{(T,W)} \mathrm{w}(T,W).$$

A vertex partition $V = (A, S, B)$ is a *balanced k-partition* if $|S| = k+1$, S separates A and B, and $\frac{1}{3}(n-k-1) \le |A|, |B| \le \frac{2}{3}(n-k-1)$. Then S is called a *balanced separator*. The following result connecting balanced partitions and treewidth is due to Kloks [9].

▶ **Lemma 1** ([9]). *Let G be a graph with n vertices and $\mathrm{tw}(G) \le k$ such that $n \ge k - 4$. Then G has a balanced k-partition.*

The treedepth $\mathrm{td}(G)$ of a graph G was introduced by Nešetřil and Ossona de Mendez as a tree-like parameter in the scope of homomorphism theory. In particular, it provides an alternative definition of bounded expansion classes [11]. Moreover, the notion of the treedepth is closely connected to the treewidth. Intuitively speaking, the treewidth of a graph G is a parameter that measures the similarity between G and a certain tree, while the treedepth of G measures how close G is to a star. In other words, the treedepth also takes into account the diameter of the tree we are comparing the graph with.

This concept of treedepth has been introduced using different names in the literature. It is equivalent to the height of an elimination tree used in Cholesky decomposition [14]. Analogous definitions can be found using the terminology of rank function [10], vertex ranking number (or ordered coloring) [3] or weak coloring number [8].

Let T be a rooted tree. The *closure* of T is the graph that has the same set of vertices and two vertices are connected if they are relatives (ancestor or predecessor) in T. Consider a *rooted forest* as the disjoint union of rooted trees whose height is the maximum of the height among all the trees. The closure of a rooted forest will consist of the disjoint union of the closures of each rooted tree. The *treedepth* of a graph G, $\mathrm{td}(G)$, is defined to be the minimum height of a rooted forest, whose closure contains G as a subgraph.

Observe that, by definition, if G is a graph with components $H_1, \dots, H_m$,

$$\mathrm{tw}(G) = \max_i \mathrm{tw}(H_i), \qquad \mathrm{td}(G) = \max_i \mathrm{td}(H_i). \tag{1}$$

This two parameters are closely related by the following inequalities:

$$\mathrm{tw}(G) \le \mathrm{td}(G) \le \mathrm{tw}(G)(\log n + 1),$$

both bounds being sharp. For example, if S is a star, $\mathrm{tw}(S) = \mathrm{td}(S) = 1$, while if P_n is a path of length n, $\mathrm{tw}(P_n) = 1$ and $\mathrm{td}(P_n) = \lfloor \log n \rfloor + 1$.

Results and organization of the paper. In this paper we study the values of $\mathrm{tw}(G)$ and $\mathrm{td}(G)$ of a random geometric graph $G = \mathcal{G}(n, r)$ for different values of $r = r(n)$. In particular, we prove the following two main theorems:

▶ **Theorem 2.** *There is some constant $0 < c_1 < c^-$, such that for any $0 < r \le c_1$, a.a.s. $\mathrm{tw}(\mathcal{G}(n,r)) = \Theta(\frac{\log n}{\log\log n})$, and also a.a.s. $\mathrm{td}(\mathcal{G}(n,r)) = \Theta(\frac{\log n}{\log\log n})$.*

▶ **Theorem 3.** *There is some constant $c_2 > c^+$, such that for any $r = r(n) \ge c_2$, a.a.s. $\mathrm{tw}(G(n,r)) = \Theta(r\sqrt{n})$, and also a.a.s. $\mathrm{td}(\mathcal{G}(n,r)) = \Theta(r\sqrt{n})$.*

▶ Remark. For $G = \mathcal{G}(n, r)$ with r constant, but $r \ge c_2$, by the results of [2], many problems such as Steiner Tree, Feedback Vertex Set, Connected Vertex Cover can be solved in time $O(poly(n)3^{\sqrt{n}})$, and Connected Dominating Set, Connected Feedback Vertex Set, Min Cycle Cover, Longest Path, Longest Cycle, Graph Metric Travelling Salesman Problem can be solved in time $O(poly(n)4^{\sqrt{n}})$.

▶ Remark. Other width parameters that are sandwiched between treewidth and treedepth will have the same asymptotic behavior in $\mathcal{G}(n, r)$. For instance, the *pathwidth* of a graph, introduced by Robertson and Seymour [16], is defined to be the similarity between a graph and a path. Since the pathwidth is bounded from below by the treewidth and bounded from above by the treedepth (see Theorem 5.3 and Theorem 5.11 of [18]), the former theorems imply that for those values of $r = r(n)$ the pathwidth of the graph is of the same order.

We point out that it is an interesting feature of $\mathcal{G}(n, r)$ that treewidth and treedepth are asymptotically of the same order for a wide range of parameters r, since this is not true for random graphs in general [13]. The similar value of treedepth and treewidth implies that $\mathcal{G}(n, r)$ is more similar to a star–shaped tree than to a path–shaped tree, which in general is not true for random graphs. Observe also that in the period before the giant component the tree-like parameters are proportionally larger respect to the order of the components than when a giant component appears. In the classical random graph model the existence of a linear number of edges slightly above the giant component already implies a linear treewidth (see [4]), whereas a random geometric graph with the same number of edges (and a giant component) only has treewidth $\Theta(\sqrt{n})$.

