

On Treewidth and Related Parameters of Random Geometric Graphs

Mitsche, Dieter; Perarnau, Guillem

DOI:

[10.1137/120874448](https://doi.org/10.1137/120874448)

License:

None: All rights reserved

Document Version

Publisher's PDF, also known as Version of record

Citation for published version (Harvard):

Mitsche, D & Perarnau, G 2017, 'On Treewidth and Related Parameters of Random Geometric Graphs', *SIAM Journal on Discrete Mathematics*, vol. 31, no. 2, pp. 1328-1354. <https://doi.org/10.1137/120874448>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

© 2017, Society for Industrial and Applied Mathematics
Made available under the self-archiving terms of the journal.

Checked 29/6/2017

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

ON TREewidth AND RELATED PARAMETERS OF RANDOM GEOMETRIC GRAPHS*

DIETER MITSCHE[†] AND GUILLEM PERARNAU[‡]

Abstract. We give asymptotically exact values for the treewidth $\text{tw}(G)$ of a random geometric graph $G \in \mathcal{G}(n, r)$ in $[0, \sqrt{n}]^2$. More precisely, let r_c denote the threshold radius for the appearance of the giant component in $\mathcal{G}(n, r)$. We then show that for any constant $0 < r < r_c$, $\text{tw}(G) = \Theta(\frac{\log n}{\log \log n})$, and for c being sufficiently large, and $r = r(n) \geq c$, $\text{tw}(G) = \Theta(r\sqrt{n})$. Our proofs show that for the corresponding values of r the same asymptotic bounds also hold for the pathwidth and the treedepth of a random geometric graph.

Key words. random geometric graphs, treewidth, treedepth, pathwidth

AMS subject classifications. 05C80, 05C62, 90B15

DOI. 10.1137/120874448

1. Introduction. Let V be a set of n points in the square $\mathcal{S}_n = [0, \sqrt{n}]^2$ and $r = r(n)$ a nonnegative real number. This choice of the square is only for convenience; by suitable scaling we could have chosen the square $[0, 1]^2$ and all the results would be still valid. We will identify each point with its position, that is, $v \in V$ refers also to the geometrical position of v in \mathcal{S}_n .

The *geometric graph* G of V with radius r is the graph constructed by connecting two points of V if their Euclidean distance in \mathcal{S}_n is smaller than r . For any two points $u, v \in \mathcal{S}_n$ we will denote by $\text{dist}_E(u, v)$ their Euclidean distance and by $\text{dist}_G(u, v)$ their distance in the graph G .

Then we define $\mathcal{G}(n, r)$ as the probability space of the geometric graphs of order n with radius r . A graph G chosen uniformly at random from $\mathcal{G}(n, r)$ will be called a *random geometric graph* and will be denoted by $G \in \mathcal{G}(n, r)$. Note that with probability one, no two vertices of $G \in \mathcal{G}(n, r)$ are placed in the same position.

Starting with the seminal paper of Gilbert [8], 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 [11, Chapter 3], and for a survey of many theoretical results, we refer to Penrose's monograph [22].

All our stated results are asymptotic as $n \rightarrow \infty$. 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., [9]), that is, there exists a constant value r_c

*Received by the editors April 20, 2012; accepted for publication (in revised form) April 14, 2017; published electronically June 22, 2017. An extended abstract of this paper appeared in *Proceedings of STACS*, 2012, pp. 408–419 [17].

<http://www.siam.org/journals/sidma/31-2/87444.html>

Funding: 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.

[†]Université Côte d'Azur, Laboratoire J-A Dieudonné, Nice Cedex 02 (dmitsche@unice.fr).

[‡]School of Mathematics, University of Birmingham, Ring Rd N, Birmingham, B15 2TS, United Kingdom (g.perarnau@bham.ac.uk).

such that for any $\varepsilon > 0$, a.a.s. the largest component of $G \in \mathcal{G}(n, r_c - \varepsilon)$ is of order $O(\log n)$, whereas in $G \in \mathcal{G}(n, r_c + \varepsilon)$, a single component of order $\Theta(n)$ is present, while the others have order $O(\log n)$ (see [22, Chapter 10]). The exact value of r_c is not yet determined, but it is known that $c^- \leq r_c \leq c^+$, where $c^- \approx 0.834$ and $c^+ \approx 1.836$ (see [22, p. 189]). Moreover, simulation studies suggest that the exact value of $r_c \approx 1.2$ (see again [22, p. 189]).

Since random geometric graphs have been heavily used for modeling communication networks, it is natural to analyze the expected complexity of different algorithms applied to this class. Courcelle's theorem [5] states that any problem that can be expressed in monadic second order logic can be solved in linear time for the class of graphs with bounded treewidth. This motivates the study of this parameter and other tree-like parameters on random geometric graphs. In this paper, we study the behavior of the treewidth and the treedepth on random geometric graphs.

The treewidth was introduced independently by Halin in [10] and by Robertson and Seymour in [26].

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

1. $\bigcup_{i=1}^s W_i = V$,
2. for any $e = uv \in E$ there exists a set $W_i \in \mathcal{W}$ such that $u, v \in W_i$,
3. for any $v \in V$, the subgraph induced by the $W_i \ni v$ is connected as a subgraph of T .

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

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

Observe that if G is a graph with connected components H_1, \dots, H_m , then

$$(1) \quad \text{tw}(G) = \max_{1 \leq i \leq m} \text{tw}(H_i).$$

The concept of treedepth has been introduced under different names in the literature. In this paper we follow the definition given by Nešetřil and Ossona de Mendez as a tree-like parameter in the scope of homomorphism theory, where it provides an alternative definition of bounded expansion classes [19]. For the sake of completeness, we note that the treedepth is also equivalent to the height of an elimination tree (used, for instance, in the parallel Cholesky decomposition [24]). Furthermore, analogous definitions can be found using the terminology of rank function [18], vertex ranking number (or ordered coloring) [7], or weak coloring number [12].

We now give the precise definition of treedepth. Let T be a rooted tree. The *height* of T is defined as the number of vertices of the longest rooted path. The *closure* of T is the graph that has the same set of vertices and a pair of vertices is connected by an edge if one is an ancestor of the other in T . We say that the tree T is an *elimination tree* of a connected graph G if G is a subgraph of the closure of T . The *treedepth* of a connected graph G , $\text{td}(G)$, is defined to be the minimum height of an elimination tree of G .

The definition of treedepth can also be extended to nonconnected graphs. If G is a graph with connected components H_1, \dots, H_m ,

$$(2) \quad \text{td}(G) = \max_{1 \leq i \leq m} \text{td}(H_i).$$

Hence, if $S \subset V(G)$ separates G into two subsets A and B , we have

$$(3) \quad \text{td}(G) \leq |S| + \max\{\text{td}(A), \text{td}(B)\} .$$

Observe that if H is a subgraph of G , then

$$(4) \quad \text{td}(H) \leq \text{td}(G) \text{ and } \text{tw}(H) \leq \text{tw}(G) .$$

Both parameters are closely connected: while the treewidth of a graph G is a parameter that measures the similarity between G and the class of trees in general, 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. The two parameters are related by the inequalities

$$\text{tw}(G) \leq \text{td}(G) \leq (\text{tw}(G) + 1) \log_2 n,$$

both bounds being sharp (see [19]). Note also that $\text{tw}(G) \geq \omega(G) - 1$, where $\omega(G)$ denotes the size of the largest clique in G .

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

THEOREM 1.1. *Let $0 < r < r_c$ and let $G \in \mathcal{G}(n, r)$. Then, a.a.s., $\text{tw}(G) = \Theta(\frac{\log n}{\log \log n})$, and also a.a.s., $\text{td}(G) = \Theta(\frac{\log n}{\log \log n})$.*

THEOREM 1.2. *Let c be a sufficiently large constant. Let $r = r(n) \geq c$ and $G \in \mathcal{G}(n, r)$, a.a.s. $\text{tw}(G) = \Theta(r\sqrt{n})$ and $\text{td}(G) = \Theta(r\sqrt{n})$.*

Remark 1.3. For $G \in \mathcal{G}(n, r)$ with r constant, but $r \geq c$, by the results of [6], many problems such as STEINER TREE, FEEDBACK VERTEX SET, CONNECTED VERTEX COVER can be solved in time $O(\text{poly}(n)3^{\sqrt{n}})$, while others like CONNECTED DOMINATING SET, CONNECTED FEEDBACK VERTEX SET, MIN CYCLE COVER, LONGEST PATH, LONGEST CYCLE, GRAPH METRIC TRAVELLING SALESMAN PROBLEM can be solved in time $O(\text{poly}(n)4^{\sqrt{n}})$.

Remark 1.4. Other width parameters that are sandwiched between the treewidth and the treedepth clearly then also have the same asymptotic behavior in $\mathcal{G}(n, r)$. For instance, the *pathwidth* of a graph, introduced by Robertson and Seymour [25], measures the similarity between a graph and a path. Since the pathwidth is well known to be bounded from below by the treewidth and bounded from above by the treedepth (see [27, Theorems 5.3 and 5.11]), the former theorems imply that for those values of $r = r(n)$ the pathwidth of the graph is of the same order.

Remark 1.5. Whereas intuitively it might be clear that around the threshold of the existence of a giant component there should be a jump for parameters like treewidth or treedepth in $\mathcal{G}(n, r)$, the orders of magnitude of these parameters are not so obvious (for us). Moreover, we point out that there are differences between $\mathcal{G}(n, r)$ and $\mathcal{G}(n, p)$: it is known that in the Erdős–Rényi random graph model $\mathcal{G}(n, p)$, as soon as the giant component appears, the graph has linear treewidth (see [14]). In contrast to this, Theorem 1.2 shows that a random geometric graph with a linear number of edges containing a giant component only has treewidth $\Theta(\sqrt{n})$. This different behavior of the two models can be explained by their different expansion properties and the connection between balanced separators and treewidth (see Lemma 4.3 below). Classical random graphs have very good expansion properties, and thus it is

difficult to find small separators of large sets of vertices. The geometric properties of the model $\mathcal{G}(n, r)$ imply the lack of large expanders. For this reason, in the latter case one can construct a tree decomposition with smaller bags. On the other hand, in the subcritical regime (with a linear number of edges, but before the existence of a giant component) the treedepth of $\mathcal{G}(n, p)$ is $\Theta(\log \log n)$ (see [23]), whereas by Theorem 1.1, for random geometric graphs it is already $\Theta(\frac{\log n}{\log \log n})$. (In fact, a lower bound of this order is very easy, since the largest clique is of that order, and an upper bound of $O(\log n)$ is also easy, since $O(\log n)$ is an upper bound for the size of the largest component.) Furthermore, in this range, in classical random graphs the treewidth is bounded by a constant (see [23]), whereas our theorems show that in $\mathcal{G}(n, r)$ both treewidth and treedepth are asymptotically of the same order for a wide range of parameters r . The fact that for random geometric graphs the treedepth and treewidth are always asymptotically of the same order 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.

Poissonization. In order to simplify calculations, we will use the well-known idea of Poissonization (see [22, section 1.7]): let V be a set of points obtained as a homogeneous Poisson point process $\mathcal{G}(P_1, r)$ of intensity 1 in \mathcal{S}_n . In other words, V consists of N points in the square \mathcal{S}_n chosen independently and uniformly at random, where N is a Poisson random variable of mean n . Exactly as in $\mathcal{G}(n, r)$, two points $u, v \in V$ are connected by an edge if their Euclidean distance in \mathcal{S}_n is at most r . The main advantage of the Poisson point process is that the number of points of V that lie in any region $A \subseteq \mathcal{S}_n$ of area a has a Poisson distribution with mean a ; and the number of points of V in disjoint regions of \mathcal{S}_n are independently distributed. Moreover, by conditioning $\mathcal{G}(P_1, r)$ upon the event $N = n$, we recover the original distribution of $\mathcal{G}(n, r)$. Therefore, since $\Pr(N = n) = \Theta(1/\sqrt{n})$, any event holding in $\mathcal{G}(P_1, r)$ with probability at least $1 - o(f_n)$ must hold in $\mathcal{G}(n, r)$ with probability at least $1 - o(f_n\sqrt{n})$. In particular, an event holding with probability $1 - o(n^{-1/2})$ in $\mathcal{G}(P_1, r)$ holds a.a.s. in $\mathcal{G}(n, r)$. We make use of this property throughout the article and perform the proofs of Theorems 1.1 and 1.2 for a graph $G \in \mathcal{G}(P_1, r)$.

