J. Phys. A: Math. Gen. **14** (1981) L169–L177. Printed in Great Britain

LETTER TO THE EDITOR

Order propagation near the percolation threshold

Robert Pike† and H Eugene Stanley‡§

† Department of Physics, California Institute of Technology, Pasadena, CA 91125 and Bell Laboratories, Murray Hill, New Jersey 07974, USA

‡ Center for Polymer Studies¶ and Department of Physics, Boston University, Boston, MA 02215, USA

Received 17 February 1981

Abstract. A new Monte Carlo method for studying bond percolation clusters is developed and used to identify new critical quantities associated with the percolation threshold. The bonds in each cluster are partitioned into three distinct connectivity classes, 'red' (singly connected backbone bonds), 'blue' (multiply connected backbone bonds) and 'yellow' (non-backbone bonds, often called dangling ends). Among the new cluster properties studied are the mean number of red bonds, a critical quantity diverging at p_c with exponent $\gamma_R \simeq 1$, and the length of the shortest connected path through the cluster which is critical with exponent $\gamma_{min} = 1.35 \pm 0.02$. For all cluster properties studied, we also compute averages over only the largest clusters; the corresponding critical exponents are found to be significantly different from those obtained by averaging over clusters of all sizes.

Introduction

Phase transitions are characterised by a distinct change in the manner in which correlations propagate through a system. How does one describe the propagation of order near a critical point? Although the classical literature focused on this question (see e.g. Zernike 1940, Ashkin and Lamb 1943), it was largely ignored for many years because intuitively plausible schemes such as the Ornstein–Zernike theory could not provide an answer consistent with experiment.

Interest in the propagation of order was rekindled a few years ago in the context of disordered systems (Stauffer 1975b, Birgeneau *et al* 1976, 1980, Cowley *et al* 1977, 1980a, b, Stanley *et al* 1976, Lubensky 1977). Consider for example, a site-dilute random magnet. When the fraction p of magnetic sites is very small, the system consists of small disconnected clusters of magnetically correlated sites. As $p \to p_c^-$, the mean cluster size increases until at $p = p_c$ a single cluster spans the entire lattice. The percolation threshold p_c is a critical point. By studying the propagation of magnetic correlations through the incipient infinite cluster that appears in the dilute magnet at its percolation threshold, one can obtain information about order propagation near this critical point.

The incipient infinite cluster dominates the behaviour of the system, and it is important to be able to describe its structure. If a cluster is considered as a network of wires carrying electrical current between two parallel bus bars, it can be decomposed

§ John Simon Guggenheim Memorial Fellow 1980–81.

¶ Supported in part by grants from ARO and ONR.

0305-4470/81/050169+09$01.50

into a conducting 'backbone' (Kirkpatrick 1978, Shlifer *et al* 1979) and many 'dangling ends' that do not contribute to the electrical conductivity (and hence order propagation) between the ends. Describing the topology of the backbone is a formidable unsolved problem. In this letter we report findings for bond percolation in two dimensions that provide insight into this problem.

Method

The backbone bonds may be divided into two classes, conveniently visualised as 'red' and 'blue' (see figure 1 of Stanley 1977). Red (blue) bonds are singly connected (multiply-connected); removing a single red bond breaks (does not break) the connection between the bus bars. The backbone is a linear chain of red bonds interrupted by blue 'blobs'. To study the connectivity distribution of bonds in the lattice, we have written a Monte Carlo computer program that generates complete percolation clusters and partitions their bonds into three distinct 'colours': red, blue and yellow (the dangling ends). To obtain correct statistics, the program must efficiently generate and analyse clusters of tens of thousands of bonds, *without introducing an artificial cut-off on cluster size*. The program implements a novel simulation procedure that includes several substantial improvements over previous algorithms. Each cluster starts as an initial 'seed' bond on a square lattice. Using a pseudorandom number generator, all adjacent bonds either join to the initial bond with probability p, or are marked as disconnected with probability $1-p$. This procedure continues recursively until the cluster is completely surrounded by a vacant perimeter. The cluster is stored as a doubly linked graph, with each cluster bond linked to each of its $z = 4d-2$ neighbours, but not to the perimeter. In the d-dimensional hypercubic lattice, the data consume $(d/2+1)(b+t)+(z+1)b$ words of computer memory, where b and t are the numbers of cluster and perimeter bonds, respectively. Since for large clusters, $t \propto b$ (see e.g. the recent reviews Stauffer 1979, Essam 1980), the storage and time to generate even the largest clusters increases only linearly with b, an important distinguishing feature of the method†.

