J. Phys. A: Math. Gen. **31** (1998) 4851–4869. Printed in the UK PII: S0305-4470(98)91261-6

On anisotropic spiral self-avoiding walks

R Brak†, A L Owczarek† and C E Soteros‡

† Department of Mathematics and Statistics, The University of Melbourne, Parkville, Victoria 3052, Australia

‡ Department of Mathematics and Statistics, University of Saskatchewan, Saskatoon SK S7N 5E6, Canada

Received 3 February 1998

Abstract. We report on a Monte Carlo study of so-called two-choice-spiral self-avoiding walks on the square lattice. These have the property that their geometric size (such as is measured by the radius of gyration) scales anisotropically, with exponent values that seem to defy rational fraction conjectures. This polymer model was previously understood to be in a universality class different to ordinary self-avoiding walks, directed walks (which are also anisotropic), and symmetric spiral walks, in two dimensions. Our Monte Carlo study concurs with those previous exact enumeration studies in that respect. However, we estimate substantially different values for the scaling exponents associated with the geometric size of the walks. We give arguments that explain this difference in terms of a turning point in the local exponent values, and in turn explain this by arguing for the existence of probable logarithmic corrections. We also supply numerical evidence supporting a conjecture concerning the angle of anisotropy in the model.

1. Introduction

The scaling behaviour of the thermodynamic, geometric and topological properties of different types of long chain polymers in solution has been described by a wide variety of models in statistical mechanics. One major group of models are lattice based and involve various types of *self-avoiding walk* (SAW) [1]. A large number of modifications, such as the addition of various interactions (e.g. surface or intra-polymer) or particular restrictions (e.g. directedness), have been made to the basic model to mimic either various physical situations or to allow for easier analysis (such as exact solution). Some of these changes in the basic model modify the scaling behaviour of system properties, and hence change the universality class. For example, it is well known that restricting SAW on the square or cubic lattices by only allowing steps in the positive axial directions, thus producing so-called directed (or rather fully directed) walks (see [2] and references therein), changes the way that the radius of gyration scales with polymer length.

In two dimensions a fairly complete study [3] of the universality classes of restricted step SAW on the square lattice, without interactions, has recently been made. These models are specified by the directions in which subsequent steps are allowed after steps in each of the four lattice directions are made. For example, one might specify that after either positive x- or y-axis steps only positive x- or y-axis steps can be made and that negative steps are disallowed—this gives fully directed walks. Such rules might model oriented polymers in complicated external fields. We note here that these 'two-step' walk models are by their nature oriented. However, the main motivation of the study [3] was to further elucidate the relationship between the symmetries of the lattice models and their universality classes.

0305-4470/98/214851+19$19.50

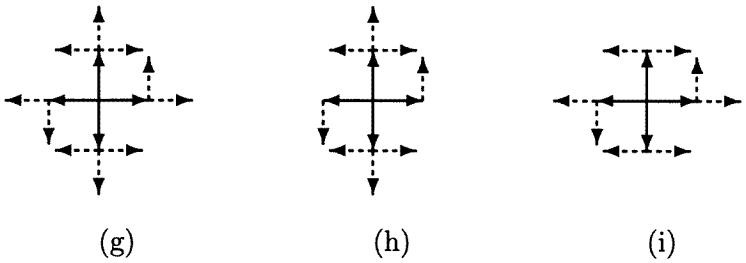

Figure 1. The three restricted two-step SAW rules which fall into the ASSAW class: rules (g), (h) and (i).

This work based its conclusions on previous analyses of various cases already considered in the literature and exact enumeration/series analyses. The universality class was delineated by mainly considering the mean square end-to-end distance $\langle R_e^2 \rangle(N)$ scaling of these walk models with walk length N; this being a measure of the size of the model polymer similar in behaviour to the radius of gyration. The associated exponent is denoted ν and one expects

$$\langle R_e^2 \rangle(N) \sim A_e N^{2\nu} \qquad \text{as } N \to \infty. \tag{1.1}$$

For SAW without restriction in two dimensions it is expected that $\nu = \frac{3}{4}$ [4]. Fully directed walks (DW) [2] are anisotropic with exponents for mutually orthogonal directions given as $\nu_\parallel = 1$ and $\nu_\perp = \frac{1}{2}$. There are various quasi-one-dimensional and zero-dimensional models with exponents $\nu_\parallel = 1$, $\nu_\perp = 0$ and $\nu_\parallel = 0$, $\nu_\perp = 0$ respectively.

