Numerical results for crossing, spanning and wrapping in two-dimensional percolation

Numerical results for crossing, spanning and wrapping in two-dimensional percolation

Gunnar Pruessner[1] **and Nicholas R Moloney**[2]

[1] Department of Mathematics, Imperial College London, 180 Queen's Gate,
London SW7 2BZ, UK
[2] Beit Fellow for Scientific Research, Condensed Matter Theory, Blackett Laboratory,
Imperial College London, Prince Consort Rd, London SW7 2BW, UK

E-mail: gunnar.pruessner@physics.org

Abstract
Using a recently developed method to simulate percolation on large clusters of
distributed machines [1], we have numerically calculated crossing, spanning
and wrapping probabilities in two-dimensional site and bond percolation with
exceptional accuracy. Our results are fully consistent with predictions from
conformal field theory. We present many new results that await theoretical
explanation, particularly for wrapping clusters on a cylinder. We therefore
provide possibly the most up-to-date reference for theoreticians working on
crossing, spanning and wrapping probabilities in two-dimensional percolation.

PACS numbers: 64.60.Ak, 05.70.Jk

1. Introduction

In the last decade, percolation has enjoyed the attention of conformal field theorists who have
sought to calculate crossing probabilities for various aspect ratios and geometries. In rough
terms, calculations generally involve mapping percolation to a 1-state Potts model, constructing
a correlation function corresponding to the boundary conditions necessitated by the crossing
cluster of interest, and finding a differential equation thereof. In 1991, Langlands *et al* [2]
were the first to investigate 'the universality of crossing probabilities in two-dimensional
percolation'. Shortly afterwards, Cardy [3] obtained an exact equation for the probability of
a crossing cluster on a rectangle for different aspect ratios, using conformal field theory.

In 1996, Watts extended Cardy's results to obtain an exact equation for the probability
of a cluster crossing both horizontally and vertically [4]. In the same year, after numerical
work by Hu and Lin [5], Aizenman proved that the probability of more than one crossing
cluster is finite in the thermodynamic limit [6]. This does not contradict the rigorous result
that the number of infinite clusters is 0, 1 or ∞ with probability 1 [7, 8]. The asymptote for

the probability of obtaining n distinct, simultaneously crossing clusters has been calculated by Cardy [9], who also considers spanning on a cylinder.

We define all these probabilities systematically below. In this paper, we provide accurate numerical data for the many analytical results now available, many of which are verified for the first time. We also provide data, particularly for wrapping on the cylinder, for which there is currently no theoretical explanation. After briefly discussing the results for some exotic cluster configurations, we present and discuss results for crossing, spanning and wrapping probabilities on rectangles and cylinders of various aspect ratios for site and bond percolation. The results are of such accuracy that we are able to make firm statements regarding the validity of various estimates, conjectures and formulae made in the past. The appendix collects together some technical notes on identifying various cluster types when simulating percolation.

2. Observables

In this section we give the various observables that we measured for site and bond percolation. We consider a square lattice in which sites are linked via bonds. To avoid confusion, we stress that boundaries are made of sites only [2]. In site percolation, sites are occupied with probability $p^{(s)}$ and all bonds are active. In bond percolation all sites are occupied, while bonds are active with probability $p^{(b)}$. Two sites belong to the same cluster if they are both occupied and if there is a path between them along active bonds and occupied sites. A cluster is the set of sites belonging to it.

2.1. Open boundary conditions

Each cluster can be characterized by the borders it touches. If the borders of the lattice are labelled N, E, S, W, as in figure 1, then each cluster is assigned a subset of these labels, indicating the borders it touches. There are 15 different combinations of cluster labels in which at least one border is touched. In the following, these combinations are called 'types', S_Ω. A cluster of type S_Ω touches all borders in the set Ω. For example, the three clusters shown in figure 1 are of types $S_{\{N,E\}}$, $S_{\{S,W\}}$ and $S_{\{W\}}$.

We distinguish between clusters touching *only* a set of borders and *at least* a set of borders. A cluster of type S_Ω is also of type $U_{\Omega'}$ if $\Omega \supseteq \Omega'$. For example, a cluster of type $S_{\{S,W\}}$ is also of types $U_{\{S\}}$, $U_{\{W\}}$ and $U_{\{S,W\}}$. A 'crossing cluster', i.e. a cluster that connects two opposite borders, is therefore either of type $U_{\{N,S\}}$ or $U_{\{E,W\}}$.[3]

Normalized histograms $\mathcal{P}_N(\mathsf{T}, n, r)$ have been generated for all 15 S and U types, where N is the size of the lattice (number of sites) and r is its aspect ratio (length over height). The histogram estimates the probability that a random realization contains n clusters of type T. It is worthwhile pointing out that even though S and U histograms are correlated, they cannot be derived from each other, because the number of clusters of a particular U-type must be determined on a per-realization basis.

The moments associated with each histogram are defined by

$$\mathcal{M}_N(\mathsf{T}, m, r) = \sum_{n=0}^{\infty} \mathcal{P}_N(\mathsf{T}, n, r) n^m. \tag{1}$$

A hat on a quantity, for example $\widehat{\mathcal{M}}$, indicates its expectation value, i.e. the value in the limit of an infinitely large ensemble and system size.

---

[3] In the following, we will apply the term 'crossing' only to systems with open boundaries, and reserve the terms 'spanning' and 'wrapping' for cylindrical boundary conditions.
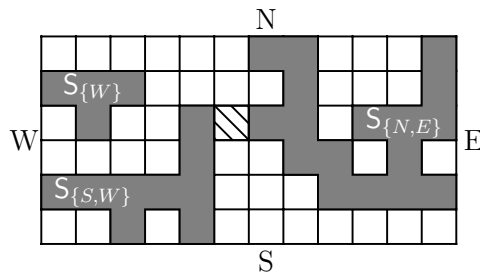
Figure 1. A realization of a lattice which is consistent with figure 8. As indicated, the clusters are of types $S_{\{N,E\}}$, $S_{\{S,W\}}$ and $S_{\{W\}}$. The hatched square in the centre will be filled to give rise to the border configuration shown in figure 9.
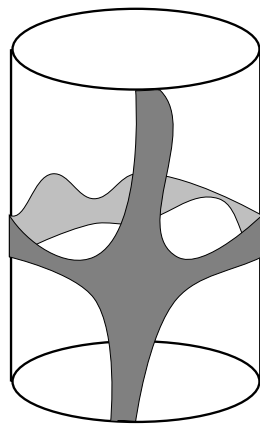

Figure 2. A realization of percolation on a cylinder that would contribute to $\mathcal{P}_N(\mathsf{O}^+, 1, r)$ and $\mathcal{P}_N(\mathsf{I}^+, 1, r)$, but not to $\mathcal{P}_N(\mathsf{O}, 1, r)$ or $\mathcal{P}_N(\mathsf{I}, 1, r)$.

