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

J. Phys. A: Math. Gen. **35** (2002) L605–L612 PII: S0305-4470(02)52694-9

LETTER TO THE EDITOR

Canonical Monte Carlo determination of the connective constant of self-avoiding walks

A Rechnitzer[1] **and E J Janse van Rensburg**[2]

[1] Department of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria 3010, Australia
[2] Department of Mathematics and Statistics, York University, North York, Ontario M3J-1P3, Canada

E-mail: a.rechnitzer@ms.unimelb.edu.au and rensburg@mathstat.yorku.ca

Received 21 August 2002
Published 8 October 2002
Online at stacks.iop.org/JPhysA/35/L605

Abstract
We define a statistic $a_n(w)$, the size of the *atmosphere* of a self-avoiding walk, w, of length n, with the property that $\langle a_n(w)\rangle \to \mu$ as $n \to \infty$, where μ is the growth constant of lattice self-avoiding walks. Both μ and the *entropic exponent* γ may be estimated to high precision from $\langle a(w)\rangle$ using *canonical* Monte Carlo simulations of self-avoiding walks. Previous Monte Carlo measurements of μ and γ have used grand canonical Monte Carlo simulations. Our simulations indicate that $\mu = 2.638\,16 \pm 0.000\,06$ and $\gamma = 1.345 \pm 0.002$. These results, based on a modest computer run, are comparable to the best estimates for μ and γ from (grand canonical) Monte Carlo simulations, and are at most two digits of the best series estimates of μ for self-avoiding walks available in the literature.

PACS number: 05.50.+q

The lattice self-avoiding walk is a hard combinatorial model of polymers with self-excluded volume [3, 4, 19]. The most important quantity in this model is c_n, which is the number of distinct self-avoiding walks of length n steps, starting from the origin in (say) the square lattice. There is overwhelming analytic and numerical evidence that

$$c_n = A\mu^n n^{\gamma-1}(1 + o(1)) \tag{1}$$

where A is an *amplitude*, $\mu = e^\kappa$ is the *growth constant of self-avoiding walks* while κ is called the *connective constant* [2], and γ is the *entropic exponent*[3]. The exponential growth of c_n with n was established decades ago in [2, 7, 8] but the power law correction to the exponential growth is a conjecture in low dimensions (see, for example, [19] and

[3] This terminology is frequently abused in the literature; μ is often called the connective constant of self-avoiding walks. However, the connective constant is $\kappa = \log\mu$, as it was originally defined by Broadbent and Hammersley [2].

references therein). This asymptotic expression for c_n has been proven for dimensions $d \geqslant 5$ [11] , but remains a conjecture if $d < 4$. In $d = 4$ dimensions the conjecture is modified by a logarithmic correction (see [17]). In this letter we focus on $d = 2$ dimensions, and we show that μ and γ can be estimated using a canonical Monte Carlo algorithm. The motivation for this is to verify the digits of μ determined by exact enumeration studies independently, using a completely different method, which will also provide statistical confidence intervals on our estimates.

It is known that the connective constant can be defined via the limit [10]

$$\kappa = \lim_{n\to\infty} (\log c_n)/n \tag{2}$$

and one would ideally like to strengthen this result to

$$\mu = \lim_{n\to\infty} \frac{c_{n+1}}{c_n} \tag{3}$$

but this remains an open question [8] and no one has proven the existence or non-existence of this limit in the square lattice (however, the limit exists in non-bipartite lattices such as the triangular lattice (see [19]). It is known that

$$\mu^2 = \lim_{n\to\infty} \frac{c_{n+2}}{c_n} \tag{4}$$

a result due to Kesten [15, 16] (and also see [18]). Showing that $c_{n+2} \geqslant c_n$ is not difficult, but the monotonicity of c_n (i.e. $c_{n+1} \geqslant c_n$ for all n) is far more challenging—this result has been proven by O'Brien [23].

