Spanning Probability in 2D Percolation

Robert M. Ziff
Department of Chemical Engineering, University of Michigan, Ann Arbor, Michigan 48109-2136
(Received 3 August 1992)

The probability $R_L(p)$ for a site percolation cluster to span a square lattice of side L at occupancy p is reexamined using extensive simulations and exact calculations. It is confirmed that $R_L(p_c) \to \frac{1}{2}$ as $L \to \infty$ in agreement with universality but not with renormalization-group theory. Many estimates of p_c that derive from $R_L(p)$ are shown to scale with L more weakly than normal finite-size scaling, and the value $p_c = 0.592\,7460 \pm 0.000\,0005$ is determined.

PACS numbers: 64.60.Ak, 05.70.Jk

The percolation model has been investigated intensely for decades, and a large variety of approaches and techniques have been developed for its study [1]. One approach has been through the function $R_L(p)$, which gives the probability that a finite system of size L percolates at occupation probability p. The percolation transition is clearly illustrated by the transformation of R_L to a step function $\Theta(p - p_c)$ about the critical threshold p_c as L goes to infinity. R_L is also central to many renormalization-group (RG) calculations (thus its symbol), and has been examined extensively [2–7].

With newer techniques based upon hull-generating methods, better values of p_c, and faster computers, it is possible to study this function much more efficiently than in the past. Here, I report on extensive numerical simulations of R_L using a method recently suggested by Grassberger [8], and on a new exact evaluation of R_L for L up to 7. The specific system studied is site percolation on a 2D square lattice, with percolation defined as spanning in one given direction with free boundaries in the other direction—rule $\mathcal{R}_1$ of Reynolds, Stanley, and Klein (RSK) [3, 4].

This work yields a number of new results. First, it confirms that $R_L(p_c) \to \frac{1}{2}$ as $L \to \infty$, in accord with universality with bond percolation [where $R_L(p_c) = \frac{1}{2}$ for all L [9]], and Cardy's recent result [10], but not in agreement with RG theory, which predicts that $R_\infty(p_c)$ should be equal to p_c rather than $\frac{1}{2}$. That is, it shows that $\mathcal{R}_1$ does not renormalize at p_c for site percolation. Second, the scaling behavior of various estimates $p(L)$ of p_c for finite systems is determined for the first time. The "cell-to-cell" RG estimate, the "average p" and the "maximum p," all introduced by RSK, and a new "median p" introduced here, are all shown to scale as

$$p(L) - p_c \sim cL^{-1-1/\nu} \tag{1}$$

as $L \to \infty$, where $\nu = 4/3$ is the correlation-length exponent and c is a constant (different for each estimate). This behavior is in contrast to the usual finite-size scaling law

$$p(L) - p_c \sim cL^{-1/\nu}\,, \tag{2}$$

which applies to the RG fixed point. The faster convergence of (1) has been seen, but not well recognized, in many previous studies of percolation. And third, these simulations yield the precise value for p_c:

$$p_c = 0.592\,7460 \pm 0.000\,0005\,. \tag{3}$$

The method of Grassberger [8] is illustrated in Fig. 1. When spanning occurs, the hull of the clusters connected to the bottom boundary of the system will reach the top boundary. This hull, however, can be generated independently of the cluster, using a hull-generating walk [11]. In such a walk, the walker looks first to the neighbor on the left, and moves directly to occupied sites, skips blocked ones (going to the next clockwise neighbor), and makes undecided sites occupied with probability p and vacant otherwise. Thus, to simultaneously simulate a system and check for spanning, the walk is carried out on a square lattice of blank (undetermined) sites, with a row of occupied sites added to the bottom, and a column of blocked sites added to the left-hand side, as shown in Fig. 1(b). The walk starts from the lower left-hand corner of the system; if an occupied site of the walk reaches the top row, then the underlying system percolates, while if a blocked site reaches the right-hand side, it does not. The added row and column account for the conducting boundary on the bottom and free boundary on the sides, and assure that the walk never stops before either the

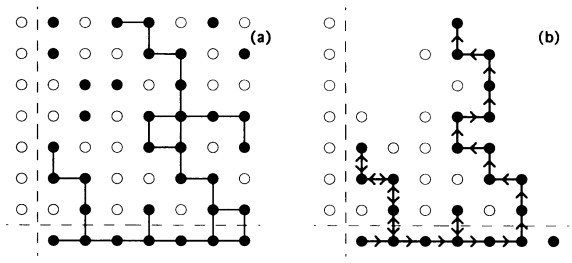

FIG. 1. (a) A typical system with $L = 7$ near the threshold, showing the percolating cluster connecting from the bottom to the top, and (b) the hull of the cluster only, as would be generated by the hull-generating walk. A row of occupied sites is added to the bottom, and a column of blocked sites is added to the left-hand side. The occupied sites are shown as solid circles and blocked ones as open circles. In (b), undecided sites are blank.

top row or right column is reached.

