On Sharp Thresholds in Random Geometric Graphs

Milan Bradonjić[1] and Will Perkins[2]

1 **Bell Labs, Alcatel-Lucent**
600 Mountain Avenue 2C318, Murray Hill, NJ, USA
milan@research.bell-labs.com

2 **School of Mathematics, Georgia Tech**
686 Cherry St, Atlanta, GA, USA
perkins@math.gatech.edu

Abstract

We give a characterization of vertex-monotone properties with sharp thresholds in a Poisson random geometric graph or hypergraph. As an application we show that a geometric model of random k-SAT exhibits a sharp threshold for satisfiability.

1998 ACM Subject Classification G.3 Probability and Statistics

Keywords and phrases Sharp thresholds, random geometric graphs, k-SAT

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2014.500

1 Introduction

A property A of a discrete random structure is said to exhibit a *sharp threshold* with respect to a parameter p if there exists a $p_c = p_c(n)$ so that for every $\epsilon > 0$, for $p > (1+\epsilon)p_c$, A holds with probability $1 - o(1)$ and for $p < (1 - \epsilon)p_c$, A holds with probability $o(1)$. The classic sharp thresholds in the Erdős-Rényi random graph $G(n, p)$ are the threshold for connectivity at $p = \log n/n$ and the threshold for a giant component at $p = 1/n$, see [2]. A property that does not exhibit a sharp threshold is that of containing a triangle: for any $c \in (0, \infty)$, when $p = c/n$ the probability that $G(n, p)$ contains a triangle is strictly bounded away from 0 and 1.

In addition to much investigation of the threshold location and behavior of specific properties of random graphs, a series of works have proved general threshold theorems. The first such result was by Bollobás and Thomason [6] showing that any monotone property A (a property closed under adding additional edges) has a threshold function: a $p^*(n)$ so that for $p \gg p^*$, $G(n, p)$ has property A with probability tending to 1, and for $p \ll p^*$, $G(n, p)$ has property A with probability tending to 0. Subsequently, Friedgut and Kalai [11] showed that every monotone property has a threshold width bounded by $O(\log^{-1} n)$: there is a function $C(\epsilon)$ so that for any $\epsilon > 0$, if $G(n, p)$ has property A with probability ϵ, then $G(n, p + C(\epsilon)/\log n)$ has property A with probability at least $1 - \epsilon$. Bourgain and Kalai [7] improved this upper bound to $O(\log^{\delta-2} n)$ for any $\delta > 0$. Nevertheless, these theorems do not imply a sharp threshold in the sense defined above unless the critical probability for the property is sufficiently high.

Friedgut [10] gave a characterization of all monotone properties of random graphs that exhibit a sharp threshold: essentially they are properties that cannot be approximated by the property of containing a subgraph from a list of constant-size subgraphs. In other words, properties with coarse thresholds are all similar to the property of containing a triangle.

17th Int'l Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'14) /
18th Int'l Workshop on Randomization and Computation (RANDOM'14).
Editors: Klaus Jansen, José Rolim, Nikhil Devanur, and Cristopher Moore; pp. 500–514
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Friedgut used his general theorem to prove that the satisfiability of a random k-SAT formula exhibits a sharp threshold, and then Achlioptas and Friedgut [1] used it to prove that the property of being k-colorable has a sharp threshold in $G(n,p)$. These properties had resisted previous analysis in part because of their complexity: determining the satisfiability of a k-SAT formula or the k-colorability of a graph are both NP-hard problems. In contrast, 2-SAT has a polynomial-time algorithm in the worst-case, and the threshold location [8] and width [5] of a random 2-SAT formula are both well understood.

In this paper we prove a general sharp threshold theorem for the *Random Geometric Graph* (RGG). The standard model of the RGG, $G_d(n,r)$, involves placing n points uniformly at random (or according to a Poisson process of intensity n) in $[0,1]^d$ (or the d-dimensional unit torus) and joining any two points at distance at most r by an edge. Unlike the edges in $G(n,p)$, the edges in the RGG are not independent. The RGG exhibits thresholds for some of the same properties as the Erdős-Rényi random graph. There is a unique giant component whose appearance occurs sharply at the threshold radius $r = \lambda_c n^{-1/d}$ [19]. The exact value of the constant λ_c is not known, but numerical simulations for $d = 2$ indicate $\lambda_c \approx 1.44$ [21] and bounds are given in [16, 14]. The RGG also has a sharp threshold for connectivity at $r = (\log n/(nV_d))^{1/d}$ [13, 18] (in the d-dimensional torus) where V_d is the volume of a unit ball in $\mathbb{R}^d$.

The RGG has been extensively studied in fields such as cluster analysis, statistical physics, hypothesis testing, and wireless sensor networks. One further application of the RGG is modeling data in a high-dimensional space, where the coordinates of the nodes of the RGG represent the attributes of the data. The metric imposed by the RGG then depicts the similarity between data elements in the high-dimensional space. See [4] or [19] for a survey of results on the RGG.

In the RGG, Goel et al. have shown that every monotone property has a threshold width (in terms of r) of $O(\log^{3/4} n/\sqrt{n})$ (for $d = 2$) and $O(\log^{1/d} n/n^{1/d})$ (for $d \geq 3$) [12]. This implies a sharp threshold in the sense described above when the critical radius of a property is sufficiently large, but not for sparser graphs, and in particular not in the connectivity or giant component regimes. For one-dimensional RGG's, McColm proved that every monotone property has a threshold function [15], in the sense of Bollobás-Thomason.

We prove a general criteria for sharp thresholds in the Poisson RGG. As an application, we introduce a geometric model of random k-SAT in which literals are placed at random in $[0,1]^d$, and prove that satisfiability exhibits a sharp threshold in this model. We also identify the location of this threshold in the case $k = 2$. Previously, a model of random k-SAT for $k = 1, 2$ with literals placed on a 2-dimensional lattice was proposed in [20], and in [17] the authors investigate a model of random k-XOR-SAT with finite interaction range, a kind of one-dimensional geometry.

The organization and main contributions of this paper are as follows:

1. In Section 2, we introduce notation, define two models of RGG's, and define a sharp threshold in each model. We then define analogous models of random geometric k-SAT.
2. In Section 3 we state our main result: a characterization of vertex-monotone properties with sharp thresholds in the Poisson RGG. We also state a result on transferring sharp thresholds from the Poisson to fixed-n model.
3. In Section 4 we state our results on random geometric k-SAT: for all $k \geq 2$, the satisfiability phase transition is sharp in the Poisson model. For $k = 2$, we find the location of this threshold.
4. Section 5 contains the proofs of the sharp threshold lemma and the sharpness of the satisfiability phase transition.
5. Sections 6–8 contain auxiliary results and proofs.

2 Models and Notation

We will denote point sets in $[0,1]^d$ by S, T and the graphs, hypergraphs or formulae formed by joining 2 (or k) points that appear in a ball of diameter r by G_S, G_T, F_S, F_T respectively.

