Exact enumeration study of self-avoiding walks on two-dimensional percolation clusters

1991 J. Phys. A: Math. Gen. 24 1355

(http://iopscience.iop.org/0305-4470/24/6/026)

J. Phys. A: Math. Gen. 24 (1991) 1355–1361. Printed in the UK

COMMENT

Exact enumeration study of self-avoiding walks on two-dimensional percolation clusters

Hisao Nakanishi† and Sang Bub Lee‡

† Department of Physics, Purdue University, West Lafayette, IN 47907, USA
‡ Department of Physics, Kyungpook National University, Taegu 702-701, Korea

Received 13 September 1990, in final form 17 October 1990

Abstract. Quenched disorder averages for the number and size of the self-avoiding walks on two-dimensional percolation clusters very near p_c are calculated by exact enumeration of all walks which start at the seed site from which clusters have been generated by Monte Carlo simulation. Our results are in good agreement with our previous work, which generated the walks by Monte Carlo simulation (rather than by enumeration) on the incipient infinite cluster. The present results show that the recently reported enumeration data by Lam are incorrect.

In this comment, we discuss the issue of whether or not the critical behaviour of the self-avoiding walks (SAW) change when they are confined to a randomly diluted cluster. This problem is often said to have implications for polymer chains trapped in a porous medium.

In particular, we treat here only the case of SAW on the percolation clusters on the two-dimensional square lattice at (or at least very near) $p_c \approx 0.59273$. This study is motivated by recent works on this problem both on analytical and numerical sides. Analytically, the most recent verdict (Meir and Harris 1989) appears to be that the universality class of SAW at p_c is different from that for $p > p_c$. The main methods in that work were a real-space renormalization and an ϵ-expansion for a replicated field-theoretic model. Numerically, a recent Monte Carlo work (Lee and Nakanishi 1988, Lee *et al* 1989) suggested a behaviour associated with the exponent ν to be very similar at p_c as for $p > p_c$. However, this work was claimed to be spurious by another more recent numerical work (Lam 1990a) which used an exact enumeration method.

While the analytic approaches cited seem to give unambiguous predictions, there were in the past equally unambiguous theoretical predictions for this problem (e.g., Chakrabarti and Kertesz 1981, Harris 1983) which were later shown to be incorrect (Kim 1983, Lyklema and Kremer 1984). For the reason that this is one of the particularly controversial issues in statistical physics, we would like to have sound and accurate numerical results to corroborate or contradict analytical (but approximate) solutions whichever the case may be. Therefore, it is all the more troublesome that the two most recent numerical results disagree drastically. It is the main purpose of this comment to show that the enumeration study of Lam (1990a) gives incorrect results and his own method, if carried out correctly, in fact gives excellent agreement with the earlier Monte Carlo work (Lee and Nakanishi 1988, Lee *et al* 1989).

0305-4470/91/061355+07$03.50

To recapitulate, the enumeration work of Lam (1989) gave an estimate of the critical index ν for SAW at p_c to be 0.81 ± 0.03 in marked contrast to the Monte Carlo calculation which suggested ν very close to the known value of $\frac{3}{4}$ for the fully occupied lattice. Lam asserted that this result corroborates the theoretical work of Meir and Harris (1989) on the universality class of SAW on diluted lattice. Lam also gave an estimate of γ but this appears to be the same as for the fully occupied square lattice, the same conclusion as in the Monte Carlo work.

It is unfortunate that the Monte Carlo work was misinterpreted by some to assert that the SAW at p_c is in the same universality class as for $p > p_c$. The work in question simply calculated ν (and γ) and showed that their values are close to the case $p > p_c$ for the square lattice while only ν is close (but probably not γ) for the simple cubic lattice. We note that the numerical estimate by Meir and Harris (1989) for ν (0.76 ± 0.08) in $d = 2$ is also close to $\frac{3}{4}$. It is of course intriguing, and may imply very unusual physics, if the exponents are very close (possibly identical) even though the universality classes are different; however, no claims were made on the universality classes in the Monte Carlo work.

It was also remarked by Lam (1990a) that a behaviour that appeared to reflect some sort of sharp crossover was observed in the Monte Carlo work (Lee and Nakanishi 1988, Lee *et al* 1989), while no such behaviour exists in his data. Our present results indicate, however, that many of the sharp features have been due to the approximation method used in computing an integral (trapezoidal rule) and another method (as in Lam 1990a) has a smoothing effect (see figure 1(*b*)).