Partitioning cluster bonds into the three colours is a problem in graph theory (Aho *et al* 1974). To identify the red bonds, the program first finds the *critical path*, which is the shortest path through the cluster from the furthest north bond (source) to the furthest south bond (sink). The critical path consists of all the red bonds and some of the blue bonds. The red bonds are then isolated by recursively descending the blue blobs until all paths converge at a critical path bond—the red bond that leads out of the blob.

We generated 40 000 clusters for the $d = 2$ (square) lattice, at each of 22 different values of p, from $p = 0.050$ to $p = 0.485$. In this system, p_c is exactly $\frac{1}{2}$, so the minimum value of $\varepsilon \equiv (p_c - p)/p_c$ is 0.03, well inside the critical region. The cluster distribution $n(b, p)$ differs by a factor of $b/2p$ from the distribution $n_0(b, p)$ generated by the more conventional method of filling a large finite box with bonds (Dean 1963, Hoshen and Kopelman 1976),

$$n(b, p) = \frac{1}{p}\frac{b}{d} n_0(b, p) \tag{1}$$

† Methods which embed clusters in a lattice, if implemented naively may consume time and storage proportional to r^d, whereas our method is proportional to b, Leath (1976a, b) and Reich and Leath (1980) used a hashing technique to limit resources to a small constant factor times b. We found clusters at $p = 0.485$ with over 60 000 bonds; the surrounding square lattice contains over 10^6 bonds.

because each cluster starts from a given bond (the factor b/p) at a fixed orientation (the factor d). The factor b makes the physically important large clusters much more numerous. For example, to generate one 60 000-bond cluster near p_c by the box-filling method requires generating 60 000 times as many one-bond clusters as required in the present method, 30 000 times as many two-bond clusters, and so forth. Thus for a given number of clusters generated, our method provides much more accurate statistics for large clusters.

The method has several other important advantages. There are no boundary problems: the lattice on which the simulation is carried out has no walls. Because the entire cluster and its perimeter are stored, and isolated from the rest of the lattice, it is simple to study directly many properties of the clusters and their perimeters. There are, however, some disadvantages. The most important is that the finite memory of our computer limits the clusters to about 100 000 bonds. It is this limit which keeps the simulation below $p = 0.485$; at higher values of p, some large clusters would have to be rejected and the complete statistical distribution would not be obtained.

Is the present procedure sufficiently accurate with 40 000 cluster realisations? Our answer is that it is in fact accurate enough with only 10 000 realisations, and our main results were checked by averaging separately four groups of 10 000 realisations apiece. Among the checks on the accuracy are the following.

(i) The critical threshold p_c, calculated several different ways, is within 0.01 of the exact value, $\frac{1}{2}$ (Sykes and Essam 1963).

(ii) The cluster distribution $n(b, p)$ agrees with exact results for $b \leqslant 13$ (Sykes *et al* 1981), and with Monte Carlo results using the box-filling method (Nakanishi and Stanley 1980 and references therein). We find, of course, fluctuations about the exact values; however the deviations are of order $\sqrt{n(b, p)}$ and there is no *systematic* discrepancy such as occurs in box-filling methods due to finite-size effects.

(iii) The functions $G_k(p)$ defined through

$$G_k(p) = \sum_{b=1}^{k} n_0(b, p) \tag{2}$$

agree with calculations for $k = 1, 2, \ldots, 13$ (Sykes *et al* 1981) and, when extrapolated to p_c, with the exact result $G_\infty(p_c) = 0.0981 - (1 - p_c)^4$ (Temperley and Lieb 1971).

Results for individual clusters

For each cluster j ($j = 1, \ldots, 40\,000$ for each value of p), we first calculated four quantities $x_m(j)$:

(1) the total number of cluster bonds $x_1(j) \equiv b(j)$,

(2) the perpendicular distance between the bus bars or 'cluster diameter' $x_2(j) \equiv r(j)$,

(3) the number of bonds on the critical path $x_3 \equiv b_{\min}(j)$ and

(4) the number of red bonds $x_4 \equiv b_R(j)$.

