COMBINATORICA
Bolyai Society – Springer-Verlag

COMBINATORICA **28** (5) (2008) 529–545
DOI: 10.1007/s00493-008-2283-3

TWO-POINT CONCENTRATION IN RANDOM GEOMETRIC GRAPHS

TOBIAS MÜLLER*

Received March 1, 2006

A random geometric graph G_n is constructed by taking vertices $X_1,\dots,X_n \in \mathbb{R}^d$ at random (i.i.d. according to some probability distribution ν with a bounded density function) and including an edge between X_i and X_j if $\|X_i - X_j\| < r$ where $r = r(n) > 0$. We prove a conjecture of Penrose ([14]) stating that when $r = r(n)$ is chosen such that $nr^d = o(\ln n)$ then the probability distribution of the clique number $\omega(G_n)$ becomes concentrated on two consecutive integers and we show that the same holds for a number of other graph parameters including the chromatic number $\chi(G_n)$.

1. Introduction

In the $G(n,p)$ or Erdős–Rényi model of random graphs, it has been known since at least the seventies that the probability distribution of a number of graph parameters becomes concentrated on two consecutive integers as n tends to ∞. To our knowledge the first such result was by Matula ([10]) who noticed that for fixed p, the clique number $\omega(G(n,p))$ satisfies

$$\mathbb{P}\Big(\omega(G(n,p)) \in \{k(n), k(n)+1\}\Big) \to 1,$$

as $n \to \infty$ for a sequence $k(n)$, which is given explicitly in the theorem. Two decades later, Łuczak ([7]) showed that the chromatic number of $G(n,p)$ exhibits similar behaviour as long as $p = p(n)$ is chosen in such a way that

Mathematics Subject Classification (2000): 05C80; 05C15, 05C69, 60D05

* The author was partially supported by EPSRC, the Department of Statistics, Bekker-la-Bastide fonds, Dr. Hendrik Muller's Vaderlandsch fonds, and Prins Bernhard Cultuur-fonds.

$p(n) < n^{-\frac{5}{6}-\epsilon}$ for some $\epsilon > 0$. This result was subsequently extended by Alon and Krivelevich ([2]) to include any choice of $p(n)$ that satisfies $p(n) \leq n^{-\frac{1}{2}-\epsilon}$. For other choices of $p(n)$, including fixed p, the problem of determining the (asymptotic) range of the distribution of $\chi(G(n,p))$ remains open.

In this paper we establish results analogous to the ones mentioned above for *random geometric graphs*. Given random points $X_1, X_2, \ldots \in \mathbb{R}^d$, generated i.i.d. according to some probability distribution ν on $\mathbb{R}^d$, and a sequence $r(n)$ of positive numbers, a sequence of random geometric graphs $(G_n)_n$ can be obtained, where G_n has vertex set $\{X_1, \ldots, X_n\}$ and an edge $X_iX_j \in E(G_n)$ if $\|X_i - X_j\| < r(n)$. Here $\|.\|$ may be an arbitrary norm and the only restriction we will be putting on ν in this paper is that it has a bounded density function. In [14], Penrose showed that if ν is uniform on $[0,1]^d$ and $r(n)$ is chosen in such a way that $nr^d = o(\ln n)$, then the maximum degree $\Delta(G_n)$ of the random geometric graph is two-point concentrated, ie.,

$$\mathbb{P}\big(\Delta(G_n) \in \{l(n), l(n)+1\}\big) \to 1,$$

for some sequence $l(n)$. He also conjectured that the same holds for the clique number $\omega(G_n)$. In the paper [13] he had already shown this to be true under the stronger assumption that nr^d is bounded. We remark here that it is natural to state our results in terms of the quantity nr^d, because this is a good measure of the average degree of G_n (see appendix A for a precise result).

A random variable that is closely related to some of these graph parameters is the (continuous) *scan statistic*. If $W \subseteq \mathbb{R}^d$ then we will denote

$$M_W := \max_{x \in \mathbb{R}^d} \big|\{X_1, \ldots, X_n\} \cap (x + rW)\big|,$$

ie., M_W is the largest number of points in any translate of rW. The random variable M_W is the scan statistic associated with W (and n, r). It can be used to devise hypothesis tests for spatial dependence in point patterns, and has been studied in connection with applications such as the spread of epidemics. See [5,6] for an overview of results and applications of the scan statistic. We remark that often a sequence of (shrinking) sets $(W_n)_n$ is considered, but we have chosen to formulate things in terms of a sequence $r(n)$ and a fixed set W because this fits better with the presentation. Månsson ([9]) noticed that the scan statistic becomes two-point concentrated when (translated into our $r(n)$, fixed W setting) nr^d is bounded by $n^{-\epsilon}$ for some $\epsilon > 0$, and this result was later extended by Penrose whose results in [13] show that two-point concentration also holds under the weaker condition that nr^d is bounded (he also imposes some regularity conditions on the probability density function of ν). In this paper we will prove that two-point concentration occurs when $nr^d = o(\ln n)$ for a range of random variables including the scan statistic and the clique number $\omega(G_n)$ and chromatic number $\chi(G_n)$ of the random geometric graph. For a brief discussion of the behaviour of some of these random variables when r is larger, see section 5.

2. Statement of results

In order to prove a two-point concentration result for a number of different graph parameters and related random variables in one go, we will use the following set-up. We assume we are given a *clustering rule*, consisting of a sequence of maps $(h_n)_n$ that assign non-negative integers to finite subsets of $\mathbb{R}^d$ in such a way that there exist $R_2 \geq R_1 > 0$ such that the following hold for all n and any finite set of points $A \subseteq \mathbb{R}^d$:

(C1) $h_n(A) \leq |A|$;
(C2) If $A \subseteq B(x; R_1 r(n))$ for some $x \in \mathbb{R}^d$ then $h_n(A) = |A|$;
(C3) If $h_n(A) > 0$ then $\|a - b\| < R_2 r(n)$ for all $a, b \in A$;
(C4) If $h_n(A) = l$ then $h_n(A \setminus \{a\}) \geq l - 1$ for all but at most one $a \in A$.

We will be interested in the maximum $M = M(n)$ of $h_n(A)$ over all subsets $A \subseteq \{X_1, \ldots, X_n\}$. If $h_n(A) = l$ then we will say that A is an *l-cluster* and we will say that A is a *$(\geq l)$-cluster* if $h_n(A) \geq l$.

We remark that our notion of a clustering rule differs from the notion used in [13], so that we can accommodate a wider variety of random variables.

Example 2.1. We get $M(n) = \omega(G_n)$ if we set $h_n(A) = |A|$ if $\operatorname{diam}(A) < r(n)$ and $h_n(A) = 0$ otherwise. We may take $R_1 = \frac{1}{2}, R_2 = 1$.

Example 2.2. If W is a bounded set with non-empty interior, then we can get $M(n) = M_W$ by setting $h_n(A) = |A|$ if $A \subseteq x + r(n)W$ for some $x \in \mathbb{R}^d$ and $h_n(A) = 0$ otherwise. We may for instance put $R_2 = \operatorname{diam}(W)$ and $R_1 = \rho$ for any $\rho > 0$ such that there exists a ball $B(x; \rho) \subseteq W$. Observe that we can also consider a sequence $(W_n)_n$ of scanning sets, as long as there is a universal upper bound on the diameter and a universal lower bound on the inradius of W_n.

