Journal of Statistical Physics, Vol. 47, Nos. 1/2, 1987

Probability Distribution for Percolation Clusters Generated on a Cayley Tree at Criticality

S. Havlin,[1,2] J. E. Kiefer,[2] F. Leyvraz,[3] and G. H. Weiss[2]

Received August 28, 1986; revised December 22, 1986

We present analytical and numerical results for the probability distributions of the number of sites S as a function of the number of shells l for several ensembles of percolation clusters generated on a Cayley tree at criticality. We find that for the incipient infinite percolation cluster the probability distribution is $P(S|l) \sim (S/l^4) \exp(-aS/l^2)$ for $S \gg l \gg 1$.

KEY WORDS: Random walks; random media; scaling behavior.

In recent years, there has been great interest in the concept of the fractal dimension of random aggregates.[1] This concept has been applied in various fields and its properties have been extensively studied.[1–6] The fractal dimension d_f of a random aggregate is defined as the exponent characterizing the relation between the average mass S of the cluster and a parameter L that characterizes its length

$$\langle S \rangle \sim L^{d_f} \tag{1}$$

Much effort has been devoted to the exploration of this relation for different types of random aggregates. However, little is known about the probability distribution $P(S|L)$ whose first moment has the scaling form shown in Eq. (1).

Very recently, probability distributions for several fractal properties have been found to be of log-normal type, for which a hierarchy of exponents characterizes the moments.[7–9] A related characterizing property

[1] Department of Physics, Bar-Ilan University, Ramat-Gan, Israel.
[2] Physical Sciences Laboratory, DCRT, National Institutes of Health, Bethesda, Maryland, 20892.
[3] Center for Polymer Studies, Boston University, Boston, Massachusetts.

0022-4715/87/0400-0173$05.00/0

of the fractal is the probability distribution of the mass as a function of L. We find normal distributions for several fractal systems generated on a Cayley tree with a single gap exponent for the set of moments. The study of the probability distribution is also important because the entropy of a system can be derived from it, allowing physically useful parameters to be calculated similar to those found for the ensemble of self-avoiding walks.[10] Several authors have addressed similar problems for percolation clusters on the Bethe lattice.[11 15] However, both the techniques and emphasis differ from those of the present work.

In this paper, using a combination of exact and numerical methods, we study the probability distributions for the total mass as a function of the chemical distance[12] from the origin for several related percolation ensembles generated on a Cayley tree at criticality. The coordination number of the Cayley tree is taken equal to 3 and the critical concentration for such a fractal is $p_c = 0.5$. The analysis to be presented assumes a percolation cluster grown on a Cayley tree at criticality using the Leath[16]–Alexandrowicz[17] algorithm. Thus, at each node, a bond has the possibility of generating 0, 1, or 2 bonds in the next generation with probabilities 1/4, 1/2, and 1/4, respectively. The number of bonds at a chemical distance l from the origin will be denoted by $B(l)$ and the cumulative number of bonds up to and including those at distance l will be denoted by

$$S(l) = \sum_{l'=1}^{l} B(l') \tag{2}$$

These quantities are illustrated in Fig. 1. Two distributions will be of interest: $P(B \mid l)$, the distribution of the number of bonds at distance l, and $P(S \mid l)$, the corresponding distribution of S.

Useful results can be obtained for four different models that define the ensembles of clusters and bonds to be counted.

1. All generated clusters, including those that terminate, are to be counted. If termination occurs at l', the resulting structure having S' bonds, then $S = S'$ and $B(l) = 0$ for $l \geqslant l'$.
2. Only clusters that ultimately grow to infinity are included in the counts.
3. Clusters that terminate at l' are taken into account for $l \leqslant l'$, but not for $l > l'$.
4. All clusters are forced to grow. If there is exactly one growth site left in the cluster, the probability that a new bond is added is equal to 1.

