Asymptotics in High Dimensions for Percolation

Harry Kesten[1]

Abstract

We prove that the critical probability for bond or site percolation on $\mathbb{Z}^d$ is asymptotically equal to $1/(2d)$ as $d \to \infty$. If the probability of a bond (respectively site) to be occupied is $\gamma/(2d)$ with $\gamma > 1$, then for the bond model the percolation probability converges as $d \to \infty$ to the strictly positive solution $y(\gamma)$ of the equation $y = 1 - \exp(-\gamma y)$. In the site model the percolation probability is asymptotically equal to $\gamma y(\gamma)/(2d)$ under these conditions. An asymptotic independence property for the random field of sites which belong to the infinite cluster is given.

1. Introduction

Broadbent and Hammersley (1957) created the theory of percolation. Largely because of Hammersley's impetus the subject has grown enormously and is at present a very lively research area in probability and statistical mechanics. Since much of my own research has been inspired by John Hammersley it is a pleasure to dedicate an article on percolation to him in this Festschrift.

Recently Aizenman, Bricmont, and Lebowitz (1987) used the behavior of the critical probability of site percolation in high dimension to obtain some interesting properties of the Ising model. For oriented percolation the asymptotic behavior of the critical probability in high dimension was derived by Cox and Durrett (1983). Here we shall consider the asymptotic properties as $d \to \infty$ of (unoriented) bond and site Bernoulli percolation on $\mathbb{Z}^d$. In (Bernoulli) bond percolation the bonds are occupied (respectively vacant) with probability p (respectively $q := 1 - p$) and all bonds are independent. The corresponding product measure on the configurations of bonds is denoted by P_p. $C(x)$ is the (occupied) *cluster* of x; it is the collection of all points which can be reached from x by an occupied path. (An *occupied path* is a path all of whose edges are occupied.) We write

[1]Research supported by the NSF through a grant to Cornell University.

$\theta(p) = \theta(p, \mathbb{Z}^d, \text{bond})$ for the percolation probability:

$$\theta(p) = P_p\{C(0) \text{ is infinite}\}. \tag{1.1}$$

It is known (Aizenman, Kesten, and Newman 1987; Gandolfi, Grimmett, and Russo 1988) that if $\theta(p) > 0$, then there exists w.p.1 a unique infinite cluster. If $\theta(p) > 0$ we say that *percolation occurs*. Broadbent and Hammersley (1957) and Hammersley (1959) proved that there exists a nontrivial critical probability $p_c = p_c(\mathbb{Z}^d, \text{bond})$ which separates the parameter domains where percolation occurs and where it does not occur. In other words, if we set

$$p_c = \sup\{p : \theta(p) = 0\}, \tag{1.2}$$

then

$$0 < p_c < 1,\ \theta(p) = 0 \text{ if } p < p_c \text{ and } \theta(p) > 0 \text{ if } p > p_c.$$

It is believed (but so far only proven when $d = 2$) that $\theta(p_c) = 0$. All of the preceding has its analogue for site percolation; we merely have to replace 'bond' by 'site' everywhere in the above description of the bond model.

The principal result of this paper gives the asymptotic behavior of p_c. The result is not unexpected, since simple results about branching processes tell us that on a tree with all vertices of degree $2d$, percolation occurs if and only if $p > (2d-1)^{-1}$. Theorem 1 says that asymptotically for large d the critical probability for such a tree and for $\mathbb{Z}^d$ are the same in first order; the circuits which exist on $\mathbb{Z}^d$ play only a small role for large d. Gordon (1988) recently also proved that $2d\, p_c(\mathbb{Z}^d, \text{bond}) \to 1$ as $d \to \infty$ by a rather different method.[2]

Theorem 1.

$$\frac{1}{2d-1} \le p_c(\mathbb{Z}^d, \text{bond}) \le p_c(\mathbb{Z}^d, \text{site}) \le \frac{1}{2d} + O\left(\frac{(\log\log d)^2}{d \log d}\right). \tag{1.3}$$

Theorem 1 can be used to show that if one takes $p = \gamma/(2d)$, then the random field of the sites which belong to an infinite cluster behaves for large d like an independent random field (with success probability converging to the $y(\gamma)$ of (1.5)). For site percolation we have a similar result after a simple modification of the statement. Such a modification is necessary for the following trivial reason. In the bond model a site is incident to

[2]Note added in proof: It seems that the forthcoming paper of Hara and Slade (1989) implicitly proves that $p_c(\mathbb{Z}^d, \text{bond}) = (2d)^{-1} + O(d^{-2})$. In addition, it has just come to our attention that asymptotic expansions (in powers of $(2d-1)^{-1}$) for $p_c(\mathbb{Z}^d, \text{site})$ and $p_c(\mathbb{Z}^d, \text{bond})$ were given on a non-rigorous basis in Gaunt, Sykes, and Ruskin (1976) and Gaunt and Ruskin (1978).

$2d$ edges, each of which can potentially connect the site to ∞. As we shall see this leads to a strictly positive limit for $\theta(\gamma/(2d), \mathbb{Z}^d, \text{bond})$ when $\gamma > 1$ is fixed. In the case of site percolation our definitions require the site x to be occupied in order for x to be connected to ∞. Consequently $\theta(\gamma/(2d), \mathbb{Z}^d, \text{site}) \le \gamma/(2d)$. In order to obtain a situation comparable to that of the bond model we should ignore the state of x itself or condition on x being occupied. Theorem 2S shows that this indeed leads to a result for the site model which is almost the same as for the bond model.

We should note that Theorems 2B and 2S (and their proofs) express the generally held belief that in high dimensions the system exhibits 'mean field behavior'. E.g. in the bond model this means that around a fixed site x the number of neighbors of x connected to ∞ is close to its expected value $2d\theta(p)$, irrespective of the states of the edges incident to x itself. Once this is accepted it is easy to derive a consistency relation for $\theta(p)$. This is the so called mean field equation; at $p = \gamma/(2d)$ the limit of this equation as $d \to \infty$ is just (1.5).

We write $|A|$ for the number of vertices in the set A.

THEOREM 2B. *In the bond model, when $\gamma > 1$ is fixed,*

$$\lim_{d\to\infty} \theta\Big(\frac{\gamma}{2d}, \mathbb{Z}^d, \text{bond}\Big) = y(\gamma), \tag{1.4}$$

where $y(\gamma)$ is the unique strictly positive solution of

$$y = 1 - e^{-\gamma y}. \tag{1.5}$$

More generally, for fixed $\gamma > 1$

$$\lim_{d\to\infty} \sup_{A,B} \Big| P_{\gamma/2d}\{\text{all sites in } A \text{ belong to the infinite occupied cluster,} \\ \text{but none of the sites in } B \text{ do}\} - \big(y(\gamma)\big)^{|A|}\big(1-y(\gamma)\big)^{|B|}\Big| = 0. \tag{1.6}$$

The supremum in (1.6) is over all pairs of finite disjoint sets A and B.

THEOREM 2S. *In the site percolation model, when $\gamma > 1$ is fixed,*

$$\lim_{d\to\infty} \frac{2d}{\gamma}\theta\Big(\frac{\gamma}{2d}, \mathbb{Z}^d, \text{site}\Big) = y(\gamma)$$

(with $y(\gamma)$ as in (1.5)). More generally, for fixed $\gamma > 1$,

$$\lim_{d\to\infty} \sup_{A,B} \Big| P_{\gamma/2d}\{\text{all sites in } A \text{ have a neighbor which belongs to} \\ \text{the infinite occupied cluster, but none of the sites in } B \text{ do}\} \\ - \big(y(\gamma)\big)^{|A|}\big(1-y(\gamma)\big)^{|B|}\Big| = 0. \tag{1.7}$$

The supremum in (1.7) is over the same A, B as in (1.6).