In Section 2 we give the proof of Theorem 2. Whereas the lower bound follows from a standard argument about the maximum clique order, the proof of the upper bound is more involved. In Section 3 we continue by proving Theorem 3. Finally, in Section 4 we conclude mentioning open problems.

2 Proof of Theorem 2

Let $r_t = \Theta(1)$ the (not yet known) threshold radius of having a giant component, i.e. a connected component H with $V(H) = \Theta(n)$. In this section we will compute the treedepth for a random geometric graph with $r < r_t$, i.e. when there is no giant component. We also assume $r = \Theta(1)$. From now on, unless otherwise stated, we will call the vertices of $\mathcal{G}(n, r)$ as *points*, since we use *vertex* for a different graph related to $\mathcal{G}(n, r)$ (see below). In [12] it is shown that the order of the largest component in this case is *a.a.s.* $\Theta(\log n)$, and we will assume this from now on. This implies directly the coarse upper bound $\text{td}(G) = O(\log n)$.

For the sake of simplicity, we assume, moreover that $r < c_1$, where c_1 is a constant chosen in such a way that the order of each component is *a.a.s.* at most $\log n$ (this value exists, see Theorem 10.3 of [12], and is only chosen to simplify calculations).

We derive a lower bound on $\text{tw}(G)$ by studying the clique number of G, $\omega(G)$. Tessellate S_n into square *cells* of side length $r/\sqrt{2}$. Note that we have a linear number of such cells and note that any two points in the same cell are connected by an edge. The distribution of the number of points inside the cells can be modeled as a *balls and bins* problem: we have n balls and $m = \Theta(n)$ bins, and each of the n balls is thrown independently and uniformly at random into one of the bins. Denoting X_i denote the number balls inside the cell C_i, classical results (see e.g. [15]) state that if $m = \Theta(n)$, then $\max_i X_i = (1 + o(1))\frac{\log n}{\log\log n}$ a.a.s..

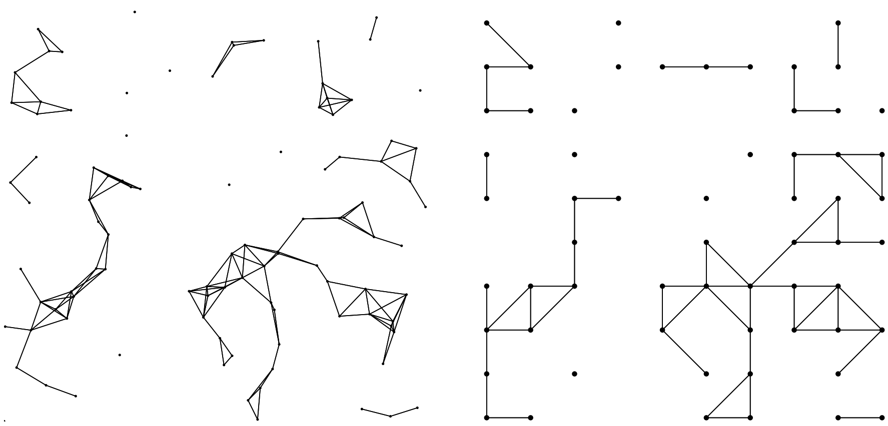
As any pair of points that belong to the same cell of the tessellation, is connected by an edge, G contains a clique subgraph formed of $\max_i X_i$ points, and therefore

$$\text{td}(G) \geq \text{tw}(G) \geq \omega(G) = \frac{\log n}{\log\log n}.$$

We will now show an upper bound on $\text{td}(G)$ which asymptotically matches this lower bound. We use the following lemma.

▶ **Lemma 4.** *Let* $X \sim \text{Po}(\lambda)$*. For any* $k \geq 2\lambda$*,* $\Pr(X \geq k) \leq 2\Pr(X = k)$*.*

Having tessellated S_n into cells, we construct a *cell-graph* C_G of G using the following criterion: each non-empty cell will be represented by a *vertex* and two vertices of C_G will be joined if there exist two points in the corresponding cells of G that share an edge (see Figure 1, where the tessellation is omitted for clarity). The cell-graph C_G has a structure similar to the original graph, but simpler.



(a) Random geometric graph **(b)** Cell-graph

Figure 1 A random geometric graph and its corresponding cell graph

Having in mind the previously established lower bound on the order of the maximum clique, set $\text{T}_{\max} = \frac{\log n}{\log\log n}$. We focus on a certain connected component H of G that will have order at most $\log n$. Note that there are at most n different components, not necessarily all of logarithmic order. Let C_H denote the cell-graph of component H. Note that since the side length is $\frac{r}{\sqrt{2}}$, each cell belongs to at most one connected component. Letting A_i be the number of points in the cell i (which will produce an A_i-clique) the number of points in H can be written as $\sum_{i \in V(C_H)} A_i$, and we have $\sum_{i \in V(C_H)} A_i \leq \log n$.

We will call a cell of the tessellation *sparse* if it contains less than $T = \frac{\sqrt{\log n}}{\log \log n}$ points, and *dense* otherwise. Observe that all the cells contain at most $\mathrm{T}_{\max}$ points.