Importantly, there were two other universality classes associated with two-step restriction models, such as considered in [3], that have been found. These force some type of spiralling behaviour on the typical configurations. First, there is the exactly solved spiral SAW (SSAW) model [5] with

$$\langle R_e^2 \rangle(N) \sim AN \log(N) \qquad \text{as } N \to \infty. \tag{1.2}$$

Secondly, there are several models [6–8, 3] that come under the general title *anisotropic spiral* walks (ASSAW) that have been argued to scale anisotropically, with $\nu_\perp = \nu_\parallel/2$, although the values of the exponents have been found not to be those of the directed class. The latest series analysis estimates [3] gave $\nu_\parallel = 0.845(5)$. These models have the distinction of being SAW models in two dimensions where no simple rational fractions have been either calculated or conjectured for the exponent values. As they are not rotationally invariant they cannot be described by a conformal field theory, and share this property, and the former one concerning the possibility of irrational exponent values, with the problem of directed percolation [9]. The fact that both these models are non-trivially anisotropic (that is, do not simply act one-dimensionally in one direction) make them interesting cases for investigation.

The first models, that are now under the title ASSAW, were introduced by Manna [6] under the names two-choice-spiral and three-choice-spiral walks. In figure 1 we give a diagrammatic representation of the step rules for two-choice-spiral walks (our rule (i)), three-choice-spiral walks (our rule (g)), and one further rule (rule (h)). Figure 1 is a subset of figure 1 of [3]. The pictures are understood as thus: the heavy curves illustrate the four possible directions for walk steps. The broken curves show the possible continuing steps from each of those initial directions. For example, in the case of rule (h) a step in the positive x-direction can only be followed by one in the positive y-direction, a step in the negative x-direction can only be followed by one in the negative y-direction, while a

step in the positive or negative y-direction can be followed by steps in the three remaining directions respectively (not the positive or negative y-direction respectively). Note that rules which are isomorphic to the three rules in figure 1 can be obtained by symmetry transformations such as rotation by 90°. Also, rule (h) can be obtained from rule (i) by reversing the direction of orientation on the walk configurations (see [3] for a proof) and so rule (h) is isomorphic to Manna's two-choice-spiral walks.

Manna [6] utilized exact enumeration data up to $N = 28$ and $N = 21$ to analyse the partition function (or number of configurations), and the average square end-to-end distance, for the two-choice-spiral and three-choice-spiral walks respectively. This gave estimates of the exponent γ and connective constant μ from the partition function scaling, and gave, from the components of $R_e^2(N)$ in various directions, the associated exponent ν in each of those directions. It was concluded that the dominant size exponent ν in each model was about 0.84 though one model was probably isotropic and the other anisotropic. This was based on extrapolation of local exponent estimates against $1/N$. The inclusion of a particular confluent (multiplicative) logarithmic correction gave smaller exponent estimates.

Whittington [8] gave a proof for the values of the connective constants of the two models which was also subsequently verified numerically to high accuracy [7]. Guttmann and Wallace [7] extended the exact enumerations to $N = 40$ and $N = 30$ for the two-choice-spiral walks and three-choice-spiral walks respectively. They [7] considered the possibility that the partition function scaling had a different form than had previously been assumed and reanalysed the global (dominant) end-to-end distance scaling by carefully considering corrections-to-scaling. They found that no consistent confluent logarithmic corrections could be attributed to the end-to-end distance scaling of the two models. They found the dominant scaling exponent ν to be 0.855(20).

Recently, it was the work of Guttmann *et al* [3] that showed that the two-choice-spiral walk and three-choice-spiral walk models of Manna [6] were in the same universality class since they were both anisotropic, with the size along the minor axis scaling at half the rate as the size along the major axis. In fact, they argued that the angle of anisotropy could be calculated exactly using a random walk argument, building on the idea of Whittington [8]. The angle they found for the two-choice-spiral case was

$$\theta_{2c} = \arctan((3 - \sqrt{5})/2) \approx 0.364\,863\,828\,113\,483\,1817 \approx 20.905^\circ. \qquad (1.3)$$

They [3] also discussed the relationship between the symmetry of these rules and their universality class. The exact enumerations were extended slightly with the maximum lengths enumerated being $N = 44$ for two-choice-spiral walks (rule (i)) and $N = 32$ for three-choice-spiral walks (rule(g)). The revised estimate for what was now denoted $\nu_\parallel$ for the ASSAW class was 0.845(5) as mentioned above.