The numerical value of μ has been estimated for the square lattice using a variety of techniques, including series analysis and grand canonical Monte Carlo simulations. However, the best estimates for μ have been obtained from series analysis of *lattice polygons* (rather than self-avoiding walks), which are known to have the same connective constant [9]. Using series analysis, μ for polygons (together with their entropic exponent, α) has been determined [13, 14] to an amazing number of digits:

$$\mu = 2.638\,158\,529\,27 \pm 0.000\,000\,000\,01 \text{ (lattice polygons)} \tag{5}$$

$$\alpha = 0.500\,000\,5 \pm 0.000\,001\,0. \tag{6}$$

Determining μ from self-avoiding walk data is not nearly this successful. The best estimate for μ and the entropic exponent γ obtained from self-avoiding walk data are

$$\mu = 2.638\,158\,7 \pm 0.000\,000\,7 \text{ (self-avoiding walks)} \tag{7}$$

$$\gamma = 1.343\,72 \pm 0.000\,10 \tag{8}$$

as determined, for example, in [6].

In two dimensions, Monte Carlo determination of the connective constant of self-avoiding walks does not even approach the precision obtained in the series enumeration above. In higher dimensions, where the finite lattice method [14] is less efficient, the precision of Monte Carlo estimates is more competitive [24]. Nevertheless, Monte Carlo simulations have the added benefit that statistical error bars can be obtained, and provide an independent means of confirming the digits found by series analysis. It is therefore of interest that the numerical determination of connective constants and critical exponents be pursued by Monte Carlo techniques, and that more efficient algorithms be developed for estimating these quantities.

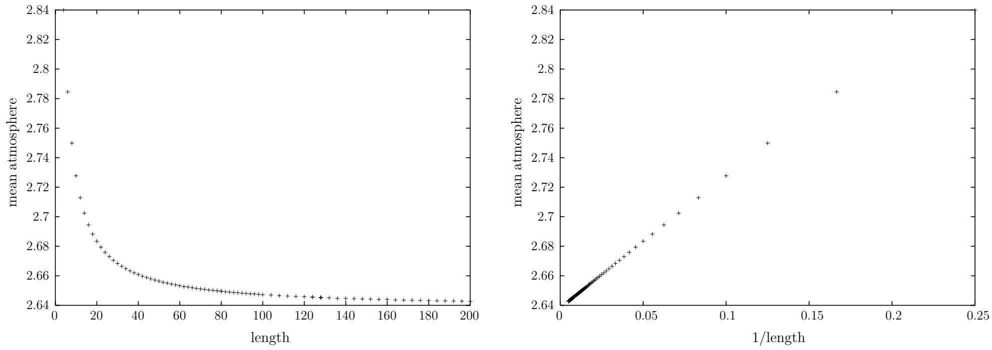

Figure 1. A plot of the mean atmosphere against length (left) and against inverse length (right).

Traditionally, estimates for μ and γ have been made using *grand canonical Monte Carlo algorithms* which sample self-avoiding walks from a distribution over their lengths. The most well known among such algorithms is the Beretti–Sokal algorithm [1]. This algorithm has produced estimates of the connective constant as follows [21]:

$$\mu = 2.638\,164 \pm 0.000\,014 \tag{9}$$

with error bars a combined 95% statistical confidence interval and an estimated systematic error due to uncertainties in the model. More recently, the PERM algorithm has been used to find precise estimates of μ and γ in three dimensions and higher [5, 24].

In this letter we show that estimates for μ and γ can also be obtained from canonical algorithms (such as the pivot algorithm for self-avoiding walks) which sample walks of fixed length from the uniform distribution [20]. The basic idea is to measure the number of ways that an edge may be added to a walk of length n to create a walk of length $n+1$. We call this statistic the *atmosphere* of the walk. Below we show how this can give estimates of μ and γ and we analyse numerical data to determine μ and γ.

The atmosphere of an (oriented) self-avoiding walk is the collection of edges which may be appended onto its last vertex to extend the walk by one step while maintaining self-avoidance. We shall denote the size of the atmosphere of a walk w by $a(w)$. It is not difficult to show that the mean atmosphere of walks of length n is given by

$$\langle a(w)\rangle_n = c_{n+1}/c_n. \tag{10}$$

Examining equation (3) suggests that $\langle a(w)\rangle_n$ can be interpreted as a *local estimate* of μ. If we assume that equation (3) is true, and that c_n has an asymptotic form given by equation (1), then

$$\langle a\rangle_n = \mu\left(1 + \frac{\gamma-1}{n} + o(1/n)\right). \tag{11}$$

Consequently, by obtaining precise estimates of the mean atmosphere at various *fixed* lengths, we are able to estimate both μ and γ. In figure 1 the mean size of the atmosphere is plotted against n and against $1/n$; the near-linearity of the second plot strongly supports the scaling form in equation (11) and furthermore suggests that the $o(1/n)$ corrections are not large.