In a moderate number ($\approx 2\times10^5$) of simulations on lattices up to $L = 1280$, Grassberger [8] found that $R_L(p_c)$ approaches the value 0.5 within about $\frac{1}{2}\%$. Here, I have carried out many more runs to verify this result and to investigate the scaling behavior. I simulated 10^7, 4×10^7, and 10^7 hulls at $p = 0.592\,745$, $0.592\,7465$, and $0.592\,75$, respectively, on lattices of 1024×1024 sites (corresponding to $L = 1023$, because of the added boundary row and column). I also recorded intermediate results at $L = 3, 7, 15, \ldots, 511$. The values of p considered here are close to our recent estimates of p_c, $0.592\,745 \pm 0.000\,002$ [12] and $0.592\,7460 \pm 0.000\,0005$ [13], determined by the hull-gradient method [12, 14]. Each hull contained on the average 1.11×10^6 occupied and vacant sites, so that a total of $\approx 6\times10^{12}$ random numbers were generated [15], requiring a few months computing time on about a dozen computer workstations (Apollo 425, IBM R/S 6000) running simultaneously.

The results of these simulations are summarized in Fig. 2, where $(L + L_0)[R_L(p) - 0.5]$ is plotted against L for the different p. This plot shows sensitively that $R_L(p_c)$ approaches the value 0.5 with a finite-size correction of order $1/L$, $R_L(p_c) \sim 0.5 + b_0/L$, with $b_0 \approx 0.319$. The constant $L_0 \approx 1.6$ in this plot is used to correct for higher-order terms; its value is determined by plotting $(L + L_0)[R_L(p) - 0.5]$ vs $1/L^2$, and adjusting L_0 to get linear behavior (also yielding b_0 as the ordinate intercept). I have not attempted to assign error bars on the result $R_\infty(p_c) = 0.5$, since there is little doubt from these simulations and from theoretical arguments that this is exact.

The data points for the largest system considered are $R_{1023}(0.592\,745) = 0.499\,85 \pm 0.000\,16$, $R_{1023}(0.592\,7465) = 0.500\,43 \pm 0.000\,08$, and $R_{1023}(0.592\,75) = 0.500\,90 \pm 0.000\,16$, where the error bars represent expected statistical fluctuations. The predicted behavior (assuming the horizontal-line fit shown in Fig. 2) is $R_{1023}(p_c) = 0.5 + b_0/1023 = 0.500\,31$. Comparing these values, I infer the estimate of p_c given in (3), which agrees with the results given in Ref. [13].

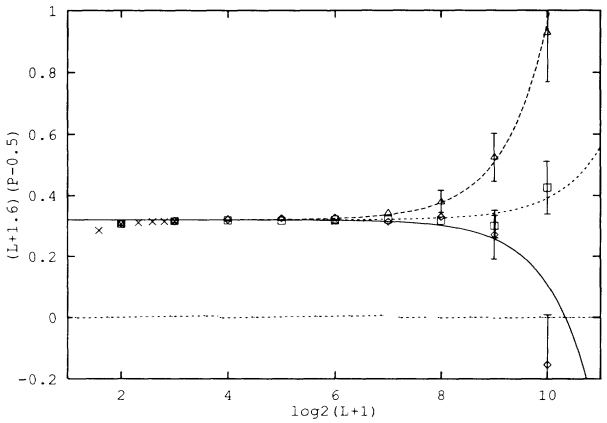

FIG. 2. A plot of $(L+1.6)[R_L(p)-0.5]$ vs $\log_2(L+1)$ of the Monte Carlo data for $p = 0.592\,745$ ($\diamond$), $0.592\,7465$ ($\square$), and $0.592\,75$ ($\triangle$), and the exact results evaluated at $p = 0.592\,746$ ($\times$). The curves represent $R_L(p)$ given by (4).