2.2. Cylinder

In addition to the type classification above, we have identified wrapping and spanning clusters on a cylinder. Our convention for cylinders is illustrated in figure 2, which defines the aspect ratio as circumference over height. A wrapping cluster is then a cluster which winds around the cylinder, i.e. it provides a path of length 2π. A spanning cluster, meanwhile, connects the bottom and top of the cylinder.

The resulting histograms are $\mathcal{P}_N(\mathsf{I}, n, r)$ for the number distribution of clusters that only span (rather than wrap) the cylinder, and $\mathcal{P}_N(\mathsf{O}, n, r)$ for the number distribution of clusters that only wrap (rather than span) the cylinder.

In addition, $\mathcal{P}_N(\mathsf{I}^+, 1, r)$ counts the number of realizations with a single spanning cluster that may also be wrapping. Similarly, $\mathcal{P}_N(\mathsf{O}^+, 1, r)$ counts the number of realizations with a single wrapping cluster that may also be spanning. For an example, see figure 2. The distinction between I and I^+ (O and O^+) disappears for more than one spanning (wrapping) cluster, because the existence of two or more simultaneously spanning (wrapping) clusters prohibits wrapping (spanning) at the same time.

3. Results

We have generated the histograms mentioned above for site ($p^{(s)} = 0.592\,746\,21$ [10]) and bond percolation ($p^{(b)} = 1/2$ [11]), each for three different system sizes, $N = 30000^2$, $3000^2, 300^2$. If not mentioned explicitly, the results presented are for $N = 30000^2$. Henceforth, site percolation is indicated by a superscript (s), and bond percolation by (b). In each of these simulations, 14 different aspect ratios were realized while keeping the area constant [12]. These aspect ratios are: 30/30, 36/25, 45/20, 50/18, 60/15, 75/12, 90/10, 100/9, 150/6, 180/5, 225/4, 300/3, 450/2 and 900/1. Three different boundary conditions were applied to each aspect ratio, corresponding to an open system, a cylinder glued vertically and a cylinder glued horizontally. 'Gluing' is our technical term here referring to the procedure for applying periodic boundaries. For the definition and explanation of these technicalities, see [1] and the appendix. The random number generator used is described in [13]. Between 10^6 and 2×10^6 independent samples were produced for each system size and percolation type, using up to 61 undergraduate computers when idle.

The different aspect ratios are derived from the same set of patches (see the appendix), by gluing them together to form rectangles in random permutations and orientations (for details see [1]), and are therefore *not* statistically independent. However, their correlations are assumed to be *very* small. When considering the outcome for several aspect ratios at the same time, one could multiply the error bars by the square root of the number of aspect ratios considered, as if each aspect ratio were based only on a subset of the original sample. Because this procedure assumes maximum correlations between the patches, even though they are randomly permuted, rotated and mirrored between different aspect ratios, this is a strong overestimation of the correlations.

We group our results as follows:

- corner clusters of types $S_{\{N,E\}}$, $S_{\{E,S\}}$, $S_{\{S,W\}}$ and $S_{\{N,W\}}$;
- 'three-legged' clusters of types $S_{\{N,E,S\}}$, $S_{\{E,S,W\}}$, $S_{\{N,S,W\}}$ and $S_{\{N,E,W\}}$;
- types $U_{\{N,S\}}$ and $U_{\{E,W\}}$, the subject of Cardy's predictions [3, 9], as well as spanning and wrapping clusters on a cylinder;
- type $S_{\{N,E,S,W\}}$, in answer to Watts' prediction.

Apart from some straightforward arguments, there are no theoretical predictions for corner and three-legged clusters, which we will refer to as 'exotic' types.

3.1. Exotic types

3.1.1. Corner clusters. For completeness, we include the following results for corner clusters, i.e. types $S_{\{N,E\}}$, $S_{\{E,S\}}$, $S_{\{N,W\}}$ and $S_{\{S,W\}}$. In a rectangle, all corners are equivalent. Naively one expects these clusters to be arranged like onion skins. Assuming scale invariance, this suggests a logarithmic dependence of their average number $\mathcal{M}_N(S, 1, r)$ on the lattice size. This, however, is not the case, as can be seen in figure 3, which should show collapsing lines if N enters only as a factor $\ln(N)$. It is clear that the graph must level off for $r \to \infty$, but this region is not yet reached, even for $r = 900$ and any N. The lines are remarkably straight, but they seemingly cannot define a universal exponent.

In the Ising model, the corner magnetization has been calculated analytically by Davies and Peschel [14] using a corner transfer matrix approach [15]. A similar approach seems to be suitable for corner clusters in percolation.
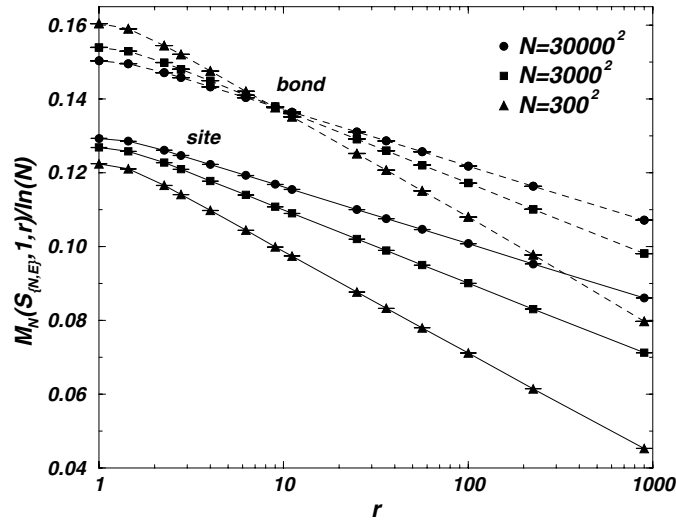
Figure 3. $\mathcal{M}_N(\mathsf{S}_{\{N,E\}}, 1, r)/\ln(N)$ for $N = 300^2, 3000^2, 30000^2$ and site and bond percolation.