The analytical tools used to calculate the appropriate probabilities fall into two classes. Moments of S and B were calculated for models 1 and 2 using

Shell number l	Number of bonds up to shell, $S(l)$	Number of bonds in shell, $B(l)$
1	1	1
2	3	2
3	5	2
4	8	3
5	12	4
6	14	2

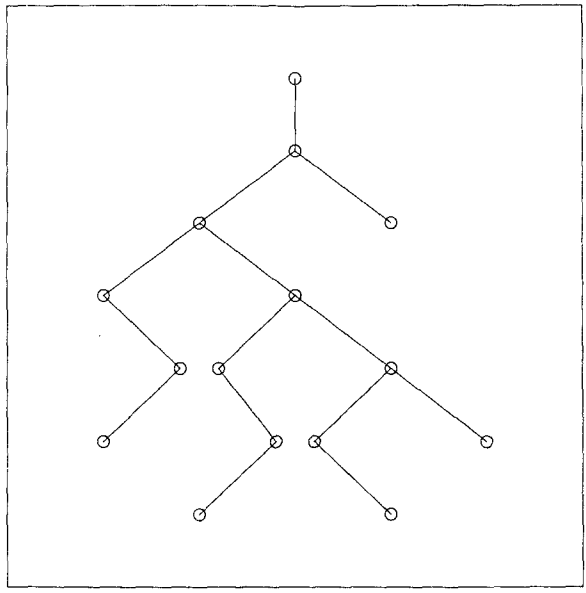

Fig. 1. An example of a cluster with coordination number 3 generated on a Cayley tree at critically.

techniques common in the theory of branching processes,[18] and from these moments the probabilities (or densities in the appropriate limit) are found. Asymptotic results for models 3 and 4 are obtained using random walk techniques.[19]

MODEL 1

The moments of S and B were calculated using generating function techniques, as is done in the theory of branching processes,[18] some details

of which are given in the Appendix. Numerical values of the moments up to order 100 were calculated using these methods and are found to have the form

$$\langle B^n(l)\rangle = \frac{n!\, l^{n-1}}{b_1^{n-1}}, \qquad \langle S^n(l)\rangle = \frac{n!\, l^{2n-1}}{c_1^{n-1}}, \qquad n = 1, 2,... \tag{3}$$

for values of $l \leqslant 4000$ with $b_1 = 5.0$ and $c_1 = 10.0$. These relations are to be supplemented by the obvious relations $\langle B^0(l)\rangle = \langle S^0(l)\rangle = 1$. In order to find an expression for the probability (or probability density in the limit of large S and B), we evaluate the characteristic function in terms of moments making the assumption that Eq. (3) is valid for all n. Since B and S are nonnegative, the most convenient form of the characteristic function is

$$C_B(z) = \int_0^\infty P_1(B|l)\, e^{-zB}\, dB = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \langle B^n\rangle\, z^n \tag{4}$$

with a similar expression for $C_S(z)$. On substituting the expression for $\langle B^n\rangle$ from Eq. (3) into this last expression and performing the summation, we find that

$$C_B(z) = \left(1 - \frac{b_1}{l}\right) + \left(\frac{b_1}{l}\right)^2 \frac{1}{z + b_1/l} \tag{5}$$

which is equivalent to

$$P_1(B|l) = \left(1 - \frac{b_1}{l}\right)\delta(B) + \left(\frac{b_1}{l}\right)^2 e^{-b_1 B/l} \tag{6}$$

where the first term obviously accounts for terminated clusters. In a similar fashion, one finds for $P_1(S|l)$

$$P_1(S|l) = \left(1 - \frac{c_1}{l}\right)\delta(S) + \frac{c_1^2}{l^3} \exp\left(-\frac{c_1}{l^2} S\right) \tag{7}$$