Example 2.3. Set $h_n(A) = 0$ if A is not contained in some ball $B(x; r)$ and otherwise let $h_n(A)$ be equal to the maximum degree + 1 of the subgraph of G_n induced by A. Then $M(n) = \Delta(G_n) + 1$ and we can put $R_1 = \frac{1}{2}$ and $R_2 = 2$.

In this paper we will show that whenever the sequence of maps $(h_n)_n$ satisfies properties (C1)–(C4) above then the following theorem holds:

Theorem 1. *If $r(n)$ satisfies $nr^d = o(\ln n)$ then there exists a sequence $k(n)$ such that*

$$\mathbb{P}\big(M(n) \in \{k(n), k(n)+1\}\big) \to 1.$$

Although the chromatic number $\chi(G_n)$ does not directly correspond to a clustering rule, theorem 1 does allow us to deduce the following result:

Corollary 2. *If $r(n)$ satisfies $nr^d = o(\ln n)$ then there exists a sequence $l(n)$ such that*

$$\mathbb{P}\big(\chi(G_n) \in \{l(n), l(n)+1\}\big) \to 1.$$

The same remarks apply to the degeneracy $\delta^*(G_n)$ (recall that the degeneracy is the maximum over all subgraphs of the minimum degree):

Corollary 3. *If $r(n)$ satisfies $nr^d = o(\ln n)$ then there exists a sequence $m(n)$ such that*

$$\mathbb{P}\big(\delta^*(G_n) \in \{m(n), m(n)+1\}\big) \to 1.$$

3. Notation and preliminaries

In this paper we often suppress the subscript or argument n for the sake of readability when no confusion can arise. We will say that a sequence of events $(A_n)_n$ holds whp. (with high probability) if $\mathbb{P}(A_n) \to 1$. All use of the notation $B(x;\rho)$ is wrt. the (arbitrary) norm $\|.\|$ that we have equipped $\mathbb{R}^d$ with (ie., $B(x;\rho) := \{y \in \mathbb{R}^d : \|x-y\| < \rho\}$). The d-dimensional volume (Lebesgue-measure) of a (measurable) set $A \subseteq \mathbb{R}^d$ will be denoted by $\mathrm{vol}(A)$. The volume of the unit ball wrt. $\|.\|$ will be denoted by $\theta := \mathrm{vol}(B(0;1))$. For $A \subseteq \mathbb{R}^d$ we will denote $\mathcal{N}(A) := |A \cap \{X_1, \dots, X_n\}|$.

Recall that the only restriction we are putting on the probability distribution ν that generates the points $X_1, X_2, \dots$ is that the density f of ν is bounded. In the rest of this paper $\nu_{\max}$ will denote the essential supremum of f, i.e.,

$$\nu_{\max} := \sup\big\{t : \mathrm{vol}(\{f > t\}) > 0\big\}.$$

Thus we have $\nu(A) \le \nu_{\max} \mathrm{vol}(A)$ for any (measurable) $A \subseteq \mathbb{R}^d$. The only other fact about ν that the proofs in the next section rely on is given by the following lemma:

Lemma 3.1. *There exists a constant $\eta > 0$, dependent only on ν, such that the following holds. For any bounded set $W \subseteq \mathbb{R}^d$ with $\mathrm{vol}(W) > 0$ and any $r > 0$ there exist $\Omega(r^{-d})$-many disjoint translates $W_1, \dots, W_K$ of rW with $\nu(W_i)/\mathrm{vol}(W_i) > \eta$ for all i.*

The proof can be found in appendix B.

The proofs in the next section will rely on the following standard elementary result (see for instance [11]).

Lemma 3.2. *Let Z be a binomial random variable and $k \geq \mu := \mathbb{E}Z$. Then*

$$\left(\frac{\mu}{ek}\right)^k \leq \mathbb{P}(Z \geq k) \leq \left(\frac{e\mu}{k}\right)^k.$$

Now and then $\mathrm{Bi}(n,p)$ will denote a binomial random variable with parameters n and p. We will also need the following result, which is due to [8]:

Lemma 3.3. *Let $(Z_1,\dots,Z_m)$ have a (joint) multinomial distribution. Then*

$$\mathbb{P}(Z_1 \leq k_1,\dots,Z_m \leq k_m) \leq \Pi_{i=1}^m \mathbb{P}(Z_i \leq k_i).$$

4. Proofs

Throughout this paper we let $j = j(n)$ denote $\ln n / \ln(\frac{\ln n}{nr^d})$. We shall assume throughout that $\ln n / nr^d \to \infty$. Thus $j/nr^d \to \infty$ and in particular $j > 0$ for large n.

Lemma 4.1. *For any $\epsilon > 0$ it holds that $(1-\epsilon)j \leq M \leq (1+\epsilon)j + 1$ whp.*

Proof. For the lower bound we observe that if $M < (1-\epsilon)j$ then *every* ball of radius $R_1 r$ contains less than $(1-\epsilon)j$ points. Let $\eta > 0$ be the constant from lemma 3.1. There are $\Omega(r^{-d})$ disjoint translates $W_1,\dots,W_K$ of $B(0;R_1 r)$ with $\nu(W_i) \geq \eta\theta R_1^d r^d$. The joint distribution of $\mathcal{N}(W_1),\dots,\mathcal{N}(W_K),\mathcal{N}(\mathbb{R}^d \setminus \bigcup_i W_i)$ is multinomial and so we can write

$$\begin{aligned}\mathbb{P}\big(M \leq (1-\epsilon)j\big) &\leq \mathbb{P}\big(\mathcal{N}(W_1) \leq (1-\epsilon)j,\dots,\mathcal{N}(W_K) \leq (1-\epsilon)j\big) \\ &\leq \Pi_{i=1}^K \mathbb{P}\big(\mathcal{N}(W_i) \leq (1-\epsilon)j\big).\end{aligned}$$

By lemma 3.2 we have

$$\begin{aligned}\mathbb{P}\big(\mathcal{N}(W_i) \geq (1-\epsilon)j\big) &= \mathbb{P}\big(\mathrm{Bi}(n,\nu(W_i)) \geq (1-\epsilon)j\big) \\ &\geq \left(\frac{\eta\theta R_1^d nr^d}{e(1-\epsilon)j}\right)^{(1-\epsilon)j} = e^{-(1-\epsilon)j(\ln(j/nr^d)+C)},\end{aligned}$$

where $C = -\ln\big(\frac{\eta\theta R_1^d}{e(1-\epsilon)}\big)$. Notice that

$$\begin{aligned}j\Big(\ln\Big(\frac{j}{nr^d}\Big) + C\Big) &= \frac{\ln n}{\ln\big(\frac{\ln n}{nr^d}\big)}\Big(\ln\Big(\frac{\ln n}{nr^d}\Big) - \ln\Big(\ln\Big(\frac{\ln n}{nr^d}\Big)\Big) + C\Big) \\ &= \ln n\Big(1 - \ln\Big(\ln\Big(\frac{\ln n}{nr^d}\Big)\Big)/\ln\Big(\frac{\ln n}{nr^d}\Big) + C/\ln\Big(\frac{\ln n}{nr^d}\Big)\Big) \\ &= \ln n(1 + o(1)),\end{aligned}$$

where we've used that $\ln n/nr^d \to \infty$. Hence

$$\mathbb{P}\big(\mathcal{N}(W_i) \geq (1-\epsilon)j\big) \geq e^{-(1-\epsilon)j(\ln(j/nr^d)+C)} = n^{-1+\epsilon+o(1)}. \tag{1}$$

As $nr^d = o(\ln n)$ and $K = \Omega(r^{-d})$ we have that $K \geq n/\ln n = n^{1+o(1)}$ if n is sufficiently large. We find that

$$\mathbb{P}\big(M \leq (1-\epsilon)j\big) \leq \big(1 - n^{-1+\epsilon+o(1)}\big)^{n^{1+o(1)}} \leq e^{-n^{\epsilon+o(1)}} = o(1),$$

where we've used the fact that $1-x \leq e^{-x}$. This takes care of the lower bound.