Theorem 1 for bond percolation and Theorem 2B are special cases of similar results for a more general cluster model in which bonds are not independent. These so called Fortuin-Kasteleyn models have an extra parameter Q. For integral $Q \geq 1$ these cluster models have a close relationship with the Potts model with Q colors. In Bricmont, Kesten, Lebowitz, and Schonmann (1989) and Kesten and Schonmann (1989), these models are described in more detail and results corresponding to the above results are proved there for integer Q. For $1 \leq Q \leq 2$ one can even obtain the exact parallels to the above results and we shall give these proofs elsewhere (Kesten 1989). The proof of Theorem 1 has to be given for percolation first and that will be done here. However, Theorem 2 is better treated for all $1 \leq Q \leq 2$ at the same time, and its proof will therefore be deferred to Kesten (1989).

Acknowledgement. The author is indebted to R. Schonmann for suggesting Theorems 2B and 2S and part of their proof.

2. Proof of Theorem 1

The first inequality in (1.3) is one of the earliest results in the subject. It was proven by means of a Peierls argument by Broadbent and Hammersley (1957). The second inequality was proven a number of times; see McDiarmid (1980), Hammersley (1961), and Oxley and Welsh (1979). The only novelty of (1.3) is therefore the last inequality and *for the remainder of this section we shall work with site percolation on* $\mathbb{Z}^d$.

As in Cox and Durrett (1983), which dealt with oriented percolation, we shall basically apply Chebyshev's inequality to the number of occupied paths which connect the origin, $\mathbf{0}$, to points at distance $n-1$ from $\mathbf{0}$ (for n large). Unfortunately, for standard percolation there is less independence among such paths than for oriented percolation, and in order to regain some independence we have to restrict ourselves to certain subclasses of paths which we now define. First, a *path* (of *length* m) on $\mathbb{Z}^d$ is a sequence $v_1, \dots, v_m$ of m vertices of $\mathbb{Z}^d$ such that v_i and v_{i+1} are neighbors. We do not insist that all the v_i are distinct; a path is not necessarily self-avoiding. The ith *step* of the path is the vector $s_i := v_i - v_{i-1}$. e_k will denote the kth unit coordinate vector. We now define for positive integers N and n the following class ($\lfloor a \rfloor$ denotes the largest integer a):

$\mathcal{C}(N, n)$ = collection of paths of length $nN - 1$ whose steps s_i satisfy
(a) $s_i \in \{e_k : k > d - \lfloor d/N \rfloor\}$ if $i = jN$ for $j = 1, \dots, n-1$ and
(b) $s_i \in \{\pm e_k : k \leq d - \lfloor d/N \rfloor\}$ if $jN < i < (j+1)N$ for $j = 0, \dots, n-1$
(this also applies to s_1, which we define as v_1). (2.1)

In the sequel we shall make the convention that $v_0 = \mathbf{0}$ and $s_1 = v_1$ for paths in $\mathcal{C}(N,n)$. We note that there are $\lfloor d/N \rfloor$ choices for each of the steps of the form (a) and $2d - 2\lfloor d/N \rfloor$ choices for each of the steps of the form (b). Thus $\#\mathcal{C}(N,n)$, the cardinality of $\mathcal{C}(N,n)$ is

$$(2d - 2\lfloor d/N \rfloor)^{n(N-1)} \lfloor d/N \rfloor^{n-1}. \tag{2.2}$$

Any path $v_1, \dots, v_{nN-1}$ in $\mathcal{C}(N,n)$ starts at a neighbor of $\mathbf{0}$ and

$$\text{for } kN \le i < (k+1)N \text{ the sum of the last } \lfloor d/N \rfloor \text{ components of } v_i \text{ equals } k. \tag{2.3}$$

We shall count paths in $\mathcal{C}(N,n)$, but not just occupied paths. Instead we define a stronger property. We attach to each vertex v of $\mathbb{Z}^d$ a sequence of 0–1 valued random variables $Y_1(v), Y_2(v), \dots$ such that

$$\text{all variables } \{Y_i(v) : i \ge 1, v \in \mathbb{Z}^d\} \text{ are independent} \tag{2.4}$$

and

$$P\{Y_i(v) = 1\} = p \text{ for all } i \text{ and } v. \tag{2.5}$$

If $r = (v_1, \dots, v_{nN-1})$ is a path of length $nN - 1$ then we define

$$\begin{aligned} k(r,v) &= \text{number of } i \ge 1 \text{ with } v_i \text{ equal to } v \\ &= \text{the number of visits by } r \text{ to } v. \end{aligned}$$

We say that the event $A(r)$ occurs if

$$Y_j(v) = 1 \text{ for } j \le k(r,v) \text{ for all } v. \tag{2.6}$$

Thus if we think of $Y_j(v)$ as the Y value sampled at the jth visit to v, then $A(r)$ occurs if and only if the Y sampled at each visit to a vertex by r is $+1$. Consequently

$$P\{A(r)\} = p^{nN-1} \tag{2.7}$$

for all paths r of length $nN - 1$.

In the proof of Lemma 1 it is explained how the event $A(r)$ is related to r being occupied. In any case we shall be interested in the number of paths r for which $A(r)$ occurs. To estimate the variance of this number we introduce some further quantities. For a pair of paths $r = (v_1, \dots, v_{nN-1})$, and $r' = (v'_1, \dots, v'_{nN-1})$, both of length $nN - 1$, we define

$$J(r,r') = \sum_v k(r,v) \wedge k(r',v) \tag{2.8}$$

($a \wedge b$ denotes $\min\{a, b\}$). Thus if r visits v_i at time i, and this is the νth visit to v_i by r, then this visit adds to the count $J(r, r')$ if and only if there exists an index j such that $v'_j = v_i$ and r' visits v_i for the νth time at time j.

Finally we introduce a probability measure on ordered pairs of paths. $\mathbb{P}$ will be the probability measure which picks a pair r,r' from $\mathcal{C}(N, n)$ with all pairs equally likely. Thus, the probability mass assigned to any pair of paths in $\mathcal{C}(N, n)$ is $[\#\mathcal{C}(N, n)]^{-2}$. Actually, at this moment $\mathbb{P}$ depends on nN, but we suppress this dependence in the notation. $\mathbb{E}$ denotes expectation with respect to $\mathbb{P}$.

LEMMA 1. *For any fixed N,*

$$\theta(p, \mathbb{Z}^d, \text{site}) \geq p \limsup_{n \to \infty} [\mathbb{E}\{p^{-J(r,r')}\}]^{-1}. \tag{2.9}$$

PROOF: Choose v occupied if $Y_1(v) = 1$ and vacant if $Y_1(v) = 0$. It is easily seen that under (2.5) the distribution of the occupancy configurations is P_p. Assume now that $A(r)$ occurs for some $r = (v_1, \dots, v_{nN-1}) \in \mathcal{C}(N, n)$. Then by 'loop-removal' we can find an occupied self-avoiding path from v_1 to v_{nN-1}. Loop-removal consists of first finding the last index k such that $v_k = v_1$. We then take out from r the vertices $v_2, \dots, v_k$. We are then left with the path $(v_1, v_{k+1}, \dots, v_{nN-1})$ which visits v_1 only at time 1. Next we find the last index $m \geq k + 1$ for which $v_m = v_{k+1}$ and we remove the vertices $v_{k+2}, \dots, v_m$ to obtain the path $(v_1, v_{k+1}, v_{m+1}, \dots, v_{nN-1})$ which visits each of v_1 and v_{k+1} exactly once. We continue this procedure until no vertex is visited more than once. Let $\overline{r}$ be the self-avoiding path which is left over after this procedure. Its first vertex is v_1 and it is easily seen that the last vertex of $\overline{r}$ must be equal to the endpoint of r, v_{nN-1} (even though v_{nN-1} may be visited several times by r, and in the loop-removal procedure the last vertex of $\overline{r}$ may appear as a v_t which equals v_{nN-1}, but with $t < nN - 1$). By (2.3) with $k = n - 1$ this endpoint of $\overline{r}$ is at least at distance $n - 1$ from $\mathbf{0}$ (the distance here is the l^1 distance, not the Euclidean one). Also $\overline{r}$ must be occupied since all its vertices had corresponding $Y_1 = 1$ if $A(r)$ occurred. Thus $A(r)$ implies that there exists an occupied self-avoiding path of length $n - 1$ starting at a neighbor of the origin. As $n \to \infty$ the probability of the last event converges to

$$P_p\{\text{a neighbor of } \mathbf{0} \text{ is connected to } \infty\} = \frac{1}{p}\, \theta(p, \mathbb{Z}^d, \text{site}).$$

