IOPscience iopscience.iop.org

Home Search Collections Journals About Contact us My IOPscience

Are the tails of percolation thresholds Gaussians?

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2004 J. Phys. A: Math. Gen. 37 3743

(http://iopscience.iop.org/0305-4470/37/12/001)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 171.66.16.90
The article was downloaded on 02/06/2010 at 17:51

Please note that terms and conditions apply.

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **37** (2004) 3743–3748 PII: S0305-4470(04)68154-6

Are the tails of percolation thresholds Gaussians?

P M C de Oliveira[1,2], **R A Nóbrega**[2] **and D Stauffer**[1,3]

[1] Laboratoire de Physique et Mécanique des Milieux Hétérogènes, École Supérieure de Physique et de Chimie Industrielles, 10, rue Vauquelin, 75231 Paris Cedex 05, France
[2] Instituto de Física, Universidade Federal Fluminense, av Litorânea s/n, Boa Viagem, Niterói, 24210-340, Brazil
[3] Institute for Theoretical Physics, Cologne University, D-50923 Köln, Euroland

E-mail: pmco@if.uff.br, rafaella@if.uff.br and stauffer@thp.uni-koeln.de

Received 27 August 2003, in final form 4 February 2004
Published 10 March 2004
Online at stacks.iop.org/JPhysA/37/3743 (DOI: 10.1088/0305-4470/37/12/001)

Abstract
The probability distribution of percolation thresholds in finite lattices was first believed to follow a normal Gaussian behaviour. With increasing computer power and more efficient simulational techniques, this belief turned to a stretched exponential behaviour, instead. Here, based on a further improvement of the Monte Carlo data, we show evidence on square lattices that this question is numerically not yet answered at all.

PACS numbers: 02.70.−c, 05.10.Ln, 64.60.Ak, 05.70.Jk

1. Introduction

In [1], the percolation on an N-site square lattice is treated with high numerical accuracy. Indeed, the best known estimate for the critical threshold, $p_c = 0.592\,746\,21(13)$, comes from this work. In order to study this kind of problem, the authors follow a very fruitful Monte Carlo approach which allows one to obtain continuous functions of p, the concentration of occupied sites, namely the canonical-like average

$$R(p) = \sum_n C_N^n p^n (1-p)^{N-n} R_n \tag{1}$$

of some quantity R. Here, R_n is a uniform average over all configurations with just n occupied sites, i.e. a microcanonical-like average. $C_N^n = \binom{N}{n}$ are the binomial factors. By filling the initially empty lattice, site by site at random, and repeating this process many times, one is able to get the discrete set of microcanonical averages R_n accumulated into an n-histogram, over the entire range, $n = 0, 1, 2, \ldots, N$. From this set of numbers, the determination of the continuous p-function $R(p)$ is straightforward.

In particular, the authors of [1] fix attention on the horizontal wrapping probability $R_L(p)$ around an $L \times L$ torus, i.e. a square lattice with periodic boundary conditions on both directions.

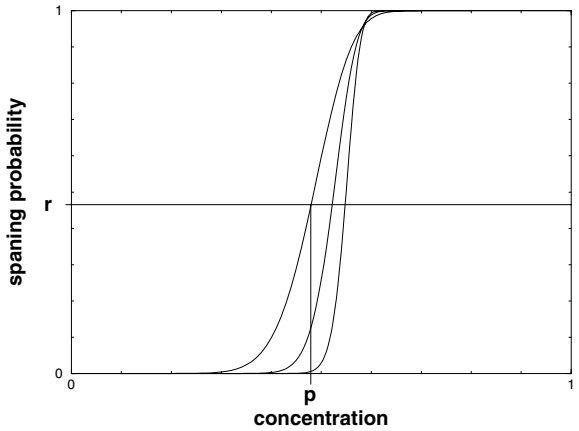

Figure 1. Spanning probability function for a fixed lattice size ($L = 18$, left curve). For larger and larger sizes ($L = 22$ and 26, from left to right), this function approaches a step. By fixing some value r at the vertical axis, one can find a sequence of values $p_L(r)$ at the horizontal axis approaching the critical threshold p_c, for increasing lattice sizes.