The paper is organized as follows. In section 2 we define the cell graph of a geometric graph and give some properties of it. The proof of Theorem 1.1 is presented in section 3. Whereas the lower bound follows from a standard argument using the clique number of $\mathcal{G}(n, r)$, the proof of the upper bound is more involved. In section 4 we continue by proving Theorem 1.2. Finally, in section 5 we conclude by mentioning some open problems.

2. Properties of deterministic geometric graphs.

2.1. The cell graph of a geometric graph. For any constant $\ell > 0$, we tessellate \mathcal{S}_n into squares of sidelength ℓ called *cells*. For the sake of simplicity of presentation, we assume that \sqrt{n}/ℓ is an integer for the values of ℓ considered in this paper. We use this tessellation to construct the *cell graph* $C_G(\ell)$ of G : each nonempty cell will be represented by a *vertex* and two different vertices of $C_G(\ell)$ will be joined if there exist two points of G in the corresponding cells that share an edge. (See Figure 1, where the tessellation is omitted for clarity.)

From now on, unless otherwise stated, we will call *points* the vertices of the geometric graph G and use the word *vertex* for the cells of $C_G(\ell)$. The cell-graph $C_G(\ell)$ simplifies the original geometric graph G while preserving the same structure. For any subgraph H of G we will denote its cell graph by $C_H(\ell)$.

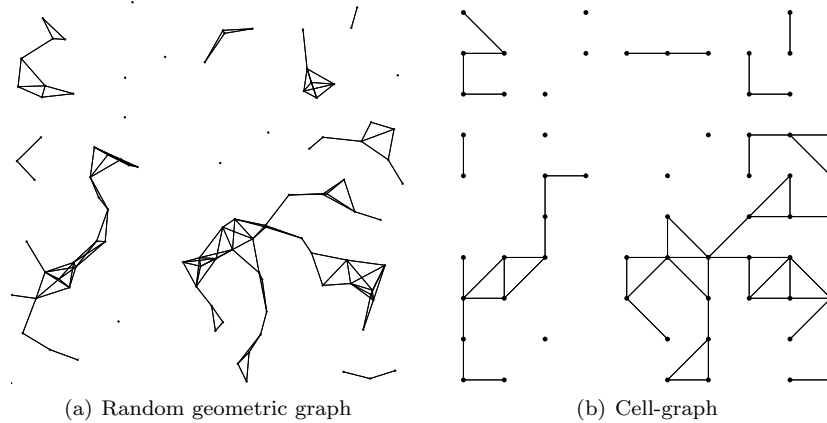


FIG. 1. A random geometric graph and its corresponding cell graph.

Remark 2.1. Notice that $C_H(\ell)$ is always a subgraph of $C_G(\ell)$. Observe that, for any $\ell \leq r/\sqrt{2}$, each nonempty cell contains points from exactly one connected component of G , since all the points inside a cell are connected. Thus, if $\ell \leq r/\sqrt{2}$, there exists a natural bijection between the connected components of G and the connected components of $C_G(\ell)$.

We need another auxiliary graph, the *grid graph* $L_{a,b}^k$, defined as follows: its vertex set is $V(L_{a,b}^k) = \{(i, j) : 1 \leq i \leq a, 1 \leq j \leq b\}$, and $(i, j)(i', j') \in E(L_{a,b}^k)$ if and only if $(i, j) \neq (i', j')$ and $\max\{|i - i'|, |j - j'|\} \leq k$. Note that by construction, for a geometric graph G in S_n with radius r we have the following relation (as subgraphs):

$$(5) \quad C_G(\ell) \subseteq L_{\lceil \sqrt{n}/\ell, \lceil \sqrt{n}/\ell \rceil}^{\lceil r/\ell \rceil}.$$

The following lemma bounds the maximal number of different connected subgraphs of a given size in $L_{a,b}^k$.

LEMMA 2.2. *The number of connected subgraphs of size s in $L_{a,b}^k$ is at most $O(ab(2k+1)^{4s})$.*

Proof. A connected subgraph is determined by a root v and any of its spanning trees, rooted at v . Observe that there are ab many ways to choose $v \in V(L_{a,b}^k)$. Moreover, the degree of a vertex in $L_{a,b}^k$ is at most $(2k+1)^2$, since for any cell (i, j) there are at most $(2k+1)^2$ cells (i', j') such that $\max\{|i - i'|, |j - j'|\} \leq k$.

One can construct at most $((2k+1)^2)^{2s-3} \leq (2k+1)^{4s}$ walks of length $2s-2$ that have both start and end points at v . In particular, these walks contain all the possible spanning trees rooted at v since a spanning tree has $s-1$ edges and each edge is traversed twice. Thus, the lemma follows. \square

Remark 2.3. Lemma 2.2 is certainly not tight. For the same problem on the integer lattice (each cell is connected to the four closest ones) the asymptotic growth is $\text{poly}(s)\lambda^s$. However the exact value of λ is not yet known. The best known lower and upper bounds for λ are 4.0025 and 4.5685, respectively (see [3, 2]).

The following proposition bounds the treedepth of a strong product of a graph and a clique. Given two graphs G_1 and G_2 , the strong product $G = G_1 \boxtimes G_2$ is defined

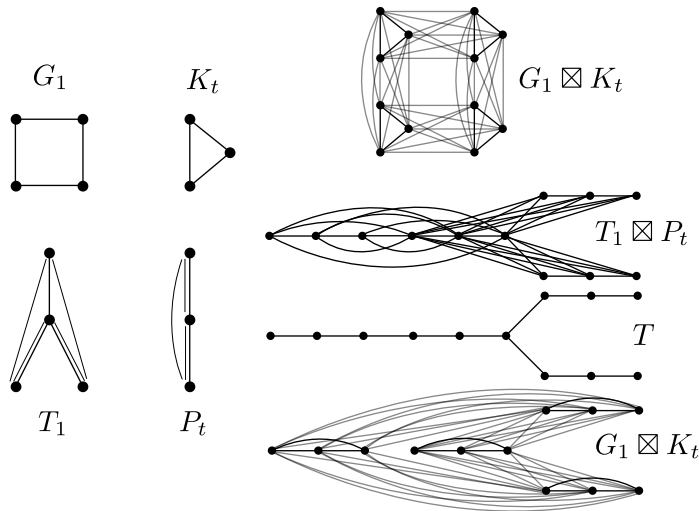


FIG. 2. Embedding of the strong product.

as $V(G) = V(G_1) \times V(G_2)$ and $(u_1, u_2)(v_1, v_2) \in E(G)$ iff for $i = 1, 2$, either $u_i = v_i$ or $u_i v_i \in E(G_i)$. Denote by K_t the complete graph on t vertices.

LEMMA 2.4. Let $G = G_1 \boxtimes K_t$. Then

$$\text{td}(G) \leq t \text{td}(G_1) .$$

Proof. Let T_1 be a tree of height $\text{td}(G_1)$ that embeds G_1 in its closure. Note also that K_t is contained in the closure of a rooted path of order t , P_t . Observe that $T_1 \boxtimes P_t$ is not a tree, but it contains a tree T , in whose closure $T_1 \boxtimes P_t$ is contained (see Figure 2). Indeed, T can be constructed in the following way: each vertex $u \in V(T_1)$ is replaced by a path of order t (call these new vertices u_1, \dots, u_t), and if there is an edge $uv \in E(T_1)$, such that u is ancestor of v , then in T , u_t is connected by an edge to v_1 (the depth of v_1 in T is exactly one more than the depth of u_t); see Figure 2. Note that T is a tree and its closure contains G as a subgraph. Since each vertex of G_1 is replaced by t vertices, $\text{td}(G) \leq t \cdot \text{td}(G_1)$. \square

Observe also that for a geometric graph G ,

$$(6) \quad G \subseteq C_G(\ell) \boxtimes K_t ,$$

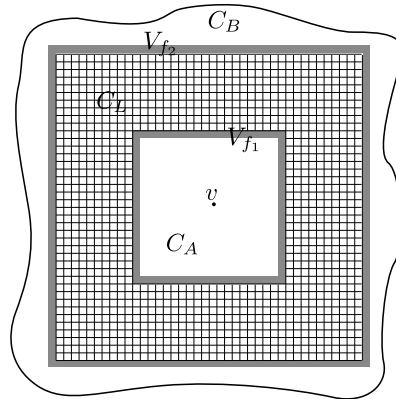
where t is the maximum number of points inside a cell of the tessellation of length ℓ .

Since we can express the treedepth of G in terms of the treedepth of its cell graph and the latter one is a subgraph of $L_{a,b}^k$, the following proposition will be useful.

PROPOSITION 2.5. Let $L_{a,b}^k$ the grid graph defined as above and suppose that $a \leq b$. Then

$$\text{td}(L_{a,b}^k) \leq O(ka \log b).$$

Proof. We describe an elimination tree for $L_{a,b}^k$ in a recursive way. First, note that $\text{td}(L_{a,k}^k) = O(ka)$, since the treedepth of a graph is always smaller than its order. Let us compute now the treedepth of $L_{a,b}^k$. By removing the central copy of $L_{a,k}^k$ in

FIG. 3. Decomposition of C_H .

$L_{a,b}^k$, we disconnect the original graph and get two copies of $L_{a,(b-k)/2}^k$. Applying this recursively and using (3), we obtain

$$\begin{aligned} \text{td}(L_{a,b}^k) &\leq O(ka) + \text{td}(L_{a,(b-k)/2}^k) \\ &\leq \underbrace{O(ka) + \cdots + O(ka)}_{\log b} + \text{td}(L_{a,k}^k) \\ &= O(ka \log b). \quad \square \end{aligned}$$

The following proposition will be very useful in the proof of Theorem 1.1 but can be applied to any sparse geometric graph.

PROPOSITION 2.6. *Let H be a geometric graph of order m such that there are no more than t points inside each cell of length $\ell = r/\sqrt{2}$.*

Then, we have

$$\text{td}(H) = O\left(\max\left\{\frac{m}{\log m}, t(\log m)^3\right\}\right).$$

Proof. Throughout this proof all cells will have length $\ell = r/\sqrt{2}$. Notice that by Remark 2.1 the connected components of the cell graph $C_H(\ell)$ are in one to one correspondence with the connected components in H . Thus, we may assume that H is connected. We will show an upper bound on $\text{td}(H)$ by providing an elimination scheme for C_H which then induces an elimination scheme for H .

Fix a vertex $v \in V(C_H)$ corresponding to a cell of the tessellation. For any integer $d \geq 0$, denote by V_d the set of vertices in the cell graph, which are at L_∞ distance d in the underlying grid graph from v (see Figure 3).

Analogously, we define P_d to be the set of points of H inside the cells of V_d .

For the sake of convenience, we define

$$K = \frac{m}{(\log m)^2}.$$

The idea of the proof is to find a separator S of H that contains at most $O(K)$ points. This separator will split the graph into some smaller subgraphs. Using (3) and applying the same procedure recursively to the remaining parts, we will get an upper bound on $\text{td}(H)$.

Let f be the largest integer for which

$$(7) \quad \sum_{d=0}^{f-1} |P_d| \leq \frac{m}{2} .$$

Let f_1 be the largest integer for which $f_1 \leq f$ and $|P_{f_1}| \leq K$ and f_2 be the smallest integer for which $f_2 \geq f$ and $|P_{f_2}| \leq K$. Since H contains m points, $f_2 - f_1 \leq \frac{m}{K} = (\log m)^2$.