MODEL 2

The same strategy, that of finding the functional form of the moments and summing the resulting characteristic function, was used. For this model we were not able to fit the moments for all n, but rather for $n > 5$ we found

$$\langle B^n\rangle \sim 1.187 n!\, l^n / b_2^{n-1} \tag{8a}$$

$$\langle S^n\rangle \sim 0.156 n!\, l^{2n} / c_2^{n-1} \tag{8b}$$

where $b_2 = 3.653$ and $c_2 = 9.375$. Equation (8) for $l \geqslant 400$ is in agreement with accurately calculated values of the moments to approximately 1% or better for $n > 5$. If we substitute Eq. (8a) into Eq. (4), we find that

$$C_B(z) = 4.337 \sum_{n=0}^{\infty} (-1)^n \frac{l^n z^n}{b_2^n} + \sum_{n=0}^{5} \frac{\langle \Delta B_n \rangle}{n!} (-1)^n z^n \tag{9}$$

where $\langle \Delta B_n \rangle$ measures the discrepancy between the numerically calculated value of $\langle B_n \rangle$ and the value given in Eq. (8a). The inverse of the first line of Eq. (9) is

$$P_2(B \mid l) = (15.843/l) \exp(-b_2 B/l) \tag{10}$$

again, in the approximation in which the lattice model can be replaced by a continuum description. In a similar fashion, one can show that for S one has

$$P_2(S \mid l) = (150.35/l^4) \exp(-c_2 S/l^2) \tag{11}$$

It can be shown from the form of the moments in Eq. (8) that they uniquely determine the distribution.[20] The remaining terms must now be taken into account. In the continuum approximation that we have adopted, the inverse of the term $(-1)^n z^n$ is the nth derivative of a delta function evaluated at the origin. Hence, if one only has a finite number of discrepant terms, as is the case in Eq. (9), the most serious errors occur at the origin. In order to incorporate this observation into our approximation to $P_2(S \mid l)$, let us assume that this density can be expressed as

$$P_2(S \mid l) = (150.35/l^4) f(S/l^2) \exp(-c_2 S/l^2) \tag{12}$$

where $f(x)$ is a smoothly varying function of x with the properties $f(0) = 0$, $f(\infty) = 1$. By examining exact numerical data, we found that an accurate form for $f(x)$ is

$$f(x) = a e^{-\lambda/x} \tag{13}$$

where $\lambda = 0.07$ and a is a constant. Thus, in this approximation the normalized form for $P_2(S \mid l)$ is

$$P_2(S \mid l) = \frac{S/l^4}{2(\lambda/c_2) K_2(2(c_2\lambda)^{1/2})} \exp\left(-\frac{\lambda l^2}{S} - c_2 \frac{S}{l^2}\right) \tag{14}$$

with moments

$$\langle S^n \rangle = l^{2n} \left(\frac{\lambda}{c_2}\right)^{n/2} \frac{K_{2+n}(2(c_2\lambda)^{1/2})}{K_2(2(c_2\lambda)^{1/2})} \tag{15}$$

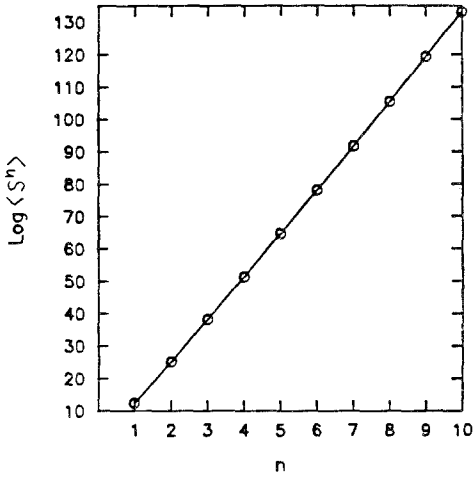

Fig. 2. Plot of $\log\langle S^n \rangle$ versus n for model 2. The circles are numerical values for $l = 1000$ and the line gives the theoretical values using Eq. (15).

where $K_m(x)$ is a modified Bessel function of the second kind of order m. The agreement between the moments calculated from Eq. (15) and those obtained numerically is illustrated in Figs. 2 and 3.