▶ **Proposition 5.** *For any component H, the number of points belonging to dense cells is a.a.s. not larger than $O(\mathrm{T}_{\max})$.*

Proof. Since $A_i \sim \mathrm{Po}(\lambda)$, for some constant $\lambda = \lambda(r)$,

$$p = \Pr(A_i \geq T) \geq \Pr(A_i = T) \sim \frac{e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^T, \tag{2}$$

using the Stirling approximation $T! \sim \sqrt{2\pi T}\left(\frac{T}{e}\right)^T$.

To count the number of points lying in dense cells, we define the following random variables:

$$Y_i = \begin{cases} t & \text{if } i \text{ is dense and has } t \text{ points inside} \\ 0 & \text{otherwise} \end{cases}$$

Our aim is to show that $Y = \sum_{i \in V(C_H)} Y_i$ is at most $O(\mathrm{T}_{\max})$. In this case (at least to us) it is not clear how a Chernoff type inequality can be used. Nevertheless, we will show that the probability that Y is larger than $8\,\mathrm{T}_{\max}$ is $o(n^{-1})$ and taking a union bound over all at most n components, *a.a.s.* no component will have more than $8\,\mathrm{T}_{\max}$ points in dense cells.

The probability of having a sparse cell is $1-p$, while the probability of having $T+j$ points inside a cell is $\Pr(\mathrm{Po}(\lambda) = T+j) \sim \frac{e^{-\lambda}}{\sqrt{2\pi(T+j)}}\left(\frac{e\lambda}{T+j}\right)^{T+j}$. Since $\frac{e^{-\lambda}}{\sqrt{2\pi(T+j)}}\left(\frac{e\lambda}{T+j}\right)^{T+j} \leq (\frac{e\lambda}{T})^T \frac{e^{-\lambda}}{\sqrt{2\pi T}} (\frac{e\lambda}{T})^j$, and using (2) we have

$$\Pr(\mathrm{Po}(\lambda) = T+j) \leq p\left(\frac{e\lambda}{T}\right)^j.$$

These observations lead to the definition of the following random variable:

$$R_i = \begin{cases} 0 & \text{with probability } 1-p. \\ T+j & \text{with probability } p\left(\frac{e\lambda}{T}\right)^j \text{ for any } j \geq 1. \\ T & \text{with probability } p\left(1 - \frac{e\lambda}{T-e\lambda}\right) \end{cases}$$

First of all, observe that R_i is a probability distribution. The random variables Y_i and R_i have similar distributions. In fact, each variable R_i stochastically dominates the corresponding random variable Y_i. Analogously we define $R = \sum_{i \in V(C_H)} R_i$. Then,

$$\Pr(Y > t) \leq \Pr(R > t) \text{ for any } t \in \mathbb{R} \tag{3}$$

and in particular this holds, if $t = O(\mathrm{T}_{\max})$.

Now we compute explicitly an upper bound for $\Pr(R > 8\,\mathrm{T}_{\max})$. We have $|V(C_H)| < \log n$ cells in H. There are n initial cells and then at most e^s different connected sets of s cells, and for this reason there are at most $ne^{\log n}$ ways to construct C_H. Assuming that i of them are dense, we have $\binom{|V(C_H)|}{i}$ ways to choose them, and after that, at most $(\log n)^i$

ways to distribute the points among these cells. The probability of having a dense cell with $R_i = T + j$ is $p\left(\frac{e\lambda}{T}\right)^j$, so that

$$\begin{aligned}\Pr(R > 8\,\mathrm{T}_{\max}) &= ne^{\log n} \sum_{i=1}^{|V(C_H)|} \sum_{S \in \binom{V(C_H)}{i}} \sum_{\sum_{j\in S} c_j \geq 8\,\mathrm{T}_{\max}} \Pr\left(\bigwedge_{j\in S} A_j = c_j\right) \\ &\leq n^2 \sum_{i=0}^{|V(C_H)|} \binom{\log n}{i} (\log n)^i \prod_{j=1}^{i} p\left(\frac{e\lambda}{T}\right)^{c_j - T}.\end{aligned}$$

We use the upper bound $\binom{\log n}{i} \leq (\log n)^i$. It must be also stressed that we have

$$\prod_{j=1}^{i} p\left(\frac{e\lambda}{T}\right)^{c_j-T} \leq (p\sqrt{2\pi T})^{\frac{\sum c_j}{T}} < (p\sqrt{2\pi T})^{8\sqrt{\log n}},$$

since $\sum c_j > 8\,\mathrm{T}_{\max}$. Moreover, since $c_j > T$ (the cells are dense), we have for $i = 8\sqrt{\log n}+k$, $\prod_{j=1}^{i} p\left(\frac{e\lambda}{T}\right)^{c_j-T} \leq p^{8\sqrt{\log n}} p^k$. Therefore, it is useful to split the former equation into two sums:

$$\begin{aligned}\Pr(R > 8\,\mathrm{T}_{\max}) &\leq n^2 \sum_{i=0}^{8\sqrt{\log n}} (\log n)^{2i} (p\sqrt{2\pi T})^{8\sqrt{\log n}} \\ &\quad + n^2 \left((\log n)^2 p\right)^{8\sqrt{\log n}} \sum_{k>0} \left((\log n)^2 p\right)^k\end{aligned}$$