For p near p_c, it is known [3] that $R'_L(p) \sim L^{1/\nu}$ for large L, so that

$$R_L(p) \sim 0.5 + b_0/L + a_1(p - p_c)L^{1/\nu}, \tag{4}$$

where $a_1 \approx 0.765$ is determined below and also agrees with the data of RSK. Plots of (4) are included in Fig. 2, and it can be seen that this equation fits the simulation data very well.

That the spanning probability for site percolation at p_c should be the same asymptotically as for bond percolation can be seen from a simple universality argument: At criticality, both systems should "look" the same when viewed on a coarse length scale, and the function R_L is a coarse-grained quantity. This result has also recently been proven by Cardy, who considers rectangular boundaries as well, using conformal field theory [10].

The result $R_\infty(p_c) = \frac{1}{2}$, however, contradicts RG theory [3, 16, 17] which maintains that the solution $p^*(L)$ of the equation $R_L(p^*) = p^*$ is a good estimate for p_c. Since $p^*(L) \to p_c$ as $L \to \infty$, the RG conditions would imply that $R \to p_c \approx 0.5927\ldots$ rather than $\frac{1}{2}$ as $L \to \infty$. Thus, the RG theory does not apply here. That is, the criterion of spanning in one direction for 2D site percolation does not renormalize at p_c. As a consequence of this lack of renormalization, $p^*(L)$ converges with the relatively slow "normal scaling" of (2). An analysis using (4) verifies this behavior, with the coefficient in (2) given by $c = (p_c - 0.5)/a_1 \approx 0.121$.

Other estimates of p_c, however, converge an order in L faster than normal scaling for this system. The simple criterion $R_L(p) = 0.5$ gives an estimate $p_{0.5}(L)$ which, according to (4), converges as (1) with $c = -b_0/a_1 \approx -0.417$. To compare these two estimates, note that for a system of size $L = 500$, p^* differs from p_c by 0.0011, while $p_{0.5}$ differs by only $-0.000\,008$.

Equation (4) also implies that RSK's cell-to-cell RG estimate $p_{c-c}(L)$, which is the solution to the equation $R_L(p) = R_{L-1}(p)$, converges as (1) with $c = \nu b_0/a_1 \approx 0.556$. This result confirms the faster convergence of this estimate compared with p^* that was observed by these authors. RSK also introduced the estimates $\overline{p}(L) = \langle p \rangle \equiv \int_0^1 pR'_L(p)dp$, and $p_{\max}(L)$, the value of p at the maximum of the distribution $R'_L(p)$. To discuss the behavior of these two estimates, we need to first develop the finite-size scaling of P further. Note that $\overline{p}$, $p_{0.5}$, and $p_{\max}$ are respectively the mean, median, and maximum of the distribution $R'_L(p)$.

To proceed, I determined $R_L(p)$ for $L \le 7$ exactly, by carrying out a binary search of all possible hulls, assigning a weight of p for every occupied site and $1 - p$ for

every blocked site. Using only the hulls is far more efficient than the conventional procedure of enumerating every realization of the lattice, and it allowed me to extend the number of known polynomials [4] by three. The enumeration for $L = 6$ (20 650 432 hulls) required 3 min of CPU time on an IBM R/S 6000, while that for $L = 7$ (18 871 178 030 hulls, compared with $2^{49} \approx 5.6 \times 10^{14}$ complete realizations) required 65 h. In Table I, I list the resulting expressions for $R'_L(p)$, showing factorization by $[p(1-p)]^{L-1}$, which implies that the first $L-1$ derivatives of $R_L(p)$ are zero at both $p = 0$ and 1. The values of $R_L(0.592\,746)$ for $L \leq 7$ determined from these polynomials are included in Fig. 2.

Now, (4) suggests that the finite-size scaling behavior of $R_L(p)$ is of the form

$$R_L(p) \sim f_1(x) + (1/L) f_2(x) + \cdots \tag{5}$$

for large L, where $x = (p - p_c)L^{1/\nu}$ and $f_1(-\infty) = 0$, $f_1(\infty) = 1$, $f_2(-\infty) = f_2(\infty) = 0$. The function $f_1(x)$ represents the scaling limit ($L \to \infty$ and $p \to p_c$ with x constant) of $R_L(p)$ [18], and is a universal function for all 2D square systems satisfying $\mathcal{R}_1$. Furthermore, by virtue of universality with bond percolation, $f'_1(x)$ must be even. For small x, therefore

$$f_1(x) \sim a_0 + a_1 x + a_3 x^3 + \cdots, \tag{6}$$

$$f_2(x) \sim b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots, \tag{7}$$

TABLE I. Exact expressions for $R'_L(p)$. The first three agree with RSK [4].