Thus (2.9) will follow if we can prove

$$P\{A(r) \text{ occurs for some } r \in \mathcal{C}(N, n)\} \geq [\mathbb{E}\{p^{-J(r,r')}\}]^{-1}. \tag{2.10}$$

However, (2.10) is almost immediate from Schwarz's inequality. Indeed if M denotes the number of r in $\mathcal{C}(N,n)$ for which $A(r)$ occurs, then

$$P\{A(r) \text{ occurs for some } r\} = P\{M > 0\} \geq \frac{(E\{M\})^2}{E\{M^2\}}.$$

Now

$$E\{M\} = \#\mathcal{C}(N,n)p^{nN-1}.$$

(see (2.7)), while

$$\begin{aligned} E\{M^2\} &= \sum_{r,r'} P\{A(r) \text{ and } A(r') \text{ occur}\} \\ &= \sum_{r,r'} p^{2nN-2-J(r,r')} = [\#\mathcal{C}(N,n)]^2 p^{2nN-2}\mathbb{E}\{p^{-J(r,r')}\}. \end{aligned}$$

The second equality here follows from the fact that the number of Y's sampled by r and r' *together* is $2nN-2-J(r,r')$, because J counts precisely the number of times r samples a Y which is also sampled by r'. (2.10) follows from these formulae. ∎

To estimate $\mathbb{E}\{p^{-J}\}$ we shall break up J into a sequence of contributions which behave more or less like a Markov chain. Before we do this it is convenient to view the paths $(v_1,\dots,v_{nN-1})$ and $(v'_1,\dots,v'_{nN-1})$ as the initial pieces of two infinite paths $r=(v_1,v_2,\dots)$ and $r'=(v'_1,v'_2,\dots)$. Accordingly we extend $\mathbb{P}$ to a measure on pairs of infinite paths which are independent under $\mathbb{P}$ and whose ith step is any one of the unit vectors in $\{\pm e_k : k \leq d-\lfloor d/N\rfloor\}$ with probability $(2d-2\lfloor d/N\rfloor)^{-1}$ when N does not divide i, and whose ith step is any one of $\{e_k : k > d-\lfloor d/N\rfloor\}$ with probability $\lfloor d/N\rfloor^{-1}$ when i is a multiple of N. Here the first step of the path r is v_1 and the first step of r' is v'_1. We also maintain our convention that $v_0=\mathbf{0}$. One easily checks that the initial pieces of length $nN-1$ of r and r' are independently uniformly distributed over $\mathcal{C}(N,n)$ as with the previous definition of $\mathbb{P}$. If necessary we shall write $J_{nN}(r,r')$ now, instead of our previous $J(r,r')$, to indicate that we are working with the initial pieces of length $nN-1$.

We define the kth *block* of r to be the path $(v_{(k-1)N}, v_{(k-1)N+1},\dots, v_{kN-1})$. By a slight abuse of notation we shall also say that the time t or the index t occurs in the kth block if $(k-1)N \leq t < kN$. We say that r has a *high density point* in the kth block if there exist t and s in the kth block such that

$$t,s \geq 1,\ |t-s| \geq 2, \text{ and } |v_t - v_s| \leq 1.$$

In particular if v_t is a double point of r in the rth block, then it is also a high density point. However, v_t is also a high density point if one of its neighbors is visited by r at any other time than $t-1$ or $t+1$. Similar definitions hold for r'. By our choice of $\mathbb{P}$, (2.3) still holds for all k so that w.p.1 $v_t = v_s$ can actually occur only if t and s belong to the same block. For the same reason $v_t = v'_s$ can w.p.1 occur only when t and s lie in the same block.

We next define *special indices*. If r has no high density point in the kth block, then $(k-1)N+i$, with $0 \le i < N$, is a special index if and only if

$$\begin{aligned}&v'_{(k-1)N+i} \text{ has not been visited by } r' \text{ at any time } 1 \le \\ &t < (k-1)N+i, \text{ and in addition } v'_{(k-1)N+i} = v_{(k-1)N+j} \\ &\text{for some } j. \text{ (If } k=1 \text{ we also require } i,j \ge 1.)\end{aligned} \tag{2.11}$$

We point out that (w.p.1) the occurrence of (2.11) depends on the kth blocks of r and r' only, since the only possible values for t and $(k-1)N+j$ at which r' or r can visit $v_{(k-1)N+i}$ are in the kth block. With each such special index t we associate a contribution $L(t) = L(t; r, r')$ of size 1 to J. Here and in the future we index a contribution L by the special index to which it corresponds. Next, when r has a high density point in the kth block, then there is either no special index in $[(k-1)N, kN)$ or exactly one. The former is the case if there are no t and s in $[(k-1)N \vee 1, kN)$ with $v_t = v'_s$ ($a \vee b$ denotes $\max\{a,b\}$). If there do exist such t and s, then the only special index in the kth block is taken to be $kN-1$ and the corresponding contribution $L(kN-1)$ is defined as

$$\begin{aligned}L(kN-1) = {} & \text{ number of } s \in [(k-1)N \vee 1, kN) \text{ for which } v'_s \\ & \qquad\qquad \text{equals a } v_t \text{ in the } k\text{th block of } r.\end{aligned}$$

Now let $t(1) < t(2) < \cdots < t(\rho)$ be all the special indices $\le nN-1$ (thus the next special index $t(\rho+1)$ occurs at or after time nN; this defines $\rho = \rho(nN)$). We claim that

$$J_{nN}(r, r') \le \sum_{k=1}^{\rho} L(t(k); r, r'). \tag{2.12}$$

To prove (2.12) consider the kth block of r. The vertices in this block can be visited only at the times $[(k-1)n, kN)$. If r has no high density points in this block and v is one of the vertices of r in this block then $k(r,v) = 1$. Therefore the only contributions to (2.8) from this block come from v's with $k(r,v) \wedge k(r',v) = 1$. Let v be such a vertex and let t be the smallest index t for which $v'_t = v$. Then t is a special index and the corresponding $L(t) = 1 = k(r,v) \wedge k(r',v)$. Thus all contributions to J from a block without high density points also appear in the right hand side of (2.12). If r has a high density point in the kth block but $v_t \neq v'_s$ for all $t,s \in [(k-1)N, kN)$, then there are no contributions from this block to either side of (2.12). If $v_t = v'_s$ for some t, s, then the contribution to J_{nN} from this block is

$$\sum_{\substack{v \in k\text{th} \\ \text{block of } r}} k(r,v) \wedge k(r',v) \leq \sum_{\substack{v \in k\text{th} \\ \text{block of } r}} k(r',v)$$
$$= L(kN-1).$$

Thus for a block with high density points the contribution to the right hand side of (2.12) is always at least as large as the one to the left hand side, and (2.12) must hold.

Finally we associate a *type* with each special index and its corresponding contribution L. We make the convention that $t(0) = 0$. For $i \geq 1$ we say that $t(i)$ is of

type 1 if t belongs to a block without high density points,
$t(i) - t(i-1) = 1$, and $t(i)$ *is not* a multiple of N,
type 2 if t belongs to a block without high density points,
$t(i) - t(i-1) = 1$, but $t(i)$ *is* a multiple of N,
type 3 if t belongs to a block without high density points,
and $t(i) - t(i-1) \geq 2$,
type 4 if t belongs to a block with high density points.