MODELS 3 AND 4

In the first two models, no account was taken of blocked sites, i.e., sites from which no further branching occurred. In the present models,

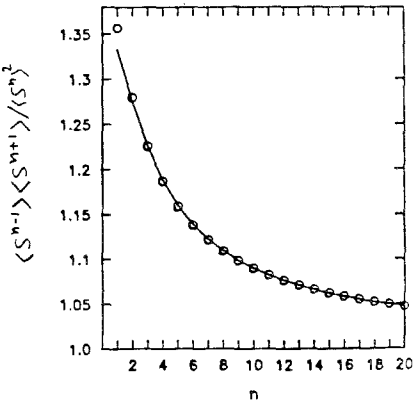

Fig. 3. The ratio $\langle S^{n+1} \rangle \langle S^{n-1} \rangle / \langle S^n \rangle^2$ versus n plotted for model 2. This ratio is independent of l. The circles are the numerical values and the line the theoretical prediction using Eq. (15).

these will be included in the count, but it will be convenient in addition to define two new variables $B'(l)$ and $S'(l)$, where $B'(l)$ is the number of growth sites (at distance l) that are adjacent to an occupied site, but are themselves neither occupied nor blocked from further growth. The variable $S'(l)$ is the total number of occupied sites included in all levels less than l. The differences between B' and B, S', and S are illustrated in Fig. 1. When B and S are much greater than 1, the relationship between the primed and unprimed variables is approximately $B' = 2B$, $S' = 2S$.

As a first step, we derive an equation for $B'(S')$, following which the dependence on the distance l will be included. The two possibilities for changing $B'(S')$ are

$$\begin{aligned} B'(S'+1) &= B'(S')+1 \qquad \text{if one adds an occupied site to the cluster} \\ B'(S'+1) &= B'(S')-1 \qquad \text{if one adds a blocked site to the cluster} \end{aligned} \tag{16}$$

These two possibilities occur with probability 1/2, so that the process just mentioned constitutes an ordinary random walk in one dimension. If $P(B' \mid S')$ denotes the conditional probability of B' given S', then in the continuum limit Eq. (16) is equivalent to

$$\frac{\partial^2 P}{\partial S'} = \frac{1}{2}\frac{\partial^2 P}{\partial B'^2} \tag{17}$$

Model 3 requires that any cluster be destroyed when it reaches $B' = 0$, which implies that Eq. (17) is to be solved subject to an absorbing boundary condition at $B' = 0$. Model 4 requires that at any time at which $B' = 0$ is reached, the process is restarted at $B' = 1$. Hence, one must solve Eq. (17) subject to a reflecting boundary condition. Thus, for large B' and S' we can assert that

$$P_3(B' \mid S') \sim \frac{aB'}{(S')^{3/2}} \exp\left(-\frac{B'^2}{2S'}\right) \tag{18}$$

$$P_4(B' \mid S') \sim \frac{b}{(S')^{1/2}} \exp\left(-\frac{B'^2}{2S'}\right) \tag{19}$$

for the resulting probability densities, where a and b are constants. Since model 3 involves the absorption of the random walk, it is not normalizable to unity, since there is a loss proportional to $(S')^{-1/2}$. On the other hand, since there is no such loss in model 4,

$$\sum_B P_4(B' \mid S') = 1 \tag{20}$$

Equations (17) and (18) give results for $P(B' \mid S')$, but the densities that are of greater interest are of the form $P(B' \mid l)$ and $P(S' \mid l)$. We will be able to calculate $P_3(B' \mid l)$ and $P_4(B' \mid l)$ exactly, but are only able to calculate $P_4(S' \mid l)$.