The exact enumeration of the SAW work in this way: first, many different *disorder* configurations are generated by a Monte Carlo method often referred to as the Leath method (Leath 1976). Then on each such configuration, *all* SAW are enumerated starting always from the seed site from which the configuration was generated. (Note that a particular shape of a cluster represents many different disorder configurations because a configuration is defined relative to a particular reference point (in the case of the Leath method, the seed site).) Then, their number C_N and the mean-square radius of gyration $\langle S_N{}^2 \rangle$, and the mean-square end-to-end distance $\langle R_N{}^2 \rangle$ are obtained. Finally, these quantities are averaged over the disorder configurations.

In other words, for the number of SAW,

$$\overline{C_N} = \frac{\sum_{\mathcal{C}} P(\mathcal{C}) C_N(\mathcal{C})}{\sum_{\mathcal{C}} P(\mathcal{C})} \sim N^{\gamma-1} \mu^N \tag{1}$$

where the overbar indicates averaging over disorder, $\mathcal{C}$ denotes a disorder configuration, and $P(\mathcal{C})$ is the probability for $\mathcal{C}$ to occur in the percolation problem. The sum over $\mathcal{C}$ is over all configurations that support one or more N-step SAW starting from the reference point. For the radius of gyration of the SAW, we have

$$\overline{\langle S_N{}^2 \rangle} = \frac{\sum_{\mathcal{C}} P(\mathcal{C}) \langle S_N{}^2 \rangle_{\mathcal{C}}}{\sum_{\mathcal{C}} P(\mathcal{C})} \sim N^{2\nu} \tag{2}$$

where

$$\langle S_N{}^2 \rangle_{\mathcal{C}} = \frac{\sum_{\gamma_N \in \mathcal{C}} S(\gamma_N)^2}{C_N(\mathcal{C})} \tag{3}$$

where γ_N denotes any N step SAW and $S(\gamma_N)$ is the radius of gyration of γ_N. The necessary averages are completely analogous for the end-to-end distance R_N. Needless to say, our previous Monte Carlo work computed such quenched averages.

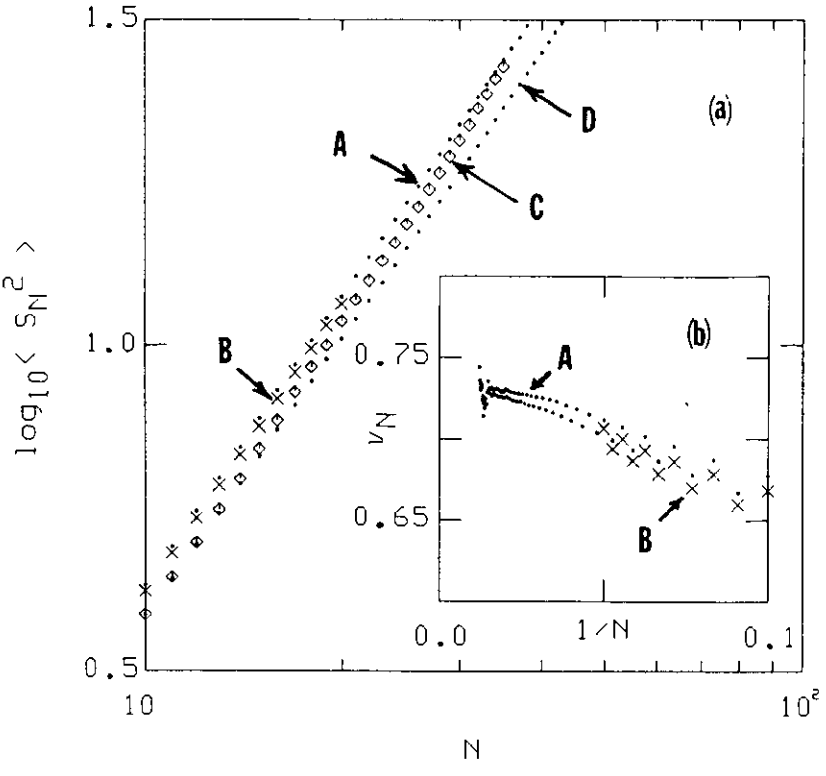

Figure 1. (*a*) Comparison of previous Monte Carlo data (A, · · · · · ·), exact enumeration by us (B, ×), exact enumeration in Lam (1990a) (C, ◇), and full lattice result (D, · · · · · ·). Standard errors of the mean are smaller than the symbols. (*b*) Lam's definition of ν_N used with the data from (*a*).