Given a graph G and $S \subset V(G)$, we will denote by $G[S]$ the subgraph of G induced by S . We decompose of C_H into the following subgraphs (see Figure 3):

$$C_S = C_H[V_{f_1} \cup V_{f_2}] , \quad C_A = C_H \left[\bigcup_{d=0}^{f_1-1} V_d \right] ,$$

$$C_L = C_H \left[\bigcup_{d=f_1+1}^{f_2-1} V_d \right] , \text{ and} \quad C_B = C_H \left[\bigcup_{d \geq f_2+1} V_d \right] ,$$

and we define accordingly

$$H_S = H[P_{f_1} \cup P_{f_2}] , \quad H_A = H \left[\bigcup_{d=0}^{f_1-1} P_d \right] ,$$

$$H_L = H \left[\bigcup_{d=f_1+1}^{f_2-1} P_d \right] , \text{ and} \quad H_B = H \left[\bigcup_{d \geq f_2+1} P_d \right] .$$

In the case $|P_f| \leq K$, we have $f_1 = f_2$ and C_L and H_L are graphs on zero vertices. Thus, suppose that this is not the case, and focus on C_L .

Since $\ell = r/\sqrt{2}$, by (5) we know that C_L is a subgraph of at most four copies of $L_{a,b}^2$ (see Figure 3), where $a = (\log m)^2$ and $b = m$, since $f_2 - f_1 \leq (\log m)^2$ and $|P_d| \leq m$ for any d . By (3) and Proposition 2.5, we get

$$\text{td}(C_L) \leq O(4a) + \text{td}(L_{a,b}^2) = O((\log m)^3) .$$

Moreover, $H_L \subseteq C_L \boxtimes K_t$. Hence, by Lemma 2.4,

$$\text{td}(H_L) = O(t(\log m)^3) .$$

By (3), now applied to H and the separator $S = V_{f_1} \cup V_{f_2}$, we have

$$(8) \quad \begin{aligned} \text{td}(H) &\leq |S| + \max\{\text{td}(H_A), \text{td}(H_L), \text{td}(H_B)\} \\ &\leq 2K + \max\{\text{td}(H_A), O(t(\log m)^3), \text{td}(H_B)\}, \end{aligned}$$

since $|S| \leq 2K$ by definition of f_1 and f_2 .

We recursively repeat this procedure for the two subgraphs H_A and H_B . By the choice of f in (7), both subgraphs contain at most $m/2$ points. Hence, the recursion depth of our procedure is at most $\log_2 m = O(\log m)$. This implies that

$$\text{td}(H) = O(\max\{K \log m, t(\log m)^3\}) = O\left(\max\left\{\frac{m}{\log m}, t(\log m)^3\right\}\right) . \quad \square$$

2.2. Separators and cells. During the rest of the section we will consider a tessellation of length $\ell = r/4$.

Given $\mathcal{S} \subseteq \mathcal{S}_n$ a set of positive measure, we denote by $\text{vol}(\mathcal{S})$ the area of \mathcal{S} and by $\partial\mathcal{S}$ its boundary in the euclidean topology. We also use $\text{vol}(\partial\mathcal{S})$ to refer to the length of $\partial\mathcal{S}$. We only consider sets \mathcal{S} that are finite unions of discs, so that the length of the boundary is well defined.

For any set $A \subseteq V(H)$, let $\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) \leq \frac{r}{2}\} \subseteq \mathcal{S}_n$, and notice that $\partial\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) = \frac{r}{2}\}$.

We will use the fact that for any cell D and for any two elements $u, v \in D$

$$(9) \quad \text{dist}_E(u, v) \leq \frac{r}{2\sqrt{2}}.$$

Also, we make use of the following isoperimetric inequality (see [20, Theorem 1.6.1]): for any connected set of positive measure $\mathcal{S} \subset \mathbb{R}^2$,

$$(10) \quad \text{vol}(\partial\mathcal{S}) \geq \Omega\left(\sqrt{\text{vol}(\mathcal{S})}\right).$$

This inequality can be extended to a nonconnected set \mathcal{S} as follows: suppose that \mathcal{S} is a union of disjoint connected sets $\mathcal{S}_1, \dots, \mathcal{S}_m$. Then, for each $i = 1, \dots, m$, we have $\text{vol}(\partial\mathcal{S}_i) = \Omega(\sqrt{\text{vol}(\mathcal{S}_i)})$, and thus

$$(11) \quad \text{vol}(\partial\mathcal{S}) = \sum_{i=1}^m \text{vol}(\partial\mathcal{S}_i) \geq \sum_{i=1}^m \Omega\left(\sqrt{\text{vol}(\mathcal{S}_i)}\right) = \Omega\left(\sqrt{\text{vol}(\mathcal{S})}\right),$$

where the last inequality follows from concavity of the square root function, that is, for any $x, y \geq 0$, we have $\sqrt{x} + \sqrt{y} \geq \sqrt{x+y}$.

Denote by $\mathring{\mathcal{S}}_n$ the interior of \mathcal{S}_n . We have the following lemma.

LEMMA 2.7. *Let $\mathcal{S} \subset \mathcal{S}_n$ be a measurable connected set. Then,*

$$\text{vol}\left(\partial\mathcal{S} \cap \mathring{\mathcal{S}}_n\right) = \Omega(\min\{\text{vol}(\partial\mathcal{S}), \text{vol}(\partial(\mathcal{S}_n \setminus \mathcal{S}))\}).$$

Proof. Consider the complement of \mathcal{S} , $\mathcal{U} = \mathcal{S}_n \setminus \mathcal{S}$. Let $\mathcal{U}_1, \dots, \mathcal{U}_m$ denote the disjoint connected sets of \mathcal{U} .

Let us focus on \mathcal{U}_i for some $i \in [m]$. Let $\mathcal{V}_i = \mathcal{S}_n \setminus \mathcal{U}_i$ denote its complement. We will show that $\text{vol}(\partial\mathcal{U}_i \cap \mathring{\mathcal{S}}_n) = \Omega(\min(\text{vol}(\partial\mathcal{U}_i), \text{vol}(\partial\mathcal{V}_i)))$. Since \mathcal{U}_i and \mathcal{V}_i are connected sets that partition \mathcal{S}_n , if $\partial\mathcal{U}_i \cap \mathring{\mathcal{S}}_n = \partial\mathcal{U}_i$, then we are done. Otherwise, there exist two points x and y in $\partial\mathcal{U}_i \cap \partial\mathcal{S}_n$ such that $\partial\mathcal{U}_i = \mathcal{C}_1 \cup \mathcal{C}_2$, where \mathcal{C}_i is a connected simple curve with endpoints x and y , $\mathcal{C}_1 \subseteq \partial\mathcal{S}_n$ and $\mathcal{C}_2 \cap \partial\mathcal{S}_n = \{x, y\}$. Let $\mathcal{C}_3 = \partial\mathcal{S}_n \setminus \mathcal{C}_1$ and notice that $\partial\mathcal{V}_i = \mathcal{C}_2 \cup \mathcal{C}_3$ and that $\mathcal{C}_1 \cup \mathcal{C}_3 = \partial\mathcal{S}_n$.

Let $\mathcal{W}_i = \mathcal{U}_i$ if $\text{vol}(\mathcal{C}_1) \leq \text{vol}(\mathcal{C}_3)$ and $\mathcal{W}_i = \mathcal{V}_i$ otherwise. This implies that $\text{vol}(\mathcal{C}_2) \geq \|x - y\|_2 = \Omega(\min\{\text{vol}(\mathcal{C}_1), \text{vol}(\mathcal{C}_3)\})$. Using that $\text{vol}(\partial\mathcal{W}_i) = \text{vol}(\mathcal{C}_2) + \min\{\text{vol}(\mathcal{C}_1), \text{vol}(\mathcal{C}_3)\}$, we have $\text{vol}(\partial\mathcal{W}_i \cap \mathring{\mathcal{S}}_n) = \text{vol}(\mathcal{C}_2) = \Omega(\min\{\text{vol}(\partial\mathcal{U}_i), \text{vol}(\partial\mathcal{V}_i)\})$.

Since each point in $\partial\mathcal{S}_n$ belongs to at most one set \mathcal{U}_i , there is at most one set \mathcal{U}_{i^*} such that $\text{vol}(\partial\mathcal{U}_{i^*}) \geq \text{vol}(\partial\mathcal{V}_{i^*})$. If this is not the case, then we have $\text{vol}(\partial\mathcal{S} \cap \mathring{\mathcal{S}}_n) =$

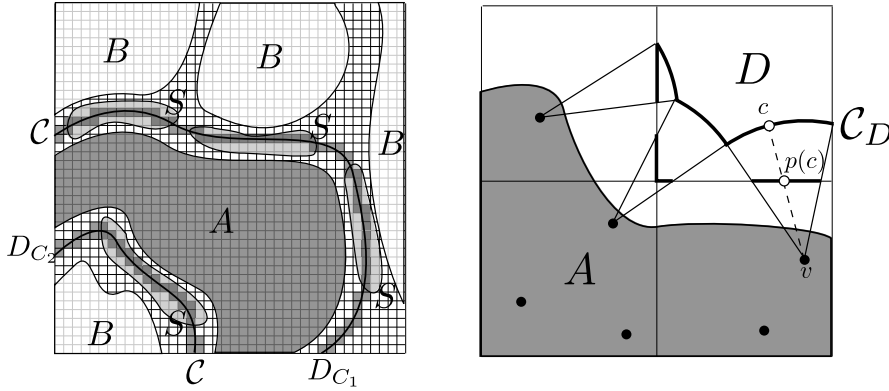


FIG. 4. Cells of D_S and the projection of C_D .

$\sum_{i=1}^m \text{vol}(\partial \mathcal{U}_i \cap \mathring{S}_n) = \sum_{i=1}^m \Omega(\text{vol}(\partial \mathcal{U}_i)) = \Omega(\text{vol}(\partial \mathcal{U}))$. Otherwise,

$$\begin{aligned} \text{vol}(\partial \mathcal{S} \cap \mathring{S}_n) &= \text{vol}(\partial \mathcal{U} \cap \mathring{S}_n) \\ &= \sum_{i=1}^m \text{vol}(\partial \mathcal{U}_i \cap \mathring{S}_n) = \sum_{i=1}^m \Omega(\min\{\text{vol}(\partial \mathcal{U}_i), \text{vol}(\partial \mathcal{V}_i)\}) \\ &= \Omega\left(\text{vol}(\partial \mathcal{V}_{i^*}) + \sum_{i \neq i^*} \text{vol}(\partial \mathcal{U}_i)\right) = \Omega(\text{vol}(\partial \mathcal{S})), \end{aligned}$$

where the last equality follows from

$$\text{vol}(\partial \mathcal{V}_{i^*}) + \sum_{i \neq i^*} \text{vol}(\partial \mathcal{U}_i) = \text{vol}(\mathcal{S}) + \sum_{i \neq i^*} \text{vol}(\partial \mathcal{U}_i \cap \partial \mathcal{S}_n).$$

□

The following lemma shows that for any separator S of a geometric graph H , we can find a large number of cells of length $\ell = r/4$, whose points are entirely contained in S (see also Figure 4, left).

LEMMA 2.8. *Let H be a connected geometric graph of order m and $S \subset V(H)$ be a separator of H . Fix a connected component H_1 of $H \setminus S$ and denote by $A = V(H_1)$.*

Consider a tessellation with side length $\ell = r/4$. There exists a set of cells D_S of size d_S , such that all points inside D_S belong to S and

$$d_S = \Omega\left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}}\right).$$

Proof. Define $B = V(H) \setminus (S \cup A)$, that is, B is the set of vertices of H that are contained neither in S nor in A .

Observe that for any pair of points $v \in A$ and $w \in B$, we have $\text{dist}_E(v, w) \geq r$, since v and w belong to different connected components of $H \setminus S$. Let $\mathcal{C} = \partial \mathcal{A} \cap \mathring{S}_n$ denote the boundary of \mathcal{A} . By definition, all points in \mathcal{C} lie at distance exactly $r/2$ from some point in A . Thus, they lie at distance at least $r/2$ from any point in B .

Let D_S be the union of cells that have nonempty intersection with \mathcal{C} . Let us point out that some of these cells may not contain any point of $V(H)$. Let us now show that $d_S = \Omega(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}})$.