We denote (hyper)graph properties by A and write $G \in A$ if graph G has property A. We say a property A holds 'with high probability' or 'whp' if $\Pr[G \in A] = 1 - o(1)$ as $n \to \infty$. We write $f(n) \sim g(n)$ if $f(n) = g(n)(1 + o(1))$.

We work with two models of random geometric graphs. For $n, d \in \mathbb{N}$ and $\mu, r \in \mathbb{R}^+$, $G_d(n, \mu, r)$ is the random graph formed by drawing a point set S according to a Poisson point process of intensity $n \cdot \mu$ on $[0,1]^d$ and then forming G_S by joining any two points at distance at most r. For the hypergraph version of this model, we form a k-uniform hyperedge on any set of k points in S that appear in a ball of diameter r. If $t > k$ points all appear in one ball of diameter r, then all $\binom{t}{k}$ possible k-uniform hyperedges are formed. The second model, $G_d(n, r)$, is the random graph drawn by placing n points uniformly and independently at random in $[0,1]^d$ to form S, then forming G_S by connecting points at distance at most r. Note that $G_d(n, r)$ has the same distribution as $G_d(n, \mu, r)$ conditioned on $|S| = n$. We use balls of diameter r instead of radius r to form the hypergraphs so as to match the definition of the RGG in the case $k = 2$.

We say a property A has *sharp threshold* in $G_d(n, \mu, r)$ if there exists a function $\mu^*(n), r(n)$ so that for any $\epsilon > 0$,

1. For $\mu > (1 + \epsilon)\mu^*$, $\Pr[G_d(n, \mu, r) \in A] = 1 - o(1)$.
2. For $\mu < (1 - \epsilon)\mu^*$, $\Pr[G_d(n, \mu, r) \in A] = o(1)$.

For $G_d(n, r)$ it is more convenient to describe a sharp threshold in terms of the probability that two random points in $[0,1]^d$ form an edge[1]. We write $r(p)$ for the radius that achieves edge probability p. With this definition, we say that a property A has *sharp threshold* in $G_d(n, r)$ if there exists a function $p^*(n)$ so that for any $\epsilon > 0$,

1. For $p > (1 + \epsilon)p^*$, $\Pr[G_d(n, r(p)) \in A] = 1 - o(1)$.
2. For $p < (1 - \epsilon)p^*$, $\Pr[G_d(n, r(p)) \in A] = o(1)$.

For the k-SAT problem, we will work with formulae on n boolean variables $x_1, \ldots, x_n$. A *literal* is a variable x_i or its negation $\overline{x}_i$. We say a formula $F \in SAT$ if F is satisfiable.

We define two random geometric distributions over k-SAT formulae, $F_k(n, \gamma)$ and $F_k(n, \mu)$:

1. $F_k(n, \gamma)$: Randomly place $2n$ points uniformly and independently in $[0,1]^d$ each labeled with the name of a unique literal in $\{x_1, \ldots, x_n, \overline{x}_1, \ldots, \overline{x}_n\}$. For any set of k literals that appear in a ball of diameter $r = \gamma n^{-1/d}$, form the corresponding k-clause and add it to the random formula.
2. $F_k(n, \mu)$: Draw independent Poisson point processes of intensity μ on $[0,1]^d$ for each of the $2n$ literals. For any set of k literals that appear in a ball of diameter $r = n^{-1/d}$, add the corresponding clause.

Note that $F_k(n, \gamma)$ with $\gamma = 1$ has the same distribution as $F_k(n, \mu)$ conditioned on each literal appearing exactly once.

In this work, we will consider k, γ, μ and d fixed with respect to n, and take asymptotics as $n \to \infty$. We use ℓ_∞ balls for simplicity in what follows, but all results hold for Euclidean balls as well, with constants involving the volume of the d-dimensional unit sphere.

---

[1] For constant dimension d, this definition is equivalent to asking for a critical threshold radius r^*, but for $d = d(n) \to \infty$, allowing r to increase by a factor $(1 + \epsilon)$ will cause a super-constant factor increase in the number of edges of the graph.

Another natural model to consider would be the following, call it $\tilde{F}(n,r)$: randomly place n points uniformly and independently in $[0,1]^d$, each labeled with the name of a variable $x_1, \ldots x_n$ (instead of the name of a literal). Then for each set of k variables appearing in a ball of diameter r, add a k-clause with the signs of the k variables chosen uniformly and independently from the 2^k possible choices. The threshold behavior of satisfiability in $\tilde{F}(n,r)$ is simpler than in the other two models: the threshold is coarse, and determined locally by large cliques of variables (see Section 7).

3 Sharp Thresholds in Random Geometric Graphs

The following theorem characterizes vertex-monotone properties with sharp thresholds in the Poissonized random geometric graph $G_d(n,\mu,r)$. It is an application of Bourgain's theorem in the appendix of Friedgut's paper on sharp thresholds in random graphs [10].

▶ **Theorem 1.** *Let A be a vertex-monotone property of a k-uniform hypergraph that does not have a sharp threshold in $G_d(n,\mu,r)$. Then there exists constants $\epsilon,\delta,K>0$ independent of n so that for arbitrarily large n there is an $\alpha\in(\delta,1-\delta)$ so that either*

1. $\Pr_{G_d(n,\mu,r)}[\exists H\subseteq S : |H|\le K, G_H\in A]\ge\epsilon$,
   or
2. *There exists a point set T in $[0,1]^d$ with $|T|\le K$, $G_T\notin A$ so that*

$$\Pr[G_d(n,\mu,r)\in A|T\subseteq S]\ge\alpha+\epsilon\,.$$

with μ chosen so that $\Pr_{G_d(n,\mu,r)}[A]=\alpha$.

In other words, if a property does not have a sharp threshold, then either there is a constant probability that a constant-size witness of A exists in the RGG or there is a point set of constant size in $[0,1]^d$ that by itself does not have property A, but by conditioning on the presence of these points significantly raises the probability of A in the RGG. To prove that a property has a sharp threshold, we rule out both of these possibilities.

We can connect sharp thresholds in $G_d(n,\mu,r)$ with those in $G_d(n,r)$. In particular, if the threshold intensity $\mu^*(n)$ has a limit, then there is a sharp threshold edge probability p^* in $G_d(n,r)$, and it too is uniform in n, up to a technical condition on the form of the threshold density.

▶ **Proposition 2.** *Let $q(n)=a\log^b(n)n^{-c}$ be a decreasing function of n for constants a,b,c. Suppose a vertex and edge-monotone property A has a uniform sharp threshold in $G_d(n,\mu,r)$: there exists a constant μ^*, independent of n, so that*

1. *For $\mu>(1+\epsilon)\mu^*$, $\Pr[G_d(n,\mu,r(q(n)))\in A]=1-o(1)$*
2. *For $\mu<(1-\epsilon)\mu^*$, $\Pr[G_d(n,\mu,r(q(n)))\in A]=o(1)$,*

then A has a sharp threshold in $G_d(n,r)$ in a uniform sense: there exists a t^*, independent of n, so that

1. *For $p>(1+\epsilon)t^*q(n)$, $\Pr[G_d(n,r(p))\in A]=1-o(1)$.*
2. *For $p<(1-\epsilon)t^*q(n)$, $\Pr[G_d(n,r(p))\in A]=o(1)$.*

Here we think of $q(n)$ as a typical threshold function, for example: $1/n, c/n^2, \log^2 n/n$, etc. The technical condition on q is required to rule out properties whose definition depends non-uniformly on n, eg. for small n, A is the property of containing a triangle, while for large n, it is the property of containing an edge.