3.1.2. Three-legged clusters. Three-legged clusters touch three borders in a T-like manner. There are four different types of three-legged clusters, but the symmetry of a rectangle splits them into two pairs, while a rotation of the system by $\pi/2$ transforms r into $1/r$ so that

$$\widehat{\mathcal{P}}(\mathsf{S}_{\{N,E,S\}}, 1, r) = \widehat{\mathcal{P}}(\mathsf{S}_{\{N,S,W\}}, 1, r) = \widehat{\mathcal{P}}(\mathsf{S}_{\{N,E,W\}}, 1, r^{-1}) = \widehat{\mathcal{P}}(\mathsf{S}_{\{E,S,W\}}, 1, r^{-1}).$$

The statistics of these clusters turns out to be universal. This is possibly not very surprising, since they involve a crossing path (see below). However, what is surprising is the asymptotic value of $\mathcal{P}(\mathsf{S}_{\{N,E,S\}}, 1, r)$ in the large r limit, when the vertical crossing probability approaches 1. At $r = 100$ and $N = 30000^2$, these asymptotic values are 0.5004(4) and 0.5002(3), for site and bond percolation respectively. At $r = 900$, the values are 0.4999(4) and 0.5004(3). Analysing smaller system sizes reveals that this quantity is very sensitive to finite-size effects: while even the smallest system size reaches a value very close to 0.5 at $r \approx 6$, for small N deviations towards higher probabilities occur at large r. Evidently, the probability of a three-legged cluster tends to the occupation probability of sites on the lattice for $r \to N$, where the system effectively becomes a one-dimensional strip of length N and height 1. The region where the probability remains close to 0.5 becomes larger as the system size increases, and for $N = 30000^2$ we were unable to detect any significant deviation from 0.5 for $r > 9$. We therefore conjecture that

$$\lim_{r \to \infty} \widehat{\mathcal{P}}(\mathsf{S}_{\{N,E,S\}}, 1, r) = \tfrac{1}{2}. \tag{2}$$

This can be understood as the probability of intersecting a vertically crossing cluster when vertically cutting a long, narrow percolating system. However, it remains unclear how to derive this limit analytically.

3.2. Results related to conformal field theory

We now present numerical results that are related to theoretical predictions based on conformal field theory. This includes measurements of many formerly unknown quantities, conjectures and comparisons with exact results, which all give further support to the conformal invariance of critical percolation.

3.2.1. Crossing probability with open boundaries. Cardy's seminal paper [3] contained a comparison between his exact result for the crossing probability and the numerical results obtained by Langlands *et al* [12]. These were based on systems with $N = 200^2$ sites and $r = 1 \cdots 7.35$. Later studies by Shchur and Kosyakov [16, 17] investigated $\mathcal{P}_N(\mathsf{U}_{\{N,S\}}, n, r)$ at $r = 1$ for $n > 1$ with very small systems, $N \leqslant 64^2$. Shchur [18] later extended these results to systems up to size $N = 256 \times 3200$, apparently still encountering finite-size corrections. Other studies, such as Sen [19, 20] and Hovi and Aharony [21], used similar system sizes, while also considering other properties of spanning clusters.

Using the data presented in this paper, it is possible to compare Cardy's prediction with much greater accuracy, based on systems with $N = 30000^2$ sites and $r = 1 \cdots 900$. However, for very large (and very small) values of r, relevant clusters become either too rare to give any reasonable estimate for the associated probabilities, or their number count becomes too broadly distributed.

The crossing probability is the probability of finding at least one crossing cluster in a particular direction. By symmetry

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, n, r) = \widehat{\mathcal{P}}(\mathsf{U}_{\{E,W\}}, n, r^{-1}) \tag{3}$$

which has been used in the data presented below. Consequently, results for $r > 1$ and $r < 1$ are not statistically independent, since they are based on the same realization (but supposedly different clusters contribute). Using the short-hand notation

$$\mathcal{P}(\mathsf{T}, \geqslant n, r) \equiv \sum_{m=n}^{\infty} \mathcal{P}(\mathsf{T}, m, r) \tag{4}$$

for arbitrary cluster type T, Cardy's exact result reads

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r) = \frac{3\Gamma\left(\frac{2}{3}\right)}{\Gamma\left(\frac{1}{3}\right)^2} \eta^{\frac{1}{3}} \, {}_2F_1\left(\frac{1}{3}, \frac{2}{3}; \frac{4}{3}; \eta\right) \tag{5}$$

where

$$\eta = \left(\frac{1-k}{1+k}\right)^2 \qquad \text{and} \qquad r^{-1} = \frac{K(1-k^2)}{2K(k^2)}. \tag{6}$$

$K(u)$ is the complete elliptic integral of the first kind and ${}_2F_1$ is the hypergeometric function.

Figure 4 shows the difference between the numerical result and the exact value from (5) in units of standard deviations. From this plot it is clear that the systematic deviation for large r observed in [3, 12] was only a finite-size problem.

The asymptotic number distribution of crossing clusters in percolation with open boundaries has been derived analytically [9]

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, n, r) \to \widehat{C_X}(n) \exp\left(-\tfrac{2}{3}\pi n\left(n - \tfrac{1}{2}\right)/r\right) \quad \text{for } n > 1 \tag{7}$$

in the limit of $N \to \infty$ and $r \to 0$. In general, the amplitude $C_X(n)$ is not known exactly, but it is universal and can be derived from numerics by fitting (7) against the numerical data in an appropriate region of r values for each n separately. The range of aspect ratios used in the fit is determined by two competing interests: the fit should include as many points as possible, but exclude aspect ratios where the asymptotic behaviour has not yet set in. This determines the largest r in the fit. The smallest r is given by the value of r for which n simultaneously crossing clusters are observed at least once. Fitting ranges are given in all tables below.

In contrast to spanning on the cylinder, Cardy's asymptotic formula (7) does not distinguish between a cluster crossing in exclusively one direction and a cluster crossing in possibly more than one direction. For consistency in our notation, $C_X(n)$ refers to amplitudes of *exclusively*
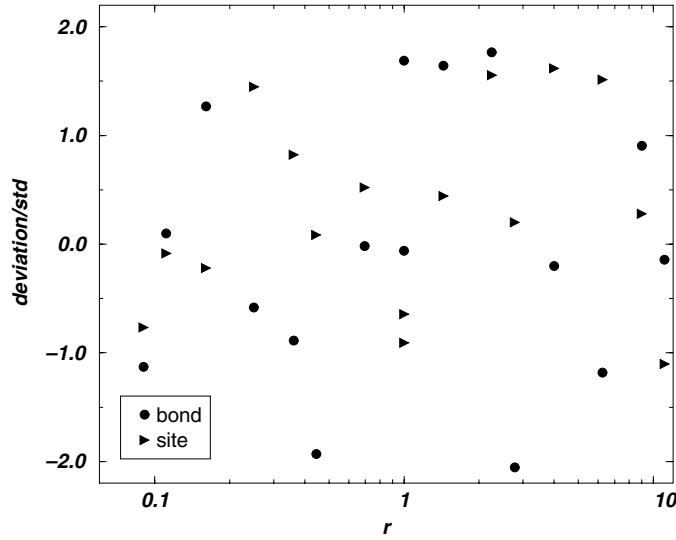
Figure 4. Deviation between analytic (5) and numerical results for the reduced crossing probability $\mathcal{P}_N(\mathsf{U}_{\{N,S\}}, \geqslant 1, r)$, in units of standard deviations of the numerical results versus aspect ratio r, for bond (circles) and site (triangles) percolation with $N = 30000^2$ sites.