By Lemma 2.7 and using the isoperimetric inequality in (10), we have that

$$\begin{aligned}
 \text{vol}(\mathcal{C}) &= \text{vol}\left(\partial\mathcal{A} \cap \overset{\circ}{\mathcal{S}}_n\right) \\
 &= \Omega(\min\{\text{vol}(\partial\mathcal{A}), \text{vol}(\partial(\mathcal{S}_n \setminus \mathcal{A}))\}) \\
 (12) \quad &= \Omega\left(\sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}}\right) .
 \end{aligned}$$

For any cell $D \in D_S$ we denote by $\mathcal{C}_D = \mathcal{C} \cap D$ the restriction of \mathcal{C} to D . We will show that the length of \mathcal{C}_D is not too large by projecting the elements of \mathcal{C}_D onto ∂D , in such a way that the length of \mathcal{C}_D does not decrease by too much.

Let $p : \mathcal{C}_D \rightarrow \partial D$ the application that sends an element $c \in \mathcal{C}_D \subset \mathcal{C}$ being at distance $r/2$ from a point $v \in A$ to the intersection of ∂D and the segment that joins c and v (see Figure 4, right). In the case where there is more than one point of A at the same distance from c , $p(c)$ chooses one of them arbitrarily.

Note that p is injective, since no two elements of \mathcal{C}_D can have the same image: indeed, suppose that there exist two different $c, c' \in \mathcal{C}_D$ with corresponding points $v, v' \in A$ such that $p(c) = p(c')$. Then, the segments cv and $c'v'$ would intersect at $p(c)$, and either $\text{dist}_E(c, v') < r/2$ or $\text{dist}_E(c', v) < r/2$ holds, contradicting the definition of \mathcal{C} .

Let us show that the application does not contract \mathcal{C}_D too much. Recall that $\text{dist}_E(c, v) = r/2$. Since $c, p(c) \in D$, by (9) we have $\text{dist}_E(c, p(c)) \leq \frac{r}{2\sqrt{2}}$, and therefore $\text{dist}_E(p(c), v) \geq \frac{\sqrt{2}-1}{2\sqrt{2}}r$ by the triangle inequality.

A simple geometric argument shows that

$$\text{vol}(p(\mathcal{S})) \geq \frac{\frac{r}{2}}{\frac{\sqrt{2}-1}{2\sqrt{2}}\pi r} \text{vol}(\mathcal{S}).$$

Since p is injective and $\text{vol}(\partial D) = 4\ell = r$,

$$\text{vol}(\mathcal{C}_D) = O(\text{vol}(\partial D)) = O(r) .$$

Using this upper bound for all cells $D \in D_S$, we obtain

$$d_S \geq \frac{\text{vol}(\mathcal{C})}{\max_{D \in D_S} \text{vol}(\mathcal{C}_D)} = \Omega\left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}}\right) .$$

Moreover, all points contained in D_S belong to S : by (9), any point u contained in D_S lies at distance at most $r/(2\sqrt{2})$ from some element $c \in \mathcal{C}$. However, all points of $A \cup B$ lie at distance at least $r/2$ from all the elements of \mathcal{C} . Thus, $u \notin A \cup B$, implying that $u \in S$. \square

We finish with some properties of the tessellation when choosing $\ell = r/4$.

LEMMA 2.9. *Let H be a geometric graph with connected components H_1, \dots, H_t . Define $A_i = V(H_i)$ and consider a tessellation with $\ell = r/4$. Then, for any cell D we have the following:*

1. *If there exists a point $v \in A_i$ such that $v \in D$, then $D \subset A_i$.*
2. *There are at most 24 curves $\mathcal{C}_i = \partial A_i$ that intersect the cell.*

Proof. For the first part, by (9), for any $u \in D$,

$$\text{dist}_E(u, v) < \frac{r}{2} ,$$

and thus $u \in A_i$.

For the second part, observe that if C_i intersects D , then there must exist a point of $v \in A_i$ at distance at most $r/2$ from some point in D . There are at most 24 cells satisfying this criterion, namely, the ones in the first and second neighborhood of D . Since all points of a cell belong to the same component (they are all connected), there are at most 24 different curves C_i intersecting D . \square

Combining Lemmas 2.8 and 2.9 we obtain the following corollary.

COROLLARY 2.10. *Let H be a connected geometric graph of order m and $S \subset V(H)$ be a separator of H . Let H_1, H_2, \dots denote the connected components of $H \setminus S$ and let \mathcal{I} index a collection of them. Let $A = V(\cup_{i \in \mathcal{I}} H_i)$. Consider a tessellation with side length $\ell = r/4$. There exists a set of cells D_S of size d_S , such that all points inside D_S belong to S and*

$$d_S = \Omega \left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}} \right) .$$

Proof. For every $i \in \mathcal{I}$, we let $A_i = V(H_i)$. By Lemma 2.8, there exists a set of cells $D_{S_{A_i}}$ containing only points in S of size

$$d_{S_{A_i}} = \Omega \left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}_i), \text{vol}(\mathcal{S}_n \setminus \mathcal{A}_i)\}} \right) .$$

Now, by Lemma 2.9 part 2, for every cell D , there are at most 24 connected components H_j of $H \setminus S$, such that ∂A_j intersects D . Hence, if $D_S = \cup_{i \in \mathcal{I}} D_{S_{A_i}}$, then D_S has size

$$\begin{aligned} d_S &\geq \frac{1}{24} \sum_{i \in \mathcal{I}} d_{S_{A_i}} \\ &= \sum_{i \in \mathcal{I}} \Omega \left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}_i), \text{vol}(\mathcal{S}_n \setminus \mathcal{A}_i)\}} \right) \\ &= \Omega \left(r^{-1} \sqrt{\min\{\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A})\}} \right) , \end{aligned}$$

where the last equality follows from $\mathcal{A} = \cup_{i \in \mathcal{I}} \mathcal{A}_i$ and from the concavity of the square root function (see also (11)). \square

3. Subcritical regime. In this section we compute the treedepth of a random geometric graph with $0 < r < r_c$, that is, below the existence of a giant component. By [22, Theorem 10.3], a.a.s. the order of each component is at most $O(\log n)$. In fact, by looking at [22, Theorem 10.3], it is easily seen that with probability at least $1 - o(n^{-3/2})$ the order of each component is $O(\log n)$.

We will use the following result several times: McDiarmid in [16] proved that for any $r = \Theta(1)$ and $G \in \mathcal{G}(n, r)$, a.a.s.

$$(13) \quad \omega(G) = \Theta \left(\frac{\log n}{\log \log n} \right) .$$

In fact, by looking at the proof of [16, Lemma 5.3], by choosing (in the notation of the proof given there) $k_1 = k_1(r)$ to be sufficiently large and $k_2 = k_2(r)$ to be sufficiently small, we can also easily see that with probability at least $1 - o(n^{-1/2})$ we have

$$(14) \quad \omega(G) = \Theta \left(\frac{\log n}{\log \log n} \right) ,$$

and by looking at [16, Lemmas 4.4 and 5.3], the same result holds for $\mathcal{G}(P_1, r)$ as well. (In fact, for Lemma 5.3, either the number of points of $\mathcal{G}(P_1, r)$ is not in the set $\{n - C\sqrt{n \log n}, n + C\sqrt{n \log n}\}$ for C large enough, which happens with probability $o(n^{-1/2})$, or the respective lower and upper bounds for the number of points can be used in the calculations of Lemma 5.3, again by choosing k_1 large enough and k_2 small enough.)

By (2), the order of the largest connected component implies a coarse upper bound, namely,

$$\text{td}(G) = O(\log n).$$

In order to find a better upper bound, more work is needed. First, we need the following simple lemma, whose proof is included for completeness.

LEMMA 3.1. *Let X be a random variable that follows a Poisson distribution with parameter λ . Then, for any $k \geq 2\lambda$,*

$$\Pr(X \geq k) \leq 2\Pr(X = k).$$

Proof. We have

$$\begin{aligned} \Pr(X \geq k) &= \sum_{i \geq k} \Pr(X = i) = \sum_{i \geq k} e^{-\lambda} \frac{\lambda^i}{i!} \\ &= e^{-\lambda} \frac{\lambda^k}{k!} \left(1 + \frac{\lambda}{k+1} + \frac{\lambda^2}{(k+1)(k+2)} + \dots \right) \\ &\leq e^{-\lambda} \frac{\lambda^k}{k!} \sum_{i \geq 0} \left(\frac{\lambda}{k} \right)^i = e^{-\lambda} \frac{\lambda^k}{k!} \cdot \frac{1}{1 - \lambda/k} \\ &\leq 2e^{-\lambda} \frac{\lambda^k}{k!} = 2\Pr(X = k), \end{aligned}$$

where the last inequality follows from the assumption $k \geq 2\lambda$. \square

Let $\nu = \nu(r)$ be a sufficiently large constant. For the sake of convenience, we define

$$T_{\max} = \frac{\nu \log n}{\log \log n} \quad \text{and} \quad T = \frac{\sqrt{2 \log n}}{\log \log n}.$$

From now on, we consider in this section the cell graph $C_G(\ell)$ of $G \in \mathcal{G}(P_1, r)$ with $\ell = r/\sqrt{2}$ and write simply C_G for $C_G(\ell)$. Notice that all points inside a cell of C_G form a clique. Hence, by (14), by choosing $\nu = \nu(r)$ sufficiently large, each cell contains less than T_{\max} points a.s. For this particular tessellation, we call a cell *sparse* if it contains less than T points, and *dense* otherwise.

PROPOSITION 3.2. *Let $0 < r < r_c$ and let $G \in \mathcal{G}(P_1, r)$. With probability at least $1 - o(n^{-1/2})$, every connected component H of G contains at most $O(T_{\max})$ points in dense cells.*

Proof. For any connected component H of G we will show that the probability that the number of points in dense cells of H is at least $2T_{\max}$ is $o(n^{-3/2})$. Since there are clearly at most n connected components in G , by taking a union bound over all them, with probability $1 - o(n^{-1/2})$ no component will have more than $2T_{\max}$ points in dense cells.

Let A_i be the number of points in the cell i . Since we are using a Poisson point process of intensity 1, A_i follows a Poisson distribution with parameter $\lambda = r^2/2$. Denote by $p = \Pr(A_i \geq T)$ the probability that cell A_i is dense.

By Lemma 3.1,

$$(15) \quad p \geq \Pr(A_i = T) = (1 - O(T^{-1})) \frac{e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^T,$$

$$(16) \quad p = \Pr(A_i \geq T) \leq 2\Pr(A_i = T) \leq \frac{2e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^T,$$

where we have used Stirling's formula $T! = (1 + O(T^{-1}))\sqrt{2\pi T} = \left(\frac{T}{e}\right)^T$.

To count the number of points lying in dense cells, we define the following random variable for each cell $i \in V(C_G)$:

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

Our aim is to show that $Y_H = \sum_{i \in V(C_H)} Y_i$ is at most $O(T_{\max})$.

Notice that the probability that the cell i is sparse is $1 - p$, while the probability of having $T + j$ points is

$$\begin{aligned} \Pr(A_i = T + j) &= (1 - O((T + j)^{-1})) \frac{e^{-\lambda}}{\sqrt{2\pi(T + j)}} \left(\frac{e\lambda}{T + j}\right)^{T+j} \\ &\leq \left(\frac{e\lambda}{T}\right)^T \frac{e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^j \end{aligned}$$

for any integer $j \geq 0$. Using (16) we have

$$\Pr(A_i = T + j) \leq 2p \left(\frac{e\lambda}{T}\right)^j.$$

These observations lead to the definition of the following independent random variable R_i for each cell $i \in V(C_G)$:

$$R_i = \begin{cases} 0 & \text{with probability } 1 - 2p, \\ T + j & \text{with probability } 2p \left(\frac{e\lambda}{T}\right)^j \text{ for any } j \geq 1, \\ T & \text{with probability } 2p \left(1 - \frac{e\lambda}{T - e\lambda}\right). \end{cases}$$

First, observe that R_i is a probability distribution. The random variables Y_i and R_i have similar distributions. In particular, 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,

$$(17) \quad \Pr(R \geq j) \geq \Pr(Y \geq j)$$

for any $j \geq 0$. In particular, this also hold if $j = O(T_{\max})$.

