Percolation in High Dimensions

Daniel M. Gordon

July 18, 1996

Abstract

Let $p_c(d)$ be the critical probability for percolation in $\mathbf{Z}^d$. In this paper it is shown that $\lim_{d\to\infty} 2d\ p_c(d) = 1$. The proof uses the properties of a random subgraph of an m-ary d-dimensional cube. If each edge in this cube is included with probability $> \frac{1}{2d(1-3/m)}$, then for large d, the cube will have a connected component of size cm^d for some $c > 0$. This generalizes a result of Ajtai, Komlós and Szemerédi.

Keywords: percolation, random graphs

1 Introduction

Consider $\mathbf{Z}^d$ as a graph, with undirected edges from each point $x \in \mathbf{Z}^d$ to each of the points distance one away from x. For any $p \in [0,1]$, we can define the graph $\mathbf{Z}^d(p)$, with the same vertex set and each edge included with probability p. The resulting graph will have either zero or one infinite connected component (see [7]). Let $p_c(d)$ be the critical probability: the smallest number such that an infinite component exists with probability one when $p > p_c(d)$.

The critical probability, particularly of 2-dimensional lattices, has been the subject of much study. For $\mathbf{Z}^1$, the critical probability is obviously 1, and for $\mathbf{Z}^2$ it was shown to be 1/2 by Kesten [10]. In higher dimensions, there are numerical estimates, but no exact results. It has been conjectured that

$$\lim_{d\to\infty} 2d\ p_c(d) = 1. \qquad (1)$$

Using a branching process argument, it is easy to show that $p_c(d) \geq 1/(2d-1)$. In this paper we will show (1), using methods from random graphs.

In [4] Cox and Durrett solve the corresponding problem for $\mathbf{Z}^d$ where the edges are oriented in the direction of increasing coordinates. Their method uses very strongly the fact that this graph has no circuits, which is not true for nonoriented edges.

In Section 2 we explain the connection between large components in d-dimensional cubes and the critical probability of $\mathbf{Z}^d$. A result of Ajtai, Komlós and Szemerédi on the largest random component of a d-cube, implies that $\lim_{d\to\infty} 2d\, p_c(d) \leq 2$. Then we show that, given a generalization of that result to m-ary d-cubes (with m vertices along each edge), we can reduce the constant to 1. In Section 3 two necessary lemmas about m-ary d-cubes are proven, and Section 4 is devoted to proving the generalization.

In the course of the proofs, many constants c_i and c'_i will be used. Their precise values are not important, but they are always positive and small enough in terms of the earlier constants. Their dependence on m, which will not be needed (unless trying to bound the error term; see the note at the end of the paper), is easily derivable. All $o(1)$ terms are understood to be as $d \to \infty$.

The Law of Large Numbers will be used several times. We will only need a weak form of it: for n independent events each happening with probability p, the expected number of events which occur is $kp = E$, and if E is sufficiently large, then

$$\text{Prob(fewer than } E/10 \text{ events occur)} < e^{-E/2}.$$

Kesten, in [11], has recently proved Theorem 5 by very different methods.

2 Large Components and Critical Probabilities

Let C^d be the d-dimensional (binary) cube, with vertices $(x_1, \ldots, x_d)$, with $x_i = 0, 1$ and edges connecting each pair of vertices with Hamming distance one (i.e. they differ in exactly one coordinate). Let $G^d(\lambda)$ be the random subgraph of C^d where each edge is included with probability λ/d. In [1], Ajtai, Komlós and Szemerédi prove the following:

Theorem 1 *In $G^d(\lambda)$, for $\lambda > 1$, there is a unique connected component of size $> c2^d$ with probability $1 - o(1)$. Moreover, all but $o(2^d)$ vertices have a neighbor in the large component.*

The methods used to prove this theorem will be used for the d-dimensional cube with m vertices on a side in Section 4. However, this theorem is sufficient to prove the following:

Theorem 2 $\lim_{d\to\infty} 2d\ p_c(d) \leq 2.$

Proof: Consider $\mathbf{Z}^d$ as a union of d-dimensional cubes

$$C^d_{a_1,\ldots,a_d} = \{(x_1,\ldots,x_d) \mid \ x_i \ = \ 2a_i, 2a_i+1, \quad \text{for } i=1,2,\ldots,d\}$$

for $a_1,\ldots,a_d \in \mathbf{Z}$. We will create a new lattice Λ, whose vertices are the cubes $C^d_{a_1,\ldots,a_d}$, with edges between vertices representing adjacent cubes. Λ is clearly isomorphic to $\mathbf{Z}^d$.