As $\left((\log n)^2 p\right)^k < 1/2$, the infinite sum is less than one. Therefore,

$$\begin{aligned}\Pr(R > 8\,\mathrm{T}_{\max}) &\leq n^2 \left(8\sqrt{\log n}+1\right) \left((\log n)^2 \sqrt{2\pi T} p\right)^{8\sqrt{\log n}} \\ &\sim \exp\left\{2\log n + 4\log\log n + 8\sqrt{\log n}\,(2\log\log n + \log p + O(\log T))\right\}\end{aligned}$$

Since $p \sim \frac{c}{\sqrt{T}}\left(\frac{e\lambda}{T}\right)^T$, by Lemma 4 and (2), $\log p \sim -\frac{1}{2}\sqrt{\log n}$. The term $\Theta(\log T) = O(\log\log n)$ is negligible and thus,

$$\Pr(R > 8\,\mathrm{T}_{\max}) \leq \exp\{-(1+o(1))2\log n\} = O(n^{-2}). \tag{4}$$

By (3), this also implies that $\Pr(Y > 8\,\mathrm{T}_{\max}) = O(n^{-2})$, and by taking a union bound over all components, this implies that *a.a.s.* there is no component having more than $8\,\mathrm{T}_{\max}$ points inside dense cells. ◀

In order to obtain the desired matching upper bound, we need to construct a representation of the shape of the connected components which simplifies the structure. We now tessellate the square $[0, \sqrt{n}]^2$ into square cells of side length r. Proposition 5 also follows for this kind of tessellation since the size of the cells differs just by a constant factor. Consider now the cell graph C_G from this tessellation. Observe that the points belonging to a cell can only be connected by an edge to points in the same cell and to points in one of the at most 8 cells adjacent to that cell. Therefore, C_G will be a subgraph of the diagonal two-dimensional grid graph $L_{\sqrt{n},\sqrt{n}}$, where each cell is adjacent to the 8 cells surrounding it. The following proposition will be useful:

▶ **Proposition 6.** *Let $L_{m,n}$ be a diagonal two-dimensional grid graph and suppose that $m \le n$. Then*

$$\mathrm{td}(L_{m,n}) \le m \log n.$$

Fix a component C_H of C_G. We know that $|V(C_H)| \le |V(H)| \le \log n$ and hence, the diameter of C_H is at most $\log n$. Without loss of generality we may assume that each vertex is connected to all its 8 neighbouring cells provided that they are non empty. Take a vertex $v \in V(C_H)$ for which there exists some other vertex at distance the diameter of C_H. The vertices of C_H at distance d from v are said to be in the d-th floor. We also refer to the points inside the cells at distance d from v as points in the d-th floor.

We provide an elimination scheme for H. We want to find a balanced separator of this component (both parts will have linear order) that contains at most $\frac{\log n}{(\log\log n)^2}$ points. In particular, the separator set will be chosen among the different floors of C_H, corresponding to points that belong to cells at some fixed distance from v in C_H. Select the last floor f such that the number of points in lower floors is at most $|V(H)|/2$. Observe that this is always a separator that splits the graph H into two smaller pieces of order at most $|V(H)|/2$. If this separator of H has order at most $\frac{\log n}{(\log\log n)^2}$, we align in the elimination tree the points of the separator in a path, and we proceed recursively for the two subgraphs. The subtrees corresponding to these subgraphs are attached as children of the last node in the separator.

Suppose now that the floor f contains more than $\frac{\log n}{(\log\log n)^2}$ points of H. Then we can have many consecutive floors, before and after f, with more than $\frac{\log n}{(\log\log n)^2}$ points. However, since the order of the component H is at most $\log n$, there can be at most $(\log\log n)^2$ such floors.

Considering C_H, this implies that we have at most $(\log\log n)^2$ such consecutive floors containing more than $\frac{\log n}{(\log\log n)^2}$ points. Let us call the cell graph of these floors L'. Right after and before these floors we have two small cuts in C_H (meaning that they contain less than $\frac{\log n}{(\log\log n)^2}$ points), call them A' and B' respectively. We will recursively repeat this procedure for the two remaining parts A (the floors before A') and B (the floors after B') (see Fig.2). Observe that both A and B contain at most $|V(H)|/2$ points each (but they could contain much less, and in fact B could be empty).

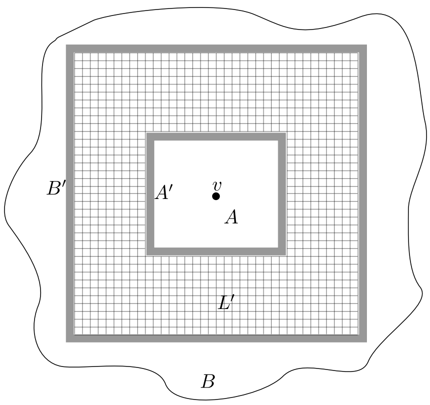


Figure 2 Decomposition of C_H

Focus now on L'. This is a subgraph of at most 4 copies of the diagonal grid $\log n \times (\log\log n)^2$ (see Fig.2), since there are at most $\log n$ points in each floor and therefore at most $\log n$ cells containing them. By cutting these 4 copies and by using Proposition 6,

$$\mathrm{td}_{C_G}(L') \le O\left((\log\log n)^3\right)$$

where td_{C_G} denotes the treedepth in the cell-graph.