Therefore, it is enough to compute an upper bound for $\Pr(R > 2T_{\max})$. Clearly, since $r < r_c$, and all connected components are of order $O(\log n)$ with probability at least $1 - o(n^{-3/2})$, with the same probability in the cell graph C_G the graph diameter of each component C_H is at most $K \log n$ for some sufficiently large constant $K = K(r)$. For the case where the graph diameter is bigger than $K \log n$, $\Pr(R > 2T_{\max})$ can be easily bounded by $o(n^{-3/2})$. For the case where it is smaller than $K \log n$, we observe the following: given a cell from C_H , all points that belong to H are contained in the

box of cells of size $(2K \log n + 1) \times (2K \log n + 1)$ centered on the first cell. Let $\eta > 0$ such that $(2K \log n + 1)^2 \leq \eta \log^2 n$.

Hence we have

$$(18) \quad \Pr(R > 2T_{\max}) \leq o\left(n^{-3/2}\right) + \sum_{m=1}^{(2K \log n + 1)^2} \sum_{S \in \binom{(\eta \log^2 n)}{m}} \sum_{\substack{c_i: i \in S \\ \sum_{i \in S} c_i \geq 2T_{\max}}} \Pr\left(\bigcap_{i \in S} R_i = c_i\right),$$

where m counts the number of dense cells in the distribution given by the R_i , S is the set of dense cells, and c_i is the number of points inside the dense cell $i \in S$. There are at most $\eta^m (\log n)^{2m}$ ways to choose the set S of size m and at most $(T_{\max})^m < (\log n)^m$ possible values for the c_i .

Recall that the variables R_i are independent and that $\Pr(R_i = T + j) = 2p\left(\frac{e\lambda}{T}\right)^j$ for any $j \geq 1$. Therefore,

$$\Pr\left(\bigcap_{i \in S} R_i = c_i\right) = \prod_{i=1}^m 2p\left(\frac{e\lambda}{T}\right)^{c_i - T}.$$

On the one hand, if $m \leq 2\sqrt{\log n}$, using (15),

$$\begin{aligned} \prod_{i=1}^m 2p\left(\frac{e\lambda}{T}\right)^{c_i - T} &\leq \prod_{i=1}^m \frac{4}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^{c_i} \\ &\leq \prod_{i=1}^m \left(\frac{e\lambda}{T}\right)^{c_i} \\ &\leq \left(2e^\lambda \sqrt{2\pi T} p\right)^{\frac{\sum c_i}{T}} \\ &\leq \left(2e^\lambda \sqrt{2\pi T} p\right)^{2\sqrt{\log n}}. \end{aligned}$$

On the other hand, if $m = 2\sqrt{\log n} + j$ for some integer $j \geq 1$,

$$\prod_{i=1}^m 2p\left(\frac{e\lambda}{T}\right)^{c_i - T} \leq (2p)^m = (2p)^{2\sqrt{\log n}} (2p)^j.$$

Therefore, by splitting the second part of (18) into two sums, we obtain

$$\begin{aligned} \Pr(R > 2T_{\max}) &\leq o\left(n^{-3/2}\right) + \sum_{m=1}^{2\sqrt{\log n}} \eta^m (\log n)^{3m} \left(2e^\lambda \sqrt{2\pi T} p\right)^{2\sqrt{\log n}} \\ &\quad + (2\eta (\log n)^3 p)^{2\sqrt{\log n}} \sum_{j \geq 1} (2\eta (\log n)^3 p)^j. \end{aligned}$$

From the bounds on p in (15) and (16), one can derive that $\eta (\log n)^3 p < 1/2$, and the infinite sum of the second term above is bounded from above by one. Thus,

$$\begin{aligned} \Pr(R > 2T_{\max}) &\leq o\left(n^{-3/2}\right) + \left(2\sqrt{\log n}\right) \left(\eta (\log n)^3 p \left(2e^\lambda \sqrt{2\pi T} + 2\right)\right)^{2\sqrt{\log n}} \\ &= o\left(n^{-3/2}\right) + \exp\left\{\log \log n / 2 + 2\sqrt{\log n} (3 \log \log n + \log p + O(\log T))\right\}. \end{aligned}$$

Moreover, by (16), we also have $p \leq \frac{2e^{-\lambda}}{\sqrt{2\pi T}} \left(\frac{e\lambda}{T}\right)^T$, and hence $\log p \leq -(1 + o(1))T \log T \leq -\sqrt{\log n}$. Thus,

$$(19) \quad \Pr(R > 2T_{\max}) < o\left(n^{-3/2}\right) + \exp\{-(1 + o(1))2 \log n\} = o\left(n^{-3/2}\right).$$

By (17), this also implies that $\Pr(Y > 2T_{\max}) = o(n^{-3/2})$, and by taking a union bound over all components, this implies that the probability of having a connected component with more than $2T_{\max}$ points inside dense cells is $o(n^{-1/2})$. \square

Proof of Theorem 1.1. The lower bound on $\text{tw}(G)$ follows easily from (14), which yields

$$\text{td}(G) \geq \text{tw}(G) \geq \omega(G) - 1 = \Omega\left(\frac{\log n}{\log \log n}\right).$$

For the upper bound, we construct an elimination tree for G . By (2) it suffices to bound from above the treedepth of each connected component. Let H be a connected component of G .

From Proposition 3.2, there are at most $O(T_{\max})$ points in dense cells of H . We temporarily remove all these points and add them at the end. Let H' be the subgraph of H that remains after removing the points in the dense cells.

Observe that now, by definition of sparse, every cell of $C_{H'}$ contains at most T points. Denoting by $m = |V(H')|$, by Proposition 2.6 we have

$$\text{td}(H') = O\left(\max\left\{\frac{m}{\log m}, T(\log m)^3\right\}\right).$$

Since, with probability at least $1 - o(n^{-3/2})$, $m = O(\log n)$, we have that for every component H of G , $\text{td}(H') = O(T_{\max})$ with probability at least $1 - o(n^{-1/2})$.

Recall that adding a new point to H can increase the treedepth by at most one unit. Thus, $\text{td}(H) \leq \text{td}(H') + O(T_{\max}) = O(T_{\max})$, and therefore, using (1), we have

$$\text{td}(G) = O\left(\frac{\log n}{\log \log n}\right)$$

with probability at least $1 - o(n^{-1/2})$. \square

4. Supercritical regime. Fix now $r = r(n) \geq c$ for some sufficiently large constant c . Recall that for any subset $\mathcal{S} \subseteq \mathcal{S}_n = [0, \sqrt{n}]^2$ of positive measure, we denote by $\text{vol}(\mathcal{S})$ the area of \mathcal{S} . We need the following standard lemma (which is a simple application of Chernoff bounds for Poisson variables, see, for example, [1, Theorem A.1.15]).

LEMMA 4.1. *For any $\mathcal{S} \subseteq \mathcal{S}_n$ and any $\delta > 0$, let $|\mathcal{S}|$ denote the number of points inside \mathcal{S} . Then, we have the following:*

1. *With probability at least $1 - (e^\delta(1 + \delta))^{-(1+\delta)\text{vol}(\mathcal{S})} \geq 1 - e^{-\frac{\delta^2}{3}\text{vol}(\mathcal{S})}$, $|\mathcal{S}| \leq (1 + \delta)\text{vol}(\mathcal{S})$.*
2. *With probability at least $1 - e^{-\frac{\delta^2}{2}\text{vol}(\mathcal{S})}$, $|\mathcal{S}| \geq (1 - \delta)\text{vol}(\mathcal{S})$.*

We will use this lemma to show that there exist separating sets with few points, and consequently, give an upper bound on $\text{td}(G)$.

PROPOSITION 4.2. *Let c be a sufficiently large constant, let $r = r(n) \geq c$, and let $G \in \mathcal{G}(P_1, r)$. With probability $1 - e^{-\Omega(r\sqrt{n})}$, $\text{td}(G) \leq O(r\sqrt{n})$.*

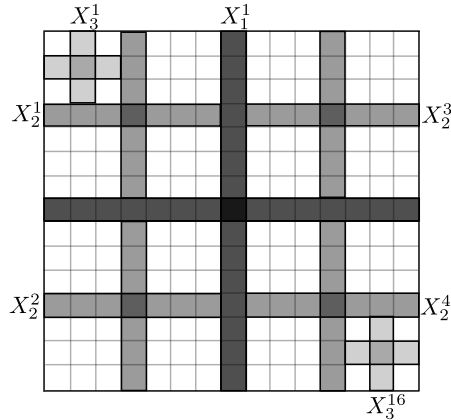


FIG. 5. Construction of the sets X_i^j .

Proof. Consider the tessellation of \mathcal{S}_n into square cells of side length $\ell = r$. Denote by $D_{(i,j)}$ the j th cell in the i th row, where $1 \leq i, j \leq a = \sqrt{n}/r$.

Define

$$X_1^1 = \left(\bigcup_{i=1}^a D_{(a/2,i)} \right) \cup \left(\bigcup_{i=1}^a D_{(i,a/2)} \right),$$

and consider the set $Y_1^1 \subset V(G)$, containing the points inside X_1^1 . Observe that Y_1^1 is a separator, since $\ell = r$, and it splits the graph into four subgraphs G_2^1, G_2^2, G_2^3 , and G_2^4 , each one consisting of a (possibly empty) union of vertex-disjoint connected components.

By (3), we have

$$\text{td}(G) \leq |Y_1^1| + \max_{1 \leq j \leq 4} \{\text{td}(G_2^j)\}.$$

We then define analogously the sets X_2^j , for all G_2^j , and using (3), we continue iteratively. Let t denote the step where all sets X_t^j have size one (see Figure 5).

The treedepth of G will be bounded from above by the maximum number of points inside any of the possible sets of cells

$$X_{j_1 j_2 \dots j_t} = X_1^{j_1} \cup X_2^{j_2} \cup \dots \cup X_t^{j_t},$$

where $1 \leq j_i \leq 4^{i-1}$.

Observe that $|X_i^{j_i}| \leq a2^{-(i-2)}$. The sets $X_{j_1 j_2 \dots j_t} = X_1^{j_1} \cup X_2^{j_2} \cup \dots \cup X_t^{j_t}$ are not disjoint, but they all have the same size

$$|X_{j_1 j_2 \dots j_t}| = \sum_{i=1}^t |X_i^{j_i}| \leq \sum_{i=1}^t a2^{-(i-2)} \leq 4a.$$

Let $Y_{j_1 j_2 \dots j_t}$ denote the set of points in $X_{j_1 j_2 \dots j_t}$. Thus, $|Y_{j_1 j_2 \dots j_t}|$ is a random variable following a Poisson distribution with mean at most $4ar^2$.

By part 1 of Lemma 4.1 applied with $\delta = 1$,

$$\Pr(|Y_{j_1 j_2 \dots j_t}| \geq 8ar^2) < e^{-4ar^2/3} = e^{-\Omega(r\sqrt{n})}.$$

Moreover, there are at most

$$\prod_{i=1}^t 4^{i-1} = e^{O(t^2)}$$

sets of the form $X_{j_1 j_2 \dots j_t}$. Observe also that, by construction, $t = O(\log a) = O(\log n)$.

Now, by a union bound over all sets,

$$\Pr(\exists j_1, j_2, \dots, j_t : |Y_{j_1 j_2 \dots j_t}| > 8ar^2) \leq e^{O(\log^2 n) - \Omega(r\sqrt{n})} = e^{-\Omega(r\sqrt{n})}.$$

Thus, we have that the treedepth of G is at most

$$\text{td}(G) \leq 8ar^2 = O(r\sqrt{n})$$

with probability at least $1 - e^{-\Omega(r\sqrt{n})}$, finishing the proof. □

For a lower bound on $\text{tw}(G)$, we need the following link between the treewidth of a graph and the existence of a vertex separator with special properties. A vertex partition $V = (A, S, B)$ is a *balanced k -partition* if $|S| = k + 1$, S separates A and B , and $\frac{1}{3}(n - k - 1) \leq |A|, |B| \leq \frac{2}{3}(n - k - 1)$. In this case, S is also called a *balanced separator*. The following result connecting balanced partitions and treewidth is due to Kloks [13].

LEMMA 4.3 (see [13]). *Let G be a graph on n vertices, and suppose that $\text{tw}(G) \leq k$ for some $n \geq k - 4$. Then G has a balanced k -partition.*

From now on and until the end of the section, we consider the tessellation of \mathcal{S}_n into square cells of side length $\ell = r/4$.