vertically crossing clusters. However, at $n = 1$, (7) applies to vertically crossing clusters irrespective of other clusters, to which we assign the amplitude $\widehat{C_{\mathsf{X}}^+}(1)$. For $n = 1$ Cardy's prediction (7) therefore reads

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, 1, r) \to \widehat{C_{\mathsf{X}}^+(1)} \exp\left(-\tfrac{1}{3}\pi/r\right). \tag{8}$$

There is no prediction for exclusive crossing, $\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, 1, r) - \widehat{\mathcal{P}}(\mathsf{U}_{\{N,E,S,W\}}, 1, r)$; however, it is consistent to fit the latter against

$$C_{\mathsf{X}}(1) \exp(-\alpha(1)/r). \tag{9}$$

The fits throughout the paper depend to some degree on the choice of the fitting interval. *A priori*, it is unknown where the asymptotic behaviour sets in (in the sense that the deviation from the asymptote is smaller than the numerical error). It is not possible to determine whether a deviation of the numerical results from the analytical value is due to statistical fluctuations or due to a wrongly chosen fitting interval. The error bars given can only reflect the former. However, the error indicated should include the exact result if it is fitted against the corresponding function in the range given.

Figure 5 shows the numerical data for the probability of finding $n = 1, 2, 3, 4$ and $n \geqslant 1$ vertically crossing clusters in reduced form, i.e. $\ln(\mathcal{P}/(1 - \mathcal{P}))$. The data are fitted to the asymptotic formulae (7), (8) and (9). As discussed below, the latter coincides asymptotically with (12), which is shown as a dotted line. For completeness, data for $\mathcal{P}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r)$ are shown together with the exact result (5).

It is clear that, asymptotically, the relative difference between the probability of more than one and exactly one crossing cluster,

$$\frac{\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r) - \widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, 1, r)}{\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r) + \widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, 1, r)} \tag{10}$$
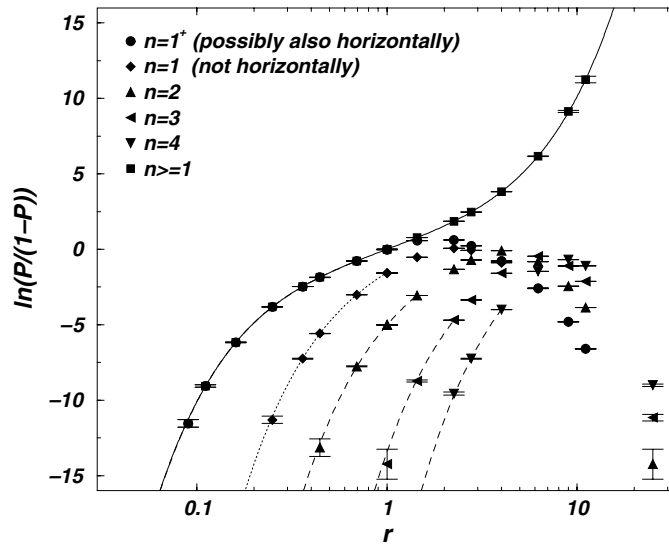
Figure 5. Numerical data ($N = 30000^2$, bond percolation, open boundaries) for the reduced probability $\ln(\mathcal{P}/(1 - \mathcal{P}))$ of $n = 1, 2, 3, 4$ and $n \geqslant 1$ vertically spanning clusters versus aspect ratio r. The data $n \geqslant 1$ refer to $\mathcal{P}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r)$. The long dashed lines are fits according to (7) ((8) and (9) for $n = 1$) and table 1, while the full line gives the exact result (5). The dotted line gives the exact result (12).

vanishes, and the crossing probability becomes dominated by the probability of a single crossing cluster. Therefore, the amplitude $\widehat{C_{\mathsf{X}}^+}(1)$ in (8) is known exactly, namely

$$\widehat{C_{\mathsf{X}}^+}(1) = 2^{4/3} \frac{3\Gamma(2/3)}{\Gamma(1/3)^2} = 1.426\,348\,255\,625\,3 \cdots \tag{11}$$

using the expansion of (5) provided by Ziff [22, 23].

Equation (10) implies that the probability of exactly one crossing cluster exclusively in one direction, $\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, 1, r) - \widehat{\mathcal{P}}(\mathsf{U}_{\{N,E,S,W\}}, 1, r)$, is the dominating term in the difference

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r) - \widehat{\mathcal{P}}(\mathsf{U}_{\{N,E,S,W\}}, 1, r) = \frac{\eta}{\Gamma(1/3)\Gamma(2/3)} \, {}_3F_2\left(1, 1, \frac{4}{3}; 2, \frac{5}{3}; \eta\right) \tag{12}$$

with η as in (6). This relation is derived from (5) and (29), and for small r it can be expanded as

$$\frac{16}{\Gamma(1/3)\Gamma(2/3)} \exp(-\pi/r) \left(1 - \frac{8}{5}\exp(-\pi/r) + \frac{4}{3}\exp(-2\pi/r)\cdots\right) \tag{13}$$

which is based on equation (16) in [22] and the definition of the generalized hypergeometric function. Therefore the two parameters appearing in (9) are known exactly:

$$\widehat{C_{\mathsf{X}}}(1) = \frac{16}{\Gamma(1/3)\Gamma(2/3)} = 4.410\,631\,163\,374\,336\,39 \cdots \tag{14}$$

and

$$\alpha(1) = \pi \tag{15}$$

which corresponds to $n = -1$ or $n = 3/2$ in (7).

The numerics are in very good agreement with the exact results. The estimate for $C_{\mathsf{X}}^+(1)$ shown in table 1 agrees perfectly with (11), while $C_{\mathsf{X}}(1)$ has a surprisingly large error.