For the upper bound, we remark that if X_i is in an l-cluster of for some $l \geq (1+\epsilon)j+1$ then $B(X_i;R_2r)$ must contain $(1+\epsilon)j$ points other than X_i by (C1) and (C3). Hence

$$\begin{aligned}\mathbb{P}\big(M \geq (1+\epsilon)j+1\big) &\leq n\mathbb{P}\big(\mathrm{Bi}(n,\nu_{\max}\theta R_2^d r^d) \geq (1+\epsilon)j\big)\\ &\leq n\Big(\frac{e\nu_{\max}\theta R_2^d nr^d}{(1+\epsilon)j}\Big)^{(1+\epsilon)j}\\ &= ne^{-(1+\epsilon)j\left(\ln\left(\frac{j}{nr^d}\right)+D\right)} = ne^{-(1+\epsilon+o(1))\ln n}\\ &= n^{-\epsilon+o(1)} = o(1),\end{aligned}$$

where $D := -\ln\big(\frac{e\nu_{\max}\theta R_2^d}{1+\epsilon}\big)$ and we've used lemma 3.2 and previous computations. ∎

We remark that when $nr^d \leq n^{-c}$ for some $c>0$ then j remains bounded ($\leq c^{-1}$ in fact). As a result the two point concentration result directly follows from lemma 4.1 in this case with $k(n) = \lfloor j(n) + \frac{1}{2} \rfloor$. What is more, when $c>1$ then lemma 4.1 (together with (C2)) even gives $M=1$ whp.

From now on we will therefore assume that $nr^d \geq n^{-\frac{3}{2}}$. We may assume this without loss of generality, because if we can prove the result for choices of $r(n)$ with $n^{-\frac{3}{2}} \leq nr^d \ll \ln n$, then it also follows for any choice of r with $nr^d = o(\ln n)$, by "splitting into two subsequences". By this we mean the following. Let us set:

$$r_1(n) := \begin{cases} r(n) & \text{if } nr^d < n^{-\frac{3}{2}}, \\ n^{-\frac{5}{2d}} & \text{otherwise.}\end{cases}, \quad r_2(n) := \begin{cases} r(n) & \text{if } nr^d \geq n^{-\frac{3}{2}}, \\ n^{-\frac{5}{2d}} & \text{otherwise.}\end{cases}$$

For $i=1,2$ let $M^{(i)}$ denote $M(n,r_i)$, the maximum of $h_n(A)$ over all $A \subseteq \{X_1,\dots,X_n\}$ if r_i is used instead of r. If we can prove two-point concentration holds for both $M^{(1)}$ and $M^{(2)}$ it must also hold for M (if $k^{(1)},k^{(2)}$ are such that whp. $M^{(i)} \in \{k^{(i)}, k^{(i)}+1\}$ then $M \in \{k,k+1\}$ setting $k(n) := k^{(1)}(n)$

if $nr^d < n^{-\frac{3}{2}}$ and $k(n) := k^{(2)}(n)$ otherwise). By the remark above $M^{(1)} = 1$ whp., so two-point concentration certainly holds for $M^{(1)}$ and we simply need to show that it also holds for $M^{(2)}$. Therefore we can indeed restrict attention to sequences r that satisfy $n^{-\frac{3}{2}} \le nr^d \ll \ln n$ in the remainder of the proofs (for the proofs of corollaries 2 and 3 notice that the previous applied to example 2.3, ie., the maximum degree plus one, shows that G_n is the empty graph whp. if $nr^d \le n^{-\frac{3}{2}}$).

For $l \in \mathbb{N}$ let N_l be the number of points in G_n that are in a $(\ge l)$-cluster. Let us denote $p_l := \mathbb{P}(X_1$ is in a $(\ge l)$-cluster), so that $\mathbb{E}N_l = np_l$. Notice that $n = N_1 \ge N_2 \ge \cdots \ge N_n \ge N_{n+1} = 0$. For fixed (but sufficiently large) n the sequence $\mathbb{E}N_l - l(\frac{j}{nr^d})^{\frac{1}{4}}$ is strictly decreasing in l, and it is negative for $l = n+1$. Hence there is a unique $k = k(n)$ such that

$$\mathbb{E}N_k - k\Big(\frac{j}{nr^d}\Big)^{\frac{1}{4}} \ge 0 > \mathbb{E}N_{k+1} - (k+1)\Big(\frac{j}{nr^d}\Big)^{\frac{1}{4}}.$$

The most important step in the proof will be to show that $\mathrm{Var}(N_k)$ is not too large, but first we must show that k is not too small.

Lemma 4.2. *$k \ge \frac{1}{2}j$ for n sufficiently large.*

Proof. Let $\eta > 0$ be the constant from lemma 3.1, and let $W_1, \ldots, W_K$ be disjoint translates of $B(0; R_1 r)$ with $\nu(W_i)/\mathrm{vol}(W_i) \ge \eta$ and $K = \Omega(r^{-d})$. By (C2) we have

$$N_{\lceil \frac{1}{2}j \rceil} \ge \Big\lceil \frac{1}{2}j \Big\rceil \cdot \Big|\Big\{ i : \mathcal{N}(W_i) \ge \frac{1}{2}j \Big\}\Big|.$$

Recall that, as $nr^d = o(\ln n)$ and $K = \Omega(r^{-d})$, we have $K \ge n^{1+o(1)}$. By computations in the proof of lemma 4.1 we also have that $\min_{i=1,\ldots,K} \mathbb{P}(\mathcal{N}(W_i) \ge (1-\epsilon)j) \ge n^{-1+\epsilon+o(1)}$ for any fixed $\epsilon > 0$ – cf. (1). Taking $\epsilon = \frac{1}{2}$ we may conclude that

$$\mathbb{E}N_{\lceil \frac{1}{2}j \rceil} \ge \Big\lceil \frac{1}{2}j \Big\rceil \cdot K \cdot \min_i \mathbb{P}\Big(\mathcal{N}(W_i) \ge \frac{1}{2}j\Big) \ge \Big\lceil \frac{1}{2}j \Big\rceil n^{\frac{1}{2}+o(1)} > \Big\lceil \frac{1}{2}j \Big\rceil \Big(\frac{j}{nr^d}\Big)^{\frac{1}{4}},$$

where in the last inequality we have used $j/nr^d \le \ln n \cdot n^{\frac{3}{2}}$ (for n sufficiently large). We must therefore have $k \ge \frac{1}{2}j$ by definition of k. ∎

Lemma 4.3. $\mathrm{Var}(N_k) = O(k\mathbb{E}N_k)$.