To examine the dependence of $x_m(j)$ on $x_n(j)$, we then make six 'scatter plots' on double logarithmic paper, with one point for each cluster j. We did this for each of the 22 values of p; an example for $p = 0.485$ is given in figure 1. The plots are remarkably linear, motivating the definition of a family of numbers $\psi_{mn}(p)$ defined by

$$x_m(j) \sim x_n(j)^{\psi_{mn}(p)} \quad \text{as } x_n(j) \to \infty. \tag{3}$$

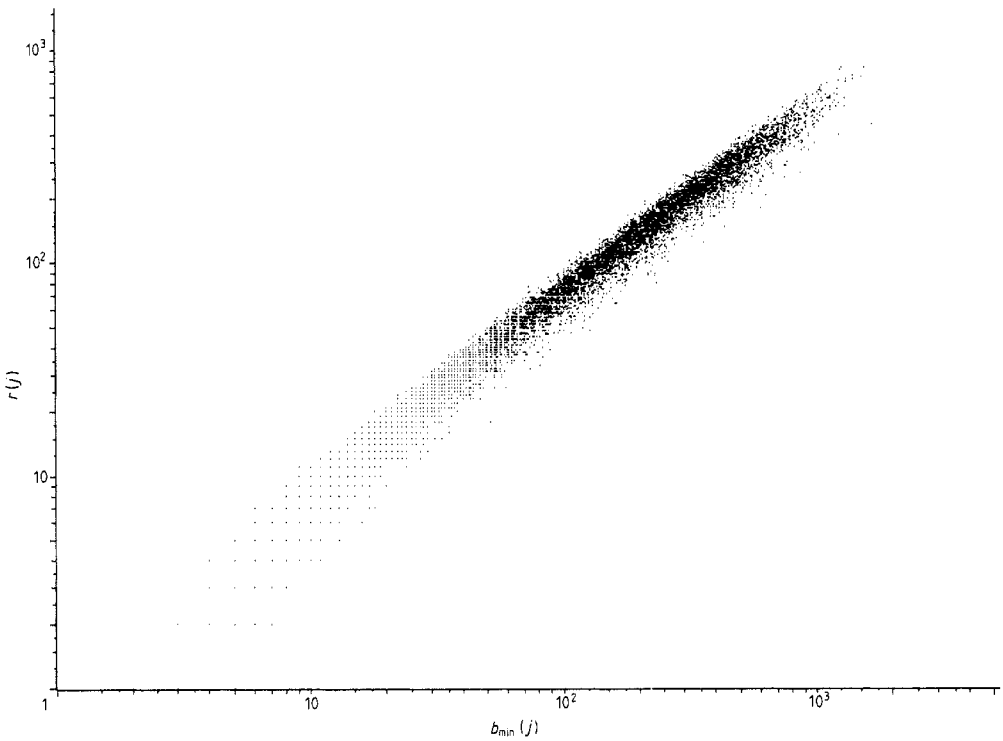

Figure 1. A double logarithmic scatter plot showing the relationship between $b_{\min}(j)$ and $r(j)$ at $p = 0.485$; $\varepsilon = 0.03$. Each point is one realisation; 10 000 realisations are shown.

Table 1 shows our results for the $\psi_{mn}(p)$, obtained using a least-squares fit to the raw data (40 000 points for each p value).

The $\psi_{mn}(p)$ appear (for unknown reasons) to vary roughly linearly with $\sqrt{\varepsilon}$, making possible reliable estimates (table 1) for the critical exponents ψ_{mn} defined through

$$\psi_{mn} \equiv \lim_{p \to p_c} \psi_{mn}(p). \tag{4}$$

Table 1. Estimates for the quantities $\psi_{mn}(p)$ defined in equation (3), based on a least-squares fit to 40 000 data points (the case $p = 0.485$ is illustrated in figure 1). Also shown are the extrapolated values ψ_{mn} defined by equation (4). A useful consistency check is the relation $\psi_{ij} = \psi_{ik}\psi_{kj}$ for any k. Because b_R has a large range among clusters of a given b, the plots involving b_R have much wider scatter than the others, and the b_R exponents are less well determined. In particular, $\psi_{34}(p)$ cannot be estimated reliably.