Table 1. Crossing with open boundaries. Amplitudes are defined as in (7)–(9), derived from numerical simulations of bond ($C_X^{(b)}$) and site ($C_X^{(s)}$) percolation with $N = 30000^2$ sites. The r range defines the fitting region (see section 3.2.1). The + marks C_X^+, i.e. not an exclusively single crossing cluster in one direction.

n	r range bond	$C_X^{(b)}(n)$	r range site	$C_X^{(s)}(n)$
1^+	$9/100 \cdots 25/36$	1.423(2)	$9/100 \cdots 25/36$	1.426(2)
1	$9/100 \cdots 20/45$	4.811(629)	$9/100 \cdots 25/36$	4.560(510)
2	$20/45 \cdots 36/25$	3.530(12)	$20/45 \cdots 36/25$	3.553(10)
3	$30/30 \cdots 45/20$	9.608(37)	$30/30 \cdots 45/20$	9.599(32)
4	$45/20 \cdots 60/15$	27.641(161)	$45/20 \cdots 60/15$	27.658(140)

Table 2. Spanning on a cylinder. Amplitudes are defined as in (18), where + marks the value of $C_I^+(1)$ (see (19)).

n	r range bond	$C_I^{(b)}(n)$	r range site	$C_I^{(s)}(n)$
1^+	$10/90 \cdots 30/30$	1.2217(4)	$10/90 \cdots 30/30$	1.2222(4)
1	$9/100 \cdots 30/30$	1.7198(11)	$9/100 \cdots 30/30$	1.7225(10)
2	$25/36 \cdots 36/25$	5.1829(256)	$25/36 \cdots 36/25$	5.2105(218)
3	$36/25 \cdots 50/18$	15.1212(764)	$30/30 \cdots 50/18$	15.0227(649)
4	$45/20 \cdots 60/15$	45.0059(3280)	$45/20 \cdots 60/15$	44.5445(2780)

However, the result still covers the exact value (14). The error is due to a narrow fitting range, forced by the late onset of asymptotic behaviour. A larger fitting range gives a much smaller error and reduced goodness-of-fit [24]. For $\alpha(1)$ it is found numerically that $\alpha^{(b)}(1) = 3.18(6)$ and $\alpha^{(s)}(1) = 3.15(5)$, in perfect agreement with (15).

The amplitudes listed in table 1 have also been used to plot the dashed lines in figure 5. They fit an exponential fairly well (using the 1^+ value in table 1):

$$C_X^{(b)}(n) \approx 0.5422(10) \exp(0.9619(14)n) \tag{16}$$

$$C_X^{(s)}(n) \approx 0.5434(9) \exp(0.9617(12)n). \tag{17}$$

3.2.2. Spanning probability on a cylinder. Cardy has also given asymptotes for spanning events on a cylinder [9]. These clusters have been investigated numerically several times [21, 16, 18]. However, this work uses system sizes three orders of magnitude larger than in former studies. The probability of obtaining n distinct spanning clusters on a cylinder is expected to behave in the limit of small r and large L like

$$\widehat{\mathcal{P}}(\mathsf{I}, n, r) \to C_\mathsf{I}(n) \exp\left(-\tfrac{2}{3}\pi\left(n^2 - \tfrac{1}{4}\right)/r\right). \tag{18}$$

The existence of a wrapping cluster prevents more than one spanning cluster, see figure 2. For $n = 1$ one can distinguish between exclusively spanning clusters and spanning clusters that may also wrap. Allowing for wrapping clusters, Cardy predicts the asymptote

$$\widehat{\mathcal{P}}(\mathsf{I}^+, 1, r) \to C_\mathsf{I}^+(1) \exp\left(-\tfrac{5}{24}\pi/r\right). \tag{19}$$

The numerical results are in full agreement with (18) and (19). The corresponding amplitudes are shown in table 2. The fact that the amplitude in (19), $C_\mathsf{I}^+(1)$, is slightly smaller than the corresponding amplitude in (18), $C_\mathsf{I}(n)$, does not contradict $\widehat{\mathcal{P}}(\mathsf{I}, 1, r) < \widehat{\mathcal{P}}(\mathsf{I}^+, 1, r)$, since the exponentials differ by a factor $\exp(\pi(7/24)/r)$ in favour of spanning without

Table 3. Wrapping on a cylinder. Results for the fits of $\mathcal{P}(\mathsf{O}, n, r)$ to (23). The r range defines the fitting region, where the right-hand value is for the largest aspect ratio in which n wrapping clusters occurred at least once. The + marks the value of $C_{\mathsf{O}}^+(1)$ that allows a spanning cluster.

n	r range (bond)	$C_{\mathsf{O}}^{(b)}(n)$	$\alpha_{\mathsf{O}}^{(b)}(n)$	r range (site)	$C_{\mathsf{O}}^{(s)}(n)$	$\alpha_{\mathsf{O}}^{(s)}(n)$
1^+	$25/36 \cdots 100/9$	1.0654(12)	$-1.0755(10)$	$25/36 \cdots 100/9$	1.0634(10)	$-1.0739(9)$
1	$25/36 \cdots 60/15$	1.7734(82)	$-3.0536(52)$	$25/36 \cdots 60/15$	1.7712(70)	$-3.0522(44)$
2	$15/60 \cdots 36/25$	1.8134(55)	$-6.9320(92)$	$15/60 \cdots 45/20$	1.7995(46)	$-6.9118(78)$
3	$12/75 \cdots 25/36$	3.6128(216)	$-17.8021(326)$	$12/75 \cdots 30/30$	3.6600(186)	$-17.8764(278)$
4	$10/90 \cdots 20/45$	6.6734(609)	$-32.6919(752)$	$10/90 \cdots 20/45$	6.6533(520)	$-32.7071(645)$

restrictions on wrapping. It would require $r > 2.67$ to suppress this factor enough to equalize both probabilities. However, at such large values of r, equations (18) and (19) are no longer valid.

The amplitudes fit an exponential extremely well (where for $n = 1$ wrapping was not allowed, i.e. the result 1^+ in table 2 was ignored):

$$C_{\mathsf{I}}^{(b)}(n) \approx 0.5788(11) \exp(1.0892(17)n) \tag{20}$$

$$C_{\mathsf{I}}^{(s)}(n) \approx 0.5816(10) \exp(1.0860(14)n). \tag{21}$$

The resulting plot of the reduced probabilities looks very similar to that in figure 5, so we omit it here.