Let $\Lambda(p,p')$ be the subgraph of Λ which includes each vertex with probability p, and each edge between two included vertices with probability p'. Then if p is the probability that a given cube has a large component, and p' is the probability that the large components of two such adjacent cubes are connected by some edge, then $\Lambda(p,p')$ is a representation of our original problem, and if $\Lambda(p,p')$ has an infinite component, so must $\mathbf{Z}^d(\lambda/d)$.

This is a mixed bond and site percolation problem (see [8] for a discussion of such problems). By Theorem 1, $p \ = \ 1-o(1)$ for $\lambda > 1$ and d sufficiently large. If pp' is greater than the critical probability for site percolation on $\mathbf{Z}^d$, then Λ will have an infinite component with probability one [8]. We will prove the theorem by showing that $p' = 1-o(1)$.

For any d sufficiently large and $\lambda > \lambda' > 1$, we will randomize first within all the cubes, forming $G^d(\lambda')$, and then perform a second ramdomization, including new edges with probability ϵ/d for $\epsilon \ = \ \lambda - \lambda' \ > \ 0$. Then we consider the edges connecting the cubes.

By Theorem 1, each cube will have a large component of size $c2^d$ with probability $1-o(1)$. For the probability that the large components in two adjacent cubes are connected, consider any two adjacent vertices, one in each cube. There are 2^{d-1} such pairs of vertices, and except for the small exceptional set mentioned in Theorem 1, each vertex has a neighbor in the large component of its cube. The probability of a vertex v being joined to the large component along the edge between v and a neighbor in that component after the second randomization is at least ϵ/d for any v, independent of any other vertex.

Thus the probability that the large components are connected through a given pair of vertices is:

$$\left(\frac{\epsilon}{d}\right)^2 \frac{\lambda}{d},$$

since the probability of the edge between them being included is λ/d. The probability of the large components being connected through some pair of vertices is:

$$p' \geq 1 - \left(1 - \frac{\epsilon^2 \lambda}{d^3}\right)^{2^{d-1} - o(2^d)} = 1 - o(1).$$

Therefore $\mathbf{Z}^d(\lambda/d)$ will have an infinite component. $\square$

This is the strongest result we can obtain by cutting $\mathbf{Z}^d$ into binary d-cubes, because each vertex in a cube has degree d, so the critical probability needed to form a giant component is too high. If we divide $\mathbf{Z}^d$ into larger cubes, say with m vertices on each side, then most vertices will have degree close to $2d$, and the probability necessary for a giant component is reduced.

Let C_m^d be the d-dimensional m-ary cube, with $n = m^d$ vertices. and $G_m^d(\mu)$ be a random subgraph of C_m^d in which each edge is included with probability $1/(2d(1 - \mu/m))$. In section 3 we will prove:

Theorem *Fix $\mu > 3$ and $m \geq 2$. There exists c_1 depending only on μ and m such that $G_m^d(\mu)$ has a component of size $> c_1 n$ with probability $1 - o(1)$. Moreover, all but $o(n)$ vertices have a neighbor in the large component.*

From this, we may deduce:

Theorem 3 $p_c(d) = \frac{1}{2d}\left(1 + O\left(\frac{1}{m}\right)\right)$.

Proof: For m a fixed integer > 3, we will show that if edges in $\mathbf{Z}^d$ are included with probability $> 1/(2d(1 - 3/m))$, then an infinite component almost certainly exists. By taking larger values of m, this probability may be taken as close to $1/(2d)$ as desired, so this will show (1).

As in the proof of Theorem 2, consider $\mathbf{Z}^d$ as a union of d-dimensional cubes, this time with m vertices on a side:

$$C^d_{a_1,\ldots,a_d} = \{(x_1, \ldots, x_d) | x_i \in \{ma_i, ma_i + 1, \ldots, ma_i + (m-1)\} \text{ for } i = 1, 2, \ldots, d\}$$

For $\mu > \mu' > 3$, we will first randomize each edge within each cube with probability $p > 1/(2d(1 - \mu'/m))$, then perform an additional randomization to increase the probability of each edge by ϵ/d, for

$$\epsilon = \frac{m}{2} \frac{\mu - \mu'}{(m - \mu)(m - \mu')}. \tag{2}$$

Then we will consider the edges connecting the cubes.