In the thermodynamic limit, this function approaches a step: $R_\infty(p) = 0$ below the critical threshold p_c, and $R_\infty(p) = 1$ above p_c. For finite sizes, $R_L(p)$ presents a sigmoid aspect similar to figure 1. A good approach to p_c is to choose some fixed value r, and solve the equation $R_L(p) = r$, getting the root p shown at the horizontal axis. Here, one can appreciate the advantage of knowing $R_L(p)$ as a *continuous function* of p. Keeping the same value r and repeating this task for a series of increasing lattice sizes (dotted lines), one gets a series of roots $p_{L_1}(r)$, $p_{L_2}(r)$, $p_{L_3}(r)$, $\cdots$ which converges to the desired threshold p_c in the thermodynamic limit.

The above reasoning is valid no matter which fixed value for r one chooses. However, for the very particular choice $r^* = 0.521\,058\,290$, a universal probability exactly known through conformal invariance arguments [2], the convergence becomes fast, i.e. the root $p(L)$ differs from p_c as $L^{-2-1/\nu} = L^{-2.75}$, where $\nu = 4/3$ is the correlation length critical exponent. The above quoted accurate value for p_c was obtained in this way. For details, see [1] and references therein. For large L and p close to the percolation threshold p_c one expects [3] the scaling form

$$R_L(p) = f[(p - p_c)L^{1/\nu}].$$

Reference [4] proposes the mathematical form

$$p_L(r) = p_c + \frac{1}{L^{1/\nu}}\left[A_0(r) + \frac{A_1(r)}{L} + \frac{A_2(r)}{L^2} + \cdots\right] \tag{2}$$

for estimators p_L obtained from quantities such as $R(p)$. The option for the wrapping probabilities around the torus and the convenient choice of Pinson's number $r = r^*$ lead to vanishing values for the first two terms $A_0(r^*) = A_1(r^*) = 0$, a lucky coincidence which accelerates the convergence very much.

The quantity $R_L(p)$ is obtained, as quoted before, by filling up the initially empty lattice site by site, at random. Clusters of neighbouring occupied sites grow. As soon as the horizontal wrapping along the torus is set, one books the corresponding value of n, the number of occupied sites so far, and stops the process. For that particular sample, the wrapping probability R_L is a step function, i.e. $R_L = 0$ below n and $R_L = 1$ above. The same routine is repeated many

times, in order to have a probability distribution for n. The various step functions are then superimposed to get the microcanonical averages R_n in equation (1), stored in an n-histogram. Finally, the continuous canonical average $R_L(p)$ can be calculated for any value of p.

Each process of filling up the lattice (one sample) yields a single value n for the statistics, i.e. just one more entry on the n-histogram. In [4], we decided to improve this point, by changing the definition from wrapping to *spanning* probability (figure 1). We fix two parallel horizontal lines on the $L \times L$ torus, separated by a distance of $L/2$, for instance lines $i = 1$ and $i = 1+L/2$. The measured quantity is now the probability of having these two lines connected by the same cluster of neighbouring occupied sites, instead of the wrapping probability along the whole torus. The advantage is that we can measure the same thing for lines $i = 2$ and $i = 2 + L/2$, for lines $i = 3$ and $i = 3 + L/2$, and so on. Moreover, vertical parallel lines can also be included in this counting. At the end, from a single sample we store L new entries into our n-histogram, instead of just one more entry. Note that this advantage even increases for larger and larger lattices.

Within the same computational effort, our approach allows the test of larger lattices. Because of that, we were able to confirm the validity of equation (2) with high precision, by sampling 27 different lattice sizes from $L = 18$ up to $L = 1594$, an eight-thousand factor in the number of sites. On the other hand, our definition does not allow the chance of both $A_0(r)$ and $A_1(r)$ vanishing at once. We can have only $A_0(\hat{r}) = 0$, for a particular universal probability $\hat{r} = 0.984\,786(11)$ numerically determined within the same work [4]. Independently, Cardy [5] tried to determine it by conformal invariance arguments; however, in looking for configurations which link two parallel lines, he was forced to disregard configurations which also wrap along the other direction. As a result of using larger lattices but a slower convergence rate of $L^{-1-1/\nu} = L^{-1.75}$, we get the same value $p_c = 0.592\,746\,21(33)$ [4] as in [1], but with a three times larger error bar. We used about two years of total CPU time on several independent Athlon processors; error bars were estimated by dividing the whole data set into, say, ten independent subsets.