We believe that the Monte Carlo method described in Lee and Nakanishi (1988) and Lee *et al* (1989) is correct and the enumeration method described by Lam (1990a) should corroborate it if both were carried out correctly. To be sure, there are various subtleties in this problem related to the difference in cluster ensembles, which we discuss later in this comment. However, we would first like to present our own exact enumeration results performed using the average (1)–(3) exactly as described above. The previous Monte Carlo data (A) are compared with these results (B) in figure 1(*a*) for the mean-square radius of gyration $\overline{\langle S_N{}^2 \rangle}$ together with the data (C) given in Lam (1990a) for the corresponding quantity (denoted by $G_N{}^2$ in that work) and the results $\langle S_N{}^2 \rangle$ for the fully occupied lattice (D). Comparison of the end-to-end distances yields a very similar picture and is thus not produced here. The actual numbers from the present work are summarized in table 1.

We enumerated SAW only up to 20 steps for (B) instead of 35 as in Lam (1990a); however, we generated 5000 configurations up to the 35th shell by the Leath method, instead of 500 as in Lam (1990a). We will comment later on why we only used 20 steps. Except for the number of clusters used in averaging, the cluster ensemble should be essentially the same for our enumeration and Lam's (1990a). This means that the two enumeration results (B) and (C) must agree up to $N = 20$. However, clearly the radii of gyration from our calculation (B) are much closer to the Monte Carlo results (A), and in rather marked disagreement with Lam's data (C).

Table 1. Quenched disorder averages from the present exact enumeration work for the number of SAW $\overline{C_N}$, mean-square radius of gyration $\overline{\langle S_N{}^2\rangle}$, and mean-square end-to-end distance $\overline{\langle R_N{}^2\rangle}$, calculated using equations (1)–(3). The error estimates are the standard error of the mean, i.e., $\sigma_{n-1}/\sqrt{n}$ where σ_{n-1} is the sample standard deviation and $n = 5000$ clusters here.

N	$\overline{C_N}$	$\overline{\langle S_N{}^2\rangle}$	$\overline{\langle R_N{}^2\rangle}$
1	2.51 ± 0.01	0.25	1.0
2	4.59 ± 0.03	0.5227 ± 0.0007	2.704 ± 0.006
3	8.41 ± 0.07	0.8340 ± 0.0013	4.858 ± 0.013
4	14.10 ± 0.14	1.1985 ± 0.0019	7.501 ± 0.021
5	24.18 ± 0.29	1.5999 ± 0.0027	10.450 ± 0.031
6	39.88 ± 0.53	2.0460 ± 0.0036	13.788 ± 0.042
7	66.6 ± 1.0	2.5233 ± 0.0046	17.362 ± 0.055
8	108.8 ± 1.8	3.0472 ± 0.0057	21.341 ± 0.070
9	178.6 ± 3.3	3.6010 ± 0.0070	25.544 ± 0.087
10	289.9 ± 5.8	4.1918 ± 0.0084	30.04 ± 0.10
11	471 ± 10	4.8066 ± 0.0099	34.70 ± 0.12
12	758 ± 18	5.454 ± 0.012	39.62 ± 0.14
13	1225 ± 31	6.128 ± 0.013	44.74 ± 0.17
14	1957 ± 54	6.828 ± 0.015	50.03 ± 0.19
15	3150 ± 93	7.556 ± 0.017	55.60 ± 0.21
16	5016 ± 160	8.314 ± 0.019	61.40 ± 0.24
17	8052 ± 277	9.101 ± 0.021	67.42 ± 0.27
18	12800 ± 473	9.918 ± 0.023	73.73 ± 0.30
19	20490 ± 808	10.761 ± 0.025	80.21 ± 0.33
20	32519 ± 1359	11.633 ± 0.028	86.95 ± 0.36

In figure 1(*b*), we show the effective exponents ν_N for the two sets of data (A) and (B) by using Simpson's rule for integrating the logarithm of the mean-square radii of gyration as in Lam (1990a) (see equation(3) and (4) in Lam (1990a)), as opposed to the method in the Monte Carlo work which used trapezoidal rule for the mean-square radii themselves. We see that (A) and (B) are consistent with the full lattice ν in contrast to figure 1 of Lam (1990a). Of course, it is of interest to confirm if the agreement will continue for larger N; for the reason we discuss below, this will have to wait for future work. Likewise, a detailed study of C_N will be deferred.

We have confirmed, by considering subsets of our data, that the discrepancy between (B) and (C) is not due to the larger number of samples we used. Then why is there a difference? In our analysis, Lam (1990a) made two serious errors: first, he calculated the following average as the mean-square radius of gyration

$$\langle S_N{}^2\rangle_o = \frac{\sum_{\mathcal{C}} P(\mathcal{C}) \sum_{\gamma_N \in \mathcal{C}} S(\gamma_N)^2}{\sum_{\mathcal{C}} P(\mathcal{C}) C(N)} \tag{4}$$

and likewise for the end-to-end distance. In other words, instead of performing the double average of (2) as required for the quenched disorder averaging, Lam (1990a) averaged the numerator and the denominator in (3) separately over disorder. This was discovered by inspecting his computer program and was confirmed by the author as well (Lam 1990b).