We conjecture that in fact all edge-monotone properties in $G_d(n,r)$ can be characterized similarly:

▶ **Conjecture 3.** *For every edge-monotone property A with a coarse threshold in $G_d(n, r(p))$ with respect to p, there are constants $K, \epsilon, \delta > 0$ so that for large n, $\alpha \in (\delta, 1-\delta)$, and p chosen so that $\Pr[G_d(n, r(p)) \in A] = \alpha$, either*

1. $\Pr_{G_d(n,r(p))}[\exists H \subseteq S : |H| \leq K, G_H \in A] \geq \epsilon$,
   or
2. *There exists a point set T in $[0,1]^d$ with $|T| \leq K$, $G_T \notin A$ so that*

$$\Pr[G_d(n, r(p)) \in A | T \subseteq S] \geq \alpha + \epsilon .$$

4 Random Geometric k-SAT

As an application of Theorem 1, we prove that in the $F_k(n, \mu)$ model, the threshold for satisfiability is sharp:

▶ **Theorem 4.** *For all k, there exists a function $\mu_k^*(n)$ so that for every $\epsilon > 0$,*

1. *For $\mu < \mu_k^*(n) - \epsilon$, $F_k(n, \mu) \in SAT$ whp.*
2. *For $\mu > \mu_k^*(n) + \epsilon$, $F_k(n, \mu) \notin SAT$ whp.*

Next, for $k = 2$ we determine the exact location of the satisfiability threshold in both models:

▶ **Theorem 5.** *For any $\epsilon > 0$,*

1. *If $\gamma < 2^{-(1+1/d)} - \epsilon$, then whp $F_2(n, \gamma) \in SAT$. If $\gamma > 2^{-(1+1/d)} + \epsilon$, then whp $F_2(n, \gamma) \notin SAT$.*
2. *If $\mu < 2^{-(d+1)/2} - \epsilon$, then whp $F_2(n, \mu) \in SAT$. If $\mu > 2^{-(d+1)/2} + \epsilon$, then whp $F_2(n, \mu) \notin SAT$.*

Note that from Proposition 8 in Section 6, both thresholds occur at $m = n$ clauses, matching the threshold for random 2-SAT. The proof of Theorem 5 is omitted in this extended abstract, but appears in the full version of the paper.

5 Proofs

5.1 Proof of Theorem 1

To prove Theorem 1, we discretize $[0,1]^d$ and place points independently at each gridpoint with a given probability. We apply Bourgain's theorem in a dual fashion, to the product space over positioned points instead of the product space of edges as in $G(n, p)$. We then show that with a fine enough discretization, the graph formed in the discrete model is identical to the graph formed in the Poisson model with high probability.

We will prove the theorem for labeled k-uniform hypergraphs, where the label set is $\{1, 2, \ldots, L(n)\}$ and the dimension $d = d(n)$ may be constant or tend to infinity with n. Points with label i will appear in $[0,1]^d$ according to a Poisson point process of intensity $n\mu/L$, with all labels appearing independently (thus the union of all labeled points is itself a Poisson point process of intensity $n\mu$). For a random geometric graph we can specialize to $k = 2$ with a single label. For random geometric k-SAT, the label set will have size $2n$, one label for each literal.

Place N^d grid points onto $[0,1]^d$ where $N = 16^d n^3$ so that gridpoint $(i_1, \ldots, i_d)$ is located at $((i_1 - 1/2)/N, \ldots (i_d - 1/2)/N)$ and each i_j ranges over $\{1, \ldots N\}$. To that gridpoint, assign the region $A_{i_1,\ldots,i_d} = ((i_1 - 1)/N, i/N] \times \cdots \times ((i_d - 1)/N, i_d/N]$. At each grid point, let each of the L possible labels appear independently with probability $p = \mu n / L N^d$ (more than one label can appear at a single grid point). For every set of k labeled points that appear in a ball of diameter r (in l_2 or l_∞ distance, depending on the model), include the corresponding hyperedge in the hypergraph. The following proposition allows us to transfer results from the discrete model to the continuous model:

▶ **Proposition 6.** *There is a coupling of the discrete and continuous model so that with probability* $1 - o(1)$*, the labeled hypergraph generated by each is identical.*

Proof. We couple as follows: If at least one point with label l falls in the region $A_{i_1,\dots,i_d}$ in the continuous model, let the label l be present on gridpoint $(i_1, \dots, i_d)$ in the discrete model. If no point with label l falls in $A_{i_1,\dots,i_d}$ in the continuous model, then flip an independent coin that is heads with probability

$$e^{\mu n/LN^d} \cdot (\mu n/LN^d - (1 - e^{-\mu n/LN^d})).$$

If the coin is heads, let l be present at $(i_1, \dots, i_d)$.

The following facts suffice to prove the proposition:

- The coupling is faithful: the probability that gridpoint (i, j) has a point with label l is:

  $$1 - e^{-\mu n/LN^d} + e^{-\mu n/LN^d} \cdot e^{\mu n/LN^d} \cdot (\mu n/LN^d - (1 - e^{-\mu n/LN^d})) = \mu n/LN^d$$

  and all gridpoints and literals are independent by construction.
- With probability $1 - o(1)$ no coins come up heads: i.e. no extra labeled points appear in the discrete model. The probability of heads for a single coin is $O((\mu n/LN^d)^2)$, and there are at most LN^d coins flipped. By the union bound whp no heads are flipped.
- With probability $1 - o(1)$ no two copies of any one label appear in the same $A_{i_1,\dots,i_d}$. The probability that label l appears at least twice in a fixed $A_{i_1,\dots,i_d}$ is $O((\mu n/LN^d)^2)$. There are N^d such boxes and L labels, so again whp no region contains more than one.
- With probability $1-o(1)$ no hyperedges disappear and no new hyperedges appear, moving from the continuous to the discrete model. In the coupling a point moves by at most $1/2N$ in each coordinate. For l_1, l_2, l_∞ norms this means the point moves at most $d/2N$ with respect to the norm. For a hyperedge to appear or disappear due to this movement, two points would need to begin at some distance $x \in [r - d/N, r + d/N]$. For a given pair of points uniformly distributed in $[0, 1]^d$, this occurs with probability that depends on the norm, but is bounded by $4^{d+1}dr/N$. Since the total number of points has a Poisson$(n\mu)$ distribution, we can condition, and whp have at most $2n\mu$ points. Taking the union bound over $\Theta(n^2)$ pairs of points gives a failure probability of $O(n^2 4^{d+1} dr/N) = o(1)$, from our choice of N and using the fact that $r \le d$ and $d^2 \le 4^d$.