The exponent γ in equation (1) is thought to have exact value $\gamma = 43/32$ [22] and series analysis strongly supports the scaling form

$$c_n = A\mu^n n^{11/32}(1 + o(1)) + B(-\mu)^n n^{-3/2}(1 + o(1)). \tag{12}$$

The second term has odd–even parity, and if this form is used instead of equation (1) to find an asymptotic form for the mean size of the atmosphere, then

$$\langle a\rangle_n = \mu\left(1 + \frac{11}{32n} + \frac{2B(-1)^{n+1}}{An^{59/32}} - \frac{231}{2048n^2} + o(1/n^2)\right). \tag{13}$$

This suggests that there should be odd–even corrections to the mean size of the atmosphere of walks (which we did observe); we compensate for this by only considering walks of even length in our analysis. Since $59/32 \approx 2$, it is very difficult to numerically distinguish between the $n^{-59/32}$ and n^{-2} terms. In other words, it would be acceptable to assume that $\langle a\rangle_n \approx \mu(1+(\gamma-1)/n+\mathrm{const}/n^2)$ in the statistical analyses of numerical data, to determine μ and γ.

Self-avoiding walks of fixed length n can be efficiently and uniformly sampled along a Markov Chain by the pivot algorithm [20]. This algorithm is a linear time algorithm, with autocorrelations which grows as a small power of n. Doubling n essentially requires only slightly more than doubling in CPU time to maintain datasets of the same size. In order to reduce the effects of corrections to scaling identified above, one should attempt to sample longer walks; but this should also be balanced against the increase in statistical error bars with increasing n. In order to measure γ from the atmosphere data we need to measure the $1/n$ corrections, and consequently we require that the error bars in our data are considerably smaller than this. In light of this we decided that it would be a more efficient use of computer time to obtain a large number of high-quality estimates at *low* and *intermediate* values of n, rather than a small number of low-quality estimates at large n.

We sampled walks of length $2m$ for $2 \leqslant m \leqslant 50$ and length $4m$ for $26 \leqslant m \leqslant 50$. At each value of m, we sampled 8×10^7 walks and computed autocorrelation times along the resulting time series. This allowed us to compute statistical confidence intervals in the estimated means of the atmosphere. We note that since each walk may be rooted at either end point, an atmosphere may be measured at both ends of the walk, and that the average of these two statistics was taken to further reduce error bars. The estimates at each walk of length between 4 and 50 were also checked against exact values using Guttmann and Conway's 51 term series for c_n [6]. The simulations took approximately 10 days of CPU time on a desktop Intel Pentium III. We obtained estimates of μ and γ by making weighted least-squares linear fits of the atmosphere data against four different scaling forms

$$\text{model 1:}\quad \langle a\rangle_n = \mu + \frac{\mu(\gamma-1)}{n} + \frac{\mathrm{const}}{n^2} \tag{14}$$

$$\text{model 2:}\quad \langle a\rangle_n^{-1} = 1/\mu - \frac{(\gamma-1)}{\mu n} + \frac{\mathrm{const}}{n^2} \tag{15}$$

$$\text{model 3:}\quad \log\langle a\rangle_n = \log\mu + \frac{\gamma-1}{n} + \frac{\mathrm{const}}{n^2} \tag{16}$$

$$\text{model 4:}\quad \log\langle a\rangle_n = \log\mu + (\gamma-1)\log\left(\frac{n+1}{n}\right) + \frac{\mathrm{const}}{n^2}. \tag{17}$$