Proof. For $i=1,\dots,n$ let us denote $A_i:=\{X_i \text{ is in some } (\geq k)\text{-cluster}\}$. We have $N_k=\sum_i 1_{A_i}$ and $N_k^2=\sum_{i,j} 1_{A_i}1_{A_j}$ and the variance becomes

$$\begin{aligned}\operatorname{Var}(N_k) &= \sum_{i,j}\mathbb{E}1_{A_i}1_{A_j} - \sum_{i,j}\mathbb{E}1_{A_i}\mathbb{E}1_{A_j} = \sum_{i,j}\mathbb{P}(A_i\cap A_j) - \sum_{i,j}\mathbb{P}(A_i)^2 \\ &= n\mathbb{P}(A_1) + 2\binom{n}{2}\mathbb{P}(A_1\cap A_2) - n^2\mathbb{P}(A_1)^2.\end{aligned}$$

Partitioning we get

$$\begin{aligned}\mathbb{P}(A_1\cap A_2) = \mathbb{P}\big(A_1\cap A_2\cap\{\|X_1-X_2\|<2R_2r\}\big)& \\ +\,\mathbb{P}\big(A_1\cap A_2\cap\{\|X_1-X_2\|\geq 2R_2r\}\big)&.\end{aligned}$$

Let us denote by $p_{k,l}$ the probability that X_1 lies in a $(\geq k)$-cluster whose nodes are a subset of $\{X_1,\dots,X_l\}$. Then we have that $p_{k,l}\leq p_{k,l+1}$ and $p_{k,n}=p_k$. Define the events B_1,B_2 by:

$$\begin{aligned}B_1 &:= \big\{\exists C\subseteq\{X_3,\dots,X_n\} \text{ such that } \{X_1\}\cup C \text{ is a } (\geq k)\text{-cluster,} \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{and } \|X_1-X_2\|<2R_2r\big\}, \\ B_2 &:= \big\{X_1 \text{ is in some } (\geq k)\text{-cluster, and } \textit{every } (\geq k)\text{-cluster} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{that contains } X_1 \text{ also contains } X_2\big\}.\end{aligned}$$

We see that

$$\mathbb{P}\big(A_1\cap A_2\cap\big\{\|X_1-X_2\|<2R_2r\big\}\big)\leq\mathbb{P}(B_1)+\mathbb{P}(B_2).$$

Clearly $\mathbb{P}(B_1)\leq p_{k,n-1}\nu_{\max}2^d\theta R_2^dr^d\leq p_k\nu_{\max}2^d\theta R_2^dr^d$. Let E be the event that $B(X_1;R_2r)$ contains no more than $3j$ points of $\{X_2,\dots,X_n\}$. Then

$$\mathbb{P}(B_2)\leq\mathbb{P}(B_2\cap E)+\mathbb{P}(E^c).$$

By lemma 3.2 and computations as in the proof of lemma 4.1:

$$\mathbb{P}(E^c)\leq\left(\frac{e\nu_{\max}\theta R_2^dnr^d}{3j}\right)^{3j} = e^{-(3+o(1))\ln n} = n^{-3+o(1)}.$$

Observe that if E holds then by (C1) and (C3) any $(\geq k)$-cluster containing X_1 cannot have more than $3j+1$ elements. Given that both A_1 and E hold, we can sample a random $C\in\binom{\{X_1,\dots,X_n\}}{3j+1}$ uniformly from all sets of cardinality $3j+1$ that contain a $(\geq k)$-cluster that contains X_1. If B_2 also holds then we must always have $X_2\in C$, hence

$$\mathbb{P}(B_2\cap E)=\mathbb{P}(B_2|A_1\cap E)\mathbb{P}(A_1\cap E)\leq\mathbb{P}(X_2\in C|A_1\cap E)\mathbb{P}(A_1\cap E).$$

By symmetry all sets $C\in\binom{\{X_1,\dots,X_n\}}{3j+1}$ that contain X_1 are equally likely. This gives that $\mathbb{P}(X_2\in C|A_1\cap E)=\binom{n-2}{3j-1}/\binom{n-1}{3j}=\frac{3j}{n-1}$, and thus

$$\mathbb{P}(B_2\cap E)\le\mathbb{P}(X_2\in C|A_1\cap E)\mathbb{P}(A_1\cap E)\le p_k\frac{3j}{n-1}.$$

Now let us consider $\mathbb{P}(A_1\cap A_2\cap\{\|X_1-X_2\|\ge 2R_2r\})$. For $x\in\mathbb{R}^d$ let us denote by $A(x)$ the event that there exists a $C\subseteq\{X_3,\dots,X_n\}$ such that $\{x\}\cup C$ is a $(\ge k)$-cluster. So,

$$\mathbb{P}\big(A_1\cap A_2\cap\{\|X_1-X_2\|\ge 2R_2r\}\big)=\int_{\mathbb{R}^d}\int_{\mathbb{R}^d\setminus B(x;2R_2r)}\mathbb{P}\big(A(x)\cap A(y)\big)\mathrm{d}\nu(y)\mathrm{d}\nu(x).$$

Our next target is to prove the (intuitively clear) result that for $x,y\in\mathbb{R}^d$ with $\|x-y\|>2R_2r$ we have $\mathbb{P}(A(x)\cap A(y))\le\mathbb{P}(A(x))\mathbb{P}(A(y))$. To this end, let $R(x)$ denote $|\{X_3,\dots,X_n\}\cap B(x;R_2r)|$. We first claim that for $x,y\in\mathbb{R}^d$ with $\|x-y\|\ge 2R_2r$ it holds that $\mathbb{P}(A(x)\cap A(y)|R(x)=i)=\mathbb{P}(A(x)|R(x)=i)\mathbb{P}(A(y)|R(x)=i)$. To see this, fix such x,y and let $A(x,i)$ be the event that $\{X_3,\dots,X_{3+i-1}\in B(x;R_2r),X_{3+i},\dots,X_n\notin B(x;R_2r)\}$. By symmetry we have:

$$\begin{aligned}\mathbb{P}\big(A(x)\cap A(y)|R(x)=i\big)&=\mathbb{P}\big(A(x)\cap A(y)|A(x,i)\big)\\&=\mathbb{P}\big(A(x)|A(x,i)\big)\mathbb{P}\big(A(y)|A(x,i)\big)\\&=\mathbb{P}\big(A(x)|R(x)=i\big)\mathbb{P}\big(A(y)|R(x)=i\big),\end{aligned}$$

where in the second equality we have used that given $A(x,i)$ the variables $X_3,\dots,X_n$ are still independent (but not identically distributed anymore though), that $A(x)$ depends only on the points inside $B(x;R_2r)$ and that $A(y)$ depends only on the points in $\mathbb{R}^d\setminus B(x;R_2r)$.