2. The tails

The non-Gaussian behaviour of the finite-lattice-threshold distribution near the infinite-lattice critical point is already established[4] [6]. Here, we profit from the same simulational data in order to investigate the distribution tails, *far* from the critical point, i.e. for large $|p - p_c|L^{1/\nu}$. Which is the mathematical form of the tail observed in figure 1, below the root p? One possible answer is a simple Gaussian form [3, 7]

$$R_L(p) \sim \exp[-K(p - p_c)^2]. \tag{3}$$

Another alternative is a stretched exponential [8–11]

$$R_L(p) = \exp[-C(p_c - p)^\nu] \tag{4}$$

where the strict equality (for large $(p_c - p)L^{1/\nu}$) is a consequence of the periodic boundary condition [1] which holds for our data. Profiting from this strict equality, one can test equation (4) by constructing a plot of $\ln[-\ln(R)]$ against $\ln(p_c - p)$. This was done in [1], and we repeat the same for our data, in figure 2. Note that our range for $\ln(p_c - p)$ (up to -3.4 for $L \sim 10^3$) is larger than in [1] (up to -4.6 for the same size). This means that we are testing more deeply the distribution tails, thanks to our trick of sampling L new entries for each run. Even so, the conclusion in favour of either equation (3) or (4) is by no means obvious. Note a further difficulty concerning equation (3), because the leading multiplicative constant in front of the exponential is not necessarily 1.

[4] This point was not yet realized in [5], but was corrected in the 1994 reprint.

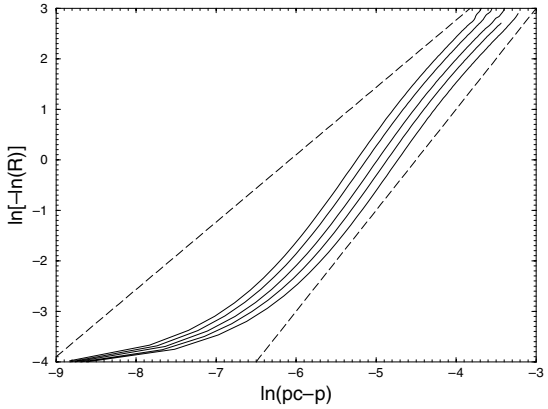

Figure 2. Test of equations (3) or (4), for tails on the left of figure 1. The five continuous curves correspond to $L = 1594$, 1354, 1126, 958 and 802, from left to right. In each case we sampled 4 million statistically independent lattice-filling processes, which correspond to 6, 5, 5, 4 and 3×10^9 entries in each n-histogram, respectively. The statistics is improved by a factor above 1000, compared to [1] for equivalent lattice sizes. The dashed lines show the alternative slopes 2 (right) or $4/3$ (left). In the authors' opinion, no definitive conclusion is possible. Smaller L, with 10^9 lattices filled, are not shown.

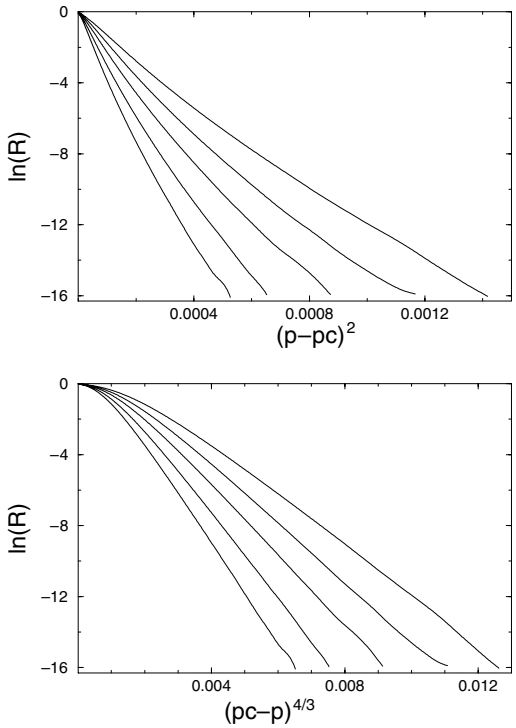

Figure 3. Alternative test of equations (3) or (4), for the same lattice sizes $L = 1594$, 1354, 1126, 958 and 802, from left to right.