Attempts to fit the data to scaling forms without the $1/n^2$ correction gave poor quality results, as measured by tracking the statistical acceptability of the weighted least-squares error. The quality of the fits can be improved by discarding data points at low values of n where corrections to scaling are stronger. Hence we performed the regressions on a subset of the data points $\{n | n \geqslant n_{\min}\}$ for $n_{\min} = 6, 8, 10, 16, 20, 26, 50$. The quality-of-fit statistic is the χ^2 statistic for the given number of degrees of freedom. We chose the best fits for each model on the basis of this statistic and whether subsequent fits lie within their 95% confidence intervals. We then used the difference between the best fit and the previous fit (in $n_{\min}$) to estimate a

Table 1. The best fits for each model. Entries are of the form: *Best estimate* (*statistical error*) (*systematic error*). Note that series analysis gives $\mu = 2.638\,158\ldots$, $1/\mu = 0.379\,052\ldots$ and $\log\mu = 0.970\,081\,1\ldots$.

Model	$N_{\min}$		$N_{\min}$	$(\gamma-1)\frac{32}{11}$
1	10	$\mu = 2.638\,156(29)(21)$	10	1.0056(24)(23)
2	10	$1/\mu = 0.379\,051\,6(41)(36)$	8	1.0017(18)(28)
3	10	$\log\mu = 0.970\,081(11)(9)$	16	1.0054(41)(39)
4	8	$\log\mu = 0.970\,078(11)(6)$	8	1.0067(23)(19)

Table 2. Fitting the mean atmosphere data. The values of parameters are given as *best estimate* (95% *statistical confidence interval*). The best estimates are given in boldface.

$N_{\min}$	μ	$(\gamma-1)\frac{32}{11}$	const	Quality of fit
	Model 1: $\langle a\rangle \sim \mu + \frac{\mu(\gamma-1)}{n} + \frac{\text{const}}{n^2}$			
6	2.638 088(23)	1.0124(14)	−0.228(4)	Unacceptable
8	2.638 177(26)	1.0033(19)	−0.137(5)	73%
10	**2.638 156(29)**	**1.0056(24)**	−0.166(24)	46%
16	2.638 155(39)	1.0056(41)	−0.162(65)	39%
20	2.638 181(46)	1.0016(57)	−0.07(11)	26%
26	2.638 173(56)	1.0030(79)	−0.11(18)	28%
50	2.638 21(12)	0.996(24)	0.2(9)	55%
$N_{\min}$	$1/\mu$	$(\gamma-1)\frac{32}{11}$	const	Quality of fit
	Model 2: $\langle a\rangle^{-1} \sim 1/\mu - \frac{(\gamma-1)}{\mu n} + \frac{\text{const}}{n^2}$			
6	0.379 0587(32)	1.0009(13)	0.0705(11)	Unacceptable
8	0.379 0480(37)	**1.0017(18)**	0.0599(21)	86%
10	**0.379 0516(41)**	1.0045(23)	0.0648(33)	49%
16	0.379 0525(55)	1.0051(41)	0.0658(90)	39%
20	0.379 0489(65)	1.0014(56)	0.053(15)	26%
26	0.379 0500(80)	1.0028(78)	0.059(26)	28%
50	0.379 046(17)	0.996(24)	0.02(13)	56%
$N_{\min}$	$\log\mu$	$(\gamma-1)\frac{32}{11}$	const	Quality of fit
	Model 3: $\log\langle a\rangle \sim \log\mu + \frac{(\gamma-1)}{n} + \frac{\text{const}}{n^2}$			
6	0.970 0585(86)	1.0110(13)	−0.1374(30)	Unacceptable
8	0.970 090(10)	1.0026(18)	−0.1059(56)	78%
10	**0.970 081(11)**	1.0051(23)	−0.1178(88)	46%
16	0.970 080(14)	**1.0054(41)**	−0.118(24)	39%
20	0.970 090(17)	1.0015(56)	−0.084(40)	25%
26	0.970 087(21)	1.0029(78)	−0.100(68)	28%
50	0.970 099(44)	0.996(24)	−0.01(35)	56%
	Model 4: $\log\langle a\rangle \sim \log\mu + (\gamma-1)\log\left(\frac{n+1}{n}\right) + \frac{\text{const}}{n^2}$			
6	0.970 0471(86)	1.0146(13)	0.0131(28)	Unacceptable
8	0.970 084(10)	1.0048(18)	0.0485(54)	62%
10	**0.970 078(11)**	**1.0067(23)**	0.0402(85)	45%
16	0.970 079(15)	1.0061(41)	0.045(23)	39%
20	0.970 089(17)	1.0020(56)	0.080(39)	26%
26	0.970 086(21)	1.0032(78)	0.066(67)	28%
60	0.970 099(44)	0.996(24)	0.18(35)	55%