◀

To complete the proof of Theorem 1, we apply the following theorem from Bourgain's appendix to Friedgut's work [10][2]. Bourgain's theorem gives a criteria for a monotone property on a product measure over the Hamming cube to have a sharp threshold, as opposed to Friedgut's result which applies only to random graphs and hypergraphs.

Consider a random subset $S \subseteq [N]$ with $i \in S$ with probability p, independently for all $1 \le i \le N$. Let A be a monotone property of subsets of $[N]$. (In the case of the random graph $G(n, p)$, $N = \binom{n}{2}$ and S is the set of present edges, A might be the property of having a triangle or connectedness.)

[2] For a recent explication of Bourgain's proof, see [3].

▶ **Theorem** (Bourgain [10]). *Assume that* $\Pr_p[A] = \alpha \in (0,1)$, $p \cdot d\Pr_p(A)/dp \leq C$ *and* $p = o(1)$. *Then there exists* $\delta(C,\alpha) > 0$ *so that either*

1. *the probability that* S *contains a subset* H *of constant size with* $H \in A$ *is greater than* δ*, or*
2. *there exists a constant-sized subset (e.g. a subgraph in* $G(n,p)$*)* $H \notin A$ *so that* $\Pr_p[Q|H \subseteq S] > \alpha + \delta$.

In other words, conditioning on the appearance of this constant-sized subset increases the probability of the property significantly. We apply this theorem directly to the discrete model above, with the product space $\{0,1\}^{LN^d}$ and $p = \mu n/LN^d$. A vertex-monotone property on random geometric graphs becomes a monotone property in this hypercube. Bourgain's theorem is applied as follows: if a property A does not have a sharp threshold, then by the mean value theorem there must be some μ so that $\Pr_\mu(A)$ is bounded away from 0 and 1, and $\mu \cdot d\Pr_\mu(A)/d\mu \leq C$, for some constant C. Then Bourgain's theorem asserts that either condition (1) or (2) must hold. The two conditions are equivalent in the discrete and continuous model since the graphs generated are identical with probability $1 - o(1)$.

5.2 Proof of Proposition 2

Let $t^* = (\mu^*)^c$. Fix $\epsilon > 0$.

First assume $p > (1+\epsilon)t^*q(n)$, and let $N = \frac{1}{\mu^*}(1+\epsilon/2)^{-c}n$. The conditions of Proposition 2 say that $\Pr[G_d(N, \mu^*(1+\epsilon/2)^{c/2}, r(q(N)) \in A] = 1 - o(1)$. From the concentration of a Poisson, with probability $1-o(1)$, the number of points drawn in $G_d(N, \mu^*(1+\epsilon/2)^{c/2}, r(q(N)))$ is bounded above by n. We also have

$$\begin{aligned} p &> (1+\epsilon)(\mu^*)^c q(n) \\ &= (1+\epsilon)(\mu^*)^c \frac{a\log^b n}{n^c} \\ &\geq \frac{a(1+\epsilon/2)\log^b(n/(\mu^*(1+\epsilon/2)^{-c}))}{(n/\mu^*)^c} \\ &= q(N) \end{aligned}$$

Since A is both vertex monotone and edge monotone, we have $\Pr[G_d(n, r(p)) \in A] = 1 - o(1)$.

Next assume $p < (1-\epsilon)t^*q(n)$, and let $N = \frac{1}{\mu^*}(1-\epsilon/2)^{-c}n$. The conditions say that $\Pr[G_d(N, \mu^*(1-\epsilon/2)^{c/2}, r(q(N)) \in A] = o(1)$. With probability $1 - o(1)$, the number of points drawn in $G_d(N, \mu^*(1-\epsilon/2)^{c/2}, r(q(N))$ is bounded below by n, and

$$\begin{aligned} p &< (1-\epsilon)(\mu^*)^c q(n) \\ &= (1-\epsilon)(\mu^*)^c \frac{a\log^b n}{n^c} \\ &\leq \frac{a(1-\epsilon/2)\log^b(n/(\mu^*(1-\epsilon/2)^{-c}))}{(n/\mu^*)^c} \\ &= q(N) \end{aligned}$$

And again since A is both vertex monotone and edge monotone, $\Pr[G_d(n, r(p)) \in A] = o(1)$.

5.3 k-SAT proofs

Proof of Theorem 4

To prove Theorem 4, we will assume that the threshold is coarse: i.e., there is some $\alpha \in (0,1)$ so that $\Pr_\mu(\text{UNSAT}) = \alpha$, for which $\mu \cdot d\Pr_\mu(\text{UNSAT})/d\mu \le C$. It then suffices to rule out both possibilities in Theorem 1 to derive a contradiction. We will show: (1) whp there is no constant-sized set of positioned literals that is by itself unsatisfiable and (2) there is no constant-sized satisfiable 'booster', one that boosts the unsatisfiability probability from α to $\alpha + \epsilon$ when conditioned on. Using Proposition 10 (Section 8) we can assume that μ is a constant bounded from above and away from 0 independent of n.

Notation

We will denote by F_H the k-SAT formula generated by a set of positioned literals $H \subset [0,1]^d$. Let $G_\mu \subset [0,1]^d$ be a random set of positioned literals chosen according to $2n$ independent Poisson processes of intensity μ, one for each of the $2n$ literals: i.e. $F_k(n,\mu)$ has the distribution F_{G_μ}. We will use l_∞ distance to simplify calculations, but everything holds for l_2 or l_1 distance as well, with α_d, the volume of the d-dimensional unit ball replacing 2^d in the calculations below.

Condition 1: For any constant R, we show that whp there is no set of R positioned literals that form an unsatisfiable formula. We will use the *implication graph* of a 2-SAT formula: the directed graph on $2n$ vertices, each representing a literal in the formula, in which $l_1 \to l_2$ if the clause $(l_2 \vee \bar{l}_1)$ is in the formula. A *bicycle* (see eg. [8, 9]) of length L in a 2-SAT formula is a sequence of clauses

$$(u, w_1), (\overline{w}_1, w_2), (\overline{w}_2, w_3), \ldots, (\overline{w}_L, v)$$

where the w_i's are literals of distinct variables and $u, v \in \{w_1, \ldots, w_L\} \cup \{\overline{w}_1, \ldots, \overline{w}_L\}$. A 2-SAT formula is satisfiable if it does not contain a bicycle. Let Y_L be the number of bicycles of length L in F_{G_μ}. Then

$$\mathbb{E}Y_L \le n^L 2^L (2L)^2 \Pr\left[(\overline{u}, w_1), (w_L, v) \in F_{G_\mu} \wedge \bigwedge_{i=1}^{L-1} (\overline{w}_i, w_{i+1}) \in F_{G_\mu}\right]. \tag{1}$$

▶ **Claim 7.** *The probability that a specified bicycle of length L appears in F_{G_μ} satisfies:*

$$\Pr\left[(\overline{u}, w_1), (w_L, v) \in F_{G_\mu} \wedge \bigwedge_{i=1}^{L-1} (\overline{w}_i, w_{i+1}) \in F_{G_\mu}\right] \le \frac{\mu^2 + 3\mu + 1}{\mu^2} \left(\frac{2^d \mu^2}{n}\right)^{L+1},$$

where w_i's are literals of distinct variables and $u, v \in \{w_1, \ldots, w_L\} \cup \{\overline{w}_1, \ldots, \overline{w}_L\}$.