$R'_2(p) = 4p(1-p)(1+p)$

$R'_3(p) = p^2(1-p)^2(9+34p+29p^2-30p^3+9p^4)$

$R'_4(p) = 4p^3(1-p)^3(4+27p+60p^2+54p^3-35p^4-99p^5+27p^6+46p^7-12p^8-4p^9)$

$R'_5(p) = p^4(1-p)^4(25+244p+910p^2+1780p^3+1783p^4-562p^5-3722p^6-5660p^7+1823p^8+8222p^9+5612p^{10}-17084p^{11}+10504p^{12}-3524p^{13}+1179p^{14}-260p^{15}+25p^{16})$

$R'_6(p) = 4p^5(1-p)^5(9+115p+605p^2+1785p^3+3360p^4+3675p^5+318p^6-6812p^7-16548p^8-17485p^9+2340p^{10}+57777p^{11}+54747p^{12}-25990p^{13}-166535p^{14}-41139p^{15}+338581p^{16}-142902p^{17}-173191p^{18}+185135p^{19}-61126p^{20}+1874p^{21}+3663p^{22}-885p^{23}+45p^{24}+9p^{25})$

$R'_7(p) = p^6(1-p)^6(49+774p+5259p^2+20744p^3+54934p^4+101688p^5+122574p^6+48920p^7-170295p^8-545926p^9-931219p^{10}-1021728p^{11}+464618p^{12}+3535668p^{13}+7371786p^{14}+1812934p^{15}-11753621p^{16}-29342232p^{17}+1566553p^{18}+46387036p^{19}+71120886p^{20}-65689076p^{21}-206506493p^{22}+171317982p^{23}+307523341p^{24}-556512164p^{25}+322800996p^{26}-32853804p^{27}-46710219p^{28}+18173946p^{29}+1550692p^{30}-2527898p^{31}+703668p^{32}-108462p^{33}+12799p^{34}-1050p^{35}+49p^{36})$

where $a_0 = \frac{1}{2}$. The a_0, a_1, and b_0 terms give (4), which represents the small-x behavior of these equations. Analyzing the exact expressions for $L \leq 7$, I find $a_3 \approx -0.50$, $b_1 \approx 0.466$, $b_2 \approx -0.176$, and $b_3 \approx -0.21$.

Equations (5)–(7) imply that both $p_{\max}$ and $\overline{p}(L)$ converge as (1) with c given by $-b_2/(3a_3) \approx -0.11$ and by $\int_{-\infty}^{\infty} x f'_2(x)\,dx$, respectively. This predicted behavior of $\overline{p}$ is consistent with the work of RSK, whose plot of $L^{-1/\nu}$ vs $\overline{p}$ (Fig. 13 of Ref. [4]; see also [19]) is essentially vertical, indicating a higher-order scaling behavior. Note that this scaling behavior is a consequence of $f'_1(x)$ being a symmetric function—if it were not, then these two estimates would scale as (2).

Thus, I find that the estimates $\overline{p}$, $p_{0.5}$, $p_{\max}$, and p_{c-c} all converge as (1), with coefficients summarized in Table II. In the scaling limit, these estimates all become identical to p_c, and (1) reflect the correction-to-scaling term. That such estimates should behave this way is suggested in the work of Binder and Heermann [18], who remark that, as a consequence of $R_L(p)$ being a function of only x in the scaling limit, the curves of $R_L(p)$ for different finite L will intersect at a common point, apart from corrections to scaling. As we have seen, that common point is at $(R, p) = (\frac{1}{2}, p_c)$, and the corrections (for site percolation) are given by (1).