p	$\psi_{12}(p)$	$\psi_{13}(p)$	$\psi_{14}(p)$	$\psi_{23}(p)$	$\psi_{24}(p)$
0.440	1.712 ± 0.036	1.526 ± 0.036	1.694 ± 0.060	—	0.987 ± 0.005
0.450	1.724 ± 0.024	1.526 ± 0.024	1.754 ± 0.047	—	0.994 ± 0.005
0.460	1.739 ± 0.015	1.538 ± 0.012	1.754 ± 0.031	—	1.005 ± 0.005
0.470	1.754 ± 0.009	1.547 ± 0.010	1.805 ± 0.033	0.800 ± 0.020	1.020 ± 0.005
0.475	1.773 ± 0.009	1.552 ± 0.005	1.818 ± 0.027	0.810 ± 0.012	1.030 ± 0.005
0.480	1.782 ± 0.006	1.572 ± 0.007	1.862 ± 0.024	0.820 ± 0.010	1.044 ± 0.005
0.485	1.795 ± 0.006	1.579 ± 0.008	1.923 ± 0.026	0.835 ± 0.007	1.060 ± 0.005
Extrapolation	1.88 ± 0.02	1.64 ± 0.02	2.20 ± 0.15	0.92 ± 0.02	1.16 ± 0.02

One exponent has been discussed previously: ψ_{12}, the fractal dimension d_f^+ (ψ_{12}^{-1} has been called ρ by Leath 1976a, Harrison *et al* 1978, Peters *et al* 1979).

Finally, we generated 8000 clusters at $p = p_c$, rejecting those that grew to over 20 000 bonds (about two thirds of the clusters grown were rejected). The well verified scaling property (Stauffer 1979, Essam 1980) of the distribution $n_0(b, b)$ for large b and small ε,

$$n_0(b, \varepsilon) = b^{-\tau} f(\varepsilon s^{\sigma}) \tag{5}$$

(with $\tau = 2 + 1/\delta$ and $\sigma = 1/\beta\delta$) result in the prediction that $n(b, p_c) \sim b^{-\tau+1}$. Our data produce the estimate $\tau = 2.050 \pm 0.005$ ($\delta = 20 \pm 2$).

Results for averaged quantities

We calculated the averages of the quantities $x_l(j)$ with respect to the cluster distribution $n(b, p)$,

$$\langle x_l \rangle = \sum_{j=1}^{40\,000} n(b, p) x_l(j). \tag{6}$$

In addition to the four quantities $x_1(j), \ldots, x_4(j)$ defined above, we included the number of cluster perimeter bonds $x_5(j) \equiv t(j)$.

First, we tested the equation

$$\frac{\langle b \rangle - 1}{\langle t \rangle} = \frac{p}{1-p} \tag{7}$$

which has a colourful history of 'proofs' of increasing degrees of rigour and generality (Stauffer 1975a, Leath 1976a, Kunz and Souillard 1978, Coniglio and Russo 1979). Our results for all 22 values of p are displayed in figure 2, while the straight line with slope unity is the prediction of equation (7). The agreement is striking, and provides the first corroboration of (7) by direct cluster simulation (see also Coniglio and Stauffer 1980, Leath and Reich 1978).

We also studied the asymptotic dependence on ε of the averaged quantities $\langle x_l \rangle$ ($l = 1, \ldots, 4$). We found that if we include in the definition (6) only the largest P% of the clusters ($P = 15, 10, 7, 5, 2, 1$), we obtain P-dependent values of the exponents which are significantly different from the exponents obtained by averaging over *all* clusters (figure 3). Moreover, we find that the P-dependent exponents vary smoothly with P and hence can be extrapolated to $P = 0$. The exponents so obtained are listed in the second column of table 2. In particular, ν is in accord with the recent conjecture of den Nijs (1979), with recent estimates using the renormalisation group (Reynolds *et al* 1980, Eschbach *et al* 1981) and with exact enumeration methods (Dunn *et al* 1975). It follows from (5) that $S_T \sim \varepsilon^{\beta\delta}$ for the largest clusters. According to the extended den Nijs conjecture (Nienhuis *et al* 1980, Pearson 1980), $\beta\delta = \beta + \gamma = 91/36$, which is comparable with our estimate 2.48 ± 0.02.