It will turn out that the main task is to estimate

$$\mathbb{E}\{p^{-L(t(i))}; i \leq \rho,\ t(i) \text{ is of type } l \mid \mathcal{F}_{i-1}\} \tag{2.13}$$

on the event

$$\{(i-1) \leq \rho,\ t(i-1) \text{ is of type } m\}, \tag{2.14}$$

where

$$\mathcal{F}_j := \text{ the } \sigma\text{-field generated by } t(j) \text{ and } \{v_t : t < \lceil t(j)/N \rceil N\} \cup \{v'_t : t \leq t(j)\}$$

($\lceil a \rceil$ denotes the smallest integer $\geq a$). It may be useful for the reader to skip Lemmas 2–6 at first reading to see how the main line of the argument runs once (2.13) has been estimated.

Note that r and r' are not treated equally in the definition of $\mathcal{F}_j$; we are forced to do this by the asymmetric definition of the special indices which involve first looking at the whole block of r to see whether it contains a high density point, while high density points of r' do not play such a role.

We have defined $\mathcal{F}_j$ in such a way that $L(t(j))$ is measurable with respect to $\mathcal{F}_j$.

To estimate (2.13) we shall need some estimates which are basically known facts about a simple random walk. For the remainder of this section we take $D = d - \lfloor d/N \rfloor$ and $\{S_u\}$ a simple random walk in $\mathbb{Z}^D$ with $S_0 = \mathbf{0}$. K_i will denote some universal constant (independent of d, N and p). Furthermore we restrict p and N to satisfy

$$\frac{1}{2d} \leq p \leq \frac{2}{2d} \quad \text{and} \quad 8 \leq N \leq \frac{\log d}{2 \log \log d}. \tag{2.15}$$

LEMMA 2. *Let*

$$\mathcal{G}_t = \sigma\text{-field generated by all } v_i \text{ and by the } v'_q \text{ with } q \leq t.$$

and let τ be a stopping time with respect to the $\mathcal{G}_t$. Then for p and N satisfying (2.15) and for any vertex w we have

$$\begin{aligned} &\mathbb{P}\{v'_i = w \text{ for some } \tau \vee (jN-1) < i < (j+1)N \mid \mathcal{G}_\tau\} \\ &\quad \leq \frac{(1 + K_1 N/d)}{2D} \text{ on the set } \{\tau = s\}, \text{ for any } jN \leq s < (j+1)N - 1. \end{aligned} \tag{2.16}$$

Moreover

$$\begin{aligned} &\mathbb{P}\{r' \text{ visits the } (j+1)\text{th block of } r \text{ at some time} > \tau \mid \mathcal{G}_\tau\} \\ &\qquad \leq \begin{cases} \frac{N}{2D}(1 + K_1 N/d) & \text{on the set } \{\tau \geq jN\} \\ \frac{2N}{d} & \text{on the set } \{\tau < jN\}. \end{cases} \end{aligned} \tag{2.17}$$

PROOF: Note that, given r and the event $\{\tau = s\}$, possibly intersected with some other event in $\mathcal{G}_s$, the conditional distribution of the steps s'_i with $i > s$ is still the same as the unconditional distribution under the extended $\mathbb{P}$ as defined above. In particular the steps s'_i for $s < i < (j+1)N$ and i not divisible by N are distributed like the steps of a simple random walk $\{S_u\}$ on $\mathbb{Z}^D$ with $S_0 = \mathbf{0}$. More precisely, this holds for the projection of the s'_i on the span of the first D coordinate vectors. We shall be somewhat cavalier about this and shall not always distinguish between s'_i and this projection. For any vector w in $\mathbb{Z}^d$ or $\mathbb{Z}^D$ we use $w(i)$ to denote the ith component of w and $\overline{w} = (w(1), \ldots, w(D))$ for the projection of w on $\mathbb{Z}^D$ if $w \in \mathbb{Z}^d$.

Now it is known for a simple random walk $\{S_u\}$ on $\mathbb{Z}^D$ with $S_0 = \mathbf{0}$ (cf. Kesten 1964, Sect. 3) that

$$\sup_{\overline{w}} P\{S_{2u+1} = \overline{w}\} \leq \sup_{\overline{w}} P\{S_{2u} = \overline{w}\} = P\{S_{2u} = \mathbf{0}\}, \tag{2.18}$$

$$\sum_{u=1}^{\infty} P\{S_{2u} = \mathbf{0}\} \leq \frac{1}{2D}(1 + K_2 D^{-1}), \ \sum_{u=2}^{\infty} P\{S_{2u} = \mathbf{0}\} \leq K_2 D^{-2}. \tag{2.19}$$

Also, by counting all possibilities (cf. (3.5) in Kesten 1964), one easily obtains

$$\sup_{\overline{w}} P\{S_1 = \overline{w}\} = \frac{1}{2D}, \ \sup_{\overline{w}\neq \mathbf{0}} P\{S_2 = \overline{w}\} \le \frac{K_3}{D^2}, \qquad \sup_{\overline{w}} P\{S_3 = \overline{w}\} \le \frac{K_3}{D^2}. \tag{2.20}$$

It follows from these observations that if $jN \le s$ and $w = (w(1), \dots, w(d))$ is such that

$$\sum_{q=1}^{D} \{w(q) - v'_s(q)\} \quad \text{is even,} \tag{2.21}$$

then we have on the set $\{\tau = s\}$

$$\begin{aligned}\mathbb{P}\{v'_i = w \text{ for some } s < i < (j+1)N \mid \mathcal{G}_\tau\} &\le P\{S_u \text{ visits } \overline{w} - \overline{v}'_s \text{ for some } u > 0\} \\ &= P\{S_u \text{ visits } \overline{w} - \overline{v}'_s \text{ at some even time} > 0\} \\ &\le \frac{1}{2D}\left(1 + \frac{K_2}{D}\right) \quad \text{(see (2.19)).}\end{aligned} \tag{2.22}$$

If the sum in (2.21) is odd instead of even then we obtain (2.22) by replacing 'even' by 'odd' in (2.22) and using (2.18), (2.19), as well as the special estimates (2.20) for the terms corresponding to $u = 1$ or 3. This proves (2.16).

Next we note that on $\{\tau = s\}$ with $jN \le s < (j+1)N$ the first case of (2.17) is immediate from (2.16) since there are only N points in the $(j+1)$th block of r and these can be visited by r' only during the $(j+1)$th block. (2.17) is also clear on $\{\tau \ge (j+1)N\}$ for then the left hand side is zero. In order to obtain (2.17) on $\{\tau = s\}$ when $s < jN$ we observe that the sum of the last $\lfloor d/N \rfloor$ coordinates is the same for all the v_t in the $(j+1)$th block of r (compare (2.3)). The same comment applies to r'. Therefore r' can visit the $(j+1)$th block of r only if the sum of the last $\lfloor d/N \rfloor$ coordinates is the same for v'_{jN} and v_{jN}. Moreover the last $\lfloor d/N \rfloor$ coordinates of v'_{jN} are w.p.1 the same as those of $v'_{(j-1)N} + s'_{jN}$. Thus if we condition on $\mathcal{G}_{\tau\vee(jN-1)}$ then on $\{\tau < jN\}$ the conditional probability that r' visits the $(j+1)$th block of r is bounded by

$$\begin{aligned}\mathbb{P}\{s'_{jN} = \big(0, \dots, 0, v_{jN}(D+1) - v'_{(j-1)N}(D+1), \dots \\ \dots, v_{jN}(d) - v'_{(j-1)N}(d)\big) \mid \mathcal{G}_{\tau\vee(jN-1)}\} \le \lfloor d/N \rfloor^{-1}.\end{aligned}$$

In the last step we used that s'_{jN} takes any given value with probability at most $\lfloor d/N \rfloor^{-1}$ by the definition of $\mathbb{P}$. This implies (2.17) on $\{\tau < jN\}$ as well, since $\mathcal{G}_\tau \subset \mathcal{G}_{\tau\vee(jN-1)}$. ■

We remind the reader of our convention that $t(0) = 0$. If we declare $t(0)$ to be a special index of type 2 then Lemmas 3–6 remain valid even for $i = 1$. In other words, for $i = 1$ the estimates in these lemmas for $m = 2$ apply also to

$$\mathbb{E}\{p^{-L(t(1))}; 1 \le \rho,\ t(1) \text{ is of type } l\}.$$

We leave most of the slight modifications necessary for $i = 1$ to the reader.