3.2.3. Wrapping on a cylinder. Aizenman's original statement [6] regarding the number of crossing clusters can also be applied to the number of wrapping clusters on a cylinder. In the limit of large r

$$\ln(\widehat{\mathcal{P}}(\mathsf{O}, n, r)) \in \mathcal{O}(rn^2) \tag{22}$$

according to a hand-waving scaling argument by Cardy [9]. There is no better estimate known, so by fitting each set of histograms to

$$\widehat{\mathcal{P}}(\mathsf{O}, n, r) = C_{\mathsf{O}}(n) \exp(\alpha_{\mathsf{O}}(n)r) \tag{23}$$

one can determine the n dependence of $\alpha_{\mathsf{O}}(n)$ and the amplitude $C_{\mathsf{O}}(n)$. Table 3 shows the corresponding results. It turns out that $\alpha_{\mathsf{O}}(n)$ very well fits a second order polynomial with coefficients

$$\alpha_{\mathsf{O}}^{(b)}(n) \approx -3.150(11)n^2 + 5.51(4)n - 5.41(3) \tag{24}$$

$$\alpha_{\mathsf{O}}^{(s)}(n) \approx -3.176(10)n^2 + 5.61(4)n - 5.48(3). \tag{25}$$

The amplitude $C_{\mathsf{O}}(n)$ again fits an exponential, but only if the very first value, $C_{\mathsf{O}}(1)$, is neglected. In this case we find

$$C_{\mathsf{O}}^{(b)}(n) \approx 0.483(5) \exp(0.662(4)n) \tag{26}$$

$$C_{\mathsf{O}}^{(s)}(n) \approx 0.473(4) \exp(0.670(3)n). \tag{27}$$

We stress again that the ambiguity in the choice of the fitting ranges introduces an error which is not reflected in the numerical error given. We find it, therefore, justified to conjecture that

$$\alpha_{\mathsf{O}}(n) = -\pi n^2 + \tfrac{7}{4}\pi(n - 1) \tag{28}$$

which is somewhat surprising, since we naively expected similar arguments [9] to those for (18) should apply, giving rise to a leading term $-\tfrac{2}{3}\pi n^2$. The numerical results are shown together with the proposed analytical behaviour according to (23) using the data in table 3 in figure 6.
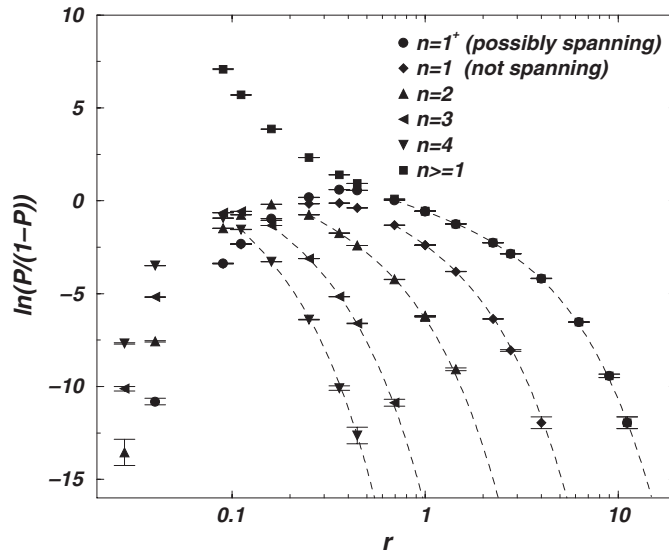
Figure 6. Numerical data ($N = 30000^2$, bond percolation) for the reduced probability $\ln(\mathcal{P}/(1 - \mathcal{P}))$ of $n = 1, 2, 3, 4$ and $n \geqslant 1$ wrapping clusters on a cylinder versus aspect ratio r. The data $n \geqslant 1$ refer to $\mathcal{P}(\mathsf{O}, \geqslant 1, r)$ and look qualitatively similar to those shown in figure 5. The long dashed lines are fits according to (23) and table 3.

3.2.4. Spanning simultaneously in both directions. Watts [4] has exactly calculated the probability of a cluster that crosses both directions simultaneously in a system with open boundaries, $\widehat{\mathcal{P}}(\mathsf{U}_{\{N,E,S,W\}}, 1, r)$, given by

$$\widehat{\mathcal{P}}(\mathsf{U}_{\{N,E,S,W\}}, 1, r)$$
$$= \frac{3\Gamma\left(\frac{2}{3}\right)}{\Gamma\left(\frac{1}{3}\right)^2} \eta^{\frac{1}{3}} \, {}_2F_1\left(\frac{1}{3}, \frac{2}{3}; \frac{4}{3}; \eta\right) - \frac{\eta}{\Gamma\left(\frac{1}{3}\right)\Gamma\left(\frac{2}{3}\right)} \, {}_3F_2\left(1, 1, \frac{4}{3}; 2, \frac{5}{3}; \eta\right) \quad (29)$$

with η as defined in (6) and ${}_3F_2$ being the generalized hypergeometric function. The first term on the right-hand side of (29) is identical to $\widehat{\mathcal{P}}(\mathsf{U}_{\{N,S\}}, \geqslant 1, r)$. According to the expansion provided by Ziff [22] and consistent with (8), for small r this term is proportional to $\exp\left(-\frac{1}{3}\pi/r\right)$, while, according to (13), the second term decays even faster, namely $\exp(-\pi/r)$. This is, of course, what one expects, because the probability of spanning in both directions is, away from $r = 1$, dominated by the probability of spanning in the longer direction.

Figure 7 shows the difference between our numerical results and Watts' prediction. As in figure 4, the plot suggests that the deviations observed in [4] are only finite-size corrections.

3.3. Finite-size corrections

Since the system sizes investigated are large compared to former studies [12, 16, 18], one might be inclined to completely ignore finite-size corrections. In order to estimate their strength, table 4 lists numerical results of various quantities for different system sizes and compares them with estimates for the values of these quantities in the thermodynamic limit found in the literature or their exact results. For large systems, our numerical results agree very well. For smaller systems, there may be some mild corrections. From the three-legged cluster data (section 3.1.2), we expect that finite-size corrections may become visible if one side of the
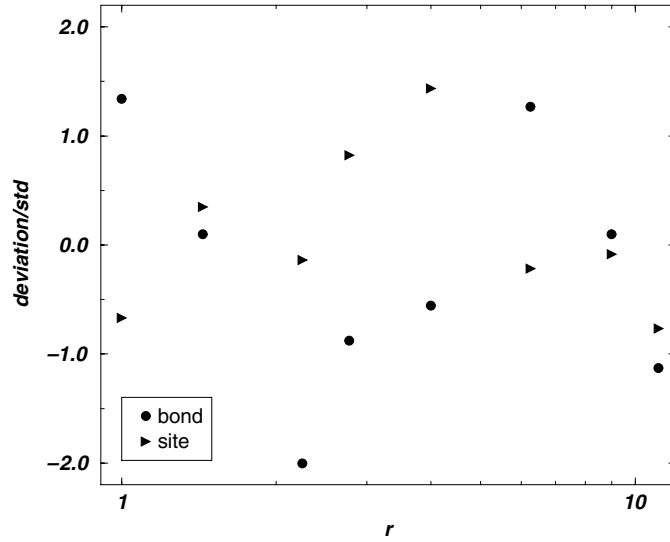
Figure 7. Deviation between the analytic result (29) and the numerical results for the reduced crossing probability $\mathcal{P}_N(\mathsf{U}_{\{N,E,S,W\}}, 1, r)$, in units of standard deviations of the numerical results versus aspect ratio r, for bond (circles) and site (triangles) percolation with $N = 30000^2$ sites.