Proof. The literals in the above event are not all distinct, and so the clauses are not all independent. There may be two literals that are repeated as $\overline{u}$ and v, and perhaps $\overline{u} = v$. We consider three different cases for the overlapping clauses:

Case 1: $u \neq v$, $(u, v) \neq (\overline{w}_i, w_{i+1})$ for any i.

Say $u = w_i$ and $v = w_j$, though the argument will be the same if either or both is a negation. For $k \neq i-1$ or $j-1$, the clauses $(\overline{w}_k, w_{k+1})$ are independent of all other clauses in the bicycle. Each has probability of appearing $\sim 2^d \mu^2/n$ for our choice of μ. Now consider the pairs of clauses $\{(u = w_i, w_1), (\overline{w}_{i-1}, w_i)\}$ and $\{(\overline{w}_L, v = w_j), (\overline{w}_{j-1}, w_j)\}$.

The clauses within each pair are not independent, but the pairs are independent of each other. Both pairs are of the form $(l_1, l_2), (l_1, l_3)$ for distinct literals l_1, l_2, l_3. Conditioning on the number of appearances of l_1, we have

$$\begin{aligned}\Pr[(l_1,l_2),(l_1,l_3)\in F] &\sim \sum_{j=0}^{\infty} \frac{e^{-\mu}\mu^j}{j!}\left(1-e^{-2^d\mu j/n}\right)^2 \\ &\sim \sum_{j=0}^{\infty} \frac{e^{-\mu}\mu^j}{j!}\frac{2^{2d}\mu^2 j^2}{n^2} \\ &= \frac{2^{2d}\mu^2}{n^2}\left(\mu+\mu^2\right) . \qquad (2)\end{aligned}$$

All together, with the $L-3$ independent clauses, this gives that a bicycle of this type appears with probability at most

$$\left(\frac{2^{2d}\mu^3(\mu+1)}{n^2}\right)^2\left(\frac{2^d\mu^2}{n}\right)^{L-3} = \frac{(\mu+1)^2}{\mu^2}\left(\frac{2^d\mu^2}{n}\right)^{L+1} .$$

Case 2: $u \neq v$, $(u, v) = (\overline{w}_i, w_{i+1})$ for some i.

For $k \neq i$, the clauses $(\overline{w}_k, w_{k+1})$ are independent of the other clauses in the bicycle. What remains is the triple $\{(u = \overline{w}_i, w_1), (\overline{w}_i, w_{i+1}), (w_L, w_{i+1})\}$. (The argument is the same if $u = w_{i+1}$ and $v = \overline{w}_i$). This triple is of the form $(l_1, l_2), (l_1, l_3), (l_4, l_3)$. We calculate the probability such a triple appears by conditioning on the number of appearances of l_1 and l_3:

$$\Pr[(l_1,l_2),(l_1,l_3),(l_4,l_3)\in F] \sim \sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{e^{-\mu}\mu^j}{j!}\frac{e^{-\mu}\mu^k}{k!}\frac{2^{3d}j^2k^2\mu^2}{n^3}$$

and so

$$\begin{aligned}\Pr[(l_1,l_2),(l_1,l_3),(l_4,l_3)\in F] &\sim \frac{2^{3d}\mu^2}{n^3}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\frac{e^{-\mu}\mu^j}{j!}\frac{e^{-\mu}\mu^k}{k!}j^2k^2 \\ &= \frac{2^{3d}\mu^2}{n^3}(\mu+\mu^2)^2 \\ &= \frac{2^{3d}\mu^4(\mu+1)^2}{n^3} .\end{aligned}$$

Again all together the probability of the particular bicycle appearing is at most

$$\frac{2^{3d}\mu^4(\mu+1)^2}{n^3}\left(\frac{2^d\mu^2}{n}\right)^{L-2} = \frac{(\mu+1)^2}{\mu^2}\left(\frac{2^d\mu^2}{n}\right)^{L+1} .$$

Case 3: $u = v$.

Say $u = v = w_i$. (The same will work for $u = v = \overline{w}_i$). The clauses $(\overline{w}_k, w_{k+1})$ for $k \neq i-1$ are again independent of all other clauses in the bicycle. What remains are the clauses $(u = w_i, w_1), (\overline{w}_{i-1}, w_i), (w_L, v = w_i)$. This is a triple of the form $(l_1, l_2), (l_1, l_3), (l_1, l_4)$ and we calculate its probability by conditioning on the number of appearances of l_1:

$$\begin{aligned}\Pr[(l_1,l_2),(l_1,l_3),(l_1,l_4)\in F] &\sim \sum_{j=0}^{\infty}\frac{e^{-\mu}\mu^j}{j!}\frac{2^{3d}\mu^3j^3}{n^3} \\ &= \frac{2^{3d}\mu^3}{n^3}(\mu^3+3\mu^2+\mu) = \frac{2^{3d}\mu^4(\mu^2+3\mu+1)}{n^3} .\end{aligned}$$

So the probability of such a bicycle is at most

$$\frac{2^{3d}\mu^4(\mu^2+3\mu+1)}{n^3}\left(\frac{\alpha_d\mu^2}{n}\right)^{L-2} = \frac{\mu^2+3\mu+1}{\mu^2}\left(\frac{2^d\mu^2}{n}\right)^{L+1}.$$

The three estimates prove the claim. ◀

Using the claim and summing from $L=1$ to R yields:

$$\sum_{L=1}^{R}\mathbb{E}Y_L \le \sum_{L=1}^{R}(2n)^L(2L)^2\frac{\mu^2+3\mu+1}{\mu^2}\left(\frac{2^d\mu^2}{n}\right)^{L+1} = O(n^{-1})$$

for any μ, R constant with respect to n. So whp there is no bicycle in the implication graph of length $\le R$ and thus no set of R literals that form an unsatisfiable formula.

For $k \ge 3$ consider an arrangement of R literals that yields an unsatisfiable k-SAT formula. The configuration of points would also induce an unsatisfiable 2-SAT formula since for each k-clause, each of the $\binom{k}{2}$ 2-clauses from the same set of literals would be present, and a satisfying assignment to the 2-SAT would also satisfy the k-SAT formula. But whp there is no set of R unsatisfiable 2-SAT literals, and so no set of R unsatisfiable k-SAT literals.

