Consistent estimation of percolation quantities

Ronald Meester   Jeffrey E. Steif*

University of Utrecht
Department of Mathematics
P.O. Box 80.010
3508 TA Utrecht
The Netherlands
meester@math.ruu.nl

Department of Mathematics
Chalmers University of Technology
S-412 96 Gothenburg
Sweden
steif@math.chalmers.se

Abstract

We consider ordinary independent percolation on the integer lattice, and construct consistent estimators for most of the important quantities associated to this model. We also present some simulation results.

1 Introduction

In this paper, we consider ordinary independent bond percolation which is defined as follows. We begin with the hypercubic lattice with vertex set $\mathbf{Z}^d$

*Research supported by grants from the Swedish National Science Foundation and from the Royal Swedish Academy of Sciences

and edge set containing all pairs $x, y \in \mathbf{Z}^d$ with

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

Let $0 \le p \le 1$ and call an edge e **open** with probability p and **closed** with probability $1-p$ independently of the states of all other edges. Considering the random graph containing the vertex set $\mathbf{Z}^d$ and the open edges, the connected components of this (random) graph are called **clusters** and we write C for the cluster containing the origin and more generally $C(x)$ for the cluster containing the x. One is interested in the percolation probability $\theta(p)$ which is defined to be the probability that C is infinite when the parameter p is used. The critical probability p_c is defined as $p_c = \inf\{p : \theta(p) > 0\}$. It is an interesting fact that $p_c \in (0,1)$ whenever $d \ge 2$. When the parameter p is used, we will denote probabilities by P_p and expectations by E_p. See [2] for an excellent treatment of this subject.

The purpose of this paper is the following. Suppose we observe a realization of the ordinary bond percolation model on the d-dimensional integer lattice restricted to the box $B(n) = [-n, n]^d$, where we shall assume throughout that $d \ge 2$. We want to estimate a number of important percolation quantities in such a way that the estimates converge almost surely to the appropriate quantity when the box size tends to infinity. Of course we do not have information about the underlying retention parameter p, though of course it would be easy to estimate p consistently. In particular we do not know a priori whether or not the process is subcritical ($p < p_c$), supercritical ($p > p_c$), or critical ($p = p_c$). In [2], a result of this type is obtained for a particular quantity although this result is not couched in the language of consistent estimators. We briefly describe this here. Let $\kappa(p)$ be $E_p[1/|C|]$

(note $|C|$ (the cardinality of C) is always at least 1). The function κ crops up frequently in calculations, see [2] for references. Given a realization in $B(n)$, we let K_n denote the number of open clusters in $B(n)$ when all edges which have at least one end point outside $B(n)$ are removed. The following result (which is of the type we consider in this paper) is contained in [2].

Theorem 1.1: *For all* $p \in [0,1]$, $K_n/|B(n)| \to \kappa(p)$ P_p *a.s.*

We will need a few more definitions. We say that the set A is connected to the set B if there is a path along open bonds from some vertex in A to some vertex in B. We write '$\to$' to denote 'is connected to'. We shall call two boxes **interior-disjoint** if they have only boundary vertices in common. Two vertices are said to be neighbors if their Euclidean distance is equal to one. The boundary of a box B is defined to be the set of all vertices in B which have at least one vertex outside B as a neighbor, and is denoted ∂B. Note that for two interior-disjoint boxes B and B' and two vertices $v \in B$ and $v' \in B'$, the events $\{v \to \partial B\}$ and $\{v' \to \partial B'\}$ are independent.

We make one final convention. Throughout this paper, all the quantities that arise will depend on the dimension d. Since we are really only interested in how these quantities depend on p, we will ignore the dependence on d in the notation throughout the paper.

2 Estimating the percolation probability

In this section, we want to find a consistent estimator for the percolation probability $\theta(p)$. The first approach is to 'partition' the observed box into interior-disjoint smaller sub-boxes and count how many centers of the smaller boxes are connected to the boundary of their individual box. The advantage of this approach is that different events are independent, which makes the analysis simple. The disadvantage is that we do not seem to make optimal use of all the available statistical information.

The second approach is to consider a box of 'medium size' in the observed box and count how many vertices in this medium sized box are connected to the boundary of the observed box. This approach seems to make better use of all available information but, not surprisingly, it will be harder to show that the corresponding estimator is consistent.