We find that the quantities $\langle x_3 \rangle = \langle b_{\min} \rangle \equiv S_{\min} \sim \varepsilon^{-\gamma_{\min}}$ and $\langle x_4 \rangle = \langle b_R \rangle \equiv S_R \sim \varepsilon^{-\gamma_R}$ are also critical (table 2). Our finding $\gamma_{\min} > \nu$ implies that the minimum path is more convoluted in large clusters than small clusters, and increases much faster than linearly with cluster radius. Kirkpatrick (1979) found a similar result for a correlated percolation model. Coniglio (1981) has independently argued that $\gamma_R = 1$, and our results (table 2) are consistent with his result.

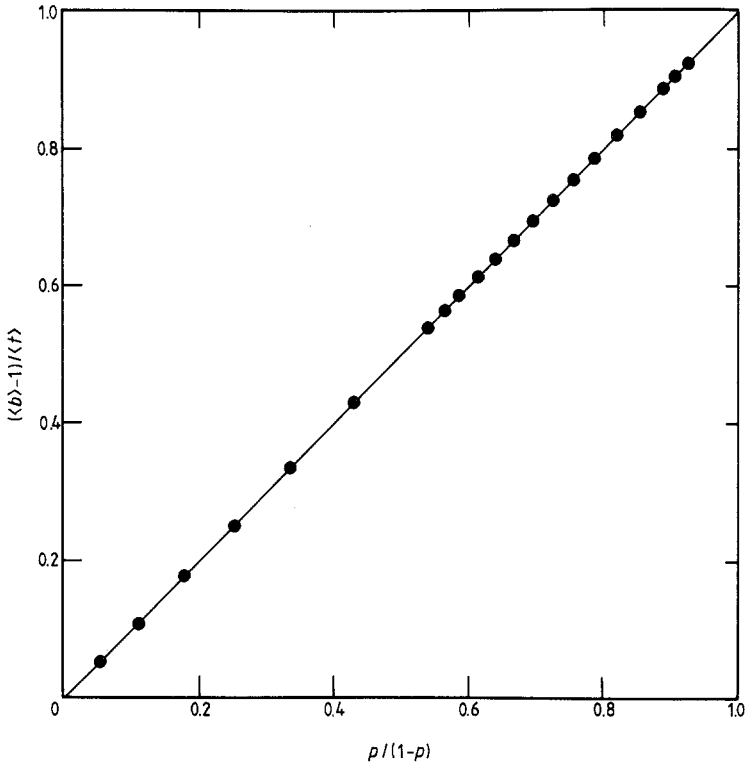

Figure 2. Test of equation (7). Points are from the simulation while the line is equation (7).

The following heuristic argument may help put the new exponent γ_R into perspective. Consider cluster growth to be controlled by a random number between 0 and 1 'attached' to each potential bond in the lattice; p is infinitesimally slowly increased from 0, and when it reaches the number stored at a bond, the bond becomes occupied. In this model, bonds form one at a time, and at the instant of formation a bond can be immediately 'coloured.' Two competing mechanisms will control the colour of a bond: (1) if a bond forms a connected path between two otherwise disjoint clusters, it is coloured red; (2) if a bond forms an alternate connected path, it is coloured blue, but may also force a previously red bond to change colour. Thus, as each bond forms, the cluster containing it is recoloured. Although the exponent for blue bonds γ_B is greater than that for red bonds (Shlifer *et al* 1979), $\gamma_R > 0$ implies that mechanism (1) is an important one and is not fully dominated by (2): a significant fraction of cluster growth occurs along singly connected subsections of the cluster.

Conclusions

In conclusion, we have developed a Monte Carlo method that does not introduce boundary effects; since it stores only the cluster and its parameter, it can be used for clusters up to 10^5 bonds. We have used this method to solve the problem of distinguishing the singly connected (red) from the multiply connected (blue) backbone bonds. Among the new findings are the following.