Condition 2: We want to show that there is no constant-sized set of positioned literals H, so that F_H is satisfiable but conditioning on the presence of H raises the probability of unsatisfiability of F_{G_μ} from α to $\alpha+\epsilon$ at the μ for which $\Pr[F_k(n,\mu) \notin SAT] = \alpha$. Assume $|H| \le R$. We will bound the conditional probability

$$\Pr[F_{G_\mu} \notin SAT | H \subseteq G_\mu] = \Pr[F_{G_\mu \cup H} \notin SAT]$$

where the equality follows from the properties of a Poisson point process. In other words, we will create a random formula by first placing the positioned literals in H in the cube, then adding each of the $2n$ literals independently on top according to a Poisson process of intensity μ, then forming the k-SAT formula from the entire set of points. Note that in the probability on the RHS H is a fixed point set, and G_μ a random point set that does not depend on H.

We now bound $\Pr[F_{G_\mu \cup H} \notin SAT]$. Let $\mathcal{X}_H$ be the set of variables of the literals in H. By assumption $|\mathcal{X}_H| \le kR$. First we show that whp the subformula of $F_{G_\mu \cup H}$ consisting of clauses entirely from $\mathcal{X}_H$ is satisfiable. By assumption, F_H is satisfiable so to create an unsatisfiable subformula on $\mathcal{X}_H$ we need the addition of G_μ to add at least one clause with variables entirely in $\mathcal{X}_H$. There are two different ways this could happen - either a clause is created entirely with randomly placed literals, or a clause is created with some literals from H and some random literals.

We bound the expected number of clauses in F_{G_μ} containing only variables from $\mathcal{X}_H$, call this $\mathbb{E}Y_{\mathcal{X}_H,\mu}$, by bounding the number of literals from $\mathcal{X}_H$ appearing within distance $n^{-1/d}$ of each other in G_μ:

$$\mathbb{E}Y_{\mathcal{X}_H,\mu} \le \binom{2kR}{2}\frac{2^d\mu^2}{n} = o(1).$$

Next, we bound the expected number of literals from $\mathcal{X}_H$ placed by G_μ within distance $n^{-1/d}$ of a literal in H. The total volume of the cube within distance $n^{-1/d}$ of H is bounded by $2^dk^2R^2/n$, and so the expected number of literals from $\mathcal{X}_H$ appearing at random in this region is bounded by $2^dk^2R^2(2kR\mu)/n = o(1)$.

The remainder of the proof follows the general plan of Section 5 of [10]. We separate the n variables into two sets $\mathcal{X}_H$ and $\mathcal{X}_H^c$, and we have shown that whp after the addition of G_μ there is an assignment to $\mathcal{X}_H$ that satisfies the subformula of clauses entirely in $\mathcal{X}_H$, call this assignment x_H. We now show that with probability at least $1 - \alpha - \epsilon/2$, we can extend this assignment on $\mathcal{X}_H^c$ to satisfy $F_{G_\mu \cup H}$. The remaining formula consists of two types of clauses: clauses which contain variables from $\mathcal{X}_H$ (overlapping clauses) and clauses that contain only variables from $\mathcal{X}_H^c$ (non-overlapping). With probability at least $1 - \alpha$, the set of non-overlapping clauses in F_{G_p} is satisfiable, from the definition of μ. We will show that adding the overlapping clauses decreases this probability by at most $\epsilon/2$.

Step 1: The overlapping clauses created with the addition of G_μ are dominated (in terms of inducing unsatisfiability) by adding a constant number of independent random unit clauses.

We can assume that F_H is maximal in the sense that it admits exactly one satisfying assignment, x_H. Adding H to G_μ has two effects: it adds the constraint that $\mathcal{X}_H = x_H$ and it may create some new clauses involving positioned literals from H and G_μ. We have shown above that whp these new clauses all contain at least one variable from $\mathcal{X}_H^c$. Consider the following modification of F_{G_μ}: call the set of literals from $\mathcal{X}_H^c$ that fall within distance $n^{-1/d}$ of a literal from $\mathcal{X}_H$ (either in H or in G_μ) L. Note that the literals in L are uniformly random over all literals in $\mathcal{X}_H^c$. Remove the set L from G_μ to form the random point set G_μ^-. Create the formula $F^*_{G_\mu^-}$ by forming k-clauses according to the usual rules for G_μ^-, but add a unit clause (l) for every literal $l \in L$ that was removed from G_μ. Critically the k-clauses of $F^*_{G_\mu^-}$ are independent of the unit clauses of $F^*_{G_\mu^-}$ since they are formed from points from disjoint regions of the cube. Note that if there is a satisfying assignment to $F^*_{G_\mu^-}$, then the same assignment satisfies F_{G_μ}. The inequality goes in the correct way: we progress to a formula which has less probability of being satisfied.

The expected number of literals from G_μ that fall within distance $n^{-1/d}$ of a literal in $\mathcal{X}_H$ is bounded by $2^d/n \cdot (\mu+1)2kR(2n\mu) = 2^{d+2}kR\mu(\mu+1)$, so with probability $1 - \epsilon/4$ the size of L is at most $2^{d+4}kR\mu(\mu+1)/\epsilon$.

Now consider the random formula F' which is formed by sampling a copy of F_{G_μ} and adding to it $2^{d+4}kR\mu(\mu+1)/\epsilon$ independent, uniformly random unit clauses from all $2n$ literals. With probability $1 - o(1)$ this is the same as adding the same number of uniformly random unit clauses chosen from $\mathcal{X}_H^c$, and F_{G_μ} stochastically dominates the k-clauses of $F^*_{G_\mu^-}$ (formed from a Poisson process on a larger region), so $\Pr[F' \in SAT] \le \Pr[F^*_{G_\mu^-} \in SAT] + \epsilon/4 \le \Pr[F_{G_\mu \cup H} \in SAT] + \epsilon/4 + o(1)$.

Step 2: $\Pr[F' \in SAT] \ge \Pr[F_{G_\mu} \wedge C_1 \wedge \cdots \wedge C_{\sqrt{n}} \in SAT]$, where the C_i's are a collection of $\sqrt{n}$ independent, uniformly random k-clauses. This is Lemma 5.7 from [10].

Step 3: $\Pr[F_{G_\mu} \wedge C_1 \wedge \cdots \wedge C_{\sqrt{n}} \in SAT] \ge \Pr[F_{G_\mu \cup G_{\mu_s}} \in SAT]$, where G_{μ_s} is an independent sprinkling of random positioned literals with intensity $\mu_s = n^{-\delta}$ for each of the $2n$ literals.

We will sprinkle literals independently, adding each literal as a Poisson process of intensity μ_s. Split the cube into n disjoint small cubes with side length $n^{-1/d}$. The probability that a single small cube has at least k sprinkled literals is $\sim (2\mu_s)^k/k! = 2^k n^{-k\delta}/k!$. The expected number of boxes with k literals is $\Theta(n^{1-k\delta})$ and whp there are at least $n^{1-2k\delta}$ such boxes. If we pick one k-clause at random from each box that has one, we will get a set of at least $n^{1-2k\delta}$ uniform and independent random k-clauses. Picking $\delta = 1/5k$ suffices.