By the above theorem, each cube will have a large component of size $c_1 n$ with probability $p = 1 - o(1)$. We did not show that there is only one large component, but uniqueness is unnecessary; if multiple large components exist in a cube, choose one arbitrarily. We will create a new lattice Λ, with vertices are the cubes $C^d_{a_1,\ldots,a_d}$, and edges between vertices representing adjacent cubes. Let $\Lambda(p, p')$ be the random subgraph of Λ where vertices are included with probability p and edges between included vertices with probability p'. As in the proof of Theorem 2, $\Lambda(p, p')$ will have an infinite component with probability one if pp' is sufficiently close to 1.

There are m^{d-1} edges connecting the $(d-1)$-dimensional faces of adjacent cubes, of which all but $o(n)$ are incident on two vertices with at least one neighbor in the large component. The second randomization connects each such vertex v to its neighbor (and so guarantees that v is in the large component) with probability ϵ/d. The edge between them is included with probability $1/(2d(1-\mu/m))$, so the probability that the two large components are connected through this edge is at least:

$$\left(\frac{\epsilon}{d}\right)^2 \frac{1}{2d(1-\mu/m)},$$

independent of any other edge. Therefore the probability that the large components are connected by at least one of these edges is

$$p' \geq 1 - \left(1 - \frac{\epsilon^2}{2d^3(1-\mu/m)}\right)^{m^{d-1}(1-o(1))} = 1 - o(1).$$

Thus, the large components in neighboring cubes are connected with probability $p' = 1 - o(1)$, and so there is an infinite component in $\Lambda(p, p')$, and so in $G^d_m(\mu)$ with probability one. $\square$

3 Properties of the m-ary d-cube

The problem with Theorem 2 is that each vertex in the d-dimensional cube has degree only d, and so half of all edges are between cubes, forcing an unnecessarily large probability to get large components in each cube. By going to cubes with m vertices on each side, more edges in $\mathbf{Z}^d$ are within the cubes, and the average degree of vertices within the cube goes up. The disadvantage of this approach is that the graph is no longer regular: corner vertices still have degree d, and various vertices have each degree between d and $2d$. To use these cubes, we must first show that most vertices have high degree.

To measure distance in m-ary cubes we will use the Lee metric rather than the Hamming metric. For any vectors $(x_1, \ldots, x_d)$ and $(y_1, \ldots, y_d)$ in $\mathbf{Z}^d$, their Lee distance is

$$\sum_{i=1}^{d} |x_i - y_i|$$

This metric is a generalization of the "taxicab metric" in two dimensions. While it is less common than Hamming distance, it has also been studied [3], and many results in Hamming metric error-correcting codes have analogues for the Lee metric.

Let C_m^d be the d-dimensional m-ary cube, with vertices $(x_1, \ldots, x_d)$, $x_i \in \{0, 1, \ldots, m-1\}$, and edges connecting each pair of vertices with Lee distance one (*i.e.* that differ in exactly one coordinate, and those coordinates differ by one). Let $n = m^d$ denote the number of vertices in the cube.

The first result we need is to show that C_m^d's not being regular is unimportant:

Lemma 1 *For any $\lambda > 1$, and $m \geq 2$ there is some $c_0 = c_0(\lambda, m) > 0$ such that for all $d \geq 1$ the number of vertices in C_m^d with degree less than $2d(1 - \lambda/m)$ is at most n^{1-c_0}.*

Proof: Consider any vertex $(x_1, \ldots, x_d)$. There are d coordinates that may be incremented or decremented by one to get to an adjacent vertex. If x_j is 0 or $m-1$, then there is only one choice, otherwise there are two. Therefore, the number of vertices of degree $2d - i$ is:

$$2^i \binom{d}{i} (m-2)^{d-i}$$

The fraction of vertices of degree at most $2d - l$, for any $l = 0, \ldots, d$ is therefore:

$$\sum_{i \geq l} \binom{d}{i} \left(\frac{2}{m}\right)^i \left(\frac{m-2}{m}\right)^{d-i} \tag{3}$$

Using a result of H. Chernoff (Equation (3.7) of [5]) on the tail of binomial distributions, (3) is at most

$$\exp\left((d-l)\log\frac{d(m-2)}{m(d-l)} + l\log\frac{2d}{ml}\right)$$

for $l > 2d/m$. Setting $l = 2d\lambda/m$, this becomes

$$\exp\left(\frac{d}{m}[(m-2\lambda)\log\frac{m-2}{m-2\lambda} \; - \; 2\lambda\log\lambda]\right). \tag{4}$$