Table 4. Comparison of our numerical values and estimates found in the literature, for different system sizes. 'Estimates' are values for those quantities cited in the literature in the thermodynamic limit.

Quantity	Site	Bond	Exact value/estimate
$\mathcal{P}_{N=3000^2}(\mathsf{U}_{\{N,S\}}, \geqslant 1, 1)$	0.4995(4)	0.5012(4)	1/2 exact [2]
$\mathcal{P}_{N=30000^2}(\mathsf{U}_{\{N,S\}}, \geqslant 1, 1)$	0.4998(3)	0.4999(4)	
$\mathcal{P}_{N=3000^2}(\mathsf{U}_{\{N,S\}}, \geqslant 2, 1)$	$6.58(6) \times 10^{-3}$	$6.69(7) \times 10^{-3}$	$6.58(3) \times 10^{-3}$ [16]
$\mathcal{P}_{N=30000^2}(\mathsf{U}_{\{N,S\}}, \geqslant 2, 1)$	$6.71(6) \times 10^{-3}$	$6.66(6) \times 10^{-3}$	
$\mathcal{P}_{N=3000^2}(\mathsf{U}_{\{N,E,S,W\}}, 1, 1)$	0.3221(4)	0.3230(4)	$0.322120455\cdots$ exact [4]
$\mathcal{P}_{N=30000^2}(\mathsf{U}_{\{N,E,S,W\}}, 1, 1)$	0.3219(3)	0.3226(4)	
$\mathcal{P}_{N=3000^2}(\mathsf{I}, \geqslant 1, 1)$	0.6360(4)	0.6368(4)	0.63665(8) [21]
$\mathcal{P}_{N=30000^2}(\mathsf{I}, \geqslant 1, 1)$	0.6361(3)	0.6364(4)	

rectangle has length $\leqslant 300$, which is not the case for any aspect ratio we have simulated in a system with $N = 30000^2$ sites.

The results also indicate that the estimate of $p_c^{(\mathrm{s})}$ by Newman and Ziff [10] is valid within numerical error. This is corroborated by figures 4 and 7, where site percolation does not seem to show stronger deviations than bond, for which $p_c^{(\mathrm{b})}$ is known exactly.

4. Conclusion and discussion

Our numerical results represent probably the most comprehensive and most up-to-date study of crossing, spanning and wrapping probabilities. We have presented results for 'exotic' cluster types, for which an analytical description is still lacking. With regard to the predictions

of conformal field theory, our numerical data afford a comparison between numerical and analytical results, and we give further support to conformal invariance in percolation and other critical phenomena [25]. We believe that deviations from the predicted behaviour observed in the literature are most likely due to finite-size effects.

We have also calculated the amplitudes listed in tables 1, 2 and 3, which might be of help to theorists. For three-legged clusters, we have conjectured the asymptote (2), for wrapping probabilities of the form (28).

Acknowledgments

Appendix

Here we list a few technicalities on how to identify cluster types, etc. The identification methods are applicable in any simulation of percolation where the boundary can be represented in the Hoshen–Kopelman (HK) [26] form.

A.1. Simulation method

Our method performs asynchronously parallelized percolation on distributed machines [1]. In principle, the method relaxes all the standard constraints in numerical simulations of percolation, such as CPU power, memory and network capacity. It is especially suitable for calculating cluster size distributions and finite-size corrections, crossing probabilities, and, by applying the corresponding boundary conditions, distributions of wrapping and spanning clusters on different topologies, e.g. cylinder, torus or the Möbius strip.

The method is based on a master/slave parallelization, where slaves send 'patches' (specially prepared borders representing the lattice) to a master node, which 'glues' these patches together. After a path compression, which is essentially of the form of 'Nakanishi label recycling' [27, 28] where bulk sites are considered inactive, the result is a single border as shown in figure 8. The term 'gluing' will be used to indicate that a configuration is updated to account for a link introduced between two sites or two boundaries. The HK algorithm [26] provides the data representation, which is the key of this method. Extremely large system sizes can be simulated.

An example of a border representation is shown in figure 8 (in the following all examples are based on site percolation). Each site on the border is indexed in a clockwise manner starting with 1 in the upper left corner. Each site also contains a label. If the label is 0 the site is not occupied. If it is positive it is interpreted as a 'pointer' to the index of another site in the same cluster. If it is negative, it is called a 'root', and the magnitude of the negative number indicates the size of the cluster the site belongs to. The HK algorithm ensures that all sites of a cluster form a single tree, with a root that carries a negative label. In this way it is possible to identify the cluster each site belongs to by its corresponding root. All information about a cluster is stored at the root site.

It is technically simple to identify the types of all clusters touching the border: the algorithm scans along a border like that shown in figure 8 and assigns a flag to the root of each
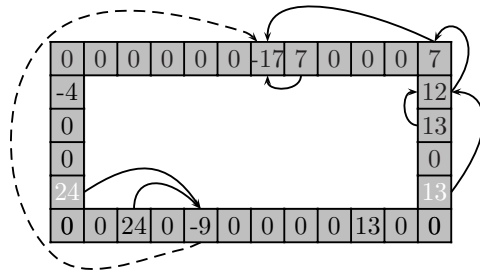
Figure 8. An example of a labelled border. A negative entry indicates a cluster size, a positive entry is a pointer to another site and a zero entry indicates that the site is not occupied. The configuration shown is a border representation of a full lattice as in figure 1. The dashed pointer applies for periodic boundary conditions along the vertical borders, joining the two highlighted sites.

(occupied) site it visits, according to the location of the site visited. The flags gathered at each root then identify the borders that the represented clusters touch. The $\mathcal{P}_N(\mathsf{T}, n, r)$ are based on the statistics of these flags.

A.2. Identifying spanning and wrapping

Starting from a configuration with open boundaries (for example figure 8), spanning clusters on a cylinder can be detected easily, by 'gluing' the appropriate borders: first the roots are identified for each pair of sites, which become nearest neighbours due to the new boundary conditions. Preferably, the root of the smaller cluster is then redirected to the other, i.e. its label is overwritten by the index of the other root, which, in turn, inherits all properties of the overwritten root, such as cluster size or border flags. The new pointer resulting from a vertical gluing is shown as a dashed line in figure 8; clusters of type $\mathsf{U}_{\{N,S\}}$ then span on the cylinder.

It is significantly more complicated to detect wrapping clusters, because these clusters cannot be defined by properties of individual sites alone.