We now have

$$\begin{aligned}\mathbb{P}\big(A(x)\cap A(y)\big)&=\sum_{i=0}^{n-2}\mathbb{P}\big(A(x)\cap A(y)|R(x)=i\big)\mathbb{P}\big(R(x)=i\big)\\&=\sum_{i=0}^{n-2}\mathbb{P}\big(A(y)|R(x)=i\big)\mathbb{P}\big(A(x)|R(x)=i\big)\mathbb{P}\big(R(x)=i\big).\end{aligned}$$

Define $f(i):=\mathbb{P}(A(x)|R(x)=i),g(i):=\mathbb{P}(A(y)|R(x)=i)$. We remark that $\mathbb{P}(A(x)\cap A(y))=\mathbb{E}f(R(x))g(R(x))$ (by the above computation) and $\mathbb{P}(A(x))=\mathbb{E}f(R(x)),\mathbb{P}(A(y))=\mathbb{E}g(R(x))$. Clearly f is increasing in i, whereas g is decreasing in i. A standard result (see for instance [15]) tells us that in such a case

$$\mathbb{E}f(R(x))g(R(x))\le\mathbb{E}f(R(x))\mathbb{E}g(R(x)).$$

In other words, we indeed have $\mathbb{P}(A(x), A(y)) \leq \mathbb{P}(A(x))\mathbb{P}(A(y))$ whenever $\|x-y\| \geq 2R_2 r$. We are now in a position to write

$$\begin{aligned}\mathbb{P}\big(A_1 \cap A_2 \cap \{\|X_1 - X_2\| \geq 2R_2 r\}\big) &= \int_{\mathbb{R}^d}\int_{\mathbb{R}^d \setminus B(x;2R_2 r)} \mathbb{P}\big(A(x) \cap A(y)\big)\mathrm{d}\nu(y)\mathrm{d}\nu(x) \\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d \setminus B(x;2R_2 r)} \mathbb{P}\big(A(x)\big)\mathbb{P}\big(A(y)\big)\mathrm{d}\nu(y)\mathrm{d}\nu(x) \\ &\leq \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbb{P}\big(A(x)\big)\mathbb{P}\big(A(y)\big)\mathrm{d}\nu(y)\mathrm{d}\nu(x) \\ &= p_{k,n-1}^2 \leq p_k^2.\end{aligned}$$

Summarising, we have

$$\begin{aligned}\mathbb{P}(A_1 \cap A_2) &\leq \mathbb{P}(B_1) + \mathbb{P}(B_2 \cap E) + \mathbb{P}(E^c) \\ &\qquad + \mathbb{P}\big(A_1 \cap A_2 \cap \{\|X_1 - X_2\| \geq 2R_2 r\}\big) \\ &\leq p_k \nu_{\max} 2^d \theta R_2^d r^d + p_k \frac{3j}{n-1} + n^{-3+o(1)} + p_k^2 \\ &\leq 4p_k \frac{j}{n-1} + p_k^2 + n^{-3+o(1)},\end{aligned}$$

using that $\nu_{\max} 2^d \theta R_2^d r^d \leq \frac{j}{n-1}$ for n large enough (recall that $\frac{j}{nr^d} \to \infty$). Hence we find that

$$\begin{aligned}\operatorname{Var}(N_k) &= n\mathbb{P}(A_1) + 2\binom{n}{2}\mathbb{P}(A_1 \cap A_2) - n^2 \mathbb{P}(A_1)^2 \\ &\leq np_k + n(n-1)\Big(p_k^2 + 4\frac{j}{n-1}p_k + n^{-3+o(1)}\Big) - n^2 p_k^2 \\ &= np_k(1 + 4j - p_k) + n^{-1+o(1)} = O(k\mathbb{E}N_k),\end{aligned}$$

where in the last line we have used that $k \geq \lceil \frac{1}{2}j \rceil \geq 1$ (so that $1+4j-p_k = O(k)$) by lemma 4.2 and the fact that $k\mathbb{E}N_k \to \infty$ (so that $n^{-1+o(1)} = O(k\mathbb{E}N_k)$) as $\mathbb{E}N_k \geq k\big(\frac{j}{nr^d}\big)^{\frac{1}{4}} \geq \big(\frac{j}{nr^d}\big)^{\frac{1}{4}} \to \infty$. ∎

Proof of Theorem 1. By lemma 4.3 we have

$$\begin{aligned}\mathbb{P}(M < k) = \mathbb{P}(N_k = 0) &\leq \mathbb{P}\big(|N_k - \mathbb{E}N_k| \geq \mathbb{E}N_k\big) \\ &\leq \frac{\operatorname{Var}(N_k)}{(\mathbb{E}N_k)^2} = O\Big(\frac{k}{\mathbb{E}N_k}\Big) = O\bigg(\Big(\frac{nr^d}{j}\Big)^{\frac{1}{4}}\bigg) = o(1),\end{aligned}$$

as $\mathbb{E}N_k \geq k\big(\frac{j}{nr^d}\big)^{\frac{1}{4}}$ and $\frac{nr^d}{j} = o(1)$.

Now define P_l to be the number of pairs $\{X_i, X_j\}$ that are contained in some $(\geq l)$-cluster. Set

$$q := \mathbb{P}\big(\{X_1, X_2\} \text{ is part of some } (\geq k+2)\text{-cluster}\big).$$

Clearly, $\mathbb{E}P_{k+2} = \binom{n}{2} q$. If $\{X_1, X_2\}$ is in some $(\geq k+2)$-cluster, then by (C4) either X_1 is part of a $(\geq k+1)$-cluster that misses X_2 or X_2 is part of a $(\geq k+1)$-cluster that misses X_1. Hence

$$\begin{aligned} q &\leq 2\mathbb{P}\big(X_1 \text{ is part of a } (\geq k+1)\text{-cluster in } \{X_1, X_3, \dots, X_n\} \\ &\qquad\qquad\qquad\qquad \text{and } \|X_1 - X_2\| < R_2 r\big) \\ &\leq 2p_{k+1}\nu_{\max}\theta R_2^d r^d. \end{aligned}$$

Hence, $\mathbb{E}P_{k+2} = \binom{n}{2} q = O(nr^d \mathbb{E}N_{k+1})$. Because $P_{k+2} > 0 \Leftrightarrow P_{k+2} \geq \binom{k+2}{2}$ (by (C1)) it now follows that

$$\begin{aligned} \mathbb{P}\big(M > k+1\big) = \mathbb{P}\big(P_{k+2} > 0\big) &= \mathbb{P}\left(P_{k+2} \geq \binom{k+2}{2}\right) \leq \frac{2\mathbb{E}P_{k+2}}{(k+2)(k+1)} \\ &= O\left(\frac{\mathbb{E}N_{k+1} nr^d}{(k+1)(k+2)}\right) = O\left(\frac{\left(\frac{j}{nr^d}\right)^{\frac{1}{4}} nr^d}{k+2}\right) = o(1), \end{aligned}$$

using that $\mathbb{E}N_{k+1} \leq (k+1)(\frac{j}{nr^d})^{\frac{1}{4}}$ and $nr^d(\frac{j}{nr^d})^{\frac{1}{4}}/(k+2) = O((\frac{nr^d}{j})^{\frac{3}{4}}) = o(1)$ (as $k = \Omega(j)$ by lemma 4.2). ∎