But using the fact that

$$z - \frac{z^2}{2} < \log(1+z) < z$$

for $z > 0$, we have

$$[(m-2\lambda)\log(m-2)/(m-2\lambda) - 2\lambda\log\lambda] < (\lambda-1)^2(\lambda-2) < 0,$$

so choosing $c_0' = (\lambda-1)^2(2-\lambda)$ and $c_0 = c_0'/(m\log m)$, the fraction of the vertices with degree less than $2d(1-\lambda/m)$ given in (4) is at most

$$\exp(-\frac{d}{m}c_0') = n^{-c_0}. \qquad \square$$

With this lemma, we can treat C_m^d as a regular graph of slightly reduced degree, with a small number of exceptional vertices. For the rest of this paper, m and λ will be fixed.

We will also need a bound on the minimal size of the boundary of a subset of C_m^d.

Let $S_k^d(v)$ be the sphere of radius k around vertex v in C_m^d, and $B_k^d(v)$ be its boundary, the ball of radius k. The size of this sphere depends on v, but the largest one is for $v = ([m/2], \ldots, [m/2])$ (centered at the middle of the cube), and the smallest one has $v = (0, \ldots, 0)$ (centered at a corner). The generating function for any $B_k^d(v)$ is d dimensions is:

$$\prod_{j=1}^{d}(1 + \delta_{1,j}x + \delta_{2,j}x^2 + \ldots + \delta_{m-1,j}x^{m-1}),$$

where $\delta_{i,j}$ is 0,1 or 2 according to how many of $v_j - i$, $v_j + i$ are between 0 and $m-1$, and the coefficient of x^k in the product is $|B_k^d(v)|$. For instance, the smallest ball has generating function:

$$(1 + x + x^2 + \ldots + x^{m-1})^d, \tag{5}$$

and the largest ball has generating function:

$$(1 + 2x + 2x^2 + \ldots + 2x^{[m/2]})^d, \tag{6}$$

for m odd.

Lemma 2 *For every $b_1 > 0$ there is a $b_2 > 0$ such that every vertex-set of C_m^d of size between $b_1 n/d^{1/4}$ and $n - b_1 n/d^{1/4}$ has a boundary of size at least $b_2 n/d^{3/4}$.*

Proof: Lemma 2 is an isoparametric inequality for C_m^d, referred to as Property C in [1]. For $m = 2$, the case considered there, this is equivalent to the theorem that the Hamming sphere is the region of smallest boundary in the d-dimensional cube. This was first shown by Harper [9].

The corresponding theorem for general m was shown by Moghadam in [12]:

Theorem 4 *For C_m^d, the region with smallest boundary is the Lee sphere around $\bar{\mathbf{0}} = (0, \ldots, 0)$.*

Lemma 2 follows from this result. The boundary of the $S_k^d(\bar{\mathbf{0}})$ is $B_k^d(\bar{\mathbf{0}})$, which has a generating function given by (5). Dividing each term by m gives:

$$\left(\frac{1}{m} + \frac{x}{m} + \ldots + \frac{x^{m-1}}{m}\right)^d, \tag{7}$$

the fraction of the cube in $B_k^d(\bar{\mathbf{0}})$. But this is also the generating function for a multinomial distribution, which for large d tends to a normal distribution, with expectation $(m-1)d/2$, and variance $d(m^2-1)/12$ (see for example [6]).

From this we can find the size of the boundary of any Lee sphere. For a sphere with radius $r = (m-1)d/2 - f(d)$, the normal approximation to the size of the boundary (the coefficient of x^r in (5)) is

$$n \cdot \sqrt{\frac{6}{(m^2-1)d\pi}} \exp\left\{-\frac{6f(d)^2}{(m^2-1)d}\right\}, \tag{8}$$

and the size of the sphere is

$$n \cdot \sqrt{\frac{(m^2-1)d}{24\pi}} \frac{1}{f(d)} \exp\left\{-\frac{6f(d)^2}{(m^2-1)d}\right\}. \tag{9}$$