LEMMA 3. *Under (2.15) the expression in (2.13) is for $l = 4$ at most*

$$K_4 p^{-N} \left[\frac{N}{2D}\right]^{N+1} \tag{2.23}$$

on the set (2.14) for any $1 \le m \le 4$.

PROOF: First observe that if

$$t(i-1) \text{ occurs in the } k\text{th block,} \tag{2.24}$$

then the next special index can be of type 4 only if it occurs in the jth block for some $j > k$ and if r has a high density point in its jth block (no matter what the type of $t(i-1)$ is). In addition v'_s must equal v_t for some s and t in the jth block for there to be any special index in the jth block. Let us assume for the rest of this proof that (2.24) occurs and let us set

$$\sigma_j = \text{smallest index } s \ge (j-1)N \text{ such that } v'_s \text{ equals some } v_t$$

($\sigma = \infty$ if no such s exists). (If $i = 1$ then we replace kN by 1 in the above definition.) Then on the event (2.24) we have

$$\begin{aligned} &\mathbb{E}\left\{p^{-L(t(i))};\ i \le \rho,\ t(i) \text{ is of type } 4 \mid \mathcal{F}_{i-1}\right\} \\ &\quad \le \sum_{j>k} \sum_{(j-1)N \le \nu < jN} \mathbb{E}\Big\{p^{-L(jN-1)};\ \sigma_j = \nu \text{ and } r \text{ has} \\ &\qquad\qquad \text{a high density point in the } j\text{th block} \mid \mathcal{F}_{i-1}\Big\}. \end{aligned} \tag{2.25}$$

We shall estimate the summands in the right hand side of (2.25) by conditioning on r. First we show that for $(j-1)N \le \nu < jN$

$$\mathbb{E}\{p^{-L(jN-1)} \mid r, \sigma_j = \nu, \mathcal{F}_{i-1}\} \le K_5 p^{-N} \left[\frac{N}{2D}\right]^{N-1}. \tag{2.26}$$

This will be seen to follow from Lemma 2. Indeed, note that $\sigma_j = \nu \in [(j-1)N, jN)$ implies $L(jN-1) \geq 1$, since there is at least the contribution to this L of the visit of r' to r at the time σ_j. For $L(jN-1)$ to be $\geq \lambda+1$, there must be at least λ further visits by r' to r, necessarily to the jth block of r and during the time interval (σ_j, jN). Thus by the first line of (2.17) (with j replaced by $j-1$)

$$\mathbb{P}\{r' \text{ visits the } j\text{th block of } r \text{ at least } \lambda \text{ times during } (\sigma_j, jN) \mid \mathcal{G}_{\sigma_j}\} \\ \leq \left\{\frac{N}{2D}(1+K_1N/d)\right\}^{\lambda}. \quad (2.27)$$

By virtue of (2.27) the left hand side of (2.26) is at most

$$p^{-1} + \sum_{\lambda=1}^{N-1} p^{-\lambda-1}\left\{\frac{N}{2D}(1+K_1N/d)\right\}^{\lambda}. \quad (2.28)$$

Note that the upper bound in the sum over λ is $N-1$ because r' cannot visit the jth block of r more than N times. With the choice of p and N restricted by (2.15) the ratio of the geometric series in (2.28) is at least 2 and (2.26) follows.

Substitution of (2.26) into (2.25) now shows that on the event (2.24)

$$\mathbb{E}\{p^{-L(t(i))};\ i \leq \rho,\ t(i) \text{ is of type } 4 \mid \mathcal{F}_{i-1}\} \\ \leq K_5 p^{-N}\left[\frac{N}{2D}\right]^{N-1} \sum_{j>k} \mathbb{P}\{r \text{ has a high density point in} \\ \text{its } j\text{th block and } r' \text{ visits the } j\text{th block of } r \mid \mathcal{F}_{i-1}\}. \quad (2.29)$$

Note that if $A \in \mathcal{F}_{i-1}$ then $A \cap \{t(i-1) \text{ occurs in the } k\text{th block}\}$ belongs to $\mathcal{H}_{kN-1}$, where

$$\mathcal{H}_t = \sigma\text{-field generated by } \{v_j, v'_j : j \leq t\}.$$

It therefore suffices to estimate the right hand side of (2.29) with $\mathcal{F}_{i-1}$ replaced by $\mathcal{H}_{kN-1}$. Now by estimates entirely analogous to those for (2.16) and (2.27) we have for $j > k$

$$\mathbb{P}\{r \text{ has a high density point in its } j\text{th block} \mid \mathcal{H}_{kN-1}\} \\ \leq \sum_{(j-1)N \leq t < jN} \Big[\mathbb{P}\{v_s = v_t \text{ for some } t < s < jN \mid \mathcal{H}_{kN-1}\} \\ + \sum_w \mathbb{P}\{v_s = w \text{ for some } t+2 \leq s < jN \mid \mathcal{H}_{kN-1}\}\Big], \quad (2.30)$$

where the inner sum over w runs over the $2d$ neighbors of v_t. The first probability in the right hand side is for each fixed t at most

$$\mathbb{E}\Big\{\sup_w \mathbb{P}\{v_s = w \text{ for some } t < s < jN \mid \mathcal{H}_t\} \mid \mathcal{H}_{kN-1}\Big\},$$

which by virtue of (2.16) (with the roles of r and r' interchanged) is at most D^{-1}. As for the second probability in the right hand side of (2.30) note that $v_s = w$ means that $v_s - v_t$ has to be a unit vector, and in fact when t and s lie in the same block this can occur only when $w \in \{\pm e_i : i \le D\}$. Therefore this probability is at most

$$\begin{aligned}\mathbb{E}\Big\{\sup_{i\le D} \mathbb{P}\{v_s - v_t = \pm e_i \text{ for some } t+2 \le s < jN \mid \mathcal{H}_t\} \mid \mathcal{H}_{kN-1}\Big\}&\\ \le \sup_{\overline{w}\neq \mathbf{0}} P\{S_u = \overline{w} \text{ for some } u \ge 2\}&\\ \le K_6 D^{-2}.&\end{aligned}$$

Since t can take at most N values and w at most $2d$ values we obtain that (2.30) is at most $K_7 N/D$. Substituting this into (2.29) we see that its right hand side, with $\mathcal{H}_{kN-1}$ instead of $\mathcal{F}_{i-1}$, is bounded by

$$K_8 p^{-N}\left[\frac{N}{2D}\right]^N \sum_{j>k} \sup_r \mathbb{P}\{r' \text{ visits the } j\text{th block of } r \mid r, \mathcal{H}_{kN-1}\}. \tag{2.31}$$

Since conditioning on r and on $\mathcal{H}_{kN-1}$ is the same as conditioning on $\mathcal{G}_{kN-1}$, the probability in the sum in (2.31) for $j = k+1$ is at most $2N/D$ (by the second line of (2.17)). The remaining sum in (2.31) is bounded by

$$\begin{aligned}\sum_{j\ge k+2} \sup_r \sum_{(j-1)N\le q<jN} \mathbb{P}\big\{v'_s = v_q \text{ for some }&\\ (j-1)N \le s < jN \mid r, \mathcal{H}_{kN-1}\big\}&\\ \le \sum_{s\ge (k+1)N} N \sup_w \mathbb{P}\big\{v'_s = w \mid \mathcal{H}_{kN-1}\big\}.&\end{aligned} \tag{2.32}$$