Proof of Corollary 2. Let us fix $\epsilon = \frac{1}{2}$, $K = 6$. Let M denote the maximum over all $x \in \mathbb{R}^d$ of the chromatic number of the subgraph $G_n[\{X_1, \dots, X_n\} \cap B(x; rK)]$ of G_n induced by the vertices inside $B(x; rK)$. Then M corresponds to the clustering rule which sets $h_n(A) = 0$ if A is not contained in some ball $B(x; rK)$ and otherwise $h_n(A)$ is equal to the chromatic number of the graph H with vertex set A and an edge $ab \in E(H)$ whenever $\|a - b\| < r$. Clearly the clustering rule satisfies (C1)–(C4) taking $R_1 = \frac{1}{2}, R_2 = 2K$, so theorem 1 applies. Applying also lemma 4.1 to the maximum number of points in a translate of $B(0; Kr)$, we see that whp. the following two statements hold:

(i) $M \in \{k, k+1\}$;
(ii) for every x, $B(x; rK)$ contains at most $(1+\epsilon)j + 1$ points,

where $k = k(n)$ is the sequence given by theorem 1 applied to M. Now consider a realisation $X_1, \dots, X_n$ such that both (i) and (ii) hold and let us prove that there exists a colouring of G_n with M colours. Let us denote $V_i := \{X_1, \dots, X_i\}$. Suppose that the subgraph $G_n[V_i]$ of G_n induced by V_i

can be M-coloured (this is certainly true when $i \le M$) and consider $G_n[V_{i+1}]$. Let A_t denote the annulus $\{x \in \mathbb{R}^d : tr < \|x - X_i\| < (t+2)r\}$. As $|B(X_{i+1}; rK) \cap V_{i+1}| \le (1+\epsilon)j + 1$, by choice of ϵ and K we must have

$$|A_t \cap V_{i+1}| \le \frac{1}{2}j \le k \le M,$$

for some $0 \le t \le K-2$ (notice that X_{i+1} is not in any of the A_t and that A_0, A_2, A_4 are disjoint), using also lemma 4.2. By (i) there exists an M-colouring f_0 of $G_n[V_{i+1} \cap B(X_{i+1}; (t+1)r)]$ and by the induction hypothesis there exists an M-colouring f_1 of $G_n[V_{i+1} \setminus B(X_{i+1}; (t+1)r)]$ (note that this is a subgraph of $G_n[V_i]$). The number of colours used by f_0 on A_t plus the number of colours used by f_1 on A_t is at most M, and hence the colourings f_0, f_1 can be permuted in such a way that the set of colours used by f_0 on A_t is disjoint from the set of colours used by f_1 on A_t. But now f_0 and f_1 fit together to give a proper M-colouring of V_{i+1}, because if $X_j \in V_{i+1} \cap B(X_{i+1}; (t+1)r)$ and $X_k \in V_{i+1} \setminus B(X_{i+1}; (t+1)r)$ are adjacent in G_n then we must have $X_j, X_k \in A_t$. ∎

Proof of Corollary 3. We proceed in a very similar manner to the proof of corollary 2. This time let us fix $\epsilon = \frac{1}{3}, K = 9$ and let M denote the maximum over all $x \in \mathbb{R}^d$ of the minimum degree plus 1 of the subgraph $G_n[\{X_1, \dots, X_n\} \cap B(x; rK)]$. Then M corresponds to the clustering rule which sets $h_n(A) = 0$ if A is not contained in some ball $B(x; rK)$ and otherwise $h_n(A)$ is equal to the minimum degree $+1$ of the graph H with vertex set A and an edge $ab \in E(H)$ whenever $\|a - b\| < r$. Again the clustering rule satisfies (C1)–(C4) with $R_1 = \frac{1}{2}, R_2 = 2K$. So again conditions (i), (ii) from the proof of corollary 2 hold whp.

Clearly $M \le \delta^*(G_n) + 1$, so it remains to show that $M \ge \delta^*(G_n) + 1$ whp. To this end, let us again consider a realisation $X_1, \dots, X_n$ satisfying (i), (ii) and let $V \subseteq \{X_1, \dots, X_n\}$ be such that $\delta(G_n[V]) = \delta^*(G_n)$. Pick $X_i \in V$ with minimum degree in $G_n[V]$, and this time let $A_t := \{x \in \mathbb{R}^d : tr < \|x - X_i\| < (t+3)r\}$. By choice of ϵ, K we have

$$|A_t \cap V| < \frac{1}{2}j \le M$$

for some $0 \le t \le K-3$. But then $\{x \in \mathbb{R}^d : (t+1)r \le \|x - X_i\| < (t+2)r\}$ cannot contain any node of V, for otherwise this node would have degree $< M - 1 \le \delta^*(G_n)$ in $G_n[V]$, contradicting the choice of V. This shows that if H is the component of $G_n[V]$ that contains X_i then the vertices of H are contained in $B(X_i; (t+1)r) \subseteq B(X_i; rK)$. In other words, $\delta^*(G_n) = \delta(H) \le M - 1$ as required. ∎

5. Discussion and further work

We have seen that a range of graph parameters of random geometric graphs, including the clique number $\omega(G_n)$, the maximum degree $\Delta(G_n)$, the degeneracy $\delta^*(G_n)$ and the chromatic number $\chi(G_n)$, all become concentrated on two consecutive integers as long as the distance thresholds $r(n)$ are chosen to satisfy $nr^d = o(\ln n)$. This proves and extends a conjecture of Penrose ([14]).

A very natural question would be to ask what happens for other choices of $r(n)$. The author was recently able to adapt an argument of Penrose ([14]) to show that when $nr^d = \Theta(\ln n)$ and ν is the uniform distribution on $[0,1]^d$, the maximum degree is not concentrated on a finite range, and that if $\ln n \ll nr^d \ll \ln^d n$ then there exist sequences $a(n), b(n)$ such that for all x:

$$\mathbb{P}\left(\frac{\Delta(G_n) - a(n)}{b(n)} < x\right) \to e^{-e^{-x}}. \tag{2}$$

In the $G(n,p)$-model a similar result to (2) holds as well (see [4]), and the results on the clique number and chromatic number in this paper are also similar to results on $G(n,p)$. There is reason to believe that the clique number $\omega(G_n)$ will display behaviour similar to (2) when $nr^d \gg \ln n$ (unlike the clique number in $G(n,p)$), but unfortunately the proof appears to be more involved than the proof for the maximum degree. Here our intuition stems from the idea that if we dissect $[0,1]^d$ into smaller hypercubes of side length Kr with K growing arbitrarily slowly and consider the maximum over all of these smaller cubes of the clique number of the subgraph induced by the points in the cube, then whp. this quantity coincides with $\omega(G_n)$. What is more, if instead of n we drop $N(n)$ points onto the unit square when constructing G_n, where $N(n)$ is a Poisson distributed random variable independent of the X_i with mean n, then the clique numbers of these induced subgraphs are independent. Thus the clique number can be approximated by the maximum of a growing number of i.i.d. discrete random variables (whose distribution changes with n). Results from extreme value theory, such as the ones in [3], suggest that something like (2) might be the case.