It was noted in [3] that the differential approximants used in the analysis to obtain $\nu_\parallel$ did not behave well. Hence the question remains as to whether the estimate quoted above for $\nu_\parallel$ encompasses the true value. To add to this uncertainty we now reiterate the point that, in the case of the SSAW, analysis of short series can lead to erroneous results: the local exponent estimates contain a *turning point*, which provides misleading extrapolates. With a relatively short series even the most sophisticated series analysis methods have difficulty predicting the correct results. The reason for this large N turning point in the SSAW case is the competition of a confluent logarithmic correction and the dominant power law. Given these considerations a Monte Carlo study that samples very long walks compared to the enumerations is one way to help settle the question of the universality class and exponent values for ASSAW. This paper describes the results of such a study: namely, we have simulated one model in the ASSAW class, that being rule (h) (which is isomorphic,

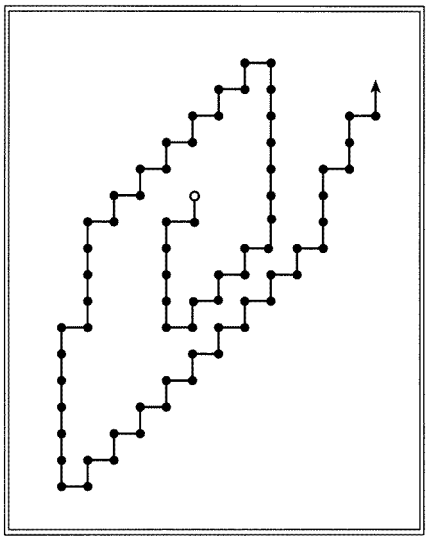

Figure 2. A typical configuration of rule (h).

as explained above, to Manna's two-choice-spiral walks). Figure 2 illustrates a typical configuration of rule (h) walks.

To simulate ASSAW we have developed a novel Monte Carlo algorithm which is an adaptation of the pivot algorithm for ordinary SAW [10, 1]. This can be simply described as an inhomogeneous pivot algorithm that obeys the rule chosen with the addition of local moves. This algorithm is explained in section 2, along with a proof that the algorithm is ergodic. This algorithm allowed us to simulate walks up to length $N = 12\,800$ with good statistics. The technicalities of this procedure are explained, and the raw results obtained are given in section 3. The equivalent total CPU time on a DEC Alpha 500/266 was about 1 year. From the statistics generated and subsequent analysis we argue that indeed ASSAW is a separate two-dimensional universality class but that the exponent estimates previously obtained are not accurate, despite their precision, and that rather

$$\nu_{\parallel} = 0.955(20). \tag{1.4}$$

A combination of strong power law corrections with a probable confluent logarithm term in the scaling form of the $R^2_{e,\parallel}(N)$ may explain the misestimation of the series since we show that their is a turning point in the local exponent estimates near $N \approx 150$. We also verify the angle of anisotropy for this model is indeed as argued in [3]. Our methods of analysis and subsequent discussion are given in section 4.

2. Algorithm

2.1. Inhomogeneous pivot algorithm

The description of the algorithm is given here for rule (h) ASSAWs but the algorithm can actually be applied to the rule (g) or (i) ASSAWs as well. Given an arbitrary size N, our Monte Carlo algorithm is designed to generate a Markov chain, $(X_t)_{t\geqslant 0}$, whose sample space is the set of all N-step rule (h) ASSAWs, $\mathcal{A}_N(h)$, and whose equilibrium distribution