We begin with an estimator based on the first approach. Consider the box $B(n^2)$ and divide this box into n^d interior-disjoint translates of $B(n)$. We define $\tilde{\theta}_{n^2}$ to be the fraction of centerpoints v of the smaller boxes (translates of $B(n)$) which are connected to the boundary $v + \partial B(n)$ of its respective box. For all $n^2 \leq k < (n+1)^2$, we define $\tilde{\theta}_k = \tilde{\theta}_{n^2}$.

Theorem 2.1: *For all $p \in [0,1]$, we have that $\tilde{\theta}_n \to \theta(p)$ P_p a.s.*

Proof: As indicated above, the independence structure of the estimator makes the analysis elementary. Fix $p \in [0,1]$, $\epsilon > 0$ and an integer n. We write

$$\begin{aligned} P_p(\tilde{\theta}_{n^2} \leq \theta(p) - \epsilon) &\leq P_p(\tilde{\theta}_{n^2} \leq P_p(0 \to \partial B(n)) - \epsilon) \\ &\leq P_p(|\tilde{\theta}_{n^2} - P_p(0 \to \partial B(n))| \geq \epsilon) \\ &\leq \frac{\text{Var}_p \tilde{\theta}_{n^2}}{\epsilon^2} \leq \frac{1}{n^d \epsilon^2}. \end{aligned}$$

It follows from this, using the Borel-Cantelli lemma, that $\liminf_{n\to\infty} \tilde{\theta}_n \geq \theta(p)$ P_p a.s.

For the other direction, it suffices to show that for all $\epsilon > 0$ and for all k we have $\limsup_{n\to\infty} \tilde{\theta}_{n^2} \leq P_p(0 \to \partial B(k)) + \epsilon$, P_p a.s., since $\lim_{k\to\infty} P_p(0 \to$

$\partial B(k)) = \theta(p)$. To do this, let $k < n$ and $\tilde{\theta}_{n^2,k}$ be the fraction of centers v of the small boxes which are connected to $v + \partial B(k)$ when we partition $B(n^2)$ into n^d interior-disjoint translates of $B(n)$ as before.

We now write

$$\begin{aligned} P_p(\tilde{\theta}_{n^2} \geq P_p(0 \to \partial B(k)) + \epsilon) &\leq P_p(\tilde{\theta}_{n^2,k} \geq P_p(0 \to \partial B(k)) + \epsilon) \\ &\leq P_p(|\tilde{\theta}_{n^2,k} - P_p(0 \to \partial B(k))| \geq \epsilon) \\ &\leq \frac{\mathrm{Var}_p \tilde{\theta}_{n^2,k}}{\epsilon^2} \leq \frac{1}{n^d \epsilon^2}, \end{aligned}$$

and the conclusion follows as before. □

Next we present an estimator based on the second approach. The proof of convergence is more involved and makes use of deep results in percolation theory. However, we shall see in the last section that this estimator performs much better than the previous one.

Denote by C_n the number of vertices in $B(n)$ which are connected to $\partial B(2n)$. We define $\hat{\theta}_{2n}$ and $\hat{\theta}_{2n+1}$ to be $C_n/(2n+1)^d$.

Theorem 2.2: *For all $p \in [0,1]$, we have that $\hat{\theta}_n \to \theta(p)$ P_p a.s.*

Proof: We distinguish three cases, and show that convergence takes place in all three situations.

CASE 1, $p < p_c$. It is well known (combine [2] and [3]) that for $p < p_c$, there exists $\gamma_p > 0$ such that $P_p(0 \to \partial B(n)) \leq \exp(-\gamma_p n)$. Denote by E_n the event that $C_n \geq 1$. Then

$$\begin{aligned} P_p(E_n) &\leq \sum_{v \in B(n)} P_p(v \to \partial B(2n)) \\ &\leq \sum_{v \in B(n)} P_p(v \to v + \partial B(n)) = (2n+1)^d P_p(0 \to \partial B(n)) \\ &\leq (2n+1)^d \exp(-n\gamma_p). \end{aligned}$$

Hence $\sum_n P_p(E_n) < \infty$ and by Borel-Cantelli this implies that P_p a.s., for all n large enough, we have $\hat{\theta}_n = 0$.