Except for the case when $nr^d \le n^{-\epsilon}$ for some $\epsilon > 0$, our proof does not tell us which are the exact values of the sequence $k(n)$. It would of course be of interest to see how much more can be said about values of $k(n)$ for the various random variables considered here. In the $G(n,p(n))$-model the precise value of the two consecutive integers that the chromatic number $\chi(G(n,p(n)))$ is concentrated on is still shrouded in mystery, but recently Achloptias and Naor ([1]) were able to pinpoint precisely which two consecutive integers the chromatic number will be concentrated on if $p(n)$ is chosen as $p(n) = \frac{c}{n}$ for some $0 < c < \infty$.

In the case when $nr^d \le n^{-\epsilon}$ for some $\epsilon > 0$, the sequence $k(n)$ is the same regardless of which exact clustering rule was chosen. A little extra effort will show that in this case it even holds that $\chi(G_n) = \omega(G_n)$ whp. (see [12]). In contrast there are sequences $r(n)$ satisfying $nr^d \le n^{-\epsilon}$ for which $\Delta(G_n) = \omega(G_n)$ with probabily tending to a constant $\in (0,1)$. Another natural line of investigation is therefore to ask what can be said about the interrelationship between the various random variables considered here. If $nr^d = o(\ln n)$ and $nr^d \gg n^{-\epsilon}$ for all $\epsilon > 0$ then we know they're all asymptotically equivalent to $j(n) = \ln n / \ln(\frac{\ln n}{nr^d})$, but do the differences $\chi(G_n) - \omega(G_n)$ etc. remain bounded or tend to infinity?

6. Acknowledgments

The author would like to thank both the anonymous referee and Colin McDiarmid for careful proofreading and many helpful suggestions which have greatly improved this paper.

References

[1] D. Achlioptas and A. Naor: The two possible values of the chromatic number of a random graph, *Ann. of Math. (2)* **162(3)** (2005), 1335–1351.

[2] N. Alon and M. Krivelevich: The concentration of the chromatic number of random graphs, *Combinatorica* **17(3)** (1997), 303–313.

[3] C. W. Anderson, S. G. Coles and J. Hüsler: Maxima of Poisson-like variables and related triangular arrays, *Ann. Appl. Probab.* **7(4)** (1997), 953–971.

[4] B. Bollobás: The distribution of the maximum degree of a random graph, *Discrete Math.* **32(2)** (1980), 201–203.

[5] J. Glaz and N. Balakrishan: *Scan Statistics and Applications*, Birkhäuser, Boston, 1999.

[6] J. Glaz, J. Naus and S. Wallenstein: *Scan statistics*, Springer, New York, 2001.

[7] T. Łuczak: A note on the sharp concentration of the chromatic number of random graphs, *Combinatorica* **11(3)** (1991), 295–297.

[8] C. L. Mallows: An inequality involving multinomial probabilities, *Biometrika* **55(2)** (1968), 422–424.

[9] M. Månsson: Poisson approximation in connection with clustering of random points, *Ann. Appl. Probab.* **9(2)** (1999), 465–492.

[10] D. W. Matula: The employee party problem, *Not. A. M. S.* **19** (1972), A–382.

[11] C. J. H. McDiarmid: Random channel assignment in the plane, *Random Structures Algorithms* **22(2)** (2003), 187–212.

[12] C. J. H. McDiarmid and T. Müller: On the chromatic number of random geometric graphs, submitted.

[13] M. D. Penrose: Focusing of the scan statistic and geometric clique number, *Adv. in Appl. Probab.* **34(4)** (2002), 739–753.

[14] M. D. PENROSE: *Random Geometric Graphs*, Oxford University Press, Oxford, 2003.
[15] S. M. ROSS: *Probability models in computer science*, Harcourt/Academic Press, 2002.
[16] W. RUDIN: *Real and Complex Analysis*, McGraw-Hill, New York, 1987.

A. The average degree and nr^d

In this section we will prove the following proposition in order to substantiate the claim that nr^d is a good measure of the average degree of G_n. Recall that θ denotes the d-dimensional volume of the unit ball $\{x : \|x\| < 1\}$ wrt. the (arbitrary) norm that we have furnished $\mathbb{R}^d$ with.

Proposition A.1. *The average degree D of G_n satisfies:*
(i) If $n^2r^d \to \infty$ then $\frac{D}{nr^d} \to \theta \int_{\mathbb{R}^d} f^2(x)dx$ in probability;
(ii) If $n^2r^d = O(1)$ then $\liminf_{n\to\infty} \mathbb{P}(D=0) > 0$.

Proof of (i). We remark that $D = 2|E(G_n)|/n$ (where $E(G)$ denotes the edge set of G). Consequently

$$\mathbb{E}D = \frac{2}{n}\binom{n}{2}\mathbb{P}\big(X_2 \in B(X_1; r)\big) = (n-1)\mathbb{E}\nu\big(B(X_1; r)\big).$$

For any fixed $x \in \mathbb{R}^d$ the sets $B(x; r(n)), n \in \mathbb{N}$ satisfy the conditions of theorem 7.10 in [16], giving $\frac{\nu(B(x;r))}{\theta r^d} \to f(x)$ almost everywhere. As $\frac{(n-1)\nu(B(x;r))}{\theta n r^d} \le \nu_{\max}$ we can apply the dominated convergence theorem to deduce that

$$\frac{\mathbb{E}D}{\theta n r^d} = \int_{\mathbb{R}^d} f(x)\frac{(n-1)\nu\big(B(x; r)\big)}{\theta n r^d}\mathrm{d}x \to \int_{\mathbb{R}^d} f^2(x)\mathrm{d}x.$$

To complete the proof it now suffices to show that $\mathrm{Var}\big(\frac{D}{nr^d}\big) \to 0$. Let I be the set of all possible candidate edges $X_iX_j, 1 \le i < j \le n$. We have that

$$\mathrm{Var}\left(\frac{D}{nr^d}\right) = \frac{4}{n^4r^{2d}}\sum_{e,f\in I}\mathbb{P}(e \in G_n, f \in G_n) - \mathbb{P}(e \in G_n)\mathbb{P}(f \in G_n).$$

If $|e \cap f| = 0$ then $\mathbb{P}(e \in G_n, f \in G_n) = \mathbb{P}(e \in G_n)\mathbb{P}(f \in G_n)$. If $|e \cap f| = 1$ then $\mathbb{P}(e \in G_n, f \in G_n) \le (\theta\nu_{\max}r^d)^2$, and if $|e \cap f| = 2$ then $\mathbb{P}(e \in G_n, f \in G_n) = \mathbb{P}(e \in G_n) \le \theta\nu_{\max}r^d$. We see that

$$\begin{aligned}\mathrm{Var}\left(\frac{D}{nr^d}\right) &\le \frac{4}{n^4r^{2d}}\left(\binom{n}{2}\theta\nu_{\max}r^d + 2(n-2)\binom{n}{2}\theta^2\nu_{\max}^2r^{2d}\right)\\ &= O\big(n^{-2}r^{-d} + n^{-1}\big) = o(1),\end{aligned}$$

as required.