The simplest solution, which is applicable to any kind of topology, such as a cylinder, torus or Möbius strip, is to assign to each site an additional set of numbers, which indicates its distance to the site it is pointing to. This distance can be measured in units of lattice spacings or, even simpler, in multiples of π: each type of border is mapped to a set of integers like

$$N \to (0, 1) \quad S \to (0, -1) \quad E \to (1, 0) \quad W \to (-1, 0) \tag{A.1}$$

and for each site an additional flag indicates its location. The distance between two sites is then the difference between the tuples associated with their location flag. For example, the distance between W and S is $(1, -1)$. In the open system, each site is assigned a tuple indicating the distance to the site it is pointing to. The periodic boundary conditions are then applied by gluing along the vertical borders. Two clusters merge in the way described above, with the additional assignment of a distance vector to the redirected root site. This distance vector indicates the distance from the redirected root site to the new root site, given by the difference between the distances of the glued sites and their roots. In this way, for any site in a cluster the distance to the root is given by the sum over all distance vectors along the tree to the root.

If the gluing procedure now comes across two sites on either border, which belong to the same cluster[4], the path length to the root is calculated for each of them. If their difference is

---

[4] This in itself does not indicate that a cluster is wrapping, as a previous encounter of the two clusters may have merged them.
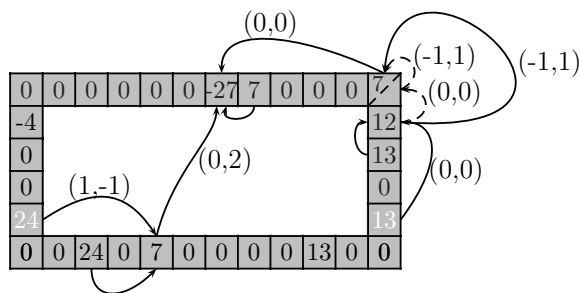
Figure 9. The HK representation of the border of the configuration shown in figure 1 with the hatched square 'activated' to allow for a wrapping cluster. The tuples indicate the distance of a site to the site it is pointing to. These numbers are only given for the two paths which are relevant for the emergence of a wrapping cluster when the system is 'glued' along the W and E borders. The dashed line in the top right-hand corner indicates the special treatment of a corner, while the dashed arrows indicate auxiliary pointers. The highlighted sites give rise to a wrapping cluster.

non-zero, a wrapping cluster has been found, the number indicating the winding number as a multiple of π. In the example in figure 9, the left-hand path has length $(1, -1) + (0, 2) = (1, 1)$, the right-hand path has length $(-1, 1)$, differing by the expected length $(0, 2)$. It is a topological fact that higher winding numbers cannot appear on a cylinder.

Corners require special attention, because they belong to two different borders at the same time. In the example shown in figure 9, all paths shown on the right-hand side have length $(0, 0)$, apart from the corner site, which is thought of as carrying an 'internal pointer' to itself, connecting the E border to the N border. The internal pointer can be understood as follows: if a site points to a corner, it has a connection to two borders at once. However, the distance vector of the pointing site can only indicate the distance to one border. A convention is required to lift this degeneracy, for example that all pointers pointing towards and away from a corner site do so with respect to the N or S boundary, never with respect to the E or W boundary. However, it might happen that another corner is glued to the corner site, or a neighbouring site is connected to it as a W or E site. In this case, auxiliary pointers are introduced: one pointer from the connected site to the corner site's W or E part and another pointer internally connecting the corner site's W or E part to the N or S part. In figure 9 these auxiliary pointers and the sub-partitioning of a corner site are shown with dashed lines.

References

[1] Moloney N R and Pruessner G 2003 *Phys. Rev.* E **67** 037701 (*Preprint* cond-mat/0211240)
[2] Langlands R, Pichet C, Pouliot P and Saint-Aubin Y 1992 *J. Stat. Phys.* **67** 553
[3] Cardy J 1992 *J. Phys. A: Math. Gen.* **25** L201 (*Preprint* hep-th/9111026)
[4] Watts G M T 1996 *J. Phys. A: Math. Gen.* **29** L363 (*Preprint* cond-mat/9603167) v2
[5] Hu C-K and Lin C-Y 1996 *Phys. Rev. Lett.* **77** 8
[6] Aizenman M 1997 *Nucl. Phys.* B **485** 551 (*Preprint* cond-mat/9609240) v4
[7] Newman C M and Schulman L S 1981 *J. Phys. A: Math. Gen.* **14** 1735
[8] Newman C M and Schulman L S 1981 *J. Stat. Phys.* **26** 613
[9] Cardy J 1998 *J. Phys. A: Math. Gen.* **31** L105
[10] Newman M E J and Ziff R M 2000 *Phys. Rev. Lett.* **85** 4104
[11] Kesten H 1980 *Commun. Math. Phys.* **74** 41
[12] Langlands R, Pouliot P and Saint-Aubin Y 1994 *Bull. Am. Math. Soc.* **30** 1
[13] Matsumoto M and Nishimura T 1998 *Monte Carlo and Quasi-Monte Carlo Methods 1998* (Berlin: Springer)
    Preprint from http://www.math.h.kyoto-u.ac.jp/˜matumoto/RAND/DC/dc.html
[14] Davies B and Peschel I 1991 *J. Phys. A: Math. Gen.* **24** 11293

[15] Baxter R J 1982 *Exactly Solved Models in Statistical Mechanics* (London: Academic)
[16] Shchur L N and Kosyakov S S 1997 *Int. J. Mod. Phys.* C **8** 473 (*Preprint* cond-mat/9702248)
[17] Shchur L N and Kosyakov S S 1998 *Preprint* cond-mat/9802082 unpublished
[18] Shchur L N 2000 *Proc. Physics* vol 85, ed D P Landau, S P Lewis and H B Schuettler (Berlin: Springer)
        (*Preprint* cond-mat/9906013) v2
[19] Sen P 1997 *Int. J. Mod. Phys.* C **8** 229 (*Preprint* cond-mat/9704112)
[20] Sen P 1999 *Preprint* cond-mat/9904009 unpublished
[21] Hovi J-P and Aharony A 1996 *Phys. Rev.* E **53** 235
[22] Ziff R M 1995 *J. Phys. A: Math. Gen.* **28** 1249
[23] Ziff R M 1995 *J. Phys. A: Math. Gen.* **28** 6479
[24] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C* 2nd edn (New York,
        NY: Cambridge University Press)
[25] Cardy J L 1987 *Phase Transitions and Critical Phenomena* vol 11, ed C Domb and J L Lebowitz (London:
        Academic)
[26] Hoshen J and Kopelman R 1976 *Phys. Rev.* B **14** 3438
[27] Nakanishi H and Stanley H E 1980 *Phys. Rev.* B **22** 2466
[28] Binder K and Stauffer D 1987 *Applications of the Monte Carlo Method in Statistical Physics (vol 36 of Topics
        in Current Physics)* 2nd edn ed K Binder (Berlin: Springer) pp 241–75