CASE 2, $p > p_c$. It is well known (combine [2] and [3]) that for $p > p_c$, there exist constants $A(p) < \infty$ and $\sigma_p > 0$ such that

$$P_p(0 \to \partial B(n), |C| < \infty) \leq A(p) n^d \exp(-n\sigma_p).$$

Denote by F_n the event that there is a vertex in $B(n)$ connected to $\partial B(2n)$ but not to infinity. Using a similar computation as in Case 1 we find that

$$P_p(F_n) \leq A(p) n^d (2n+1)^d \exp(-n\sigma_p).$$

Hence it follows that $P_p(F_n \text{ i.o.}) = 0$. This means that if we define $\bar{\theta}_n$ to be the fraction of vertices in $B(n)$ which belong to an infinite cluster (which is known to be unique, see [2]), then P_p a.s. we have that for n sufficiently large,

$$\bar{\theta}_n = \hat{\theta}_n.$$

(Of course, $\bar{\theta}_n$ is not observable but that does not matter.) Now the ergodic theorem implies that $\bar{\theta}_n$ converges P_p a.s. to $\theta(p)$, and so we are done.

CASE 3, $p = p_c$. This case is slightly more interesting. We shall need the fact that $\theta(p)$ is continuous from the right (see [2]). First of all, observe that $\bar{\theta}_n \leq \hat{\theta}_n$ where $\bar{\theta}_n$ is as in Case 2. Hence P_{p_c} a.s.

$$\theta(p_c) = \lim_{n\to\infty} \bar{\theta}_n \leq \liminf_{n\to\infty} \hat{\theta}_n \leq \limsup_{n\to\infty} \hat{\theta}_n.$$

(The first equality follows from the ergodic theorem applied to P_{p_c}.) Let $\epsilon > 0$ and let μ be any probability measure on $X \times X$ which is a coupling of P_{p_c} and $P_{p_c+\epsilon}$ satisfying

$$\mu\{(\eta, \eta') : \eta(x) \leq \eta'(x)\ \forall x \in \mathbf{Z}^d\} = 1.$$

Let $\hat{\theta}_n^1$ be the map from $X \times X$ to the reals given by $\hat{\theta}_n^1(\eta, \eta') = \hat{\theta}_n(\eta)$. Similarly, let $\hat{\theta}_n^2$ be the map from $X \times X$ to the reals given by $\hat{\theta}_n^2(\eta, \eta') = \hat{\theta}_n(\eta')$. It follows that μ a.s. $\hat{\theta}_n^1(\eta, \eta') \leq \hat{\theta}_n^2(\eta, \eta')$. Since the distribution of η' under μ is $P_{p_c+\epsilon}$, Case 2 above (with $p = p_c + \epsilon$) implies that for μ a.e. (η, η'), $\lim_{n\to\infty} \hat{\theta}_n^2(\eta, \eta') = \theta(p_c + \epsilon)$. Therefore for μ a.e. (η, η'), $\limsup_{n\to\infty} \hat{\theta}_n^1(\eta, \eta') \leq \theta(p_c + \epsilon)$. Since the distribution of η under μ is P_{p_c}, we obtain the fact that for P_{p_c} a.e. η, $\limsup_{n\to\infty} \hat{\theta}_n(\eta) \leq \theta(p_c + \epsilon)$. Taking the limit as ϵ goes to zero and using the right-continuity of $\theta(p)$ now completes the proof. $\square$

In the last section of this paper, we shall present some simulation results based on both estimators. It turns out (not surprisingly) that the second estimator behaves better than the first.

3 Estimating the indicator function $I_{\{\theta>0\}}$

Even though we know how to estimate θ, it is not obvious (since $I_{\{\theta>0\}}$ is not a continuous function of θ) that it is in fact possible to estimate $I_{\{\theta>0\}}$ consistently. Rather than boxes $B(n)$ and $B(2n)$ as in the previous section, we consider here boxes $B(n)$ and $B(n + n^n)$. We define $\hat{\beta}_{n+n^n}$ to be 1 if there is a vertex in $B(n)$ which is connected to $\partial B(n + n^n)$, and defined to be 0 otherwise. For k satisfying $n+n^n \leq k < (n+1)+(n+1)^{n+1}$, we define $\hat{\beta}_k$ to be $\hat{\beta}_{n+n^n}$.

Theorem 3.1: *For all $p \in [0, 1]\backslash p_c$, we have $\hat{\beta}_n = I_{\{\theta(p)>0\}}$ for all large n P_p a.s. If either (i) $\theta(p_c) > 0$ or (ii) $\theta(p_c) = 0$ and there exists $\rho, C \in (0, \infty)$ such that $P_{p_c}(0 \to \partial B(n)) \leq Cn^{-1/\rho}$ for all n, then we have $\hat{\beta}_n = I_{\{\theta(p_c)>0\}}$ for all large n, P_{p_c} a.s.*

While it is not even known that at least one of (i) and (ii) above necessarily holds, it is a widely accepted fact that (ii) should hold. This has in fact been proved for d sufficiently high (see [2]).