Recall that for any set $A \subset V(H)$, we define $\mathcal{A} = \{x \in \mathcal{S}_n : \min_{v \in A} \text{dist}_E(x, v) \leq r/2\}$. Observe that in a geometric graph, no direct relation exists between the size of A and the volume of \mathcal{A} . In the case of a random geometric graph and for a set A of linear size, however, $\text{vol}(\mathcal{A})$ can be bounded from below using the size of A , as the following lemma shows.

LEMMA 4.4. *Let c be a sufficiently large constant and let $r = r(n) \geq c$. Let $G \in \mathcal{G}(P_1, r)$ and let $\alpha \in (0, 1)$. Then there exists $\beta = \beta(\alpha) > 0$, such that with probability $1 - e^{-\Omega(n)}$, for any set $A \subseteq V(G)$ with $|A| \geq \alpha n$, we have*

$$\text{vol}(\mathcal{A}) \geq \beta n.$$

Proof. Let $\lambda = r^2/16$. Set $m = m(\alpha)$ to be the smallest constant such that $m\lambda$ is integer,

$$\frac{e^{-1}}{[m]!} \left(\frac{m^2}{m-1} + \frac{m}{(m-1)^2} \right) \leq \frac{\alpha}{4} \text{ and } m \geq 4e,$$

which exists for any $\alpha > 0$, since the left-hand side of the first condition tends to zero, when $m \rightarrow +\infty$.

Recall that the number of points inside a cell D follows a Poisson distribution with mean λ . Suppose that D contains $t \geq 0$ points. Define then Z_D to be the random variable

$$Z_D = \begin{cases} t & \text{if } t \geq m\lambda, \\ 0 & \text{otherwise,} \end{cases}$$

and let $Z = \sum Z_D$ be the sum of these random variables over all cells of the tessellation.

We may consider $r \geq \sqrt{32}$, since by hypothesis $r \geq c$ for some c large enough. This implies that $\lambda \geq 2$. Using that for every $N \geq 1$, $\sqrt{2\pi}N^{N+\frac{1}{2}}e^{-N} \leq N! \leq e\sqrt{2\pi}N^{N+\frac{1}{2}}e^{-N}$, we obtain

$$\begin{aligned} \Pr(Z_D = m\lambda) &= e^{-\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!} \\ &\leq \frac{e^{-\lambda}}{\sqrt{2\pi}} \left(\frac{e}{m}\right)^{m\lambda} \frac{1}{\sqrt{m\lambda}} \\ &\leq \frac{e^{-2}}{\sqrt{2\pi}} \left(\frac{e}{[m]}\right)^{[m]} \frac{1}{\sqrt{[m]\lambda}} \\ &\leq \frac{e^{-1}}{[m]!}. \end{aligned}$$

Also, for any $i \geq 1$,

$$\begin{aligned} \Pr(Z_D = m\lambda + i) &= e^{-\lambda} \frac{\lambda^{m\lambda+i}}{(m\lambda+i)!} \\ &= e^{-\lambda} \frac{\lambda^{m\lambda+(i-1)}}{(m\lambda+(i-1))!} \cdot \frac{\lambda}{m\lambda+i} \\ &\leq \frac{1}{m} \Pr(Z_D = m\lambda + (i-1)). \end{aligned}$$

Hence,

$$\begin{aligned} \mathbb{E}(Z_D) &= \sum_{t \geq m\lambda} t \Pr(Z_D = t) \\ &\leq \frac{e^{-1}}{[m]!} \sum_{i \geq 0} (m\lambda + i) m^{-i} \\ &\leq \frac{e^{-1}}{[m]!} \left(\frac{m^2\lambda}{m-1} + \frac{m}{(m-1)^2} \right) \\ &\leq \frac{\alpha\lambda}{4}, \end{aligned}$$

where the last inequality follows from the definition of m . Since $\lambda = r^2/16$ and there are $16n/r^2$ cells in the tessellation, we have

$$\mathbb{E}(Z) \leq \frac{\alpha n}{4}.$$

By Hoeffding bounds for unbounded random variables (the precise version we use here is [4, Theorem 1], applied with $X_D = \epsilon_D = Z_D$, and thus $S = T = Z$, $Y = Po(\lambda)$, $m_k = m = \mathbb{E}(Z_D)$ for any k , and $b = m\lambda - 1$, so that $m(b) = m$ and the measure $\mu^{[m]}$ is exactly our probability distribution of Z_D , and $x = 2\mathbb{E}(Z)$)

$$\Pr(Z > 2\mathbb{E}(Z)) < \inf_{h < x} e^{-h2\mathbb{E}(Z)} \mathbb{E}(e^{hZ}) \leq e^{-2\mathbb{E}(Z)} \mathbb{E}(e^Z).$$

Now, observe that

$$e^{2\mathbb{E}(Z_D)} \geq e^{2m\lambda} \Pr(Z_D = m\lambda) \geq e^{(2m-1)\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!}$$

and

$$\mathbb{E}(e^{Z_D}) = \Pr(Z_D = 0) + \sum_{i \geq 0} e^{m\lambda+i} \Pr(Z_D = m\lambda + i) \leq 1 + e^{(m-1)\lambda} \frac{\lambda^{m\lambda}}{(m\lambda)!} \sum_{i \geq 0} \left(\frac{e}{m}\right)^i.$$

Since by assumption on m , $e/m \leq 1/4$, we have

$$\mathbb{E}(e^{Z_D}) \leq 1 + \frac{4}{3} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda} \leq \frac{3}{2} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda}.$$

The random variables Z_D are mutually independent. Thus,

$$e^{2\mathbb{E}(Z)} = \prod e^{2\mathbb{E}(Z_D)} \geq \left(\frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(2m-1)\lambda}\right)^{\frac{16n}{r^2}}$$

and

$$\mathbb{E}(e^Z) \leq \left(\frac{3}{2} \frac{\lambda^{m\lambda}}{(m\lambda)!} e^{(m-1)\lambda}\right)^{\frac{16n}{r^2}},$$

and therefore

$$\Pr(Z > 2\mathbb{E}(Z)) \leq e^{-2\mathbb{E}(Z)} \mathbb{E}(e^Z) \leq \left(\frac{3}{2} e^{-m\lambda}\right)^{\frac{16n}{r^2}} = e^{-\Omega(n)}.$$

Thus, with probability at least $1 - e^{-\Omega(n)}$, there are at most $\alpha n/2$ points of G contained in cells with at least $m\lambda$ points, and thus with the same probability there are at least $\alpha n/2$ points of A contained in cells with less than $m\lambda$ points.

Therefore, with this probability, there are at least

$$\frac{\alpha n/2}{m\lambda} = \frac{8\alpha n}{mr^2}$$

different cells D that contain at least one point from A . By part 1 of Lemma 2.9, $D \subset \mathcal{A}$, and

$$\text{vol}(\mathcal{A}) \geq \frac{8\alpha n}{mr^2} \cdot \text{vol}(D) = \beta n$$

with probability at least $1 - e^{-\Omega(n)}$. □

Using the previous lemmata, we are able to provide a lower bound for $\text{tw}(G)$.

THEOREM 4.5. *Let c be a sufficiently large constant, and let $r = r(n) \geq c$. Let also $G \in \mathcal{G}(P_1, r)$. Then, $\text{tw}(G) = \Omega(r\sqrt{n})$ with probability at least $1 - e^{-\Omega(r\sqrt{n})}$.*

Before proving the theorem we sketch its proof. We are going to show that any balanced separator S of the giant component contains many points. Observe that if $\text{vol}(S)$ is large, then the probability of containing few points is exponentially small. We show that in general, any such separator has a large volume. Here we strongly use the condition that S is balanced. The conclusion will then follow by taking a union bound over all potential separators.

Proof. Fix $\gamma > 0$ to be a sufficiently small constant. Let H be the largest component of G . Note that for $r \geq c$ with c sufficiently large, by [21, Theorem 3.3],

$$(20) \quad |V(H)| = \Omega(n)$$

with probability at least $1 - e^{-\Omega(n)}$. We will for now assume deterministically that $|V(H)| = \Omega(n)$ holds and only in the end add the probability $e^{-\Omega(n)}$ that $|V(H)| = o(n)$ holds. By choosing c sufficiently large, to simplify calculations, we may even assume $|V(H)| \geq 0.9n$. We will show that there exists no balanced separator of size $\gamma r\sqrt{n}$ for H . Then, by Lemma 4.3, this implies that $\text{tw}(H) \geq \gamma r\sqrt{n} = \Omega(r\sqrt{n})$, and by (1), $\text{tw}(G) \geq \text{tw}(H) = \Omega(r\sqrt{n})$.

For any balanced separator $S \subset V(H)$ of H , denote by t be the number of connected components of the graph induced by S and let S_1, \dots, S_t denote the subsets inducing connected components within H . We may assume that S is minimal, and hence each component of S contains at least one point of H . Therefore we can assume that $t \leq \gamma r\sqrt{n}$, as otherwise there is nothing to prove. We may assume that $r \leq 2\sqrt{n}$, since for $r = 2\sqrt{n}$, $\mathcal{G}(P_1, r)$ is already the complete graph. If S is a balanced separator of size at most $\gamma r\sqrt{n} \leq 2\gamma n$, there exist two not necessarily connected sets $A, B \subset V(H)$ of size $\frac{1-2\gamma}{3}|V(H)| \leq |A|, |B| \leq \frac{2(1-2\gamma)}{3}|V(H)|$, such that $H \setminus S$ contains no edges from A to B .

Since γ is a sufficiently small constant and $|V(H)| \geq 0.9n$, $|A|, |B| \geq n/4$. By Lemma 4.4, with probability at least $1 - e^{-\Omega(n)}$, for all balanced separators S , $\text{vol}(\mathcal{A})$ and $\text{vol}(\mathcal{B})$ are linear in n . In particular, if $\beta = \beta(1/4)$ is the constant provided by Lemma 4.4 for $\alpha = 1/4$, we have

$$(21) \quad \beta n \leq \text{vol}(\mathcal{A}) \leq (1 - \beta)n$$

with probability at least $1 - e^{-\Omega(n)}$.

Since $\text{vol}(\mathcal{A}), \text{vol}(\mathcal{S}_n \setminus \mathcal{A}) \geq \beta n$, by Corollary 2.10, there exists $\eta = \eta(\beta) > 0$ such that with probability at least $1 - e^{-\Omega(n)}$, for each balanced separator S there is a set of cells D_S of size

$$d_S = \Omega\left(r^{-1}\sqrt{\text{vol}(\mathcal{A})}\right) \geq \frac{\eta\sqrt{n}}{r},$$

such that all points inside D_S belong to S . Recall that some cells in D_S may not contain any point. We will assume this deterministically for now and add the failure probability at the very end.

Now it suffices to show that with high probability, for any balanced separator S , with D_S denoting the set of cells provided by Corollary 2.10 of size at least $\eta\sqrt{n}/r$, there are at least $\gamma r\sqrt{n}$ points contained in D_S . Denote by Y_{D_S} the random variable counting the number of points inside such a set D_S . Since $\text{vol}(D_S) = \frac{r^2}{16}d_S$, by part 2 of Lemma 4.1 applied with $\delta = 1/2$, we obtain

$$(22) \quad \Pr\left(Y_{D_S} < \frac{r^2}{32}d_S\right) \leq e^{-\frac{r^2}{128}d_S}.$$

We will choose γ small enough such that $\gamma \leq \eta/32$. We will show that with high probability every balanced separator that occupies at least $\eta\sqrt{n}/r$ cells contains at least $\frac{r^2}{32}d_S \geq \gamma r\sqrt{n}$ points. We will do it by combining the inequality in (22) with a union bound over all separators S together with the corresponding sets of cells D_S of size $d_S \geq \eta\sqrt{n}/r$.

By definition of the cell graph, D_S has at most t connected components. (Some connected components of the graph induced by S can merge in D_S .) We will assume that D_S has exactly t connected components denoted by D_{S_1}, \dots, D_{S_t} and with sizes d_{S_1}, \dots, d_{S_t} .