The decomposition of $C_H = (A, A', L', B', B)$ gives the following inequality:

$$\mathrm{td}_{C_G}(C_H) \le \frac{2\log n}{(\log\log n)^2} + \max\{\mathrm{td}_{C_G}(A), O\left((\log\log n)^3\right), \mathrm{td}_{C_G}(B)\}, \tag{5}$$

since, as A' and B' were two floors with few points inside, $|A' \cup B'| \le \frac{2\log n}{(\log\log n)^2}$.

Observe that there exists $\alpha, \beta \le 1/2$ such that $|A| \le \alpha|V(H)|$ and $|B| \ge \beta|V(H)|$, and therefore, since the diameter of C_H is at most $\log n$, we can repeat this procedure at most $\log_2 |V(H)| = O(\log\log n)$ times. The constants α and β may change in each step but they are uniformly bounded by 1/2. Hence, $\mathrm{td}_{C_G}(C_H) \le O\left(\frac{\log n}{\log\log n}\right) = O(\mathrm{T}_{\max})$.

Now we are able to finish the proof of Theorem 2. By Proposition 5, we know that there are at most $O(\mathrm{T}_{\max})$ points in dense cells. We temporarily remove all these points, and add them at the end. Any of the remaining cells now has at most T points. We apply the previously described strategy of decomposition, the only difference being that each cell of L' contains now at most T points of G since there are no dense cells. Therefore, for the subgraph corresponding to L' in H we have $\mathrm{td}(L') \le O\left(T(\log\log n)^3\right)$.

Since $T(\log\log n)^3 = o\left(\frac{2\log n}{(\log\log n)^2}\right)$, the upper bound on $\mathrm{td}(H)$ that arises from the formula (5) applied on the original graph, is not affected. Therefore, the treedepth of the component after removing the dense cells is at most $O(\mathrm{T}_{\max})$. Finally, taking into account all the points corresponding to the dense cells by attaching them all in a path above the root of the elimination tree for the non dense cells, we still have

$$\mathrm{td}(H) \le O\left(\frac{\log n}{\log\log n}\right),$$

since adding a point increases the treedepth by at most 1. Using Equation (1), we have proven Theorem 2.

3 Proof of Theorem 3

Fix now $r = r(n) \ge c_2$, for some sufficiently large constant c_2 above r_t, the threshold radius of having a giant component. We will first give a strategy to construct an elimination tree for G, thus giving an upper bound on $\mathrm{td}(G)$.

Given $A \subseteq [0, \sqrt{n}]^2$, we denote by $\mathrm{vol}(A)$ the area of A. We need the following lemma:

► **Lemma 7.** *For any $A \subseteq [0, \sqrt{n}]^2$ such that $\mathrm{vol}(A) \ge c\log n$ and any $\delta > 0$, the number of points inside A is a.a.s. at most $(1+\delta)\,\mathrm{vol}(A)$.*

► **Proposition 8.** *For any $r \ge c_2$, $\mathrm{td}(G) \le r\sqrt{n}$.*

Proof. We tessellate the square $[0, \sqrt{n}]^2$ into square cells of length r. Denote, moreover, by $C_{(i,j)}$ the j-th such cell in the i-th row, for $1 \le i, j \le m = \sqrt{n}/r$.

We provide some tree decomposition, such that G can be embedded as a subgraph of the closure of the tree. Define $X_1 = \cup_{i=0}^{m} C_{(\lfloor m/2 \rfloor, i)}$ and denote by $Y_1 = \{y_1, \ldots, y_s\}$ the points inside the cells of X_1 (in arbitrary order). We start constructing the tree by putting the root into y_1 and by attaching the path $y_1 - \cdots - y_s$. Next, we define $X_2^1 = \cup_{i=0}^{\lfloor m/2 \rfloor - 1} C_{(i, \lfloor m/2 \rfloor)}$

and $X_2^2 = \cup_{i=\lfloor m/2 \rfloor+1}^{m} C_{(i,\lfloor m/2 \rfloor)}$, and let $X_2 = \cup_{i=1}^{2} X_2^i$. Define X_i and X_i^j in the same way. At the end of the path $y_1 - \cdots - y_s$, we attach now two disjoint paths constructed with the points of Y_2^1 and Y_2^2, respectively (again in arbitrary order). This process will then be iteratively repeated until all the points are added to the tree (see Figure 3). Every two steps the number of cells in X_i grows by a factor of 2. If k is the number of steps, the construction ends when $2^{k/2} = \sqrt{n}/r$, that is, when $k = \log n - 2\log r$.