In turn the last sum can be estimated by the arguments used in Lemma 2. We can condition on all steps s'_i with i divisible by N in addition to $\mathcal{H}_{kN-1}$. Then $v'_s - v'_{kN-1}$ still contains $s - kN - \lfloor (s-kN+1)/N\rfloor$ simple random walk steps independent of these conditions, so that

$$\sup_w \mathbb{P}\{v'_s = w \mid \mathcal{H}_{kN-1}\} \le \sup_{\overline{w}} P\big\{S_{s-kN-\lfloor (s-kN+1)/N\rfloor} = \overline{w}\big\}$$

and (2.32) is therefore at most

$$K_8 N \sum_{u=4}^{\infty} \sup_{\overline{w}} P\{S_u = \overline{w}\} \le K_9 \frac{N}{D^2}.$$

It follows that (2.31) is bounded by $K_{10} p^{-N}(N/(2D))^{N+1}$. The lemma follows because (2.13) is bounded by the conditional expectation (given $\mathcal{F}_{i-1}$) of (2.31). ∎

LEMMA 4. *Under (2.15) the expression (2.13) is for $l = 3$ at most*

$$K_4 \frac{N^2}{pd^2} \tag{2.33}$$

on the set (2.14) for any $1 \le m \le 4$.

PROOF: If $t(i)$ is of type 3, then $L(t(i)) = 1$. In addition, $v'_{t(i)}$ cannot have been visited before by r', so that $v'_{t(i)} \ne v'_{t(i-1)}$ (cf. (2.11)). Finally $t(i) \ge t(i-1) + 2$. Therefore

$$\begin{aligned} \mathbb{E}\{p^{-L(t(i))}; i \le \rho,\ t(i) \text{ is of type 3} \mid \mathcal{G}_{t(i-1)}\} \\ \le p^{-1}\mathbb{P}\big\{v'_s = v_t \text{ for some } s \ge t(i-1)+2 \text{ and some } t \text{ with} \\ v_t \ne v'_{t(i-1)} \mid \mathcal{G}_{t(i-1)}\big\}. \end{aligned} \tag{2.34}$$

Next we note that for given r and a time s there are at most N possible v_t which can equal v'_s, since t and s must belong to the same block for this to be possible. The right hand side of (2.34) is therefore bounded by

$$p^{-1} N \sup_{\tau} \sum_{m=2}^{\infty} \sup_{w \ne \mathbf{0}} \mathbb{P}\big\{v'_{\tau+m} - v'_{\tau} = w\big\}. \tag{2.35}$$

This sum can be estimated by almost the same method as used for (2.32). First consider the terms with $2 \le m < 5$. If τ is such that there are no i divisible by N in $(\tau, \tau+m]$, then $v'_{\tau+m} - v'_{\tau}$ has the same distribution as S_m. In particular $v'_{\tau+m} - v'_{\tau} = w$ is possible only if the last $\lfloor d/N \rfloor$ coordinates of w are zero. Also for $\overline{w} \ne \mathbf{0}$

$$P\big\{S_m = \overline{w}\big\} \le K_5 D^{-2}, \tag{2.36}$$

by virtue of (2.18)–(2.20). If $m < 5$ then there may also be exactly one i_0 in $(\tau, \tau+m]$ which is divisible by N. In this case $v'_{\tau+m} - v'_{\tau} = w$ forces w to be the sum of at most $m-1$ vectors from $\{\pm e_k : k \le D\}$ plus exactly one vector from $\{e_k : D < k \le d\}$. The step s'_{i_0} has to equal this last vector and the other $m-1$ steps s'_i with $\tau < i \le \tau+m$ have to add up to a vector determined by w. Since the probability of s'_{i_0} having a prescribed value is at most $2N/d$, we obtain that in this case

$$\sup_{w\neq\mathbf{0}} \mathbb{P}\{v'_{\tau+m} - v'_\tau = w\} \le K_5 N/(dD). \tag{2.37}$$

For $m \ge 5$ we simply observe that there are at least $(m-1-\lfloor m/N \rfloor) \ge 4$ values of i in $(\tau, \tau+m]$ which are not divisible by N so that

$$\sup_\tau \sup_{w\neq\mathbf{0}} \sum_{m=5}^\infty \mathbb{P}\{v'_{\tau+m} - v'_\tau = w\} \le K_6 \sum_{u=2}^\infty P\{S_{2u} = \mathbf{0}\} \le K_7 D^{-2} \tag{2.38}$$

(cf. (2.18) and (2.19)). (2.35)–(2.38) show that the right hand side of (2.34) is at most $K_8 p^{-1} N^2 d^{-2}$. Since $\mathcal{F}_{i-1} \subset \mathcal{G}_{t(i-1)}$ this same estimate holds for the expression in (2.13).

For $i = 1$, (2.34) should be replaced by

$$\begin{aligned}
&\mathbb{E}\{p^{-L(t(i))}; 1 \le \rho,\ t(1) \text{ is of type } 3\} \\
&\quad \le p^{-1} \sum_{s=2}^{N-1} \mathbb{P}\{v'_s = \text{some } v_t \text{ with } 1 \le t < N\} + (\text{expression in (2.35)}) \\
&\quad \le p^{-1} \sum_{s=2}^{N-1} \sum_{t=1}^{N-1} \mathbb{P}\{v'_s = v_t = \mathbf{0}\} + p^{-1} N \sum_{s=2}^{N-1} \sup_{w\neq\mathbf{0}} \mathbb{P}\{v'_s = w\} \\
&\qquad\qquad + (\text{expression in (2.35)}) \\
&\quad \le p^{-1} \sum_{s=2}^{N-1} \sum_{t=1}^{N-1} \mathbb{P}\{v'_s = \mathbf{0}\}\mathbb{P}\{v_t = \mathbf{0}\} + K_9 p^{-1} N^2 d^{-2} \\
&\quad \le K_{10} p^{-1} N^2 d^{-2}
\end{aligned} \tag{2.39}$$

(by (2.18)–(2.20)). ∎

For $l = 1$ or 2 our estimate for (2.13) on the set (2.14) does depend on m.

LEMMA 5. *Under (2.15), on the set (2.14) we have*

$$\mathbb{E}\{p^{-L(t(i))}; i \ge \rho,\, t(i) \text{ is of type } 1 \mid \mathcal{F}_{i-1}\} \le \begin{cases} (2pD)^{-1} & \text{if } m = 1 \\ (pD)^{-1} & \text{if } m = 2 \text{ or } 3 \\ 0 & \text{if } m = 4. \end{cases} \tag{2.40}$$

PROOF: As in the last lemma $L(t(i)) = 1$ if $t(i)$ is of type 1. First consider the case $m = 1$, i.e., let $t(i-1)$ be of type 1 as well. Let $t(i-1)$ belong to the kth block. We must then have that $t(i)$ also belongs to the kth block, and in fact $t(i-1)+1 = t(i) < kN$ (since $t(i)$ is not divisible by N). Also $v'_{t(i-1)}$ must equal some v_t with t in the kth block. Since $t(i-1)$ is also of type 1, $v'_{t(i-2)} = v'_{t(i-1)-1}$ is one of the neighbors of v_t and also equals some point of r. Moreover $t(i-1)$ is not divisible by N, so that $t(i-1)-1 = t(i-2) \geq (k-1)N$ also belongs to the kth block. Since r does not have a high density point in the kth block if $t(i-1)$ is of type 1, r does not visit any other neighbors on $\mathbb{Z}^d$ of v_t than v_{t-1} and v_{t+1} during $[(k-1)N, kN)$. One of these is $v'_{t(i-2)}$. But also $v'_{t(i)}$ must be equal to a neighbor of v_t which is visited during the kth block (recall that $v'_{t(i)}$ and $v'_{t(i-2)}$ can only visit points of the kth block of r, by (2.3)). Thus $v'_{t(i)}$ must be either v_{t-1} or v_{t+1}. However, it cannot equal $v'_{t(i-2)}$ because at time $t(i)$, r' must be at a point which it had not visited before (see (2.11)). Since all of r and $v'_{t(i-2)}$ are known when we condition on $\mathcal{F}_{i-1}$, there is only one choice for $v'_{t(i)}$, namely the one point of $v_{t\pm 1}$ which is not $v'_{t(i-2)}$. The probability that r' moves to this prescribed site at the $(t(i-1)+1)$th step is $(2D)^{-1}$. This proves the case $m = 1$ of (2.40).