We have

$$\begin{aligned}
 (23) \quad & \Pr(\exists S \text{ balanced, } d_S \geq \eta\sqrt{n}/r, |S| \leq \gamma r\sqrt{n}) \\
 & \leq \sum_{d \geq \eta\sqrt{n}/r} \sum_{t \leq \gamma r\sqrt{n}} \sum_{d_{S_1} + \dots + d_{S_t} = d} \\
 & \quad \Pr(\exists S \text{ balanced with } t \text{ components, } d_S = d, |S| \leq \gamma r\sqrt{n}).
 \end{aligned}$$

Denote by C_H , by C_A , and by C_B the set of cells that contain at least one point of H , A , and B , respectively. Recall that $\beta n \leq \text{vol}(\mathcal{A}), \text{vol}(\mathcal{B}) \leq (1 - \beta)n$. By Lemma 4.4, there exists an $\varepsilon > 0$ such that $|C_A|, |C_B| \geq \varepsilon n/2r^2$ with probability at least $1 - e^{-\Omega(n)}$. Since C_A and C_B are disjoint, and $C_A \cup C_B \subseteq C_H \setminus D_S$, we have $|C_H \setminus D_S| \geq \varepsilon n/r^2$ with probability at least $1 - e^{-\Omega(n)}$. Once more, we will assume this deterministically for now and add the failure probability at the very end. Let ν be a small constant.

Our aim for the rest of the proof is to show that each summand of (23) can be bounded from above by an exponentially small term. We will do it by splitting the proof into four cases:

- *Case 1*, $r > 32\sqrt{\log n}$. Observe that $t \leq d$, since $d_{S_j} \geq 1$ by definition. Since by assumption $r \geq c$ for c sufficiently large, we may assume that $r \geq 4$. Then, by setting $a = b = 4\sqrt{n}/r \leq \sqrt{n}$, $k = 4$, and $s = d_{S_j}$ in Lemma 2.2, we conclude that there are at most $n^t 9^{4(d_{S_1} + \dots + d_{S_t})} \leq n^t e^{9d}$ ways to construct possible sets of cells D_S corresponding to all balanced separators S with t components. Using t nonnegative numbers, there are at most $d^t \leq n^t$ ways to add up to d , and thus the right-hand side of (23) can be bounded from above by

$$(24) \quad \sum_{d \geq \eta\sqrt{n}/r} \sum_{t \leq \gamma r\sqrt{n}} n^{2t} e^{9d} e^{-r^2 d/128}.$$

Since $t \leq d$ and $r \geq 32\sqrt{\log n}$, it follows that

$$n^{2t} e^{9d} e^{-r^2 d/128} = e^{(2 \log n + O(1) - r^2/128)d} \leq e^{-r^2 d/256} = e^{-\Omega(r\sqrt{n})}.$$

- *Case 2*, $c \leq r \leq 32\sqrt{\log n}$ and $d \geq \sqrt{n}(\log n)^{3/2}/r$. We start with (24) as before. Note that for c sufficiently large, since $c \leq r$, $e^{9d} < e^{r^2 d/256}$. Note also that $e^{r^2 d} \geq e^{r\sqrt{n}(\log n)^{3/2}} \geq e^{c\sqrt{n}(\log n)^{3/2}}$. Thus, since $t \leq \gamma r\sqrt{n}$, we have $n^{2t} = e^{2t \log n} \leq e^{2\gamma r\sqrt{n} \log n} \leq e^{64\gamma\sqrt{n}(\log n)^{3/2}} \leq e^{\frac{64\gamma}{c} r^2 d}$. Provided that c is a large enough constant, we obtain

$$n^{2t} e^{9d} e^{-r^2 d/128} \leq e^{-r^2 d/512} = e^{-\Omega(r\sqrt{n})}.$$

- *Case 3*, $c \leq r \leq 32\sqrt{\log n}$, and $t \leq \frac{\nu r\sqrt{n}}{\log n}$. Once more, we start with (24). If ν is small enough, since $d \geq \frac{\eta\sqrt{n}}{r}$, we have $n^{2t} < e^{2\nu r\sqrt{n}} < e^{r^2 d/512}$. If c is sufficiently large, we have $e^{9d} < e^{r^2 d/512}$. Thus, in such case the summand in (24), and therefore also the right-hand side of (23), is bounded from above by $e^{-r^2 d/256} = e^{-\Omega(r\sqrt{n})}$.
- *Case 4*, $c \leq r \leq 32\sqrt{\log n}$, $t > \frac{\nu r\sqrt{n}}{\log n}$, and $d < \sqrt{n}(\log n)^{3/2}/r$. Since this is the most complex case, we will first highlight the main steps of the argument:
 1. We first show that most cells of $C_H \setminus D_S$ are in components of the cell graph that are relatively large.

2. We then show that these large components occupy a positive fraction of \mathcal{S}_n , implying that they participate in many cells of D_S . We will restrict the counting on the number of points in S to the boundaries of these components.
3. Although there are not many large components, each of them may induce many connected components in D_S . Thus, we temporarily extend the tessellation from \mathcal{S}_n to \mathbb{R}^2 . After the extension, the boundary of these components consists of an “exterior” connected part and a number of “internal” parts (corresponding to the “holes” of each component).
4. We next fill in these holes to obtain a connected set of cells corresponding to the boundary for each large component. In order to avoid having a unique filled component that spans all \mathcal{S}_n , we exclude the component of H having the largest intersection with the boundary of \mathcal{S}_n . (In fact we already exclude it at step (3).) We will use the fact that the boundaries of the filled components are connected to count the number of possible boundaries via counting lattice animals.
5. Finally, to show that there are many points inside D_S , we restrict the boundaries of the filled components to \mathcal{S}_n again. Since all the considered components have a relatively small intersection with the boundary of \mathcal{S}_n , using Lemma 2.7, we show that the number of boundary cells inside \mathcal{S}_n is of the same order as the total number of boundary cells, which is of order at least \sqrt{n}/r .

Let us now go into detail. We start with step (1) and show that almost all cells of $C_H \setminus D_S$ are contained in components of the cell graph of order at least $\frac{\sqrt{n} \log n}{\nu r^3}$ (called *large* components). Assume that this is not the case. Recall that $|C_H \setminus D_S| \geq \varepsilon n/r^2$, and assume that there exists an $\varepsilon' \leq \varepsilon$ such that at least $\varepsilon' n/r^2$ cells of $C_H \setminus D_S$ are in components of order at most $\frac{\sqrt{n} \log n}{\nu r^3}$. As in the concavity argument of (11), d is minimized if there are at most $\frac{\varepsilon' r \sqrt{n}}{\log n}$ components of order $\frac{\sqrt{n} \log n}{\nu r^3}$. Recalling that each cell has area $\Theta(r^2)$, by Corollary 2.10 we obtain

$$d = \Omega \left(r^{-1} \sqrt{r^2 \cdot \frac{\sqrt{n} \log n}{\nu r^3}} \cdot \frac{\varepsilon' r \sqrt{n}}{\log n} \right) = \Omega \left(\frac{n^{3/4}}{r^{1/2} \sqrt{\log n}} \right) > \frac{\sqrt{n} (\log n)^{3/2}}{r},$$

which contradicts our assumption on d .

Our final goal is to show that for any such balanced separator, there are many points inside. In this particular case, we will show that with high probability there are already many points in a subset of S . This will be particularly convenient since there are fewer candidates for subsets than for S , and a simple union bound will suffice to finish the proof. In particular, to form such a subset we will only count the boundary contribution of some connected components in $H \setminus S$.

We now proceed with step (2). We will get rid of all components in $H \setminus S$ of size at most $\frac{\sqrt{n} \log n}{\nu r^3}$, while at the same time still keeping a large number of cells occupied by the separator. Recall that by our deterministic hypothesis we have $d \geq \eta \sqrt{n}/r$. Denote by H_0, \dots, H_p the connected components of the graph $H \setminus S$.

For every $0 \leq i \leq p$, let D_{H_i} be the set of cells of size d_{H_i} provided by Lemma 2.8. (Note that the union of them is the set of cells provided by

Corollary 2.10.) Some of the points in D_S may appear in more than one set D_{H_i} , but recall that by Lemma 2.9 part (2) each cell can be counted for at most 24 components. Thus, $Y_{D_S} \geq \frac{1}{24} \sum_{i=0}^p Y_{D_{H_i}}$, and hence

$$\Pr(Y_{D_S} \leq \gamma r \sqrt{n}) \leq \Pr\left(\frac{1}{24} \sum_{i=0}^p Y_{D_{H_i}} \leq \gamma r \sqrt{n}\right),$$

and we can focus on the individual random variables $Y_{D_{H_i}}$.

Since S has size at most $2\gamma n$, we have that $\sum_{i=0}^p |V(H_i)| \geq n/2$. Thus, by Lemma 4.4, there exists some $\delta > 0$ such that $\sum_{i=0}^p \text{vol}(\mathcal{H}_i) \geq \delta n$. Without loss of generality, let H_0, \dots, H_q be the connected components of $H \setminus S$ whose cell graphs contain at least $\frac{\sqrt{n} \log n}{\nu r^3}$ cells. Since almost all cells of $C_H \setminus D_S$ are in such cell components, $\sum_{i=0}^q \text{vol}(\mathcal{H}_i) \geq \delta n/2$; that is, large components of $H \setminus S$ occupy a constant proportion of \mathcal{S}_n .

We continue with step (3). Recall that $\partial\mathcal{S}_n$ denotes the boundary of the square \mathcal{S}_n . Without loss of generality, let H_0 be the connected component that maximizes $\text{vol}(\mathcal{H}_i \cap \partial\mathcal{S}_n)$. Since the curves $\mathcal{H}_i \cap \partial\mathcal{S}_n$ are disjoint for different values of i , for every $1 \leq i \leq q$, we have

$$(25) \quad \text{vol}(\mathcal{H}_i \cap \partial\mathcal{S}_n) \leq \frac{1}{2} \text{vol}(\partial\mathcal{S}_n).$$

Since S is balanced, we also have that $\sum_{i=1}^q |V(H_i)| \geq n/4$ (otherwise $V(H_0)$ is too large, and the cut is not balanced), which by Lemma 4.4 implies $\sum_{i=1}^q \text{vol}(\mathcal{H}_i) \geq \delta' n$ for some $\delta' > 0$. The same argument also implies that, for every $1 \leq i \leq q$, we have $\text{vol}(\mathcal{H}_i) < (1 - \delta')n$ (provided that δ' is small enough).

Similarly as before, Corollary 2.10 implies that

$$\sum_{i=1}^q d_{H_i} = \Omega\left(\frac{\sqrt{n}}{r}\right).$$

Let us now show that q cannot be too large. Since we use a tessellation with side length $\ell = r/4$, there are at most $16n/r^2$ cells. Moreover, each cell induces a clique which implies that each cell belongs to at most one connected component. Since each large cell component contains at least $\frac{\sqrt{n} \log n}{\nu r^3}$ cells, we have that

$$q \leq \frac{16\nu r \sqrt{n}}{\log n}.$$

Although the number of large components is small, each separator set D_{H_i} may induce a lot of different connected components. This can cause problems when bounding (24). In order to fix this issue, we consider a tessellation of \mathbb{R}^2 and extend the random geometric graph from \mathcal{S}_n to \mathbb{R}^2 . Recall that C_{H_i} denotes the set of cells that contain at least one point of H_i . For every $1 \leq i \leq q$, we consider the set $D'_{H_i} \subseteq \mathbb{R}^2$ of size d'_{H_i} defined as follows: a cell D belongs to D'_{H_i} if either $D \in D_{H_i}$, or $D \in \mathbb{R}^2 \setminus \mathring{\mathcal{S}}_n$ and D intersects C_{H_i} . In other words, D'_{H_i} contains in addition to the cells already present in D_{H_i} also all cells touching the boundary of \mathcal{S}_n containing at least one point

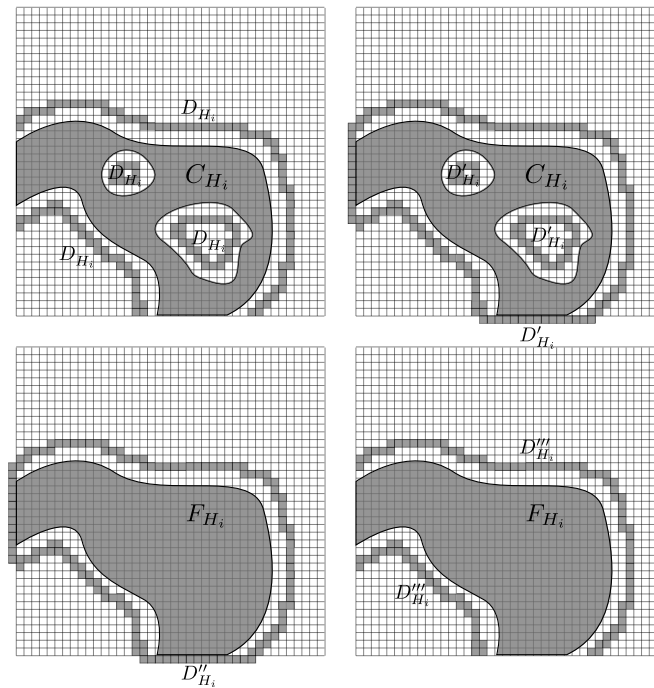


FIG. 6. Examples of sets D_{H_i} , D'_{H_i} , D''_{H_i} , and D'''_{H_i} .

of C_{H_i} . (See also Figure 6 for an example.) One can imagine D'_{H_i} to be the extension of D_{H_i} to the tessellation of \mathbb{R}^2 . Observe that the number of connected components in D'_{H_i} is at most the number of those in D_{H_i} , but this number can still be large (for instance, if the component contains some holes). Clearly, $d'_{H_i} \geq d_{H_i}$.