Proof of (ii). Let $\lambda < \infty$ be such that $n^2 r^d \leq \lambda$ ($\forall n$) and set $r' := \lambda^{\frac{1}{d}} n^{-\frac{2}{d}}$ (so that $r \leq r'$). If G'_n is the graph we get by taking $X_i X_j \in E(G'_n)$ if $\|X_i - X_j\| < r'$, then G_n is a subgraph of G'_n. By theorem 6.3 in [14] we have:

$$\mathbb{P}(D=0) = \mathbb{P}\big(\varDelta(G_n)=0\big) \geq \mathbb{P}\big(\varDelta(G'_n)=0\big) = \big(1+o(1)\big)e^{-c},$$

for some constant $c = c(\lambda)$. ∎

B. Proof of lemma 3.1

Lemma 3.1 is a direct consequence of part (ii) of the following proposition, which we will now proceed to prove. A similar result already appeared in [11], but the result and proof given there were formulated in terms of the (Euclidean) unit disk in two dimensions.

Proposition B.1. *The maximum density $\nu_{\max}$ satisfies the following:*

- *(i) $\nu_{\max} = \sup_C \frac{\nu(C)}{\mathrm{vol}(C)}$ where the supremum is over all cubes $C \subseteq \mathbb{R}^d$;*
- *(ii) Fix $\epsilon > 0$ and let $W \subseteq \mathbb{R}^d$ be bounded with $\mathrm{vol}(W) > 0$. Then there exist $\varOmega(r^{-d})$-many disjoint translates $W_1, \dots, W_N$ of rW such $\nu(W_i)/\mathrm{vol}(W_i) > (1-\epsilon)\nu_{\max}$ for all i.*

Proof of (i). Let m be the supremum over all cubes C of $\frac{\nu(C)}{\mathrm{vol}(C)}$. By definition of $\nu_{\max}$ we have $\nu(C) = \int_C f \leq \nu_{\max}\,\mathrm{vol}(C)$ giving $m \leq \nu_{\max}$, so it remains to show $m \geq \nu_{\max}$. To this end, let $A \subseteq \mathbb{R}^d$ satisfy $\nu(A) \geq (1-\epsilon)\nu_{\max}\,\mathrm{vol}(A)$. Let $\mathcal{B}$ be the collection of all "boxes" $\varPi_{i=1}^d (a_i, b_i]$ with $a_1 < b_i \in \mathbb{Q}, i = 1, \dots, d$. Then $\mathcal{B}$ generates the Lebesgue sigma field, and by the outer-measure construction (see for instance [16]):

$$\mathrm{vol}(A) = \inf\left\{\sum_{i=1}^{\infty} \mathrm{vol}(B_i) : B_1, B_2, \dots \in \mathcal{B} \cup \{\emptyset\},\ A \subseteq \bigcup_{i=1}^{\infty} B_i\right\}.$$

Thus we can find a countable $B_1, B_2, \dots \in \mathcal{B}$ with $A \subseteq \bigcup_{i=1}^{\infty} B_i$ such that $\sum_{i=1}^{\infty} \mathrm{vol}(B_i) \leq (1+\sqrt{\epsilon})\,\mathrm{vol}(A)$. Note that we may assume that the B_i are cubes (to see this note that $\varPi_{i=1}^d (a_i, b_i]$ can be dissected into finitely many cubes if $a_i, b_i \in \mathbb{Q}$ for $i = 1, \dots, d$). We now claim that at least one B_i must satisfy $\frac{\nu(B_i)}{\mathrm{vol}(B_i)} \geq (1-\sqrt{\epsilon})\nu_{\max}$. To see this suppose that $\frac{\nu(B_i)}{\mathrm{vol}(B_i)} < (1-\sqrt{\epsilon})\nu_{\max}$ for all i. Then we would have

$$\begin{aligned}\sum_i \nu(B_i) = \sum_i \frac{\nu(B_i)}{\mathrm{vol}(B_i)}\mathrm{vol}(B_i) &< \big(1-\sqrt{\epsilon}\big)\nu_{\max}\sum_i \mathrm{vol}(B_i)\\ &\leq (1-\epsilon)\nu_{\max}\,\mathrm{vol}(A) < \nu(A),\end{aligned}$$

a contradiction. Hence for some i we have $(1-\sqrt{\epsilon})\nu_{\max} \leq \frac{\nu(B_i)}{\mathrm{vol}(B_i)} \leq m$, where the last inequality follows by definition of m and by the fact that as B_i is a cube. Taking $\epsilon \to 0$ we find that $m \geq \nu_{\max}$ as required.

Proof of (ii). We first consider the special case when W is a cube. By (i) there exists a cube C with $\nu(C) \geq (1-\frac{\epsilon}{2})\nu_{\max}\,\mathrm{vol}(C)$. Let s be the side of W and t be the side of C. Note that C can be covered by $N = (\lceil \frac{t}{rs} \rceil)^d$ disjoint translates of rW, let's call them $W_1, \dots, W_N$. Now suppose that a proportion $> \frac{3}{4}$ of the translates W_i satisfies $\nu(W_i) < (1-\epsilon)\nu_{\max}\,\mathrm{vol}(W_i)$. Then, setting $C' := \bigcup_i W_i$, and noticing that $\mathrm{vol}(C') = (1+o(1))\,\mathrm{vol}(C)$ we see that (for r small enough):

$$\begin{aligned}\nu(C) \leq \nu(C') &< \frac{3}{4}(1-\epsilon)\nu_{\max}\,\mathrm{vol}(C') + \frac{1}{4}\nu_{\max}\,\mathrm{vol}(C') \\ &= \left(1 - \frac{3\epsilon}{4}\right)\big(1+o(1)\big)\nu_{\max}\,\mathrm{vol}(C) < \left(1 - \frac{\epsilon}{2}\right)\nu_{\max}\,\mathrm{vol}(C) \leq \nu(C),\end{aligned}$$

a contradiction. Hence there must be at least $\frac{1}{4}(\lceil \frac{t}{rs} \rceil)^d = \Omega(r^{-d})$ of the W_i that satisfy $\nu(W_i) \geq (1-\epsilon)\nu_{\max}\,\mathrm{vol}(W_i)$, proving the lemma for the special case when W is a cube.

Now let W be a general bounded set. Since W is bounded it is contained in some cube C. Let $p = \frac{\mathrm{vol}(W)}{\mathrm{vol}(C)}$ be the proportion of C covered by W. By the previous there exist points $x_1, \dots, x_N$ with $N = \Omega(r^{-d})$ such that the sets $x_i + rC$ are disjoint and satisfy $\nu(x_i + rC) \geq (1-p\epsilon)\nu_{\max}\,\mathrm{vol}(C)r^d$. By construction the sets $x_i + rW$ are disjoint and $x_i + rW \subseteq x_i + rC$. We now observe that $\nu(x_i + rW)$ must be $\geq (1-\epsilon)\nu_{\max}\,\mathrm{vol}(W_j)r^d$, because otherwise $\nu(x_i + rC) < (1-p)\nu_{\max}\,\mathrm{vol}(C)r^d + (1-\epsilon)\nu_{\max} p\,\mathrm{vol}(C)r^d = (1-p\epsilon)\nu_{\max}\,\mathrm{vol}(x_i + rC) \leq \nu(x_i + rC)$, a contradiction. ∎

Tobias Müller

Department of Statistics
University of Oxford
1 South Parks Road
Oxford OX1 3TG
United Kingdom
muller@stats.ox.ac.uk