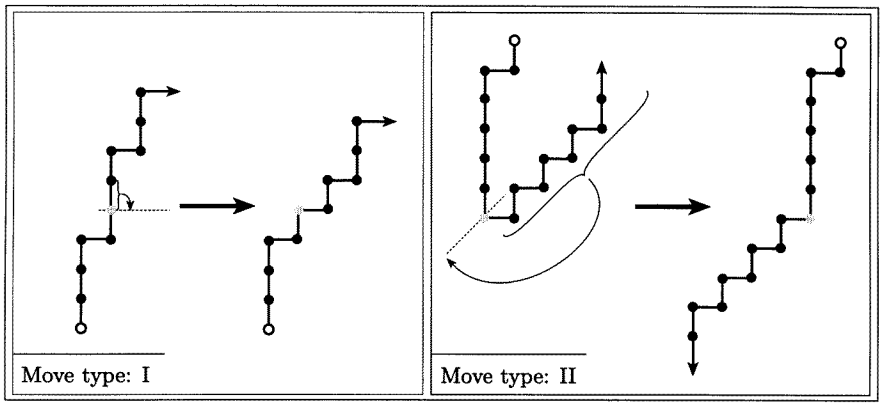

Figure 3. The two-allowed pivot-like Monte Carlo moves.

is given by

$$\pi_\omega = \frac{1}{|\mathcal{A}_N(h)|} \tag{2.1}$$

for any $\omega \in \mathcal{A}_N(h)$. The ergodicity of this algorithm is proved in section 2.3.

Note that we assume each element of $\mathcal{A}_N(h)$ starts at the origin. We label the steps of a walk $\omega \in \mathcal{A}_N(h)$ with the integers $0, 1, \ldots, N-1$ and each of these steps corresponds to one of the elements of the set $\{(0,1), (1,0), (0,-1), (-1,0)\}$ (i.e. north, east, south, west). The vertices of the walk are labelled with the integers $0, 1, 2, \ldots, N$.

The Markov chain, $(X_t)_{t \geqslant 0}$, starts at an arbitrary element $X_0 = \omega_0$ in $\mathcal{A}_N(h)$. Then for arbitrary $t \geqslant 0$ with $X_t = \omega$, $\omega \in \mathcal{A}_N(h)$:

(i) a *step label* i is chosen uniformly at random from the set $\{0, 1, \ldots, N-1\}$,

(ii) a *move type label* M is chosen uniformly at random from the set $\{I, II\}$,

(iii) if $M = I$, then the walk ω' is constructed such that the zeroth to $(i-1)$th and the $(i+1)$th to $(N-1)$th steps of ω' are the same as those of ω, and if the ith step of ω is $(\Delta x_i, \Delta y_i)$ for $(\Delta x_i, \Delta y_i) \in \{(0,1), (1,0), (0,-1), (-1,0)\}$ then the ith step of ω' is $(\Delta y_i, \Delta x_i)$ (i.e. the ith step of ω is reflected through the line $y = x$ to obtain ω'),

(iv) if $M = II$, then the walk ω' is constructed such that the zeroth to $(i-1)$th steps of the walk are the same as those of ω, and if the ith to $(N-1)$th steps of ω are $(\Delta x_i, \Delta y_i), (\Delta x_{i+1}, \Delta y_{i+1}), \ldots, (\Delta x_{N-1}, \Delta y_{N-1})$ then the ith to $(N-1)$th steps of ω' are $(-\Delta x_i, -\Delta y_i), (-\Delta x_{i+1}, -\Delta y_{i+1}), \ldots, (-\Delta x_{N-1}, -\Delta y_{N-1})$ (i.e. the ith to $(N-1)$th steps of ω are each rotated 180° to obtain ω'),

(v) in either case (iii) or (iv) above, if $\omega' \in \mathcal{A}_N(h)$, that is if ω' is self-avoiding and rule (h) is satisfied, then the next walk in the chain is taken to be $X_{t+1} = \omega'$; otherwise $X_{t+1} = \omega$.

In order to determine, in (v) above, whether ω' follows rule (h) it is necessary to consider at most the $(i-1)$, i, and $(i+1)$th steps, $(\Delta x_{i-1}, \Delta y_{i-1}), (\Delta x_i, \Delta y_i), (\Delta x_{i+1}, \Delta y_{i+1})$, of ω (only two steps need to be considered if $i = 0$ or $i = n-1$). Thus one can simplify the calculation by tabulating *a priori* which of these *three-step configurations* allow type I and/or type II moves. Then, given the three-step configuration of ω at i, move M can be rejected immediately if it is not one of the allowed move types for the three-step configuration. Since the moves of types I and II change the location of the $(i+1)$th to Nth

vertices of the walk, all of these vertices must be considered in the self-avoidance check. Figure 3 illustrates the types of allowed Monte Carlo described above.