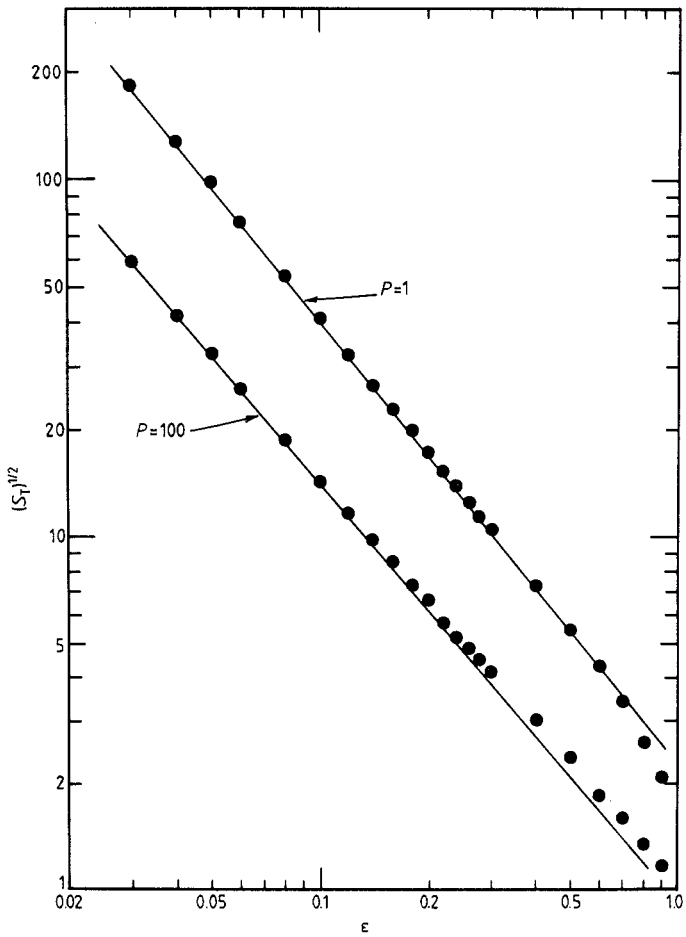

Figure 3. Double logarithmic plot of $\sqrt{S_T}$ against ε for all clusters (lower curve) and for the largest 1% of the clusters (upper curve). When only the largest P% of the clusters in definition (6) are used, the exponents increase, approaching a limiting value as $P \to 0$, in accord with the expectation of the cluster number scaling hypothesis (equation (5)).

Table 2. Critical exponents for averages over all clusters, and over the largest clusters.

Variable	Exponents (all clusters)	Exponent (largest clusters)
Size $\langle x_1 \rangle = \langle b \rangle = S_T$	$\gamma = 2.35 \pm 0.05$	$\beta\delta = 2.48 \pm 0.02$
Spanning distance $\langle x_2 \rangle = \langle r \rangle$	$\nu_a = 1.21 \pm 0.03$	$\nu = 1.33 \pm 0.01$
Critical path length $\langle x_3 \rangle = \langle b_{\min} \rangle$	$\gamma_{\min,a} = 1.35 \pm 0.02$	$\gamma_{\min} = 1.49 \pm 0.02$
Red bonds $\langle x_4 \rangle = \langle b_R \rangle$	$\gamma_{R,a} = 1.0 \pm 0.1$	$\gamma_R = 1.0 \pm 0.1$

(i) S_R, the number of singly connected 'red' bonds, is a critical quantity that diverges as $S_R \sim \varepsilon^{-\gamma_R}$, with $\gamma_R \simeq 1$, consistent with very recent independent work of Coniglio (1981).

(ii) $S_{\min}$, the number of bonds in the shortest connected path through the cluster, diverges as $S_{\min} \sim \varepsilon^{-\gamma_{\min}}$ with $\gamma_{\min} \simeq 1.35 \pm 0.02$.

(iii) Averages of critical quantities over the largest $P\%$ of the clusters result in exponents exhibiting a smooth (approximately linear) dependence upon P; the limiting $P = 0$ exponents are consistent with predictions of the cluster-number scaling hypothesis (equation (5)).

(iv) No deviations whatsoever occur as $p \to p_c^-$ in the simple relation (7) between cluster area and perimeter.

Acknowledgments

The authors gratefully acknowledge the assistance of D Hugh Redelmeier in developing the computer program, and helpful interactions with R J Birgeneau, J H Condon, A Coniglio, T D S Duff, F Family, S Kirkpatrick, W Klein, R B Leighton, H Nakanishi, P Pfeuty, S Redner, G R Reich, P J Reynolds, G Shlifer, D Stauffer and S Wolfram. We also thank D S Gaunt for communicating his results to us prior to their publication (Sykes *et al* 1981).