The case $m = 2$ or 3 is very similar, except that there now may be two choices for $v'_{t(i)}$. Again, if $t(i-1)$ belongs to the kth block, then $t(i) = t(i-1)+1$ also belongs to the kth block and $v'_{t(i-1)}$ equals some v_t of the kth block of r, $v'_{t(i)}$ must be one of the neighbors of v_t which are visited by the kth block of r. This allows at most the choices v_{t-1} or v_{t+1} for $v'_{t(i)}$. This takes care of $m = 2$ or 3 when $i \geq 2$.

For $i = 1$ we have by (2.16) (with the roles of r and r' interchanged)

$$\begin{aligned}
&\mathbb{E}\{p^{-L(t(i))}; 1 \leq \rho,\ t(1) \text{ is of type } 1\} \\
&\qquad \leq p^{-1}\mathbb{P}\{v'_1 = v_t \text{ for some } 1 \leq t < N\} \\
&\qquad \leq p^{-1} \sup_w \mathbb{P}\{v_t = w \text{ for some } 1 \leq t < N\} \\
&\qquad \leq (pD)^{-1}.
\end{aligned}$$

Finally, if $m = 4$, then $t(i-1) = kN - 1$ for some k. Then $t(i)$ cannot be of type 1, for this would require on the one hand that $t(i) = t(i-1)+1$, and on the other hand that $t(i)$ is not divisble by N. ■

LEMMA 6. *Under (2.15), on the set (2.14) we have*

$$\begin{aligned}
&\mathbb{E}\{p^{-L(t(i))}; i \leq \rho,\ t(i) \text{ is of type } 2 \mid \mathcal{F}_{i-1}\} \\
&\qquad \leq \begin{cases} 4N(pd)^{-1} & \textit{if } m = 1, 3, \textit{ or } 4 \\ 0 & \textit{if } m = 2. \end{cases}
\end{aligned} \tag{2.41}$$

PROOF: $t(i)$ can be of type 2 only if $t(i) = kN$ for some k and if $t(i-1) = kN - 1$. This rules out that $t(i-1)$ is of type 2, so that the second case of (2.41) is trivial. For $i \geq 2$ and $m = 1, 3$, or 4, on the set $\{t(i-1) = kN - 1\}$ the left hand side of (2.41) is bounded by

$$p^{-1}\mathbb{P}\{v_t = v'_{kN} \text{ for some } kN \leq t < (k+1)N \mid \mathcal{H}_{kN-1}\}.$$

This is bounded by $p^{-1}(2N/D)$ by the second case of (2.17). For $i = 1$, $t(0) = kN - 1$ is impossible. ■

We are now ready to carry out the principal estimate for

$$\mathbb{E}\big\{p^{-J_{nN}(r,r')}\big\}.$$

By (2.12) this expression is for all n at most

$$\begin{aligned}
&\sum_{u=0}^{\infty} \mathbb{E}\left\{p^{-\sum_{k=1}^{u} L(t(k))}; \rho = u\right\} \\
&\leq 1 + \sum_{u=1}^{\infty} \mathbb{E}\left\{p^{-\sum_{k=1}^{u} L(t(k))}; t(u) < nN\right\} \\
&\leq 1 + \sum_{u=1}^{\infty}\sum_{\tau} \mathbb{E}\left\{p^{-\sum_{k=1}^{u} L(t(k))}; t(u) < nN,\ t(k) \text{ has type } \tau(k),\ k \leq u\right\}.
\end{aligned} \tag{2.42}$$

The sum over τ here is over all possible sequences of types $(\tau(1), \dots, \tau(u))$ for $(t(1), \dots, t(u))$. For fixed u and τ the summand here can be written as

$$\begin{aligned}
&\mathbb{E}\Big\{p^{-\sum_{k=1}^{u-1} L(t(k))}\mathbb{E}\big\{p^{-L(t(u))}; u \leq \rho,\ t(u) \text{ is of type } \tau(u) \mid \mathcal{F}_{u-1}\big\}; \\
&\qquad\qquad u - 1 \leq \rho,\ t(k) \text{ is of type } \tau(k),\ k \leq u-1\Big\} \\
&\leq \mathbb{E}\Big\{p^{-\sum_{k=1}^{u-1} L(t(k))}\Gamma(\tau(u-1), \tau(u)); u - 1 \leq \rho, \\
&\qquad\qquad t(k) \text{ is of type } \tau(k),\ k \leq u-1\Big\},
\end{aligned}$$

where $\Gamma(m, l)$ is an upper bound for (2.13) on the set (2.14). From Lemmas 3–6 we see that we can take for Γ the following matrix:

$$\begin{pmatrix}
\frac{1}{2pD} & \frac{4N}{pd} & K_4\frac{N^2}{pd^2} & K_4 p^{-N}\left[\frac{N}{2D}\right]^{N+1} \\
\frac{1}{pD} & 0 & K_4\frac{N^2}{pd^2} & K_4 p^{-N}\left[\frac{N}{2D}\right]^{N+1} \\
\frac{1}{pD} & \frac{4N}{pd} & K_4\frac{N^2}{pd^2} & K_4 p^{-N}\left[\frac{N}{2D}\right]^{N+1} \\
0 & \frac{4N}{pd} & K_4\frac{N^2}{pd^2} & K_4 p^{-N}\left[\frac{N}{2D}\right]^{N+1}
\end{pmatrix}.$$

By iteration of this argument we now obtain

$$\mathbb{E}\left\{p^{-\sum_{k=1}^{u} L(t(k))}; t(u) < nN,\ t(k) \text{ has type } \tau(k),\ k \le u\right\}$$
$$\le \mathbb{E}\big\{p^{-L(t(1))}; t(1) \text{ has type } \tau(1)\big\} \prod_{k=1}^{u-1} \Gamma(\tau(k), \tau(k+1)).$$

As pointed out before, the estimates in Lemmas 3–6 with $m = 2$ apply to

$$\mathbb{E}\big\{p^{-L(t(1))}; 1 \le \rho,\ t(1) \text{ has type } \tau(1)\big\},$$

so that finally

$$\mathbb{E}\left\{p^{-\sum_{k=1}^{u} L(t(k))};\ t(u) < nN,\ t(k) \text{ has type } \tau(k),\ k \le u\right\}$$
$$\le \Gamma(2, \tau(1)) \prod_{k=1}^{u-1} \Gamma(\tau(k), \tau(k+1)).$$

Substituting this into (2.42) we find

$$\mathbb{E}\big\{p^{-J_{nN}(r,r')}\big\} \le 1 + \sum_{u=1}^{\infty} \sum_{\tau} \Gamma(2, \tau(1)) \prod_{k=1}^{u-1} \Gamma(\tau(k), \tau(k+1)). \tag{2.43}$$