To obtain results for $P(B' \mid l)$, we first parametrize S' in terms of l. This is easily done by considering the following identity:

$$S'(l+1) - S'(l) = B'(l) \tag{21}$$

This means that in one "l-step," the walk performs $B(l)$ ordinary steps, i.e., the diffusion constant is proportional to B. Hence, using the continuum limit, we obtain

$$\frac{\partial P(B' \mid l)}{\partial l} = \frac{\partial}{\partial B'} B' \frac{\partial}{\partial B'} P(B' \mid l) \tag{22}$$

so that l is a time parameter. If we substitute $r^2/2 = B'$ into this equation, we find

$$\frac{\partial P}{\partial l}(r \mid l) = \frac{1}{2r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} P(r \mid l) \tag{23}$$

so that $P(r \mid l)$ is the solution to the radial equation describing two-dimensional diffusion. Hence, one finds

$$P_3(r \mid l) = r^2 e^{-r^2/l}/l^3 \qquad \text{or} \qquad P_3(B' \mid l) = B' e^{-B'/l}/l^3 \tag{24}$$

and

$$P_4(r \mid l) = e^{-r^2/l}/l \qquad \text{or} \qquad P_4(B' \mid l) = e^{-B'/l}/l \tag{25}$$

To obtain $P(S' \mid l)$ is somewhat less simple. We start from the identity

$$S'(l) = \sum_{l' \leqslant l} B'(l') = \sum_{l' < l} r^2(l') \tag{26}$$

where $r(l)$ is the distance from the origin of the two-dimensional random walk describing the growth of the cluster as a function of the number of shells occupied. If we pass to the continuum limit, then what is required is clearly the distribution of the random variable

$$I(l) = \int_0^l |\mathbf{r}(l')|^2 \, dl' \tag{27}$$

In the following, we disregard the issue of absorption at the origin, i.e., we work entirely with Neumann boundary conditions (model 4). This is

because otherwise the calculations are complicated. To solve this problem, we effectively use the Feynman–Kac formula[21] by calculating the characteristic function

$$G(q, l) = \langle \exp[-qI(l)] \rangle \tag{28}$$

where the average is taken over all two-dimensional random walks. This is the Laplace transform of the distribution function. If one knows $G(q, l)$ for all q, then by inversion one can find the underlying density. To find this function, we proceed as follows. Define the conditional characteristic function

$$G(q, l|\mathbf{r}) = \int P(\boldsymbol{\rho}, l) \langle \exp[-qI(l)] | \mathbf{r} \rangle \, d^2\boldsymbol{\rho} \tag{29}$$

where $\langle \cdots | \mathbf{r} \rangle$ denotes the average taken over all random walks terminating at $\mathbf{r}$. We also define

$$\begin{aligned} F(q; \mathbf{r}, l) &= \langle \exp[-qI(l)] | \mathbf{r} \rangle \, P(\mathbf{r}, l) \\ &= \sum_{\substack{\text{RW ending} \\ \text{at } \mathbf{r} \text{ at time } t}} \exp\left[-q \int_0^l |\mathbf{r}(l')|^2 \, dl' \right] \end{aligned} \tag{30}$$

This can vary with time due either to a change in the upper limit of the integral when one holds the random walk fixed, or holding t fixed and considering a different random walk. From this consideration, or by direct application of the Feynman–Kac formalism,[21] it follows that

$$\partial F/\partial l = -q \, |\mathbf{r}|^2 \, F(q; \mathbf{r}, l) + \Delta F \tag{31}$$

where Δ is the Laplacian. This equation can be solved by making the *ansatz*

$$F(q; \mathbf{r}, l) = A(l) \exp[-B(l) \, |\mathbf{r}|^2] \tag{32}$$

For Eqs. (31) and (32) to be consistent, A and B must satisfy

$$\partial A/\partial l = -4AB, \qquad \partial B/\partial l = q - 4B^2 \tag{33}$$

These equations must be solved subject to the initial condition that $F(q; \mathbf{r}, 0) = \delta(\mathbf{r})$. The solution to Eq. (32) having this property is

$$A(l) = \operatorname{csch}(2l\sqrt{q}), \qquad B(l) = \frac{\sqrt{q}}{2} \coth 2(l\sqrt{q}) \tag{34}$$