To confirm these predictions for the behavior of the various estimates for p_c, I have calculated their values using the exact expressions for $L \leq 7$. The results are plotted in Fig. 3 as a function of $(L + L_1)^{-1.75}$, where the constant $L_1 \approx 1.4$ was chosen to correct (roughly) all the curves' higher-order terms. Even for such small L, the data for all estimates (except, as expected, the RG estimate p^*) are nearly linear. The slopes of these lines for $p_{0.5}$, $p_{\max}$, and p_{c-c} are close to the predicted values of c given above, and, for $\overline{p}$, I find $c \approx -0.55$. Note $\overline{p} < p_{0.5} < p_{\max}$, indicating that $R'_L(p)$ is skewed to the left. Also, it can be seen that $p_{\max}$ gives the best estimate of p_c. From these results for $L \leq 7$, one can easily estimate p_c within an error of about ± 0.0005. The value of ν can also be confirmed to fairly high accuracy.

The results presented here pertain only to systems that are square and satisfy $\mathcal{R}_1$. For other systems, $f_1(x)$ may not be symmetric, $R_\infty(p_c)$ may not be equal to $\frac{1}{2}$, and consequently some or all of the estimates may not scale as (1). This explains why $\overline{p}$ was found to scale as (2)

TABLE II. Values of the coefficient c in (1) or (2) for the various estimates $p(L)$ of p_c.

Estimate	Equation	c
p^* (RG)	(2)	$(p_c - 0.5)/a_1 \approx 0.121$
$\overline{p}$	(1)	$\int_{-\infty}^{\infty} x f'_2(x) dx \approx -0.55$
$p_{0.5}$	(1)	$-b_0/a_1 \approx -0.417$
$p_{\max}$	(1)	$-b_2/(3a_3) \approx -0.11$
p_{c-c}	(1)	$\nu b_0/a_1 \approx 0.556$

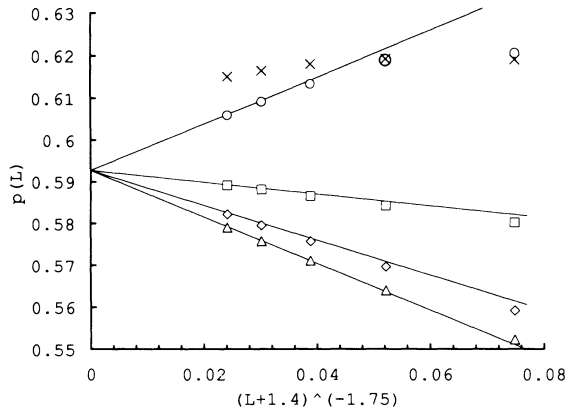

FIG. 3. Plots of the estimates p^* ($\times$), p_{c-c} ($\bigcirc$), $p_{\max}$ ($\square$), $p_{0.5}$ ($\diamond$), and $\overline{p}$ ($\triangle$) vs $(L+1.4)^{-1.75}$ for $3 \le L \le 7$, verifying that all but p^* follow (1).

rather than (1) in Ref. [6], where a triangular lattice whose boundary is effectively a rhombus was studied, in Ref. [4], where different spanning conditions for a square were considered, and in Ref. [20], where a square system with periodic boundary conditions on one pair of sides was investigated. In the latter paper, the authors remark that obtaining more precise results for free boundary conditions than for periodic ones is unusual; here we see that this is a consequence of free side boundaries on a square system being a very special case.

Likewise, the convergence of the RG estimate for the critical point will vary from system to system; for p^* to converge faster than normal scaling, $R_L(p_c)$ must coincide with p_c, at least asymptotically for large L. For site percolation on the square lattice and a square system shape with rule $\mathcal{R}_1$, this coincidence does not occur. While there is undoubtedly some set of conditions (shape and rule) where the RG equation is asymptotically followed for site percolation, those conditions will be specific for that system. There can be no universal spanning rule that renormalizes for all lattices and percolation types, because $R_\infty(p_c)$ is universal while p_c is specific for each system. This uncertainty in having proper renormalization makes application of the RG method to this percolation problematic.

I thank P. Grassberger for suggesting this problem and for sending preprints of his work, and gratefully acknowledge support from NSF Grant No. DMR-9122341 and the IBM/UM Distributed Computing Initiative.

Note added.—In a recent paper studying spanning for general rectangular systems, Langlands *et al.* [21] have also shown that $R_L(p_c) \to \frac{1}{2}$ for a square system. Also, I have carried out additional simulations at $p = 0.592\,746$ (2×10^7 hulls), and find that the data for large L fall close to the straight line of Fig. 2. For example, the last data point $R_{1023}(0.592\,746) = 0.500\,34 \pm 0.000\,11$ is in agreement with the predicted value 0.500 31. These results confirm that (3) is a good estimate for p_c.