On the other hand, if all possible configurations $\mathcal{C}$ are averaged over as in (4), the resulting average is independent of p. This was proved by Harris (1983), who, however, argued incorrectly that all disorder averages are trivially independent of p

on this basis. Of course, this is now known to be incorrect (Lyklema and Kremer 1984) and Meir and Harris (1989) uses the correct average as in (2).

Since Lam's method uses only those configurations that grow to a preset number of shells from the seed, his average is not exactly the same as this Harris average. However, we have confirmed that the difference to be extremely small and the results from a correct calculation of this averaging procedure (4) in the present enumeration context give essentially the same radius of gyration and the end-to-end distance as the fully occupied lattice. We show this using our own calculation, deliberately using the incorrect averaging (4), in figure 2. Almost identical results were also obtained by Lam (1990b) when he made the same calculation for 20-step SAW on 5000 configurations of 20 or more shells.

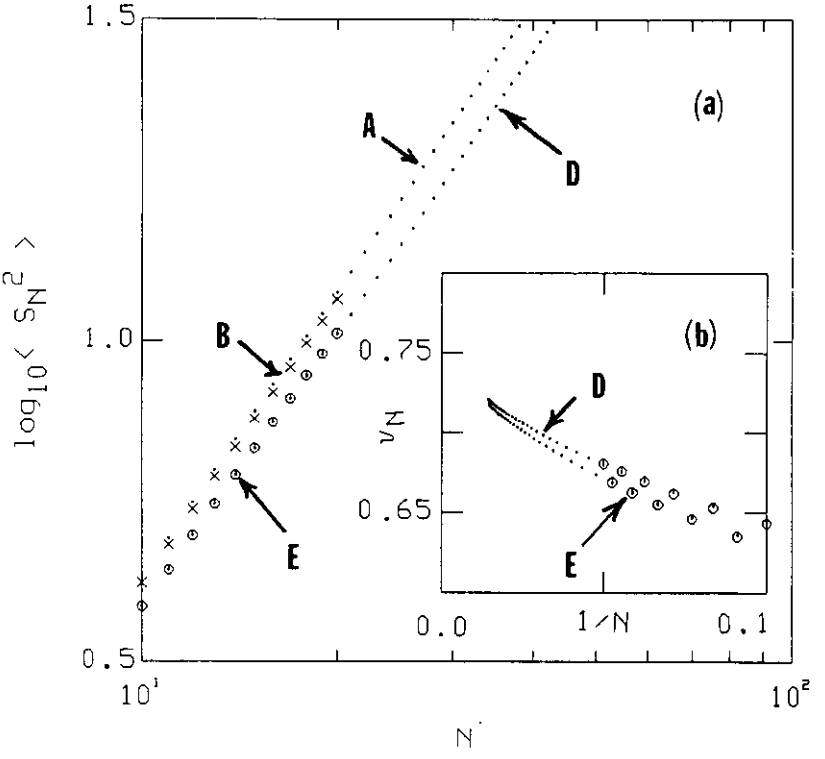

Figure 2. (*a*) Comparison of the exact enumeration result (E, O) which uses the *incorrect* average (equation (4)) and the full lattice result (D, · · · · · ·) for the radius of gyration. The data E are from the 20-step SAW on 5000 configurations of at least 35 shells. For reference, the Monte Carlo data (A, · · · · · ·) and the enumeration result (B, ×) using the *correct* average in (equation (2)) are reproduced from figure 1. (*b*) Lam's definition of ν_N is shown for E and D.

However, this error alone would have given $\nu = \frac{3}{4}$, the same as the fully occupied lattice. Lam (1990a) presented the result $\nu = 0.81 \pm 0.03$ from such a calculation. This means that another serious error must have been made in his published data, while there is no trace of it in the more recent data from 20-step walks referred to above (Lam 1990b). We think that the cause of this second error is in the extremely wide distribution of C_N from configuration to configuration. We have not presented our preliminary results from 35-step SAW primarily because we have been unable to collect

enough statistics for them. At this length, it is easy to generate configurations with C_N varying by several orders of magnitudes, and consequently, the computer time required to enumerate all SAW varies a great deal from cluster to cluster. At $N = 20$ steps, the distribution is sufficiently narrow that the averages from 500 configurations is essentially already the same as those from 5000 configurations.