2.2. Implementation

The algorithm described above was implemented in C and compiled in a 64-bit environment using both Digital's and Sun's C compilers with maximum optimizations. Comparison with unoptimized executables was made to check that the optimizations had not introduced any errors.

To implement the algorithm described above we had to consider both memory management and the generation of random numbers. For values of $N \leqslant 1600$, an array of size $(2N+1) \times (2N+1)$ (one element for each possible lattice site which can be occupied by an N-step ASSAW) was used to store the information about an ASSAW configuration. In particular, the array element corresponding to a particular lattice site can be taken to be 1 if the ASSAW visits that site and 0 otherwise. Having the ASSAW configuration stored in this way allows the self-avoidance check that is needed at step (v) of the algorithm to be done efficiently; however, for $N > 1600$ the memory cost associated with this becomes too great.

For $N > 1600$ we used a hashing algorithm to save on memory requirements, so that the coordinates of the configuration were stored in a hash table. The basic algorithm chosen was a modified UNIX ELF hash routine (see acknowledgments). Even with a hash table size of 20 times the length of the walk the maximum memory allocation for the whole program was about 27 MB.

In our simulations we began by using two different random number generators for several lengths. The first was the routine `ranf()` from the random number generator package `ranlib.c` which utilizes the base code from [11]. This code is the implementation of a mixed linear congruential generator algorithm [12], and is a package which has proven to be comparatively reliable [13]. The second algorithm was written by Janse van Rensburg and Gruner and is a mixed Weyl and lagged Fibbonaci generator of the type that was studied by Margsaglia and Zaman [14]. It has a period of 2^{1407} and has proven to be comparatively reliable in various recent Monte Carlo simulations [15, 16]. As a further test we compared answers using both generators and also using different seeds. These gave the same answers within error bars and so we used the faster generator (the Marsaglia one) for the more time-consuming longer length simulations.

2.3. Proof of ergodicity

In this section we prove that the algorithm described in section 2 is irreducible, aperiodic and reversible and hence ergodic for all three of the classes of ASSAWs. Although the algorithm in section 2 is only described for rule (h), the algorithm for the other rules is just obtained by replacing (h) by either (g) or (i) in the description of section 2.

To show reversibility (i.e. 'detailed balance') for the Markov chain, X_t, generated by the algorithm, we first note that each of the moves, type I or type II, is its own inverse. Next consider two distinct n-step ASSAWs ω_1 and ω_2 in the same class. X_t is reversible if the 'detailed balance condition' is satisfied, that is if the one step transition probabilities satisfy

$$\pi_{\omega_1} P(X_{t+1} = \omega_2 | X_t = \omega_1) = \pi_{\omega_2} P(X_{t+1} = \omega_1 | X_t = \omega_2) \tag{2.2}$$

where π_ω is given by equation (2.1) (with (h) replaced by (g) or (i) depending on which class is of interest). We show that this condition is satified next. Compare the sequence

of steps which define each walk and look for the first step where they differ. Suppose this step is the $(i+1)$th step. If this is the only step where the two sequences differ and if the steps can be made the same by reflecting one step through the line $y = x$ then ω_1 can be obtained from ω_2 by a type I move at vertex i and similarly ω_2 can be obtained from ω_1 by the same move. If instead the sequences differ at each step after the ith vertex and it is possible to make the two sequences the same by, for one of the walks, rotating each step after the ith vertex 180°, then ω_1 can be obtained from ω_2 by a type II move at vertex i and similarly ω_2 can be obtained from ω_1 by the same move. These are the only ways to go from one walk to the other in one step. In the above mentioned cases, the probability to go from one walk to the other in a single move is given by

$$P(X_{t+1} = \omega_2 | X_t = \omega_1) = P(X_{t+1} = \omega_1 | X_t = \omega_2) = \frac{1}{2N} \tag{2.3}$$

and clearly equation (2.2) is satisfied. In all other circumstances it is not possible to get from one walk to the other by a single move and hence

$$P(X_{t+1} = \omega_2 | X_t = \omega_1) = P(X_{t+1} = \omega_1 | X_t = \omega_2) = 0 \tag{2.4}$$

and equation (2.2) is satisfied.