For $f(d) = \sqrt{\beta d}$,

$$|S_r^d(\bar{\mathbf{0}})| \approx n \cdot \sqrt{\frac{(m^2-1)}{24\pi\beta}} \exp\left\{-\frac{6\beta}{(m^2-1)}\right\} = \beta_2 n$$

and

$$|B_r^d(\bar{\mathbf{0}})| \approx n \cdot \sqrt{\frac{6}{(m^2-1)d\pi}} \exp\left\{-\frac{6\beta}{(m^2-1)}\right\} = \beta_3 n/\sqrt{d},$$

so a Lee sphere of radius between $(m-1)d/2 - \sqrt{\beta d}$ and $(m-1)d/2 + \sqrt{\beta d}$ has size between $\beta_2 n$ and $(1-\beta_2)n$ and a boundary of size at least $\beta_3 n/\sqrt{d}$. This is the result used in [1] to prove the existence of the large component.

To show further that every vertex has a neighbor in the large component, we will need to use sets of size $o(n)$. If we let $f(d) = \sqrt{bd \log d}$ for some constant $b > 0$, we get a sphere with size

$$n \cdot \sqrt{\frac{(m^2-1)}{24\pi b \log d}} \exp\left\{-\frac{6b \log d}{(m^2-1)}\right\} = n \cdot \sqrt{\frac{(m^2-1)}{24\pi b \log d}} \cdot d^{-6b/(m^2-1)}$$

and boundary

$$n \cdot \sqrt{\frac{6}{(m^2-1)d\pi}} \cdot d^{-6b/(m^2-1)}.$$

Letting $|b| \leq (m^2-1)/24$, we get that a Lee sphere with size between $b_1 n d^{-1/4}/(\log d)^{1/2}$ and $n - (b_1 n d^{-1/4}/(\log d)^{1/2})$ has a boundary of size at least $b_2 n/d^{3/4}$. By Moghadam's result, the same is true for any vertex set with size in this range. □

We will also need an estimate for $|S_r^d(v)|$ when $r = cd$ for some small constant c. The approximation (9) does not hold for $f(d) = O(d)$, but the following crude estimate will be good enough. In a sphere of radius cd around v, at most cd coordinates are different than those of v, so for $k > 0$ and $c = m^{-k}$,

$$|S_{cd}^d(v)| < \binom{d}{cd} m^{cd} < m^{(k+2)cd} = n^{(k+2)c} \tag{10}$$

using Stirling's formula.

4 Existence of a Large Component

Let $G_m^d(\mu)$ be the random subgraph of C_m^d where each edge is included with probability $1/(2d(1-\mu/m))$. It remains to prove the following generalization of Theorem 1:

Theorem 5 *Fix $\mu > 3$ and $m \geq 2$. There exists c_1 depending only on μ and m such that $G_m^d(\mu)$ has a component of size $> c_1 n$ with probability $1 - o(1)$. Moreover, all but $o(n)$ vertices have a neighbor in the large component.*

The proof of Theorem 5 will closely follow the proof of the theorem in [1]. Parts of the proof which are similar will be sketched here, and the differences will be emphasized. All the c_i's will be positive constants taken as small as needed based on the earlier constants and m, but not on d.

As in [1], the proof will follow from a series of "blowing-up" lemmas, which show the existence of progressively larger connected components:

Lemma 3 *Let $\mu > \lambda + 2 > 3$. For all but n^{1-c_0} vertices v of C_m^d, the probability that v is in a component of size $> 2d/m$ in $G_m^d(\mu)$ is at least c_2.*

Define a *cell* to be the set of vertices of a connected subgraph of $G_m^d(\mu)$, *i.e.* a connected subcomponent.

Lemma 4 *In $G_m^d(\mu)$, with probability $1 - o(1)$, all except for at most n^{1-c_3} vertices v have the following property:*

Property 1: There are $c_4 d$ disjoint cells neighboring v, each of size $c_5 d$.

Lemma 5 *In $G_m^d(\mu)$, with probability $1 - o(1)$, all except for at most $n^{1-\bar{c}_3}$ vertices v have the following property:*

Property 2**:** The vertices of any neighboring cell of size $c_5 d$ have Property 1.

Lemma 6 *In $G_m^d(\mu)$, with probability $1 - o(1)$, all except for at most n^{1-c_6} vertices v have the following property:*

Property 3: There are $c_7 d$ neighbors of v belonging to components of size $> c_8 d^2$.