We are ready for step (4). For every $1 \leq i \leq q$, we delete some cells from D'_{H_i} to create a *connected* set of cells D''_{H_i} of size d''_{H_i} . We define the *fill-up* of C_{H_i} , denoted by F_{H_i} , as follows: a cell D belongs to F_{H_i} if either $D \in C_{H_i}$ or D belongs to a finite connected component of $\mathbb{R}^2 \setminus C_{H_i}$. We construct D''_{H_i} from D'_{H_i} by removing the cells that belong to F_{H_i} . Since there is just one infinite connected component in $\mathbb{R}^2 \setminus C_{H_i}$, F_{H_i} does not contain holes. Hence, D''_{H_i} is connected. Observe that it may be the case that there exists i_1, i_2 such that $C_{H_{i_2}} \subseteq F_{H_{i_1}}$. In this case, we disregard the contribution of $D''_{H_{i_2}}$ to the final sum.

Let $1 \leq i_1 < \dots < i_s \leq q$ be the indices of large cell components that are not contained in the fill-up of any other cell component. Since the components are now considered as subsets of \mathbb{R}^2 , by the standard isoperimetric inequality and its extension to nonconnected sets as in (11), we obtain

$$d'' := \sum_{j=1}^s d''_{H_{i_j}} = \Omega \left(\frac{1}{r} \sqrt{\sum_{i=1}^q \text{vol}(\mathcal{H}_i)} \right) = \Omega \left(\frac{\sqrt{n}}{r} \right),$$

where for the last equality we have used that $\sum_{i=1}^q \text{vol}(\mathcal{H}_i) \geq \delta' n$.

Now, consider $D'''_{H_i} := D''_{H_i} \cap \mathcal{S}_n$ to be the restriction of the previous sets to the square \mathcal{S}_n and let d'''_{H_i} be their size. Observe that $D'''_{H_i} \subseteq D_{H_i} \subseteq D_S$,

and thus, it suffices to prove that $\sum_{j=1}^s Y_{D''_{H_{i_j}}} \leq \gamma r \sqrt{n}$ with sufficiently small probability. Recall that by (25), for every $1 \leq i \leq q$, the component H_i satisfies $\text{vol}(\mathcal{H}_i \cap \partial \mathcal{S}_n) \leq \frac{1}{2} \text{vol}(\partial \mathcal{S}_n)$. Using Lemma 2.7, this implies

$$\text{vol}(\partial F_{H_i} \cap \mathring{\mathcal{S}}_n) = \Omega(\text{vol}(\partial F_{H_i})) = \Omega(d''_{H_i}),$$

and therefore

$$d''' := \sum_{j=1}^s d'''_{H_{i_j}} = \sum_{j=1}^s \Omega(\text{vol}(\partial F_{H_{i_j}} \cap \mathring{\mathcal{S}}_n)) = \Omega(d'').$$

Finally, we arrive at step (5): to bound the number of lattice animals that are candidates for $D'''_{H_{i_1}}, \dots, D'''_{H_{i_s}}$, we will look at $D''_{H_{i_1}}, \dots, D''_{H_{i_s}}$, which are connected and whose sizes add up to d'' . To show that the probability of having few points is small, we will look at the $d''' = \Omega(d'')$ cells that are inside some set $D''_{H_{i_j}}$. Any point there is also in D_S , and thus it suffices to restrict the counting to them. Therefore, each summand of (24) that satisfies the hypothesis of Case 4 is bounded by $n^{2s} e^{9d''} e^{-r^2 d''' / 128}$. Since $s \leq q \leq \frac{16\nu r \sqrt{n}}{\log n}$, $d''' = \Omega(d'')$ and $d'' = \Omega(\sqrt{n}/r)$, the previous term can be bounded by $e^{-\Omega(r\sqrt{n})}$, provided that c is large enough.

We showed that each term of (24) can be bounded from above by $e^{-\Omega(r\sqrt{n})}$ if $d \geq \eta \sqrt{n}/r$. Since all properties which we have assumed deterministically throughout the proof hold with probability at least $1 - e^{-\Omega(n)}$, we have together with (20),

$$\Pr\left(\exists S \text{ balanced, } |S| \leq \frac{\eta r \sqrt{n}}{32}\right) \leq e^{-\Omega(n)} + \sum_{d \geq \eta \sqrt{n}/r} \sum_{t \leq \gamma r \sqrt{n}} e^{-\Omega(r\sqrt{n})} = e^{-\Omega(r\sqrt{n})}.$$

Having chosen γ sufficiently small such that $\gamma \leq \eta/32$, the theorem follows. □

Proof of Theorem 1.2. Theorem 1.2 follows directly by recalling that $\text{tw}(G) \leq \text{td}(G)$ and combining Proposition 4.2 with Theorem 4.5. □

5. Conclusion. Given a random geometric graph $G \in \mathcal{G}(n, r)$ we showed that if $0 < r < r_c$, $\text{tw}(G) = \Theta(\frac{\log n}{\log \log n})$ and that if $r \geq c$, for some sufficiently large c , $\text{tw}(G) = \Theta(r\sqrt{n})$. We conjecture that the latter can be extended to the whole supercritical regime, that is, we conjecture that for every $r > r_c$, $\text{tw}(G) = \Theta(r\sqrt{n})$. This is a natural thing to expect since r_c is already the threshold radius for the existence of a giant component. The conjecture is equivalent to the existence of a sharp threshold width of order $o(1)$ at $r = r_c$. We remark that the general result on sharp thresholds of monotone properties of [9] implies only a sharp threshold width of order $\log^{3/4} n$. Our methods, however, require the knowledge of the exact threshold value r_c of the appearance of the giant component in a random geometric graph, which at the moment is not known.

Acknowledgments. We thank the anonymous referees for their careful reading. While this paper was under review, Li and Müller [15] proved the conjecture we stated in the conclusion: let $r > r_c$, then a.a.s. $\text{tw}(G) = \Theta(r\sqrt{n})$.

REFERENCES

- [1] N. ALON AND J. SPENCER, *The Probabilistic Method*, John Wiley and Sons, New York, 2008.
- [2] G. BAREQUET AND R. BAREQUET, *An improved upper bound on the growth constant of polyominoes*, in Proceedings of EUROCOMB, 2015.
- [3] G. BAREQUET, G. ROTE, AND M. SHALAH, $\lambda > 4$, in Proceedings of the 23rd Annual Symposium on Algorithms (ESA), 2015, pp. 83–94.
- [4] V. BENTKUS, *An extension of the hoeffding inequality to unbounded random variables*, Lith. Math. J., 48 (2008), pp. 137–157.
- [5] B. COURCELLE, *The monadic second-order logic of graphs. I. Recognizable sets of finite graphs*, Inform. and Comput., 85 (1990), pp. 12–75.
- [6] M. CYGAN, J. NEDERLOF, M. PILIPCZUK, J. M. VAN ROOIJ, AND J. O. WOJTASZCZYK, *Solving connectivity problems parameterized by treewidth in single exponential time*, in Proceedings of the IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS), IEEE, 2011, pp. 150–159.
- [7] J. S. DEOGUN, T. KLOKS, D. KRATSCHE, AND H. MÜLLER, *On vertex ranking for permutation and other graphs*, in Proceedings of STACS 94, Lecture Notes in Comput. Sci. 775, Springer, Berlin, 1994, pp. 747–758.
- [8] E. N. GILBERT, *Random plane networks*, J. Soc. Indust. Appl. Math., 9 (1961), pp. 533–543.
- [9] A. GOEL, S. RAI, AND B. KRISHNAMACHARI, *Sharp thresholds for monotone properties in random geometric graphs*, in Proceedings of STOC, 2004, pp. 580–586.
- [10] R. HALIN, *S-functions for graphs*, J. Geom., 8 (1976), pp. 171–186, <https://doi.org/10.1007/BF01917434>.
- [11] R. HEKMAT, *Ad-hoc Networks-Fundamental Properties and Network Topologies*, Springer, Berlin, 2006.
- [12] H. A. KIERSTED AND D. YANG, *Orderings on graphs and game coloring number*, Order, 20 (2003), pp. 255–264.
- [13] T. KLOKS, *Treewidth: Computations and Approximations*, Lecture Notes in Comput. Sci. 842, Springer, Berlin, 1994.
- [14] C. LEE, J. LEE, AND S.-I. OUM, *Rank-width of random graphs*, J. Graph Theory, 70 (2012), pp. 339–347, <https://doi.org/10.1002/jgt.20620>.
- [15] A. LI AND T. MÜLLER, *On the treewidth of random geometric graphs and percolated grids*, Adv. Appl. Probab., 49 (2017), pp. 49–60, <https://doi.org/10.1017/apr.2016.78>.
- [16] C. MCDIARMID, *Random channel assignment in the plane*, Random Structures Algorithms, 22 (2003), pp. 187–212.
- [17] D. MITSCHE AND G. PERARNAU, *On the treewidth and related parameters of random geometric graphs*, in Proceedings of STACS, 2012, pp. 408–419.
- [18] J. NEŠETŘIL AND S. SHELAH, *On the order of countable graphs*, European J. Combin., 24 (2003), pp. 649–663.
- [19] J. NEŠETŘIL AND P. OSSONA DE MENDEZ, *Tree-depth, subgraph coloring and homomorphism bounds*, European J. Combin., 27 (2006), pp. 1022–1041.
- [20] J. OPREA, *Differential Geometry and its Applications*, MAA, Washington, D.C., 2007.
- [21] M. PENROSE, *A central limit theorem with applications to percolation, epidemics and boolean models*, Ann. Probab., 29 (2001), pp. 1515–1546.
- [22] M. PENROSE, *Random Geometric Graphs*, Oxf. Stud. Probab. 5, Oxford University Press, Oxford, 2003.
- [23] G. PERARNAU AND O. SERRA, *On the tree-depth of random graphs*, Discrete Appl. Math., 168 (2014), pp. 119–126.
- [24] A. POTHEN AND S. TOLEDO, *Elimination structures in scientific computing*, in Handbook on Data Structures and Application, Chapman and Hall, London, 2004.
- [25] N. ROBERTSON AND P. D. SEYMOUR, *Graph minors. I. Excluding a forest*, J. Combin. Theory, 35 (1983), pp. 39–61.
- [26] N. ROBERTSON AND P. D. SEYMOUR, *Graph minors. II. Algorithmic aspects of tree-width*, J. Algorithms, 7 (1986), pp. 309–322.
- [27] R. SASAK, *Comparing 17 Graph Parameters*, Master thesis, University of Bergen, 2010.