Thus, the probability $P(S'|l)$ is the inverse Laplace transform of $G(q, l)$, where

$$\begin{aligned} G(q, l) &= \int F(q; \mathbf{r}, l) \, d^2\mathbf{r} = A(l) \int e^{-r^2 B(l)} \, d^2\mathbf{r} \\ &\sim A(l)/B(l) \sim \operatorname{sech}(2l\sqrt{q}) \end{aligned} \tag{35}$$

For the case $l \ll \sqrt{S'}$ or, equivalently, $\sqrt{q}\, l \ll 1$,

$$G(q, l) = (1 + 2ql^2)^{-1} \tag{36}$$

whose inverse transform is

$$P_4(S' \mid l) \sim (1/l^2)\, e^{-S/l^2} \tag{37}$$

In the opposite limit $S \ll l^2$, Eq. (35) implies that

$$P_4(S' \mid l) \sim (l/S^3)\, e^{-l^2/S} \tag{38}$$

In Table I, we summarize the various results we obtained for the moments as well as our results for the asymptotic distributions. It is seen that by varying the models the distribution form is affected only in the prefactors of the exponent. The exponential forms $\exp(-B'/l)$, $\exp(-S'/l^2)$, and $\exp(-B^2/S')$ are independent of the particular ensemble used. These results apply to percolation clusters in high dimensions ($D > 6$). An analog of this form of the distribution was suggested for $P(R \mid N)$ of self-avoiding walks, where R is the end-to-end distance and N is the number of steps.[10] It was also suggested recently[22] for the chemical distance distribution on two-dimensional percolation clusters. An interesting question arises about the form of $P(S \mid R)$ and $P(S \mid l)$ for percolation clusters in low dimensions. The above results suggest that the general form of $P(S \mid l)$ in D dimensions is

$$P(S \mid l) \sim (S/l^{d_l})^u \exp[-(S/l^{d_l})^v] \tag{39}$$

where $d = 2$ and $v = 1$ for $d > 6$. As far as the moments are concerned, we see that a single gap exponent characterizes the set of moments, a result that follows immediately from the scaling law obeyed by these distributions.

Table I. Moments and Asymptotic Forms for the Distributions Derived for Models 1–4

Model number	$\langle B^n \rangle_i$	$\langle S^n \rangle_i$	$P_i(B \mid l)$	$P_i(S \mid l)$	$P_i(S \mid B)$
1	l^{n-1}	l^{2n-1}	$l^{-2} \exp(-B/l)$	$l^{-3} \exp(-S/l^2)$	—
2	l^n	l^{2n}	$l^{-1} \exp(-B/l)$	$Sl^{-4} \exp(-S/l^2)$	—
3	l^{n-1}	l^{2n-1}	$Bl^{-3} \exp(-B/l)$	—	$BS^{-3/2} \exp(-B^2/S)$
4	l^n	l^{2n}	$l^{-1} \exp(-B/l)$	$l^{-2} \exp(-S/l^2)$	$S^{-1/2} \exp(-B^2/S)$

APPENDIX. METHODS FOR CALCULATING $\langle B^n(l)\rangle$ AND $\langle S^n(l)\rangle$ FOR MODELS 1 AND 2

The Cayley tree of this paper is an example of a branching process in which a node in a given generation becomes 0, 1, or 2 nodes in the next generation with respective probabilities 1/4, 1/2, 1/4. Hence, the corresponding probability generating function is

$$f(x) = \tfrac{1}{4} + \tfrac{1}{2}x + \tfrac{1}{4}x^2 = \tfrac{1}{4}(1+x)^2 \tag{A1}$$

If $U_n(x)$ is the generating function for the number of nodes at distance l from the vertex of the tree, then this quantity can be obtained from the recursion

$$U_1(x) = f(x); \qquad U_{l+1}(x) = \tfrac{1}{4}[1 + U_l(x)]^2 \tag{A2}$$

Since $U_l(x)$ can be used to generate moments by differentiation, the recursion of Eq. (A2) allows us to calculate moments recursively for the successive generations.