Another, perhaps better way to address the same question is by plotting $\ln(R)$ twice, against $(p - p_c)^2$ and $(p_c - p)^{4/3}$. Figure 3 shows the result for our data. Note that our range for $\ln(R)$ (down to -16) doubles the one presented in [1]. The would-be Gaussian case

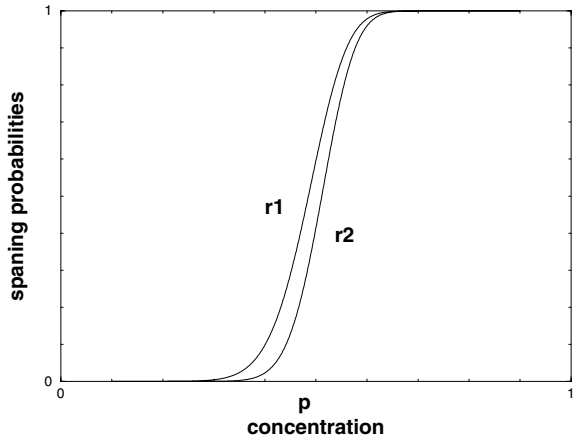

Figure 4. Spanning probability function $\langle r_2 \rangle$ (alternative to $\langle r_1 \rangle$ repeated from figure 1), where connection between two perpendicular pairs of lines is required (instead of just one pair). Here, $L = 18$ for both plots.

(up) presents clear positive curvatures, whereas the would-be stretched exponential (down) presents negative curvatures, although not so pronounced. The exponent $4/3$ seems to fit better, but one cannot extract a clear conclusion from these data.

Still more undefined is the situation of the other tails on the right of figure 1, above p_c. In this case (not shown), our accuracy limit for $\ln(1 - R)$ (down to -16) is reached much closer to p_c than the case shown in figures 2 and 3, below p_c.

Concluding, we present new Monte Carlo data concerning the probability distribution of percolation thresholds on a finite square lattice. We address the question of the mathematical form of the distribution tails, equation (3) against (4). Even considering that our statistics is more than 1000 times larger than previous works, no definitive conclusion can be extracted from our data, as far as the asymptotic tail exponent is concerned.

Acknowledgments

This work is partially supported by Brazilian agencies FAPERJ and CNPq.

Appendix

The spanning probability shown in figure 1 corresponds to two parallel lines separated by a distance of $L/2$, along the $L \times L$ torus. For each lattice filling-up process, we jump from 0 to 1 as soon as these lines become connected. Let us call this function r_1 (figure 1 shows in fact its average $\langle r_1 \rangle$ over many lattice filling-up processes). Consider now another pair of lines, perpendicular to the first pair, also separated by a distance of $L/2$. Another spanning probability r_2 can be defined: we jump from 0 to 1 a little later than for r_1, namely as soon as the second pair of lines also becomes connected. The resulting plot is shown in figure 4, for $L = 18$.

The same procedure described in the caption to figure 1 can be applied to r_2 in order to get the threshold concentration p_c. In this case, the universal value $\hat{r}_2$ is a little bit smaller than $\hat{r}_1 = 0.984\,786(11)$ [4], valid for r_1.