References

Aho A V, Hopcroft J E and Ullman J D 1974 *The design and analysis of computer algorithms* (Reading, Mass.: Addison-Wesley)
Ashkin J and Lamb W E 1943 *Phys. Rev.* **64** 159
Birgeneau R J, Cowley R A, Shirane G and Guggenheim H J 1976 *Phys. Rev. Lett.* **37** 940
Birgeneau R J, Cowley R A, Shirane G, Tarvin J A and Guggenheim H J 1980 *Phys. Rev.* B **21** 317
Coniglio A 1981 *Phys. Rev. Lett.* **46** 250
Coniglio A and Russo L 1979 *J. Phys. A: Math. Gen.* **12** 545
Coniglio A and Stauffer D 1980 *Nuovo Cimento Lett.* **28** 33
Cowley R A, Shirane G, Birgeneau R J, Svensson E C and Guggenheim H J 1977 *Phys. Rev. Lett.* **39** 894
Cowley R A, Birgeneau R J, Shirane G, Guggenheim H J and Ikeda H 1980a *Phys. Rev.* B **21** 4038
Cowley R A, Shirane G, Birgeneau R J, Svensson E C and Guggenheim H J 1980b *Phys. Rev.* B **22** 4412
Dean P 1963 *Proc. Camb. Phil. Soc.* **59** 397
Dunn A G, Essam J W and Ritchie D S 1975 *J. Phys. C: Solid State Phys.* **8** 4219
Eschbach P D, Stauffer D and Hermann J 1981 *Phys. Rev.* B **23** 422
Essam J W 1980 *Rep. Prog. Phys.* **43** 833
Frisch H L, Gordon S B, Vyssotsky V A and Hammersley J M 1962 *Bell. Sys. Tech. J.* **41** 909
Hammersley J M 1964 *Monte Carlo Methods* (London: Methuen)
Harrison R J, Bishop G H and Quinn G D 1978 *J. Stat. Phys.* A **13** L189
Hoshen J and Kopelman R 1976 *Phys. Rev.* B **14** 3438
Kirkpatrick S 1978 *AIP Conf. Proc.* **40** 99
—— 1979 *AIP Conf. Proc.* **58** 79
Kunz H and Souillard B 1978 *J. Stat. Phys.* **19** 77
Leath P L 1976a *Phys. Rev. Lett.* **37** 940
—— 1976b *Phys. Rev.* B **14** 5046
Leath P L and Reich G R 1978 *J. Phys. C: Solid State Phys.* **11** 4017
Lubensky T C 1977 *Phys. Rev.* B **15** 311
Nakanishi H and Stanley H E 1980 *Phys. Rev.* B **22** 2466
Nienhuis B, Riedel E K and Schick M 1980 *J. Phys. A: Math. Gen.* **13** L189
den Nijs M P M 1979 *J. Phys. A: Math. Gen.* **12** 1857
Pearson R B 1980 *Phys. Rev.* B **22** 2579
Peters H P, Stauffer D, Holters H P and Loewenich K 1979 *Z. Phys.* B **34** 399
Reich G R and Leath P L 1980 *Comput. Phys. Commun.* **10** 29
Reynolds P J, Stanley H E and Klein W 1980 *Phys. Rev.* B **21** 1223
Shlifer G, Klein W, Reynolds P J and Stanley H E 1979 *J. Phys. A: Math. Gen.* **12** L169

Stanley H E 1977 *J. Phys. A: Math. Gen.* **10** L211
Stanley H E, Birgeneau R, Reynolds P J and Nicoll J F 1976 *J. Phys. C: Solid State Phys.* **9** L553
Stauffer D 1975a *J. Phys. C: Solid State Phys.* **8** L172
—— 1975b *Z Physik.* B **22** 161
—— 1979 *Phys. Rept.* **54** 1
Sykes M F and Essam J W 1963 *Phys. Rev. Lett.* **10** 3
Sykes M F, Gaunt D S and Glen M 1981 *J. Phys. A: Math. Gen.* **14** 287
Temperley H N V and Lieb E H 1971 *Proc. R. Soc.* A **332** 251
Zernike F 1940 *Physica* **7** 565