Proof: CASE 1, $p < p_c$. It is easy to adapt the arguments of the second approach in the previous section to also show that for $p < p_c$ we have $P_p(\hat{\beta}_n = 1 \text{ i.o.}) = 0$.
CASE 2, $\theta(p) > 0$ (i.e., $p > p_c$ or $p = p_c$ together with (i)). In this case, it is completely obvious that we have that $P_p(\hat{\beta}_n = 0 \text{ i.o.}) = 0$. (This is because if $\theta(p) > 0$, then some vertex percolates P_p a.s.)
CASE 3, $p = p_c$ and (ii) holds. By a computation similar to the proof of Theorem 2.2, we obtain

$$P_{p_c}(\hat{\beta}_n = 1) \leq C(2n+1)^d n^{-n/\rho}.$$

Borel-Cantelli as before now completes the proof. □

There is an analogue to the easy first approach for a consistent estimator for $\theta(p)$, though things are a little more complicated here. To explain this, suppose that our estimator is based on $k(n)$ interior-disjoint translates of $B(n)$. As a first try, one might use the estimator which is equal to 1 if there is at least one centervertex v of these boxes which is connected to $v+\partial B(n)$, and 0 otherwise. This would be the analogue of $\tilde{\theta}_n$ in the previous section. However, this approach does not work as can be seen as follows. For $p > p_c$, the probability that there is no such centervertex is equal to $(1-P_p(0 \to \partial B(n)))^{k(n)} \leq (1-\theta(p))^{k(n)}$. Our estimator can then be shown to be consistent if $\sum_n (1-\theta(p))^{k(n)} < \infty$, using Borel-Cantelli as usual. Note that $1-\theta(p)$ can in principle be as close to 1 as desired, so in fact the requirement for the sequence $k(n)$ is that $\sum_n c^{k(n)} < \infty$ for all $c < 1$. On the other hand, under assumption (ii) above, the probability at p_c that the estimator is equal to 1 is bounded from above by $k(n)n^{-1/\rho}$ so in order to apply Borel-Cantelli we need to require that $\sum_n k(n)n^{-1/\rho} < \infty$. These two requirements are incompatible, because the first requirement implies that $k(n)$ tends to infinity, which is incompatible with the second requirement if ρ were to take the value 2, for instance. On the other hand, in high dimensions, ρ is believed to be equal to $1/2$. If this is the case, it is easy to see that if we take $k(n) = n^{1/2}$ then both requirements above are in fact satisfied.

However, there is an estimator based on interior-disjoint translates of $B(n)$ which does work. Consider the box $B(n^2)$ containing n^d interior-disjoint translates of $B(n)$. Let $\tilde{\beta}_{n^2}$ be 1 if at least a fraction $1/\log n$ of the centervertices of the translates of $B(n)$ are connected to their boundary, and 0 otherwise. Intermediate values of $\tilde{\beta}_n$ are defined as usual.

Theorem 3.2: *The statements in Theorem 3.1 remain true if we replace $\hat{\beta}_n$ by $\tilde{\beta}_n$ throughout.*

Proof: Suppose first that $\theta(p) > 0$ and write r_n for $1/\log n$. Let, for all n, Y_n be a random variable with a binomial distribution with parameters n^d and $\theta(p)$. We can write, using the fact that $P_p(0 \to \partial B(n)) \geq \theta(p)$ and Markov's inequality,

$$\begin{aligned} P_p(\tilde{\beta}_{n^2} = 0) &\leq P(Y_n \leq r_n n^d) \\ &\leq e^{r_n n^d} E(e^{-Y_n}) \\ &= e^{r_n n^d}\{e^{-1}\theta(p) + 1 - \theta(p)\}^{n^d} \\ &= (e^{r_n} c(p))^{n^d}, \end{aligned}$$

for some constant $c(p) < 1$. Using the fact that $\lim_{n\to\infty} r_n = 0$, it is easy to see that this is summable over n. Hence it follows from the Borel-Cantelli lemma that $P_p(\tilde{\beta}_{n^2} = 0 \text{ i.o.}) = 0$, as desired.

For $p < p_c$, it is easy to see that $P_p(\tilde{\beta}_{n^2} = 1 \text{ i.o.}) = 0$, using the exponential decay of the radius distribution as before.