[1] See D. Stauffer, Phys. Rep. **54**, 1 (1979); D. Stauffer and A. Aharony, *An Introduction to Percolation Theory* (Taylor and Francis, London, 1985), 2nd ed., and references therein.
[2] J. Roussenq, J. Clerc, G. Giraud, E. Guyon, and H. Ottavi, J. Phys. (Paris), Lett. **37**, L99 (1976).
[3] P. J. Reynolds, H. E. Stanley, and W. Klein, J. Phys. A **11**, L199 (1978).
[4] P. J. Reynolds, H. E. Stanley, and W. Klein, Phys. Rev. B **21**, 1223 (1980).
[5] S. Kirkpatrick, in *Proceedings of the Les Houches Summer School on Ill-Condensed Matter* (North-Holland, Amsterdam, 1980).
[6] P. D. Eschbach, D. Stauffer, and H. J. Herrmann, Phys. Rev. B **23**, 422 (1981).
[7] Y. Yuge, Phys. Rev. B **18**, 1514 (1978).
[8] P. Grassberger, J. Phys. A (to be published).
[9] J. Bernasconi, Phys. Rev. B **18**, 2185 (1978).
[10] J. L. Cardy, J. Phys. A **25**, L201 (1992).
[11] R. M. Ziff, P. T. Cummings, and G. Stell, J. Phys. A **17**, 3009 (1984).
[12] R. M. Ziff and B. Sapoval, J. Phys. A **19**, L1169 (1986).
[13] R. M. Ziff and G. Stell, Laboratory for Scientific Computing, University of Michigan, Report No. 88-4, 1988, Footnote 26.
[14] M. Rosso, J. F. Gouyet, and B. Sapoval, Phys. Rev. B **32**, 6053 (1986).
[15] The random number generator we use is a shift-register sequence [S. W. Golumb, *Shift Register Sequences* (Holden-Day, San Francisco, 1967)] similar to that used by S. Kirkpatrick and E. Stoll, J. Comput. Phys. **40**, 517 (1981). We use the recursion $x_n = x_{n-157} \otimes x_{n-314} \otimes x_{n-471} \otimes x_{n-9689}$ (where $\otimes$ is the bit-wise exclusive-or operation) which is based upon tripling the largest primitive trinomial $x^{471} + x^{9689} = 1$ in the table of N. Zierler, Inform. Control **15**, 67 (1969). This generator has been found to give consistent results in many studies we have carried out; a similar generator was investigated and shown to be satisfactory by A. Compagner and A. Hoogland, J. Comput. Phys. **71**, 391 (1987). Its cycle length is $2^{9689} - 1$.
[16] A. P. Young and R. B. Stinchcome, J. Phys. C **8**, L535 (1975).
[17] A. B. Harris, T. C. Lubensky, W. K. Holcomb, and C. Dasgupta, Phys. Rev. Lett. **35**, 327 (1975).
[18] K. Binder and D. W. Heerman, *Monte Carlo Simulations in Statistical Physics* (Springer-Verlag, Berlin, 1988).
[19] T. Gebele, J. Phys. A **9**, L535 (1975).
[20] D. W. Heermann and D. Stauffer, Z. Phys. B **40**, 133 (1980).
[21] R. P. Langlands, C. Pichet, P. Pouliot, and Y. Saint-Aubin, J. Stat. Phys. **67**, 553 (1992).

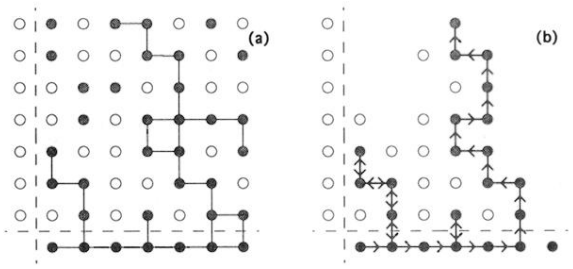

FIG. 1. (a) A typical system with $L = 7$ near the threshold, showing the percolating cluster connecting from the bottom to the top, and (b) the hull of the cluster only, as would be generated by the hull-generating walk. A row of occupied sites is added to the bottom, and a column of blocked sites is added to the left-hand side. The occupied sites are shown as solid circles and blocked ones as open circles. In (b), undecided sites are blank.