To find corresponding moments of clusters that grow to infinity, we notice that the probability that there are no live nodes at distance l is

$$d_l = U_l(0) \tag{A3}$$

As $l \to \infty$, a well-known result in the theory of branching processes[18] suffices to show that $d_l \to d$, where d is the smallest root of $f(d) = d$. In the present case, it is easy to show that $d = 1$ because of the assumption of criticality. If one is above criticality so that $d < 1$, it follows that the generating function conditional on survival is

$$V_l(x) = \frac{U_l(x) - U_l(dx)}{1 - U_l(d)} \tag{A4}$$

In the limit $d = 1$, this goes over into

$$V_l(x) \to \frac{\partial}{\partial d} U_l(dx)\bigg|_{d=1} \tag{A5}$$

This is not normalized, but it is easy to convert to normalized form. When the mean number of bonds generated at a live node is $\langle b\rangle$, then since every node generates an average of $\langle b\rangle$ bonds, it follows that the average

number of bonds at l is $\langle b \rangle^l$. It therefore follows that the normalized function corresponding to Eq. (A5) is

$$\bar{V}_l(x) = \frac{1}{\langle b \rangle^l} x \frac{d}{dx} U_l(x) \Big|_{x=1} \qquad \text{((A6)}$$

This relation, together with the recursion step in Eq. (A2), was used to generate moments for model 2.

REFERENCES

1. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, San Francisco, 1982).
2. F. Family and D. P. Landau, eds., *Kinetics of Aggregation and Gelation* (North-Holland, Amsterdam, 1984).
3. H. E. Stanley and N. Ostrowsky, eds., *On Growth and Form* (Martinus Nijhoff, The Hague, 1985).
4. L. Pietronero and E. Tosatti, eds., *Fractals in Physics* (Elsevier, Amsterdam, 1986).
5. H. J. Herrmann, *Phys. Rep.* **136**:154 (1986).
6. A. Bunde, in *Advances in Solid State Physics* (1986).
7. L. de Arcangelis, S. Redner, and A. Coniglio, *Phys. Rev. B* **31**:4725 (1985).
8. P. Meakin, H. E. Stanley, A. Coniglio, and T. A. Witten, *Phys. Rev. A* **32**:2364 (1985).
9. T. Halsey, P. Meaken, and I. Proccacia, *Phys. Rev. Lett.* **56**:854 (1986).
10. P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, New York, 1979).
11. M. E. Fisher and J. W. Essam, *J. Math. Phys.* **2**:609 (1961).
12. J. W. Essam and K. M. Gwilyn, *J. Phys. C* **4**:L228 (1971).
13. J. W. Essam, K. M. Gwilyn, and J. M. Loveluck, *J. Phys. C* **9**:365 (1976).
14. G. R. Reich and P. L. Leath, *J. Stat. Phys.* **19**:611 (1978).
15. S. Havlin and R. Nossal, *J. Phys. A* **17**:L427 (1984).
16. P. L. Leath, *Phys. Rev. B* **14**:5046 (1976).
17. Z. Alexandrowicz, *Phys. Lett.* **80A**:284 (1980).
18. T. E. Harris, *The Theory of Branching Process* (Prentice-Hall, Englewood Cliffs, New Jersey, 1963).
19. G. H. Weiss and R. J. Rubin, *Adv. Chem. Phys.* **52**:363 (1983).
20. J. A. Shohat and J. D. Tamarkin, *The Problem of Moments* (American Mathematical Society, Providence, Rhode Island, 1960).
21. M. Kac, *Probability and Related Topics in the Physical Sciences* (Interscience, New York, 1959).
22. S. Havlin, B. L. Trus, G. H. Weiss, and D. Ben-Avraham, *J. Phys. A* **18**:L247 (1985).