Finally, suppose $p = p_c$ and (ii) in Theorem 3.1 holds. Let, for all n sufficiently large, Z_n be a random variable with a binomial distribution with parameters n^d and $Cn^{-1/\rho}$. We write

$$\begin{aligned} P_p(\tilde{\beta}_{n^2} = 1) &\leq P(Z_n \geq r_n n^d) \\ &\leq e^{-r_n n^d} E(e^{Z_n}) \\ &= \{e^{-r_n}(1 + Cn^{-1/\rho}(e-1))\}^{n^d}. \end{aligned}$$

Using $1 + Cn^{-1/\rho}(e-1) \leq \exp(Cn^{-1/\rho}(e-1))$, it is not hard to show that this is again summable over n and the conclusion follows as before. □

4 Estimating the expected cluster size

A third quantity of interest is the expected cluster size $\chi(p) := E_p(|C|)$. It is well known (see again [2]) that $\chi(p) = \infty$ for $p \geq p_c$ and that $\chi(p) < \infty$ for $p < p_c$. To estimate $\chi(p)$ we shall again use boxes $B(n)$ and $B(n + n^n)$. Let $\hat{\chi}_{n+n^n}$ be the average over the vertices in $B(n)$ of the size of the cluster of that vertex when all edges outside $B(n + n^n)$ are removed. For k satisfying $n + n^n \leq k < (n+1) + (n+1)^{n+1}$, we define $\hat{\chi}_k$ to be $\hat{\chi}_{n+n^n}$.

Theorem 4.1: *For all $p \in [0,1] \backslash p_c$, we have that $\hat{\chi}_n \to \chi(p)$ P_p a.s. If either (i) $\theta(p_c) > 0$ or (ii) $\theta(p_c) = 0$ and there exists $\rho, C \in (0, \infty)$ such that $P_{p_c}(0 \to \partial B(n)) \leq Cn^{-1/\rho}$ for all n, then we have $\hat{\chi}_n \to \chi(p_c)$ P_{p_c} a.s.*

Proof: Case 1, $p < p_c$. Using the same Borel-Cantelli argument as before, we see that after a random time, no vertex of $B(n)$ will be connected to the boundary of $B(n + n^n)$ and then the estimator is exactly the average over the vertices in $B(n)$ of the size of its respective cluster in the whole space. Then consistency follows immediately from the ergodic theorem.

Case 2, $\theta(p) > 0$ (i.e., $p > p_c$ or $p = p_c$ together with (i)). As already noted in the previous section, there is P_p a.s. some vertex which percolates which means that for n large enough, there will be at least one vertex in $B(n)$ which is connected to the boundary of $B(n + n^n)$. The cluster of this vertex inside the bigger box necessarily has size at least n^n, so the average over all vertices of $B(n)$ is at least $(2n+1)^{-d}n^n$ which tends to infinity when $n \to \infty$.

Case 3, $p = p_c$ and (ii) holds. In this case, the probability that there is a vertex in $B(n)$ connected to the boundary of $B(n + n^n)$ is at most $C(2n + 1)^d n^{-n/\rho}$. Hence for n large enough there will be no such vertices a.s. and the result follows from the ergodic theorem as in Case 1 above. □

Some readers might be bothered by the fact that $\hat{\chi}_n$ never takes the value ∞ even though $\chi(p)$ can be infinity. It is easy to adapt the above estimator to achieve this: define $\hat{\chi}_{n+n^n}$ to be infinity if there is at least one vertex inside $B(n)$ which is connected to $\partial B(n + n^n)$. We are no longer concerned with finding consistent estimators which are analogous to the first approach used for estimating $\theta(p)$.

5 Estimating limits

Consider the following general setup: Let $\mathcal{F}_n$ be the σ-field generated by all edges which have at least one end point in $B(n-1)$, and let A_n be a sequence of events such that $A_n \in \mathcal{F}_n$ for all n. For example, we can take A_n to be the event $\{0 \to \partial B(n)\}$. Let $\alpha > 0$. We are interested in estimating

$$\lim_{n\to\infty} \frac{-\log P_p(A_n)}{n^\alpha},$$

whenever this limit exists. This is not always known to be the case, and therefore in order to make the results as general as possible, we have chosen the following formulation for our next result. While this may appear overly general or abstract, the applications following it will justify this formulation.