Step 4: Increasing μ to $\mu' = \mu + \mu_s$ lowers the probability of satisfiability by at most $Cn^{-\delta} = Cn^{-1/5k}$, from the assumption of a coarse threshold (bounded derivative of the probability with respect to μ, $\mu \cdot d \Pr_\mu(\text{UNSAT})/d\mu \le C$).

All together we have:

$$\begin{aligned}\Pr[F_{G_\mu} \in SAT | H \subseteq G_p] &\geq \Pr[F_{G_\mu \cup H} \in SAT]\\ &\geq \Pr[F^*_{G^-_\mu} \in SAT] + o(1)\\ &\geq \Pr[F' \in SAT] - \epsilon/4 + o(1)\\ &\geq \Pr[F_{G_\mu} \wedge C_1 \wedge \cdots \wedge C_{\sqrt{n}} \in SAT] - \epsilon/4 + o(1)\\ &\geq \Pr[F_{G_\mu \cup G_{\mu_s}} \in SAT] - \epsilon/4 + o(1)\\ &\geq \Pr[F_{G_\mu} \in SAT] - Cn^{-\delta} - \epsilon/4 + o(1)\end{aligned}$$

This contradicts condition 2 in Theorem 1, leading to the conclusion that the threshold must in fact be sharp.

6 Clause Density

The clause density in each k-SAT model is as follows:

▶ **Proposition 8.** *The number of clauses in* $F_k(n,\gamma)$ *is* $\frac{2^k\gamma^{d(k-1)}k^d}{k!}n + o(n)$ *whp. The number of clauses in* $F_k(n,\mu)$ *is* $\frac{(2\mu)^k k^d}{k!}n + o(n)$ *whp.*

Proof. Let X be the number of clauses in the random formula. To compute $\mathbb{E}X$, note that the probability that k given points, distributed uniformly at random in $[0,1]^d$ lie in an ℓ_∞-ball of diameter $\gamma n^{-1/d}$ is the probability that the smallest and largest of k independent uniform $[0,1]$ random variables differ by at most $\gamma n^{-1/d}$, raised to the dth power. This probability, p_k, can be computed by conditioning on the position of the smallest value:

$$\begin{aligned}p_k &= \int_0^1 k(1-t)^{k-1} \min\left\{1, \left(\frac{\gamma n^{-1/d}}{1-t}\right)^{k-1}\right\} dt\\ &= k(\gamma n^{-1/d})^{k-1}\int_0^{1-\gamma n^{-1/d}} dt + k\int_{1-\gamma n^{-1/d}}^1 (1-t)^{k-1} dt\\ &= \frac{k\gamma^{k-1}}{n^{(k-1)/d}}\left(1 - \frac{k-1}{k}\gamma n^{-1/d}\right) = \frac{k\gamma^{k-1}}{n^{(k-1)/d}}(1+o(1)) .\end{aligned}$$

So in the $F_k(n,\gamma)$ model,

$$\mathbb{E}X = \binom{2n}{k} p_k^d \sim \frac{2^k\gamma^{d(k-1)}k^d}{k!} n .$$

Standard estimates show that $\mathrm{var}(X) = O(n)$, and so Chebyshev's inequality gives

$$X = \frac{2^k\gamma^{d(k-1)}k^d}{k!} n + o(n)$$

whp.

The result for the $F_k(n,\mu)$ model follows from conditioning on the total number of literals that appear in the cube and applying the result for $F_k(n,\gamma)$. As this number is concentrated around its expectation, $2\mu n$, we have, whp,

$$X = \frac{(2\mu)^k k^d}{k!} n + o(n) .$$

◀

7 A Coarse Threshold for $\tilde{F}(n,r)$

Here we show that the model $\tilde{F}(n,r)$ in which variables are placed in $[0,1]^d$ and signs of clauses drawn uniformly at random has a coarse threshold.

▶ **Proposition 9.** *Let $r = \gamma n^{-\frac{U(k)}{d(U(k)-1)}}$, where $U(k)$ is the minimal number of variables u so that there exists an unsatisfiable k-SAT formula on u variables so that no two clauses share the same set of k variables. Then*

$$\lim_{n\to\infty} \Pr[\tilde{F}(n,r) \in SAT] = g(\gamma)$$

for a function $g(\gamma) \in (0,1)$. Further, $\lim_{\gamma\to 0} g(\gamma) = 1$ and $\lim_{\gamma\to\infty} g(\gamma) = 0$.

Proof. Claim: $U(k) \le (\ln 2)^{1/(k-1)}(2k)^{k/(k-1)}$. In particular, $U(k)$ is finite.

Proof: Let $u \ge (\ln 2)^{1/(k-1)}(2k)^{k/(k-1)}$. Now consider a random formula formed by taking a clause for each of the $\binom{u}{k}$ distinct sets of k variables from the set of variables $x_1, \ldots x_u$, and then assigning signs uniformly at random. The expected number of satisfying assignments is:

$$2^u(1-2^{-k})^{\binom{u}{k}} < 1$$

for our choice of u (using basic estimates). So there exists some unsatisfiable formula on u variables in which each clauses has a distinct set of variables.

Now we show that satisfiability undergoes a coarse threshold at $r = n^{-\frac{U(k)}{d(U(k)-1)}}$. The general idea of the proof is that for $r = \gamma n^{-\frac{U(k)}{d(U(k)-1)}}$, the probability that there is a set of $U(k)$ variables in a ball of diameter r is bounded away from 0 and 1. The probability that each such set forms an unsatisfiable formula is also bounded away from 0 and 1. We then show that for this choice of r, if there is no such set of variables, the formula is satisfiable whp.

For $r = \gamma n^{-\frac{U(k)}{d(U(k)-1)}}$ the expected number of sets of $U(k)$ variables that form an unsatisfiable formula tends to a constant as $n \to \infty$. To see this note that the expected number of sets of $U(k)$ variables that fall in a ball of diameter r is a constant, and that any such set of variables is unsatisfiable with probability at least $2^{-U(k)}$ from the definition of $U(k)$. To see that it is at most a constant, note that the expected number of connected components of $U(k)$ variables is constant. A modification of Theorem 3.4 of [19] shows that the number of such unsatisfiable sets of variables has a Poisson distribution asymptotically. The mean of this Poisson random variable tends to ∞ as $\gamma \to \infty$ and to 0 as $\gamma \to 0$. Finally, if there is no such set, then the formula is satisfiable whp, since whp the RGG for this r consists of connected components of size at most $U(k)$. For a component of size $< U(k)$, there must be a satisfying assignment, by the definition of $U(k)$. ◀

8 Bounds on the Satisfiability Threshold

For $k \ge 3$ we give bounds on the satisfiability threshold, showing in particular that the transition from almost certain satisfiability to almost certain unsatisfiability occurs when the number of clauses is linear in the number of variables:

▶ **Proposition 10.** *For all $k \ge 3$ there exist functions $\overline{\gamma}(k), \underline{\gamma}(k), \overline{\mu}(k), \underline{\mu}(k)$ so that for any $\epsilon > 0$,*

1. *For $\gamma < \underline{\gamma}(k) - \epsilon$, whp $F_k(n,\gamma) \in SAT$. For $\gamma > \overline{\gamma}(k) + \epsilon$, whp $F_k(n,\gamma) \notin SAT$.*
2. *For $\mu < \underline{\mu}(k) - \epsilon$, whp $F_k(n,\mu) \in SAT$. For $\mu > \overline{\mu}(k) + \epsilon$, whp $F_k(n,\mu) \notin SAT$.*

We can take $\underline{\gamma}(k) = 2^{-(1+1/d)}$, $\underline{\mu}(k) = 2^{-(d+1)/2}$, $\overline{\gamma}(k) = (k-1)^{1/d}$, *and* $\overline{\mu}(k) = k + \ln 2$. *In particular, all functions are independent of* n *and so the threshold for satisfiability occurs with a linear number of clauses.*

For $F_k(n, \gamma)$, the lower bound follows from the lower bound in Theorem 5. For the same set of points in the cube, form both the corresponding 2-SAT formula and the k-SAT formula. For each k-clause the 2-SAT formula will include each of the $\binom{k}{2}$ subclauses of length 2. If there is a satisfying assignment to the 2-SAT formula, the same assignment will satisfy the k-SAT formula.

For an upper bound, we will show that the probability that any assignment is satisfying is 0. Fix an assignment σ, and consider the set of n false literals under σ. Set $\gamma > (k-1)^{1/d} + \epsilon$. Tile $[0,1]^d$ by $(\lceil n^{1/d}/\gamma \rceil)^d$ boxes of side length $\gamma n^{1/d}$ (with boxes along the boundary possibly smaller). For large enough n (depending on ϵ), the number of boxes is strictly less than $n/(k-1)$. By the pigeonhole principle there must be a box with at least k points, and so with probability 1, an unsatisfied clause is formed. This is true for any set of n literals, and so with probability 1 there is no satisfying assignment.

For $F_k(n, \mu)$, the lower bound again follows from the $k = 2$ case and Theorem 5. For the upper bound, we bound the expected number of satisfying assignments. There are 2^n possible assignments, so it is enough to show that the probability a given assignment σ is satisfying is at most q^n for some $q < 1/2$ independent of n. Tile $[0,1]^d$ by n boxes of side length $n^{-1/d}$. The probability that there is no k-clause of negative literals under σ is bounded by the probability that none of these boxes contain k negative literals. The nodes in the different boxes are independent, so we need to show that for large enough μ, the probability there are fewer than k negative literals in a single cube of side length $n^{-1/d}$ is strictly less than $1/2$. The number of negative literals in a single such cube has distribution $\mathrm{Poiss}(\mu)$. The median of a Poisson with mean λ is at least $\lambda - \ln 2$, so if we pick $\overline{\mu}(k) > k + \ln 2$, then $\Pr[\mathrm{Poiss}(\mu) < k] < 1/2$ and via a first-moment argument whp $F_k(n, \mu)$ is unsatisfiable.

Acknowledgements. Will Perkins was supported in part by an NSF postdoctoral fellowship. The authors thank Alfredo Hubard for many interesting conversations on this topic and the anonymous referees for several helpful suggestions.

References

1 Dimitris Achlioptas and Ehud Friedgut. A sharp threshold for k-colorability. *Random Structures and Algorithms*, 14(1):63–70, 1999.

2 Noga Alon and Joel H Spencer. *The probabilistic method*, volume 57. Wiley-Interscience, 2004.

3 Deepak Bal. On sharp thresholds of monotone properties: Bourgain's proof revisited. *arXiv preprint arXiv:1302.1162*, 2013.

4 Paul Balister, Béla Bollobás, and Amites Sarkar. Percolation, connectivity, coverage and colouring of random geometric graphs. In *Handbook of Large-Scale Random Networks*, pages 117–142. Springer, 2008.

5 B. Bollobás, C. Borgs, J. T. Chayes, J. H. Kim, and D. B. Wilson. The scaling window of the 2-sat transition. *Random Structures & Algorithms*, 18(3):201–256, 2001.

6 Béla Bollobás and A. G. Thomason. Threshold functions. *Combinatorica*, 7(1):35–38, 1987.

7 Jean Bourgain and Gil Kalai. Threshold intervals under group symmetries. *Convex Geometric Analysis MSRI Publications Volume 34, 1998*, page 59, 1998.

8 V. Chvátal and B. Reed. Mick gets some (the odds are on his side). In *Foundations of Computer Science, 1992. Proceedings., 33rd Annual Symposium on*, pages 620–627. IEEE, 1992.

9 C. Cooper, A. Frieze, and G. B. Sorkin. A note on random 2-sat with prescribed literal degrees. In *Proceedings of the thirteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 316–320. Society for Industrial and Applied Mathematics, 2002.

10 E. Friedgut. Sharp thresholds of graph properties, and the k-sat problem. *Journal of the American Mathematical Society*, 12(4):1017–1054, 1999.

11 E. Friedgut, G. Kalai, et al. Every monotone graph property has a sharp threshold. *Proceedings of the American mathematical Society*, 124(10):2993–3002, 1996.

12 Ashish Goel, Sanatan Rai, and Bhaskar Krishnamachari. Sharp thresholds for monotone properties in random geometric graphs. In *STOC'04: Proceedings of the 36th Annual ACM Symposium on Theory of computing*, pages 580–586, New York, NY, USA, 2004. ACM Press.

13 P. Gupta and P. R. Kumar. Critical power for asymptotic connectivity. In *Proceedings of the 37th IEEE Conference on Decision and Control*, volume 1, pages 1106–1110, 1998.

14 Zhenning Kong and Edmund M Yeh. Analytical lower bounds on the critical density in continuum percolation. In *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops, 2007. WiOpt 2007. 5th International Symposium on*, pages 1–6. IEEE, 2007.

15 Gregory L. McColm. Threshold functions for random graphs on a line segment. *Combinatorics Probability and Computing*, 13(3):373–387, 2004.

16 Ronald Meester and Rahul Roy. *Continuum percolation*. Cambridge tracts in mathematics. Cambridge University Press, Cambridge, New York, 1996.

17 Andrea Montanari and Antoine Sinton. A simple one dimensional glassy kac model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08004, 2007.

18 Mathew D. Penrose. The Longest Edge of the Random Minimal Spanning Tree. *The Annals of Applied Probability*, 7(2):340–361, 1997.

19 Mathew D. Penrose. *Random Geometric Graphs*. Oxford University Press, 2003.

20 J. M. Schwarz and A. Alan Middleton. Percolation of unsatisfiability in finite dimensions. *Physical Review E*, 70(3):035103, 2004.

21 S. Torquato and M. D. Rintoul. Effect of the interface on the properties of composite media. *Physical Reviw Letters*, 75(22):4067–4070, 1996.