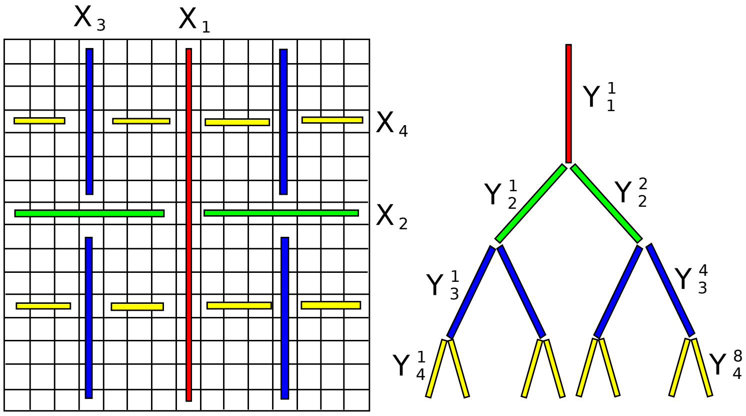


Figure 3 Sketch of the construction

Now we need to know the height of this elimination tree. Since $\mathrm{vol}(X_i)$ is at least of logarithmic size, by Lemma 7 we can always ensure the concentration on the number of points inside X_i^j.

Observe that in X_i^j there are $\frac{\sqrt{n}}{r}2^{-\lceil (i+1)/2 \rceil}$ cells of the tessellation. Then, $\mathrm{vol}(X_i^j) = r^2|X_i^j| = r\sqrt{n}2^{-\lceil (i+1)/2 \rceil}$. For a sufficiently large c, if $i \le \ell = \log n - 2\log\log n + 2\log r - \log c$, $\mathrm{vol}(X_i^j) \ge c\log n$ and by Lemma 7 together with a union bound over all j and $i \le \ell$, we have *a.a.s.*

$$|Y_i^j| = O\left(r\sqrt{n}2^{-\lceil (i+1)/2 \rceil}\right) \tag{6}$$

After this point $\mathrm{vol}(X_i)$ is too small to show concentration, but we have at most $k - \ell = 2\log\log n - 4\log r + \log c$ steps remaining. Since $\mathrm{vol}(X_i^j)$ beyond ℓ is smaller than $c\log n$, we will have at most the number of points inside an area of size $c\log n$ containing it. Thus, *a.a.s.*, for any j and $\ell \le i \le k$, $|Y_i^j| \le O(\log n)$, and *a.a.s.*

$$\sum_{i=\ell}^{k} \max_j \left|Y_i^j\right| \le O(\log n \log\log n).$$

Hence, the height of this elimination tree is *a.a.s.*

$$\begin{aligned} \mathrm{td}(G) &\le \sum_{i=0}^{\ell} \max_j \left|Y_i^j\right| + \sum_{i=\ell+1}^{k} \max_j \left|Y_i^j\right| \\ &\le O\left(r\sqrt{n}\left(\textstyle\sum_{i\ge 0} 2^{-\lceil (i+1)/2 \rceil}\right)\right) + O(\log n \log\log n) \\ &= O(r\sqrt{n}). \end{aligned}$$ ◀

For convenience, tessellate the square $[0,\sqrt{n}]^2$ into small squares of size $r/4$. Given a set $A \subseteq V$ (identified with the corresponding geometric positions in $[0,\sqrt{n}]^2$), define by ∂A the boundary of A as $\partial A = \left\{x \in [0,\sqrt{n}]^2 : \min_{u\in A} d(x,u) = \frac{r}{2}\right\}$. We use $\mathrm{vol}(\partial A)$ to refer to the length of the boundary of A.

▶ **Lemma 9.** *Let S be a separator of the giant component. Let A be a connected component of $G \setminus S$. Then there exists a connected set of cells C_S containing $|C_S| = c_S = \Theta(\mathrm{vol}(\partial A)/r)$ cells, such that all the points inside C_S are from the giant component and in the separator.*

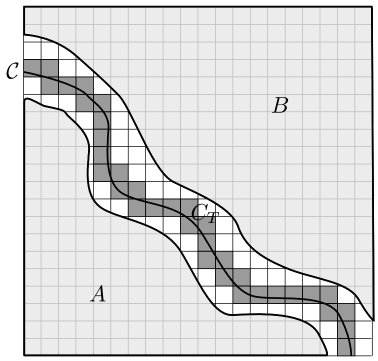


Figure 4 Cells of C_S

▶ **Theorem 10.** *There exists a constant c_2 such that for any $r \geq c_2$, a.a.s., $\mathrm{tw}(G) \geq \Omega(r\sqrt{n})$.*

Proof. We will show that there exists no balanced separator of size $o(r\sqrt{n})$ for the giant component H. Then, by Lemma 1, this implies that $\mathrm{tw}(H) = \Omega(r\sqrt{n})$, and therefore $\mathrm{tw}(G) \geq \mathrm{tw}(H) = \Omega(r\sqrt{n})$.

Let S be a fixed balanced separator of H. Let $S_1, \dots, S_m$ the different connected components of S. If $m = \Omega(r\sqrt{n})$, for each component of S there is at least one point, since H is connected. This point belongs to S and to H and therefore the separator contains at least $m = \Omega(r\sqrt{n})$ points. Therefore we can assume that $m < r\sqrt{n}$.