Theorem 5.1: *Let A_n be a sequence of events such that $A_n \in \mathcal{F}_n$ for all n. Suppose that there exists $S \subseteq [0,1]$ and $b_n \in (0,1)$ such that for all $p \in S$, $P_p(A_n) \geq b_n$ for all n large enough (possibly depending on p). Let $\alpha > 0$ and let $k(n)$ be the smallest number of the form m^d for some integer m such that*

$$\frac{1}{k(n)b_n(\exp(n^{\alpha-1})-1)^2} + \frac{1}{k(n)b_n(1-\exp(-n^{\alpha-1}))^2} \leq \left(\frac{1}{2}\right)^n.$$

Let D_n be the fraction of the interior-disjoint translates of $B(n)$ in the box $B(k(n)^{1/d}n)$ where the appropriate translate of the event A_n occurs. Then for all $p \in S$,

$$\left|\frac{-\log D_n}{n^\alpha} + \frac{\log P_p(A_n)}{n^\alpha}\right| \to 0, \quad P_p \text{ a.s.}$$

Proof: We write $q_n = P_p(A_n)$. (Note that q_n depends on p but we won't make this dependence explicit in the notation.) We then have for all $p \in S$,

$$
\begin{aligned}
&P_p\left(\left|\frac{-\log D_n}{n^\alpha} + \frac{\log q_n}{n^\alpha}\right| \geq \frac{1}{n}\right) = P_p(|\log(D_n/q_n)| \geq n^{\alpha-1}) \\
&= \quad P_p(D_n \geq q_n \exp(n^{\alpha-1})) + P_p(D_n \leq q_n \exp(-n^{\alpha-1})) \\
&\leq \quad P_p(|D_n - q_n| \geq q_n(e^{n^{\alpha-1}} - 1)) + P_p(|D_n - q_n| \geq q_n(1 - e^{-n^{\alpha-1}})) \\
&\leq \quad \frac{q_n(1-q_n)}{k(n)q_n^2(\exp(n^{\alpha-1}) - 1)^2} + \frac{q_n(1-q_n)}{k(n)q_n^2(1 - \exp(-n^{\alpha-1}))^2} \\
&\leq \quad \frac{1}{k(n)b_n(\exp(n^{\alpha-1}) - 1)^2} + \frac{1}{k(n)b_n(1 - \exp(-n^{\alpha-1}))^2},
\end{aligned}
$$

where the last inequality is true for n large enough. It follows that

$$P_p\left(\left|\frac{-\log D_n}{n^\alpha} + \frac{\log P_p(A_n)}{n^\alpha}\right| \geq \frac{1}{n}\right) \leq \left(\frac{1}{2}\right)^n,$$

and the result follows form the Borel-Cantelli lemma. □

We shall now give a number of applications of Theorem 5.1. It is well known (see e.g. [2]) that there exists a unique map $\varphi : [0,1] \to [0,\infty]$ and positive constants ρ and σ such that

$$\rho n^{1-d} e^{-n\varphi(p)} \leq P_p(0 \to \partial B(n)) \leq \sigma n^{d-1} e^{-n\varphi(p)}.$$

Note that it follows that $\varphi(p)$ can be identified as

$$\varphi(p) = \lim_{n\to\infty} \frac{-\log P_p(0 \to \partial B(n))}{n}.$$

It is well known (see [2], Theorem 5.14) that φ is continuous and non-increasing on $(0,1]$, strictly decreasing and positive on $(0, p_c)$ and satisfies $\varphi(p_c) = 0$ and $\lim_{p\to 0} \varphi(p) = \infty$.

Corollary 5.2: *Let $k(n)$ be the smallest integer of the form m^d for some integer m such that*

$$\frac{n^n}{k(n)(e-1)^2} + \frac{n^n}{k(n)(1-e^{-1})^2} \leq \left(\frac{1}{2}\right)^n.$$

Let A_n be the event $\{0 \to \partial B(n)\}$. Let D_n be the fraction of the $k(n)$ interior-disjoint translates of $B(n)$ in $B(k(n)^{1/d}n)$ where the appropriate translate of A_n occurs. Then for all $p \in [0,1]$. we have

$$\frac{-\log D_n}{n} \to \varphi(p), \quad P_p \text{ a.s.}$$

Proof: Let $S = (0,1]$, $b_n = (\frac{1}{n})^n$ and $\alpha = 1$. Note that for all $p \in S$, $P_p(A_n) \geq p^n \geq b_n$ for n sufficiently large. Using the fact that

$$\frac{-\log P_p(A_n)}{n} \to \varphi(p),$$

for all $p \in [0,1]$, it follows from Theorem 5.1 that for all $p \in (0,1]$,

$$\frac{-\log D_n}{n} \to \varphi(p), \quad P_p \text{ a.s.}$$

Since $\varphi(0) = \infty$, this also holds P_0 a.s. □

Another quantity of interest is

$$\sigma(p) = \lim_{n\to\infty} \frac{-\log P_p(0 \to \partial B(n), |C| < \infty)}{n}.$$