systematic error in this procedure. The summary of results is given in table 1. We list the results of the regressions in determining μ and γ in table 2.

In this way we obtain the following estimates of μ:

$$\mu = \begin{cases} 2.638\,156 \pm 0.000\,050 & \text{model 1} \\ 2.638\,163 \pm 0.000\,054 & \text{model 2} \\ 2.638\,158 \pm 0.000\,053 & \text{model 3} \\ 2.638\,150 \pm 0.000\,045 & \text{model 4.} \end{cases} \tag{18}$$

These results are all consistent, and the error bars are the sums of the statistical and systematic errors. Rounding up the error and taking the average of these four estimates give our best estimate

$$\mu = 2.638\,16 \pm 0.000\,06. \tag{19}$$

Taking the logarithm of μ shows that

$$\kappa = 0.970\,082 \pm 0.000\,023 \tag{20}$$

is our best estimate for the connective constant of self-avoiding walks in the square lattice.

In table 1 we also listed best estimates for $32(\gamma - 1)/11$. The exact value of this is expected to be 1, and the estimates are close to that value. Of the best estimates, the value of $n_{\min}$ is the highest for model 3 at $n_{\min} = 16$, and this fit is by our criteria the least reliable. The remaining three regressions all have $n_{\min} \leqslant 10$, and we cannot really distinguish them from one another. Thus, the best estimates are

$$(\gamma - 1)\frac{32}{11} = \begin{cases} 1.0056 \pm 0.0047 & \text{model 1} \\ 1.0017 \pm 0.0046 & \text{model 2} \\ 1.0067 \pm 0.0042 & \text{model 4.} \end{cases} \tag{21}$$

Taking the average of these estimates and rounding up the errors then give our best estimate

$$\gamma = 1 + \tfrac{11}{32}(1.005 \pm 0.005) \tag{22}$$

a result that includes the exact value 43/32 within its error bars:

$$\gamma = 1.345 \pm 0.002. \tag{23}$$

The best series results for μ and γ are in equation (8) [6]. The error bar in the estimate of γ above is roughly a factor of 20 larger as compared to the series result; that is a factor of 400 in computer time. Our result was obtained in about a week's CPU time on a desk-top computer; and this indicates that a comparable estimate to the series result can be obtained if the Monte Carlo determination is performed, say, on a large cluster of computers. The error bar in μ is a factor of ± 100 in favour of the best series results, and much more when compared to the best series results obtained by enumerating lattice polygons, as in equation (6). We confirm the digits 2.6381 as the first five digits for μ independently, but more sophisticated simulations on more powerful processors will be needed to confirm the digits in μ following these, and obtained by series analysis. Moreover, our estimate is only a factor of 3 larger than that of Nidras in equation (9) and we can easily outperform that result on a more powerful computer and with somewhat longer simulations.

In this letter we have estimated μ from canonical Monte Carlo simulations to four decimal places. Our confidence intervals on μ are somewhat conservative; to see this, take the average of the four estimates from the even numbered models in equation (18), which gives

$$\mu \approx 2.638\,157. \tag{24}$$

This gives agreement with series results to five decimal places, with some uncertainty in the sixth decimal place. This deviation is 60 times smaller than the error bar stated on μ in the abstract. The best estimate of γ can similarly be compared to its exact value. Taking averages of the estimates in equation (21) gives

$$\gamma \approx 1.3454 \tag{25}$$

while the exact value has digits 1.343 75. In other words, it deviates from the expected exact value by 2×10^{-3}. This is comparable to the error bar stated in the abstract.