Since S is balanced, there exist two sets A and B (not necessarily connected) with $|A| = \alpha n$, $|B| = \beta n$ for some $\frac{1}{3} < \alpha, \beta < \frac{2}{3}$ such that $G \setminus S$ contains no edges from A to B. By an isoperimetric inequality given a set A, $\mathrm{vol}(\partial A) = \Omega(\sqrt{\mathrm{vol}(A)})$. If $\mathrm{vol}(A) = \alpha n$ for $0 < \alpha < 1$, then even if A touches the boundary of $[0, \sqrt{n}]^2$, this is still true since at least a constant fraction of the perimeter is inside the square. Therefore we know that $\mathrm{vol}(\partial A) = \Omega(\sqrt{n})$, and by applying Lemma 9 for each connected component of S, we have a set of cells C_S with $c_S = \Omega(\sqrt{n}/r)$ such that all the points inside C_S are in S and in H.

Now we need to show that a.a.s. there are a lot of points inside C_S. Denote by Y the random variable counting the number of points inside C_S. The following simple claim shows that Y is concentrated around its expected value with very high probability.

▶ **Claim 11.** *The number of points Y inside C_S satisfies*

$$\Pr\left(Y < (1-\delta)\mathbf{E}\left(Y\right) = (1-\delta)\frac{r^2}{16}c_S\right) \leq e^{-\frac{\delta^2 r^2}{32}c_S}.$$

To show that no separator can have $o(r\sqrt{n})$ points we will use a union bound over all the possible balanced separators of H. Write $C_S = \cup C_{S_i}$ where C_{S_i} are the cells given by Lemma 9 for the separator S_i. Letting $c_{S_1}, \dots, c_{S_m}$ the sizes of these separator components, there are at most $n^m e^{c_{S_1}+\cdots+c_{S_m}}$ ways to construct C_S: for each component C_{S_i} we have n places to choose where to start and then at most $e^{c_{S_i}}$ connected set of cells of size c_{S_i}.

Combining the previous upper bound from Claim 11 with a union bound over all separators of size $c_S \geq \Omega(\sqrt{n}/r)$, the probability of having such a bad balanced separator is at most

$$\sum_{c_S \geq \Omega(\sqrt{n}/r)} \sum_{m \leq O(r\sqrt{n})} \sum_{c_{S_1}+\cdots+c_{S_m}=c_S} n^m e^{c_S} e^{-\gamma r^2 c_S}, \tag{7}$$

where $\gamma = \delta^2/32$ for any $0 < \delta < \frac{1}{3}$. The number of ways to sum i using m non-negative numbers is $\binom{i+m}{m-1} \leq (i+m)^m \leq n^m$, and thus, (7) can be bounded from above by

$$\sum_{c_S \geq \Omega(\sqrt{n}/r)} \sum_{m \leq O(r\sqrt{n})} n^{2m} e^{c_S} e^{-\gamma r^2 c_S} \tag{8}$$

Observe that if $m \leq c\frac{r\sqrt{n}}{\log n}$ for some small constant $c > 0$, then $n^{2m} < e^{2cr\sqrt{n}} = o(e^{\gamma r^2 c_S})$, for sufficiently large γ. Therefore assume that $m > c\frac{r\sqrt{n}}{\log n}$.

Suppose that there is a constant fraction of cells in $C_G \setminus C_S$ contained in components of size at least $\frac{\sqrt{n}\log n}{cr}$. We restrict our separator to these big components. For this (sub)separator we have $m \leq c\left(\frac{\sqrt{n}}{r \log n}\right)$ (there are at most n/r^2 cells), and by the previous arguments, for this (sub)separator, the probability of having few points is at most $e^{-\gamma r^2 c_S}$ for some $\gamma > 0$, and hence the probability of having few points in S is also at most $e^{-\gamma r^2 c_S}$.

Thus, there is at least a constant fraction of vertices of $C_G \setminus C_S$ in components of order at most $\frac{\sqrt{n}\log n}{cr}$. Then, by the same isoperimetric inequality as before,

$$c_S \geq \frac{n^{1/4}\sqrt{\log n}}{\sqrt{cr}} \times c\frac{\sqrt{n}}{r \log n} = \Omega\left(\frac{n^{3/4}}{r^{3/2}\sqrt{\log n}}\right),$$

since all the components have order at least $\frac{\sqrt{n}\log n}{cr}$.

We distinguish two cases. First, we consider the case $c_2 \leq r = O(\sqrt{\log n})$. Since $m = O(r\sqrt{n})$, $n^{2m} = e^{2m \log n} \leq e^{2r\sqrt{n}\log n} \leq e^{2\sqrt{n}\log^{3/2} n}$ and $e^{\gamma r^2 c_S} \geq e^{\gamma \frac{n^{3/4}\sqrt{r}}{\sqrt{\log n}}} \geq e^{\gamma \frac{n^{3/4}}{\sqrt{\log n}}}$,

$$n^{2m} e^{c_S} e^{-\gamma r^2 c_S} \leq e^{-\gamma' r^2 c_S}$$

for some $0 < \gamma' < \gamma$. Otherwise, $r = \omega(\sqrt{\log n})$. Observe that $m \leq c_S$ since $c_{S_i} \geq 1$ by definition. Therefore,