(It is shown in [2] that this limit exists for all p.) At first sight it seems that $\sigma(p)$ might be harder to estimate than $\varphi(p)$ because the events involved require information of the configuration in the whole space and therefore are not $\mathcal{F}_n$ measurable. But $\sigma(p)$ can also be identified as

$$\begin{aligned}\sigma(p) &= \lim_{n\to\infty} \frac{-\log P_p(\text{diam}(C) = n)}{n} \\ &= \lim_{n\to\infty} \frac{-\log P_p(\text{diam}(C) = n-1)}{n},\end{aligned}$$

where $\text{diam}(C)$ is the side length of the smallest box which contains C. (See [2] for a discussion of these matters.) The last equality in the displayed equation is included to make sure that the event involved is $\mathcal{F}_n$ measurable so that we can apply Theorem 5.1.

Corollary 5.3: *Let $k(n)$ be defined as in Corollary 5.2, and let A_n be the event $\{diam(C) = n-1\}$. Let D_n be defined as in Corollary 5.2 (with the new events A_n). Then for all $p \in [0,1]$ we have*

$$\frac{-\log D_n}{n} \to \sigma(p), \quad P_p \text{ a.s.}$$

Proof: This time we choose $S = (0,1)$, $b_n = (\frac{1}{n})^n$ and $\alpha = 1$. We then have that for all $p \in S$, $P_p(\mathrm{diam}(C) = n-1) \geq p^{n-1}(1-p)^{2(nd-n+1)} \geq b_n$ for n sufficiently large. Now Theorem 5.1, together with the fact that for all $p \in (0,1)$,

$$\lim_{n\to\infty} \frac{-\log P_p(\mathrm{diam}(C) = n-1)}{n} = \sigma(p),$$

implies that the required convergence takes place P_p a.s., for all $p \in (0,1)$. Since $\sigma(0) = \sigma(1) = \infty$, this is also the case for $p = 0$ and $p = 1$. □

Finally, we discuss limits which involve the events $\{|C| = n\}$. Note that this event is $\mathcal{F}_n$ measurable since the origin always belongs to C. It is known that

$$\xi(p) = \lim_{n\to\infty} \frac{-\log P_p(|C| = n)}{n}$$

exists for all $p \in [0,1]$ and satisfies $\xi(p) = 0$ for $p > p_c$ and $\xi(p) > 0$ for $p < p_c$. In order to obtain a nontrivial limit in the supercritical phase, one has to use a different scaling. It is conjectured that

$$\delta(p) = \lim_{n\to\infty} \frac{-\log P_p(|C| = n)}{n^{(d-1)/d}}$$

exists for all p and is positive and finite for all $p > p_c$. (Of course $\delta(p)$ is equal to ∞ whenever $p < p_c$.)

Corollary 5.4: *Let $k(n)$ be the smallest number of the form m^d for some integer m such that*

$$\frac{n^n}{k(n)(\exp(n^{-1/d})-1)^2} + \frac{n^n}{k(n)(1-\exp(-n^{-1/d}))^2} \leq \left(\frac{1}{2}\right)^n.$$

Let A_n denote the event $\{|C| = n\}$ and define D_n as in Corollary 5.2. Then for all $p \in [0,1]$ we have

$$\frac{-\log D_n}{n} \to \xi(p), \quad P_p \text{ a.s.}$$

and

$$\left|\frac{-\log D_n}{n^{(d-1)/d}} + \frac{\log P_p(A_n)}{n^{(d-1)/d}}\right| \to 0, \quad P_p \text{ a.s.}$$

In particular, if $\delta(p)$ exists, we have for all $p \in [0,1]$,

$$\frac{-\log D_n}{n^{(d-1)/d}} \to \delta(p), \quad P_p \text{ a.s.}$$

Proof: The proof is similar to the proof of the preceding corollaries and is therefore omitted. (Note that the choice of $k(n)$ is not optimal for the estimation of $\xi(p)$.) □

6 Simulations

In this last section we present some simulation results. It turns out that the programming is a little easier with *site* percolation rather than with bond percolation. In site percolation, we declare the *vertices* of $\mathbf{Z}^d$ open and closed with probabilities p and $1-p$ rather than the edges. The cluster C of the origin is now defined as the set of vertices which can be reached from the origin by walking along open vertices only. Working with site percolation has some (unimportant) consequences for the way we define our estimators. For instance, in our definition of $\tilde{\theta}_n$ we used the fact that n^d interior-disjoint translates of $B(n)$ fit exactly in $B(n^2)$. For site percolation, interior-disjointness is not enough to guarentee that the different subboxes behave independently; the subboxes need to be completely disjoint. Our simulations are for $d = 2$. We used $B(126)$ in which 121 disjoint translates of $B(11)$ fit exactly. The medium size box needed for $\hat{\theta}_{126}$ is just $B(63)$.