Perusing table 1, we obtain, e.g., $\overline{C_N}$ for $N = 20$ to be about 3×10^4, while Lam (1990a) reported the corresponding quantity to be $(2215 \pm 167) \times 10^1$ in his table. This is almost one third smaller than our figure, and points to the likelihood that he missed configurations with very large numbers of SAW on them, which are rare but important. In response to our query on whether Lam encountered aborted computer runs due to exceeding CPU time limits, we were told that indeed this occurred a number of times during the collection of the published data. Lam restarted the run with a different random number seed each time this occurred. We believe that this is a strong indication that the cluster ensemble used in his data were biased by systematically excluding those clusters that were much fuller and thus time consuming to enumerate SAW on.

In our view, the two errors we have discussed above are such that the data given in Lam (1990a) are of no scientific significance. This is not to say, however, his enumeration method is useless or that it does not present interesting scientific questions. The utility of the method is clear in that already with the data we present here the two different numerical methods, enumeration and Monte Carlo, are brought to yield good agreement with each other. As we discuss below this has a consequence on the use of different cluster ensembles. In addition, the extraordinary width of the distribution for C_N (already discussed by, e.g., Derrida 1984) is shown to have a disastrous effect on simulations. Of course, a systematic study of this distribution is desirable, and such a work is under way, as well as the efforts to extend the present method to longer SAW and to clusters in other dimensions.

We finally discuss some of the subtle features in the problem which relate to the existence of the multitude of seemingly different cluster ensembles. First, the previous Monte Carlo work sought to calculate the incipient infinite cluster averages (IC). The ensemble implied in (2) is, however, the one (AC) in which all configurations supporting one or more N-step SAW are included in the calculation of the radii of gyration for N-step SAW. It was argued (Lee and Nakanishi 1988, Lee *et al* 1989, Kim 1990) that, since asymptotically long SAW exist only on the incipient infinite cluster (at p_c), the two ensembles are likely to yield identical critical behaviour. However, since the AC ensemble diminishes as N grows, there may be some systematic trend that modifies the critical behaviour. (For the exponent γ a third ensemble is often discussed, for which γ trivially stays the same for all p.)

As we discussed, Lam (1990a) and also this comment use yet another ensemble (35-shell), intermediate between IC and AC, consisting of all configurations grown by the Leath method up to 35 shells or greater from a seed (or its generalization (n-shell)).

Although their differences are potentially important, we believe that IC and (35-shell) are already very similar based on our data shown in figure 1. In addition, IC and AC also show little difference according to a recent Monte Carlo work of Woo and Lee (1990). Since (n-shell) is intermediate between IC and AC, we expect the actual values of the mean-square radii to be intermediate, and this also seems to be verified. The corroboration of the results from IC and AC appear to imply that among the clusters included in AC only those that have scaling properties of the IC contribute significantly toward the averages.

Of course the same can be said about n-shell, but operationally, we can consider this in terms of the attrition rate of the clusters: in Leath method, a percolation cluster is grown from a seed, and it stops growing if all of its perimeter sites are determined to be vacant. This attrition rate of clusters decreases rapidly on the square lattice, and by the time the 35th shell is reached, the rate is already very low. In other words, most of the clusters that grow to the 35th shell will grow much further. The corresponding rate in higher dimensions, however, may decrease more slowly, possibly implying different behaviour for different ensembles.

Acknowledgments

We thank Dr P M Lam for private correspondence and exchanges of computer programs and data. We are also grateful to J W Halley, D Jacobs and J Moon for discussions. This work was supported in part by a grant from the US Office of Naval Research.

References

Chakrabarti B K and Kertesz J 1981 *Z. Phys.* B **44** 211
Derrida B 1984 *Phys. Rep.* **103** 29–39
Harris A B 1983 *Z. Phys.* B **49** 347–9
Kim Y 1983 *J. Phys. C: Solid State Phys.* **16** 1345
Kim Y 1990 *Phys. Rev.* A **41** 4554–7
Lam P M 1990a *J. Phys. A: Math. Gen.* **23** L831–6
Lam P M 1990b private communication
Leath P L 1976 *Phys. Rev. Lett.* **36** 921
Lee S B and Nakanishi H 1988 *Phys. Rev. Lett.* **61** 2022–5
Lee S B, Nakanishi H and Kim Y 1989 *Phys. Rev.* B **39** 9561–72
Lyklema J W and Kremer K 1984 *Z. Phys.* B **55** 41–4
Meir Y and Harris A B 1989 *Phys. Rev. Lett.* **63** 2819–22
Woo K Y and Lee S B 1990 to be published