A third alternative definition is to jump from 0 to 1 as soon as the first pair of lines becomes connected, jumping back from 1 to 0 as soon as the second pair of lines also becomes

connected. Let us call the corresponding function r_3. Instead of the sigmoid aspect of $\langle r_1\rangle$ and $\langle r_2\rangle$, the average $\langle r_3\rangle$ would present a peak for some concentration in between 0 and 1, going to 0 in both these limits. The larger the lattice size, the narrower this peak, and its position approaches p_c. For $L \to \infty$, $\langle r_3\rangle$ vanishes for all concentrations except p_c, where its value is a third universal constant $\hat{r}_3$. Thus, one can obtain p_c with no need to know any universal value $\hat{r}$, just by extrapolating the successive peak positions of $\langle r_3\rangle$, for larger and larger lattices. These functions are obviously related by

$$\langle r_3\rangle = \langle r_1(1 - r_2)\rangle.$$

Unfortunately, we have not stored data for r_2 or r_3 within our big runs for 27 different lattice sizes from $L = 18$ up to $L = 1594$. However, we can adopt the approximation

$$\langle r_1\rangle = \langle r_2\rangle + C$$

supposed to be valid for large enough lattices, where C is a constant, based on the following reasoning. At the thermodynamic limit $L \to \infty$, both $\langle r_1\rangle$ and $\langle r_2\rangle$ are step functions, $\langle r_1\rangle = \langle r_2\rangle = 0$ below p_c, whereas $\langle r_1\rangle = \langle r_2\rangle = 1$ above. At p_c, however, one has $\langle r_1\rangle = \hat{r}_1 = 0.984\,786(11)$ and $\langle r_2\rangle = \hat{r}_2$, both universal values.

Indeed, inspired by the comments of the referees who asked for some further numerical evidence in favour of our method, we determined $\langle r_2\rangle$ for $L = 18, 22, 26, 30, 38, 46, 52$ and 62, as functions of p, by new simulations within the same statistics, namely 10^9 samples for each size, with the further factor due to the different positions of the parallel lines (in this case $L^2/4$, better yet than simply L for r_1). In all cases, the maximum difference $\langle r_1\rangle - \langle r_2\rangle$ is approximately the same, near $C \approx 0.19$.

Now, by combining the last two equations, we can approximate $\langle r_3\rangle$ further by

$$\langle r_3\rangle = (1 + C)\langle r_1\rangle - \langle r_1^2\rangle.$$

Note that, now, we need only data for r_1, the same we have already for all 27 different lattice sizes, obtained from our original big runs. By determining the peak positions p_L of this $\langle r_3\rangle$ (above equation), and applying the mathematical form (2) for them, we can obtain p_c for different choices of the constant C. Indeed, the goodness-of-fit Q (see [4]) is very bad, unless we choose C very close to 0.19. In particular, we get $p_c = 0.592\,746\,18(61)$ and $Q = 0.73$, in complete agreement with [1, 4], for $C = 0.19$ with four terms in equation (2). Similar results were obtained for other values of C near 0.19. We believe the error bar can be much smaller than that, by computing $\langle r_3\rangle$ directly for all lattice sizes, without the approximations made here.

References

[1] Newman M E J and Ziff R M 2001 *Phys. Rev.* E **64** 016706
Newman M E J and Ziff R M 2000 *Phys. Rev. Lett.* **85** 4104
[2] Pinson H T 1994 *J. Stat. Phys.* **75** 1167
[3] Levinshtein M E, Shklovskii B I, Shur M S and Efros A L 1975 *Zh. Eksp. Teor. Fiz.* **69** 386
Levinshtein M E, Shklovskii B I, Shur M S and Efros A L 1976 *Sov. Phys.–JETP* **42** 197 (Engl. Transl.)
[4] de Oliveira P M C, Nóbrega R A and Stauffer D 2003 *Braz. J. Phys.* **33** 616 also in xxx.lanl.gov (*Preprint* cond-mat/0308525)
[5] Cardy J 2002 *J. Phys. A: Math. Gen.* **35** L565
[6] Stauffer D and Aharony A 1992 *Introduction to Percolation Theory* 2nd edn (London: Taylor and Francis)
[7] Wester F 2000 *Int. J. Mod. Phys.* C **11** 843
[8] Berlyand L and Wehr J 1995 *J. Phys. A: Math. Gen.* A **28** 7127
Berlyand L and Wehr J 1997 *Commum. Math. Phys.* **185** 73
[9] Ziff R M 1994 *Phys. Rev. Lett.* **72** 1942
[10] Haas U 1995 *Physica* A **215** 247
[11] Hovi J-P and Aharony A 1996 *Phys. Rev.* E **53** 235