The critical probability for two-dimensional independent site percolation is not known rigorously, though it has been proved that it is contained in $[0.556, 2/3]$; see [1] and [4]. Experiments indicate that the value is approximately 0.593. We tested the behavior of the estimators $\tilde{\theta}_n$ and $\hat{\theta}_n$ for $p = 0.5$, $p = 0.55$ and $p = 0.8$. Since 0.5 and 0.55 are both subcritical, we know that $\theta(p)$ is equal to zero for these cases. Therefore, testing these values gives us some information about the quality of the estimators. We did 100 experiments and computed the sample mean of the estimators, and also the sample variance. Here are our results:

p	mean of $\tilde{\theta}_n$	variance of $\tilde{\theta}_n$	mean of $\hat{\theta}_n$	variance of $\hat{\theta}_n$
0.50	0.173	$1.0 \cdot 10^{-3}$	0.000	0
0.55	0.357	$2.1 \cdot 10^{-3}$	0.007	$2.3 \cdot 10^{-4}$
0.80	0.797	$1.8 \cdot 10^{-3}$	0.798	$4.3 \cdot 10^{-6}$

It seems that the convergence of $\tilde{\theta}_n$ is particularly slow, and it is clear from this table that $\hat{\theta}_n$ performs much better.

A standard technique to improve an estimator is to use the theorem of Rao-Blackwell. We now apply this technique to $\hat{\theta}_n$. Suppose we are faced with a realization in the box $B(n)$, that is, with $0-1$ random variables X_z, $z \in B(n)$, where $X_z = 1$ iff the vertex z is open. The number of open vertices Y in the box $B(n)$ is a sufficient statistic for p, and therefore also for $\theta(p)$. The theorem of Rao-Blackwell now tells us that

$$h(Y) = E(\hat{\theta}_n \,|\, Y)$$

has a smaller variance than $\hat{\theta}_n$. We would like to use $h(Y)$ as an estimator for $\theta(p)$, but we cannot compute this conditional expectation. However, given Y, to estimate $h(Y)$ we can repeat the experiment k times, each time conditioned on the number of open vertices to be Y. (Note that conditioned on the number of open vertices to be y, a realization with the correct conditional distribution can be obtained by distributing these y open vertices uniformly over the box.) The average value of $\hat{\theta}_n$ corresponding to these k new experiments is of course an estimate of $h(Y)$ and therefore can be used as our new estimator for $\theta(p)$.

We can try to get an idea of the numerical improvement provided by this procedure. We did this for $\hat{\theta}_n$ with $p = 0.80$. First, we did 25 independent experiments, each time recording the number of open vertices. Denote these numbers by Y_i, $i = 1, \ldots, 25$. We then performed, for each i, 30 further experiments conditioned on the number of open points to be Y_i. This led, for each i, to an estimate of $\theta(p)$. Finally, we computed the sample mean and variance of these 25 estimates. The sample mean turned out to be 0.798 and the sample variance $6.7 \cdot 10^{-7}$. The variance has to be compared with the variance of $\hat{\theta}_n$ in a sample of size 25. (The table above is with sample size 100.) This turns out to be $4.1 \cdot 10^{-6}$. The Rao-Blackwell method gives an improvement by roughly a factor 10.

Acknowledgement: We thank Li Chi Wang for her help in the last section.

References

[1] Berg, J.v.d. and Ermakov, A., (1996) A new lower bound for the critical probability of site percolation on the square lattice, *Random Structures and Algorithms*, 199-212.

[2] Grimmett, G.R., (1989), *Percolation*, Springer-Verlag.

[3] Grimmett, G. R. and Marstrand, J. M., (1990) The supercritical phase of percolation is well behaved. *Proceedings of the Royal Society of London A*, **430**, 439–457.

[4] Kesten, H., (1982), *Percolation Theory for Mathematicians*, Birkhäuser.