$$n^{2m} e^{c_S} e^{-\gamma r^2 c_S} \leq n^{2c_S} e^{c_S} e^{-\gamma r^2 c_S} = e^{(2\log n + O(1) - \gamma r^2)c_S} \leq e^{-\gamma'' r^2 c_S}$$

for some $0 < \gamma'' < \gamma$. We showed that each term of (8) can be bounded by an exponentially small term. Hence, there exist constants $\nu, \nu' > 0$, such that with probability at most

$$\sum_{c_S \geq \Omega(\sqrt{n}/r)} \sum_{m \leq O(r\sqrt{n})} n^{2m} e^{c_S} e^{-\nu r^2 c_S} \leq O\left(rn^{3/2} e^{-\nu' r\sqrt{n}}\right) = o(1)$$

there exists a separator S containing less than $(1-\delta)\frac{r^2}{16}c_S = \Omega(r\sqrt{n})$ points connected to the giant component, completing the proof. ◀

4 Conclusion

We have shown that for random geometric graphs with $0 < r \leq c_1$ and for $r \geq c_2$ the parameters of treewidth and treedepth are asymptotically of the same order. The immediate natural question that remains open is whether for all values of $r = \Theta(1)$, including the values of $c_1 \leq r \leq c_2$, this happens to be true. For either of the parameters it would be interesting to know whether there is a sharp threshold width of order $o(1)$, in the sense that there exists some critical value of the radius r_c such that the treewidth (treedepth, respectively) of a graph with radius of at most $r_c - o(1)$ is of order $\Theta(\frac{\log n}{\log\log n})$ with probability at least $1-\epsilon$, and the treewidth (treedepth, respectively) of a graph with radius at least $r_c + o(1)$ is of order $\Theta(\sqrt{n})$ with probability at least $1-\epsilon$, for any $\epsilon > 0$. We remark that the general result on sharp thresholds of monotone properties of [6] implies only a sharp threshold width of order $\log^{3/4} n$. Needless to say, in case of the existence of such a sharp threshold, it would be nice to find this exact threshold value for any of the two parameters (they might coincide). Using our methods, this, however, among other problems, requires the knowledge of the exact threshold value r_t of the appearance of the giant component in a random geometric graph, which at the moment is not known.

References

1 Bruno Courcelle. The monadic second-order logic of graphs. I. Recognizable sets of finite graphs. *Inform. and Comput.*, 85(1):12–75, 1990.

2 Marek Cygan, Jesper Nederlof, Marcin Pilipczuk, Michal Pilipczuk, Johan van Rooij, and Jakub Onufry Woitaszczyk. Solving connectivity problems parameterized by treewidth in single exponential time. *ArXiv e-prints*, March 2011.

3 Jitender S. Deogun, Ton Kloks, Dieter Kratsch, and Haiko Müller. On vertex ranking for permutation and other graphs. In *STACS 94 (Caen, 1994)*, volume 775 of *Lecture Notes in Comput. Sci.*, pages 747–758. Springer, Berlin, 1994.

4 Yong Gao. On the threshold of having a linear treewidth in random graphs. In *COCOON*, pages 226–234, 2006.

5 Edward N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9:533–543, 1961.

6 Ashish Goel, Sanatan Rai, and Bhaskar Krishnamachari. Sharp thresholds for monotone properties in random geometric graphs. In *STOC*, pages 580–586, 2004.

7 Ramin Hekmat. *Ad-hoc networks-fundamental properties and network topologies*. Springer, 2006.

8 Hal A. Kierstead and Daqing Yang. Orderings on graphs and game coloring number. *Order*, 20(3):255–264 (2004), 2003.

9 Ton Kloks. *Treewidth*, volume 842 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1994. Computations and approximations.

10 Jaroslav Nešetřil and Saharon Shelah. On the order of countable graphs. *European J. Combin.*, 24(6):649–663, 2003.

11 Jaroslav Nešetřil and Patrice Ossona de Mendez. Tree-depth, subgraph coloring and homomorphism bounds. *European J. Combin.*, 27(6):1022–1041, 2006.

12 Mathew Penrose. *Random geometric graphs*, volume 5 of *Oxford Studies in Probability*. Oxford University Press, Oxford, 2003.

13 Guillem Perarnau and Oriol Serra. On the tree-depth of Random Graphs. *ArXiv e-prints*, April 2011.

14 Alex Pothen and Sivan Toledo. *Handbook on Data Structures and Applications*, chapter 59, Elimination structures in scientific computing. Chapman and Hall, 2004.

15 Martin Raab and Angelika Steger. "Balls into bins"—a simple and tight analysis. In *Randomization and approximation techniques in computer science (Barcelona, 1998)*, volume 1518 of *Lecture Notes in Comput. Sci.*, pages 159–170. Springer, Berlin, 1998.

16 Neil Robertson and Paul D. Seymour. Graph minors. I. Excluding a forest. *Journal of Combinatorial Theory*, 35(1):39–61, 1983.

17 Neil Robertson and Paul D. Seymour. Graph minors. II. Algorithmic aspects of tree-width. *J. Algorithms*, 7(3):309–322, 1986.

18 Robert Sasak. Comparing 17 graph parameters. Master Thesis, Univ. of Bergen. 2010.