Proof of Lemma 3: This lemma follows from a branching process argument. For a Galton-Watson process where a vertex has i offsets with probability $p_i = \binom{l}{i}\alpha^i(1-\alpha)^{l-i}$, and the expected number of offsets $\sum_{i\geq 0} i p_i = l\alpha \geq 1+\epsilon$, it is known [2] that the probability of the process not terminating is some $q \geq \delta$, where δ depends on ϵ, but not l or α. We will use this result with $l = 2d(1 - (\lambda + 2)/m)$, and $\alpha = 1/(2d(1 - \mu/m))$, so that

$$l\alpha = 1 + \frac{\mu - (\lambda + 2)}{m - \mu} > 1. \tag{11}$$

Start with any vertex v of degree at least $2d(1-\lambda/m)$. By Lemma 1, this excludes only n^{1-c_0} vertices. From v, pick m of its neighbors and randomize the edges leading to them. The probability of connecting v to i neighbors this way is p_i. Call the number of connected neighbors X_1. If $X_1 > 2d/m$, we are done. Otherwise, denote the neighbors by $d_1, \ldots, d_{X_1}$.

Pick l of the neighbors of d_1 (not including v) and randomize the edges leading to them. If still not enough are connected, continue with $d_2, d_3, \ldots, d_{X_1}$. If this is still not enough, go to the next level (offsets of the $d_i' s$), and so on, until $2d/m$ vertices are found or the process terminates.

If we can indeed always pick l neighbors for each vertex, none of which have been used before, then by the Galton-Watson process results there is probability $c_2 > 0$ that the process will not terminate before we have a component of size at least $2d/m$. But the degree of v is at least $2d(1-\lambda/m)$, and the degree of a vertex in C_m^d is at most one less than any of its neighbors. Therefore even if the component is a path, the smallest degree of any vertex in it will be at least $\deg(v) - 2d/m = 2d(1-(\lambda+1)/m)$. At any point before it terminates there are at most $2d/m$ vertices already in the component, so there will be more than l neighboring vertices left. □

It can be shown, using the techniques from the proof of Claim 3 in [2], that $c_2 \gg 1/m$.

Proof of Lemma 4: If $v = (v_1, \ldots, v_d)$, let v^i be a neighbor of v of the form $(v_1, \ldots, v_i \pm 1, \ldots, v_d)$, where the ith coordinate is 1 if $v_i = 0$, is $m - 2$ if $v_i = m - 1$, and is arbitrarily incremented or decremented otherwise. For $1 \leq i \leq c_9 d$, consider the $(d - i)$-dimensional cube gotten by fixing the first i coordinates:

$$C_m^{d-i}(v) = \{(x_1, \ldots, x_d) | x_1 = v_1, \ldots, x_{i-1} = v_{i-1}; x_i = v_i \pm 1\}$$

These cubes are disjoint. Pick c_9 so that

$$(1 - c_9)(1 - (\lambda + 2)/m) > (1 - \mu/m).$$

Then since $d - i \geq (1 - c_9)d$, we can apply Lemma 3 to these smaller cubes and still have probability c_2 of each v^i being in a connected component of size $> c_5 d$ in its cube, where $c_5 \leq 2(1 - c_9)/m$.

By the Law of Large Numbers, the probability that less than $c_4 d$ of the v^i's are in components of this size is less than $e^{-c_{10}d} = n^{-c_{11}}$. Therefore every v of degree at least $2d(1 - \lambda/m)$ has Property 1 with probability $> 1 - n^{-c_{11}}$, and the expected number of vertices not having Property 1 is at most $n^{1-c_{11}}$.

Markov's inequality states that for any random variable X with expectation E and positive number t,

$$\mathrm{Prob}(X > Et) \leq 1/t$$

and so the probability of more than n^{1-c_3} not having Property 1 is at most $n^{c_3 - c_{11}} = o(1)$. $\square$

We will chose c_5 sufficiently small so that $|S^d_{c_5 d}(v)| < n^{c_3/2}$ for any v. This is possible by (10).

Proof of Lemma 5: Let N be the set of vertices not having Property 1. Then the set of vertices not having Property 2 are in N or within Lee distance $c_5 d$ of a vertex in N. But the number of such vertices is at most

$$|N||S^d_{c_5 d}| < n^{1-c_3} n^{c_3/2} = n^{1-c_3/2} = n^{1-\bar{c}_3}. \qquad \square$$

Proof of Lemma 6: Choose a μ' such that $3 < \mu' < \mu$. We will first examine $G^d_m(\mu')$, and then add new edges with probability ϵ/d, with ϵ as in (2).