It will not do to take the sum here over *all* sequences $(\tau(1), \dots, \tau(u))$ with values in $\{1, 2, 3, 4\}$ because the largest eigenvalue of the matrix Γ is much bigger than 1 (in fact $\Gamma(1,2)\Gamma(2,1)$ is of order N under the restrictions (2.15) and this will grow with d; see below). However, as we saw in (2.42) we only have to sum over the sequences which are possible sequences of types for $(t(1), \dots, t(u))$. In particular, if $\tau(k) = 2$ for some k, then either all $\tau(j)$ with $k - N < j < k$ equal 1 or one of these $\tau(j)$ equals 3 or 4 and the τ's between $\tau(j)$ and $\tau(k)$ equal 1. We use this to replace Γ in (2.43) by the matrix Δ defined as

$$\begin{pmatrix} \frac{N^{3/(N-1)}}{2pD} & \frac{4}{N^2pd} & K_4\frac{N^5}{pd^2} & K_4p^{-N}\left[\frac{N}{2D}\right]^{N+1} N^3 \\ \frac{N^{3/(N-1)}}{pD} & 0 & K_4\frac{N^5}{pd^2} & K_4p^{-N}\left[\frac{N}{2D}\right]^{N+1} N^3 \\ \frac{N^{3/(N-1)}}{pD} & \frac{4}{N^2pd} & K_4\frac{N^5}{pd^2} & K_4p^{-N}\left[\frac{N}{2D}\right]^{N+1} N^3 \\ 0 & \frac{4}{N^2pd} & K_4\frac{N^5}{pd^2} & K_4p^{-N}\left[\frac{N}{2D}\right]^{N+1} N^3 \end{pmatrix}.$$

Δ is obtained from Γ by multiplying the first column by $N^{3/(N-1)}$, and the third and fourth columns by N^3, while dividing the second column by N^3. (2.43) with Γ replaced by Δ is a valid estimate because for each $\tau(k+1) = 2$ for which we lose a factor N^3 in the right hand side of (2.43) we gain a factor of at least N^3 from the $\tau(j)$ which equal 1, 3, or 4 between $\tau(k)$ and the preceding τ which equals 2 (or in all the preceding τ if $\tau(k)$ is the first τ which equals 2).

After the replacement of Γ by Δ we do sum over *all* sequences $(\tau(1), \dots, \tau(u))$ with values in $\{1, 2, 3, 4\}$ to obtain, uniformly in n,

$$\mathbb{E}\{p^{-J_{nN}(r,r')}\} \leq 1 + \sum_{u=1}^{\infty}\sum_{i=1}^{4} \Delta^u(2,i). \tag{2.44}$$

(1.3) is contained in the following stronger lemma.

Lemma 7. *The largest eigenvalue of Δ is at most*

$$\frac{N^{3/(N-1)}}{2pD} + \frac{12}{pdN^2} + 3K_4\frac{N^5}{pd^2} + 3K_4p^{-N}\left(\frac{N}{2D}\right)^{N+1} N^3. \tag{2.45}$$

(1.3) holds. Moreover, for fixed $\gamma > 1$,

$$\liminf_{d\to\infty} \theta\left(\frac{\gamma}{2d}, \mathbb{Z}^d, \text{bond}\right) \geq \liminf_{d\to\infty} \frac{2d}{\gamma}\,\theta\left(\frac{\gamma}{2d}, \mathbb{Z}^d, \text{site}\right) \geq K_5[(\gamma-1)\wedge 1]. \tag{2.46}$$

Proof: The largest eigenvalue of Δ is the same as the largest eigenvalue of $A^{-1}\Delta A$, where A is the diagonal matrix with entries 1, 3, 3, 3 along the diagonal. $A^{-1}\Delta A$ is obtained from Δ by multiplying the second, third, and fourth columns by 3 and then dividing the corresponding rows by 3. The largest row sum of the resulting matrix occurs in the first row and equals the expression in (2.45). Thus (2.45) is indeed an upper bound for the largest eigenvalue of Δ (Ostrowsky 1973, Theorem 19.1).

For $p = \gamma/(2d)$ with $1 \leq \gamma \leq 2$ and $N = \lfloor(\log d)/(2\log\log d)\rfloor$, (2.45) is bounded above by

$$\frac{d}{\gamma D} + K_6\frac{(\log\log d)^2}{\log d} \leq \frac{1}{\gamma} + K_7\frac{(\log\log d)^2}{\log d}.$$

In particular the largest eigenvalue of Δ is strictly less than one for

$$p = \frac{1}{2d}\left(1 + 2K_7\frac{(\log\log d)^2}{\log d}\right) \tag{2.47}$$

and d large. Thus for large d and p as in (2.47) the right hand side of (2.44) is finite and percolation occurs by Lemma 1. This implies (1.3). Also if we take $p = \gamma/(2d)$ for some fixed $\gamma > 1$, then for large d the right hand side of (2.44) is at most

$$\begin{aligned}1 + K_8 \sum_{u=1}^{\infty} \max_{i \le 4} \sum_{j=1}^{4} (A^{-1}\Delta A)^u(i,j) \\ &\le 1 + K_9 \sum_{u=1}^{\infty} (\text{expression in } (2.45))^u \\ &\le 1 + K_9 \left\{ 1 - \frac{1}{\gamma} - K_7 \frac{(\log\log d)^2}{\log d} \right\}^{-1} \\ &\le K_{10}\gamma(\gamma-1)^{-1}.\end{aligned}$$

The second inequality in (2.46) now follows from Lemma 1. The first inequality can be found in any one of Hammersley (1961), McDiarmid (1980), and Oxley and Welsh (1979). ■

References

Aizenman, M., Bricmont, J., and Lebowitz, J. (1987). Percolation of the minority spins in high dimensional Ising models. *Journal of Statistical Physics* 49, 859–865.

Aizenman, M., Kesten, H., and Newman, C.M. (1987). Uniqueness of the infinite cluster and continuity of connectivity functions for short and long range percolation. *Communications in Mathematical Physics* 111, 505–531.

Bricmont, J., Kesten, H., Lebowitz, J., and Schonmann, R. (1989). A note on the large dimensional Ising model. *Communications in Mathematical Physics*, to appear.

Broadbent, S.R. and Hammersley, J.M. (1957). Percolation processes. *Proceedings of the Cambridge Philosophical Society* 53, 629–641 and 642–645.

Cox, J.T. and Durrett, R. (1983). Oriented percolation in dimensions $d \ge 4$: bounds and asymptotic formulas. *Mathematical Proceedings of the Cambridge Philosophical Society* 93, 151–162.

Gandolfi, A., Grimmett, G., and Russo, L. (1988). On the uniqueness of the infinite cluster in the percolation model. *Communications in Mathematical Physics* 114, 549–552.

Gaunt, D.S. and Ruskin, H. (1978). Bond percolation processes in d dimensions. *Journal of Physics A: Mathematical and General* 11, 1369–1380.

Gaunt, D.S., Sykes, M.F., and Ruskin, H. (1976). Percolation processes in d dimensions. *Journal of Physics A: Mathematical and General* 9, 1899–1911.

Gordon, D.M. (1988). Percolation in high dimensions. Preprint.

Hammersley, J.M. (1959). Bornes supérieures de la probabilité critique dans un processus de filtration. In *Le Calcul des Probabilités et ses Applications*, 17–37, CNRS, Paris.

——— (1961). Comparison of atom and bond percolation. *Journal of Mathematical Physics* 2, 728–733.

Hara, T. and Slade, G. (1989). Mean-field critical phenomena for percolation in high dimensions. Preprint.

Kesten, H. (1964). On the number of self-avoiding walks II. *Journal of Mathematical Physics* 5, 1128–1137.

——— (1989). Asymptotics in high dimension for the Fortuin-Kasteleyn cluster model. In Festschrift for T.E. Harris, to appear.

Kesten, H. and Schonmann, R. (1989). Behavior in large dimensions of the Potts and Heisenberg models. *Reviews in Mathematical Physics*, to appear.

McDiarmid, C. (1980). Clutter percolation and random graphs. *Mathematical Programming Study* 13, 17–25.

Ostrowsky, A.M. (1973). *Solutions of Equations in Euclidean and Banach Spaces.* 3rd edition. Academic Press, New York.

Oxley, J.G. and Welsh, D.J.A. (1979). On some percolation results of J.M. Hammersley. *Journal of Applied Probability* 16, 526–540.

Department of Mathematics
Cornell University
Ithaca
New York 14853.