The atmosphere statistic for walks can be generalized to other models, such as lattice trees [12] and related models. In that case its definition is more complicated, but simulations suggest that it will be effective in computing growth constants in any number of dimensions. Further generalizations would include the calculation of extended atmospheres, where instead of just a single edge, two or even more edges are appended to the walk, producing an enumeration process on the MC process that counts walks extended by adding more than one edge. Such statistics should improve the estimates in this letter, and further investigations along these lines are in progress.

Acknowledgments

We would like to thank Aleks Owczarek and Stu Whittington for their many helpful discussions. EJJvR is the recipient of an NSERC (Canada) Operating Research Grant, from which this research was supported.

References

[1] Berretti A and Sokal A D 1985 New Monte Carlo method for the self-avoiding walk *J. Stat. Phys.* **40** 483–531
[2] Broadbent S R and Hammersley J M 1957 Percolation processes: I. Crystals and mazes *Math. Proc. Camb. Phil. Soc.* **53** 629–41
[3] Flory P J 1955 Statistical mechanics of semi-flexible chain molecules *Proc. R. Soc.* A **234** 60–73
[4] Flory P J 1969 *Statistical Mechanics of Chain Molecules* (New York: Wiley Interscience)
[5] Grassberger P 1997 Pruned-enriched Rosenbluth method: simulation of θ-polymers of chain length up to 1 000 000 *Phys. Rev.* E **56** 3682–93
[6] Guttmann A J and Conway A R 2001 Square lattice self-avoiding walks and polygons *Ann. Comb.* **5** 319–45
[7] Hammersley J M 1957 Percolation processes: II. The connective constant *Math. Proc. Camb. Phil. Soc.* **53** 642–5
[8] Hammersley J M 1960 Limiting properties of numbers of self-avoiding walks *Phys. Rev.* **118** 656
[9] Hammersley J M 1961 The number of polygons on a lattice *Math. Proc. Camb. Phil. Soc.* **57** 516–23
[10] Hammersley J M and Morton K W 1954 Poor man's Monte Carlo *J. R. Soc.* B **16** 23–38
[11] Hara T and Slade G 1992 Self-avoiding walk in five or more dimensions: I. The critical behaviour *Commun. Math. Phys.* **147** 101–36
[12] Janse van, Rensburg E J and Rechnitzer A 2002 High precision canonical Monte Carlo determination of the growth constant of square lattice trees *Preprint*
[13] Jensen I and Guttmann A J 1998 Self-avoiding walks, neighbour-avoiding walks and trials on semi-regular lattices *J. Phys. A: Math. Gen.* **31** 8137–45
[14] Jensen I and Guttmann A J 1999 Self-avoiding polygons on the square lattice *J. Phys. A: Math. Gen.* **32** 4867–76
[15] Kesten H 1963 On the number of self-avoiding walks *J. Math. Phys.* **4** 960–9
[16] Kesten H 1964 On the number of self-avoiding walks II *J. Math. Phys.* **5** 1128–37
[17] Larkin A I and Khmel'Nitskiĭ D E 1969 Phase transitions in uniaxial ferroelectrics *Sov. Phys–JETP* **29** 1123–8
[18] Madras N 1988 End patterns of self-avoiding walks *J. Stat. Phys.* **53** 689–701
[19] Madras N and Slade G 1993 The self-avoiding walk *Probability and its Applications* (Boston: Birkhauser)
[20] Madras N and Sokal A D 1988 The pivot algorithm: a highly efficient Monte Carlo method for the self-avoiding walk *J. Stat. Phys.* **47** 573–95

[21] Nidras P 1996 Grand canonical simulations of the interacting self-avoiding walk model *J. Phys. A: Math. Gen.* **29** 7929–42

[22] Nienhuis B 1982 Exact critical points and critical exponents of $O(n)$ models in two dimensions *Phys. Rev. Lett.* **49** 1062–5

[23] O'Brien G L 1990 Monotonicity of the number of self-avoiding walks *J. Stat. Phys.* **59** 969–79

[24] Prellberg T and Owczarek A L 2000 Four-dimensional polymer collapse: pseudo-first-order transition in interacting self-avoiding walks *Phys. Rev.* E **62** 3780–9