By Lemma 5, all vertices in $G^d_m(\mu')$, except for a set of size at most $n^{1-\bar{c}'_3}$ which we will again denote by N, have Property 2. For each v not in N let R_v be $c'_4 d$ disjoint cells of size $c'_5 d$ neighboring v such that all their vertices have Property 1. When we perform the second randomization, all of these cells, except for at most n^{1-c_6}, will melt into components of size $> c_8 d^2$, proving Lemma 6.

Define the *parity* of a vertex to be the parity of the sum of its coordinates: $(x_1, \ldots, x_d)$ is even if and only if $\sum_{i=1}^d x_i \equiv 0 \bmod 2$. Let $R \subset C^d_m - N$ be an arbitrary cell of $G^d_m(\mu')$ of size $\geq c'_5 d$, and let $T = \{t_1, t_2, \ldots\}$ be the set of even vertices or odd vertices, whichever is larger, so that $|T| \geq c'_5 d/2$.

t_1 has Property 1, and in the second randomization it gets connected to a random number of its $c'_4 d$ disjoint neighboring cells. The probability that it is connected to at least one is at least

$$1 - (1 - \epsilon/d)^{c'_4 d} > c_{12}.$$

Let A_1 be the union of these newly connected neighboring cells. Repeat the process with t_2. Since the t_i all have the same parity, the edges going out from them are distinct. Of the $c'_4 d$ disjoint cells neighboring t_2, at least $c'_4 d/2$ intersect in less than half with A_1 (otherwise $|A_1| \geq c'_4 c'_5 d^2/2$, and letting $c_8 = c'_4 c'_5/2$, we are done). The probability that t_2 is connected to at least one of them is at least

$$1 - (1 - \epsilon/d)^{c'_4 d/2} > c'_{12}.$$

Denote by A_2 the union of these cells newly connected to t_2, and let $B_2 = A_1 \cup A_2$. Now we pass to t_3, and continue. Let B be the union of all the A_i's. The expected size of B is at least

$$c'_{12}|T|(c'_5 d/2) \geq \frac{(c'_5)^2 c'_{12} d^2}{4} = c_{13} d^2$$

Therefore, by the Law of Large Numbers,

$$\text{Prob}(|B| < c_8 d^2) < e^{-c_{14} d} = n^{-c_{15}}. \qquad (12)$$

Therefore, any of the cells in R_v will melt into a component of size $> c_8 d^2$ with probability $> 1 - n^{-c_{15}}$. There are at most nd of these cells, so the expected number not melting into components of size $> c_8 d^2$ is less than $ndn^{-c_{15}} < n^{1-c_6}$. □

Proof of Theorem 5: As in the proof of Lemma 6, we will start with $G_m^d(\mu')$ for $\mu > \mu' > 3$, and then use a second randomization to get the result. By Lemma 6, all vertices except for a set N of size $n^{1-c'_6}$ have Property 3. Following [1], we will call the components of size $> c'_8 d^2$ in $G_m^d(\mu')$ *atoms*. At least $c'_7 n/2$ points belong to atoms, and all but $n^{1-c'_6}$ vertices neighbor them, by Lemma 6. We will show that almost all of these atoms melt together during the second randomization: that no union of atoms of size $c_{16} n/d^{1/4}$ or more is separated from the rest of the atoms, with probability $1 - o(1)$.

Suppose that this does not happen. Then we can partition the atoms into two disconnected sets, A and B, such that the smaller set has size at least $c_{16}n/d^{1/4}$. There are at most $n/(c_8'd^2)$ atoms, so the number of choices for A is at most $2^{n/(c_8'd^2)}$. If we show that for any A and B, the probability that no edges connect them is at most $e^{-K_1 n/d^2}$, for K_1 large, then the probability of such a partition exists is $o(1)$, and we are done.

Let $\Gamma(A)$ be the set of all vertices of distance ≤ 1 from some point in A, and $\gamma(A) = \Gamma(A) - A$. Define $D = \Gamma(A) \cap \Gamma(B)$, and $F = C_m^d - \Gamma(A) - \Gamma(B)$. We will deal separately with the case where D is large and where D is small.

Suppose D is large, say $|D| > K_2 n/d$. Let $D' = D - N$. Then $D' > (K_2/2)n/d$, and all points $x \in D'$ neighbor A and B, and have $> (c_7'/2)d$ neighbors in one of the sets. Choose $(K_2/4)n/d$ points x of the same parity. Then

$$\text{Prob}(x \text{ connects A and B}) > c_{17}/d.$$

Since these events are independent for different x,

$$\text{Prob}(\text{no } x \in D' \text{ connects A and B}) < (1 - c_{17}/d)^{(K_2/2)n/d} < e^{-K_1 n/d^2}$$

if K_2 is chosen large enough.

In the other case D is small, $|D| < K_2 n/d$. $\Gamma(A)$ satisfies the conditions of Lemma 2, and so has a boundary of size $> c_{18}n/d^{3/4}$. Its edge-boundary is obviously at least as big, so even after removing edges with an endpoint in D or N, we still have at least $(c_{18}/2)n/d^{3/4}$ disjoint edges from $\gamma(A)$ to $\gamma(B)$. For these edges $e = (x, y)$, with $x \in \gamma(A)$ and $y \in \gamma(B)$, there are $c_7'd$ neighbors of x in A and $c_7'd$ neighbors of y in B, and so

$$\text{Prob}(A \text{ and } B \text{ are connected through } (x, y)) > c_{19}/d.$$

Since these events are independent for different e, we have

$$\text{Prob}(A \text{ and B are not connected through any } (x, y)) < (1 - c_{19}/d)^{c_{18}n/d^{3/4}} < e^{-K_1 n/d^2}.$$

This proves that all but at most $c_{16}n/d^{1/4}$ of the vertices in atoms are in one large component. All but at most n^{1-c_6} vertices have at least $c_7 d$ neighbors in atoms, and therefore the number of vertices with no neighbors in this component is at most

$$2d \cdot \frac{c_{16}n}{d^{1/4}} \frac{1}{c_7 d}$$

which is $O(n/d^{1/4}) = o(n)$. $\square$

Note: From Theorem 3, (1) follows by choosing m arbitrarily large. To obtain an estimate for the error term, we need to have m increase as a function of d. The proof breaks down if m grows too fast, but by replacing the constants with functions of m, and estimating the error in the normal approximation in Lemma 2, it can be shown that for $m = O(d^{1/64})$, the lemmas and Theorem 5 are still valid. It follows that

$$p_c(d) = \frac{1}{2d} + O\left(\frac{1}{d^{65/64}}\right).$$

This is better than Kesten's error bound in [11] of $O((\log\log d)^2/(d\log d))$, but still undoubtedly far from the correct value.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, Largest random component of A k-cube, *Combinatorica*, 2 (1982), pp. 1-7.

[2] M. Ajtai, J. Komlós and E. Szemerédi, The longest path in a random graph, *Combinatorica*, 1 (1981), pp. 1-12.

[3] E.R. Berlekamp, *Algebraic Coding Theory*, McGraw-Hill, New York, 1968.

[4] J. Theodore Cox and Richard Durrett, Oriented percolation in dimensions $d \geq 4$: bounds and asymptotic formulas, *Math. Proc. Camb. Phil. Soc.,* 93 (1983), pp. 151-162.

[5] P. Erdös and J. Spencer, *Probabilistic Methods in Combinatorics*, Academic Press, New York, 1974.

[6] W. Feller, *An Introduction to Probability Theory and its Applications, Vol. 1*, John Wiley and Sons, New York, 1959.

[7] A. Gandolfi, G. Grimmett and L. Russo, On the uniqueness of the infinite open cluster in the percolation model, *Comm. Math. Phys.*, 114 (1988), pp. 549-552.

[8] J.M. Hammersley, A generalization of McDiarmid's theorem for mixed Bernoulli percolation, *Math. Proc. Camb. Phil. Soc.*, 88 (1980), pp. 167-170.

[9] L.H. Harper, Optimal numberings and isoperimetric problems on graphs, *J. Comb. Theory*, 1 (1966), pp. 385-393.

[10] H. Kesten, The critical probability of bond percolation on the square lattice equals 1/2, *Comm. Math. Phys.*, 74 (1980), pp. 41-59.

[11] H. Kesten, Asymptotics in high dimensions for percolation, preprint.

[12] H.S. Moghadam, *Compression operators and a solution to the bandwidth problem of the product of n paths*, Ph.D. Thesis, UC Riverside, 1983.

[13] A.N. Shiryayev, *Probability*, Springer-Verlag GTM 95, New York, 1984.

Department of Computer Science
University of Georgia
Athens, GA 30602