To show that the Markov chain is aperiodic we first consider $P(X_{t+1} = \omega | X_t = \omega)$ for $\omega \in \mathcal{A}_N$. If ω contains a vertex i such that the direction of the step leaving vertex $(i-1)$ is east (or east, or west, or west, or south, or north) and the direction of the step leaving vertex i is north (or respectively east, west, south, south, north) then it is not possible to perform a type II move at vertex i and there is a probability of at least $\frac{1}{2N}$ that a move attempted on walk ω will fail, that is $P(X_{t+1} = \omega | X_t = \omega) \geqslant \frac{1}{2N}$. Since every walk of length at least three must contain at least one of the six possible 'two configurations' listed above at some vertex i in the walk then for $N \geqslant 3$, $P(X_{t+1} = \omega | X_t = \omega) \geqslant \frac{1}{2N}$ for all $\omega \in \mathcal{A}_N$ and the Markov chain is thus aperiodic.

Theorem 2.1. A Markov chain generated by type I and type II moves is

(a) irreducible for the state space of all anisotropic walks in class g,

(b) irreducible for the state space of all anisotropic walks in class h, and

(c) irreducible for the state space of all anisotropic walks in class i

where the classes are as in figure 1 of this paper.

Proof. We focus on the proof of (b) and note that the proof of (a) is exactly the same except that one starts with a walk in class (g). The modifications needed to make the proof work for (c) are put in square brackets where appropriate.

To prove irreducibility for class (h) (class (i)) we show that it is possible to start with any n-step self-avoiding walk in class (h) (class (i)), ω, and by a sequence of moves, $m_1, m_2, \ldots, m_p$, of type I or II obtain a sequence of class (h) (class (i)) walks, $\omega_1, \omega_2, \ldots, \omega_p$ with the last walk, ω_p, equal to the straight walk, s_N [s'_N], having all its steps in the north direction (east direction). For any walk, ω, the ith vertex of the walk is defined to be the vertex at the end of the ith step and its coordinates in Z^2 are given by (x_i, y_i). The origin is the zeroth vertex.

It is thus necessary to consider several cases. □

Case 1. Suppose ω only has north and east steps. If ω has no east (north) steps then $\omega = s_N$ [$= s'_N$] and we are done; otherwise assume there is an east (a north) step at vertex $i \in [0, N-1]$. Since ω contains only north and east steps, all the vertices in the

walk with labels bigger than i are located in the quarter plane $\{(x, y)|x \geqslant x_i + 1, y \geqslant y_i\}$ $[\{(x, y)|x \geqslant x_i, y \geqslant y_i + 1\}]$ and all the vertices in the walk with labels less than i are in the quarter plane $\{(x, y)|x \leqslant x_i, y \leqslant y_i\}$. Thus we can successfully perform a type I move at vertex i by changing the east (north) step at vertex i to a north (an east) step. We thus obtain a new walk in class (h) (class (i)) with one less east (north) step. Continuing this process, one can eliminate all the east (north) steps in ω and eventually obtain s_N $[s'_N]$.

Case 2. Suppose ω has only south and west steps. Perform a type II move at the origin and then we have a walk which satisfies case (1).

Case 3. Suppose ω is not in case (1) or case (2). Define the top (bottom) vertex of an n-step walk as follows. Let S_0 be the set of vertices reached by the walk (including the origin). Construct the subset $S_1 \subset S_0$ such that the y coordinate of every vertex in S_1 has the maximum (minimum) value over all vertices in S_0. The vertex in S_1 which has the maximum (minimum) x coordinate over all vertices in S_1 is called the top (bottom) vertex of the walk. For the walk ω we denote the label of its top vertex by v_t and the label of its bottom vertex by v_b. We note also that if $v_t \neq N$ then it is always possible to perform a successful type II move at vertex v_t. In particular, since $v_t \neq N$, the step leaving vertex v_t is in the west direction and the step leaving vertex $v_t - 1$ (if this is not less than zero) is a north step; one can perform a type II move at such a vertex provided the move leaves the walk self-avoiding. To see that the resulting walk will always be self-avoiding we note that the vertices of the walk other than vertex v_t are confined to the set $R_1 = \{(x, y)|x < x_{v_t}, y = y_{v_t}\} \cup \{(x, y)|y < y_{v_t}\}$. A type II move results in the vertices of the walk after vertex v_t being moved into the set $\{(x, y)|x > x_{v_t}, y = y_{v_t}\} \cup \{(x, y)|y > y_{v_t}\}$ which is disjoint from the set R_1. Thus there can be no self-intersections as a result of a type II move at vertex v_t. Similarly, it is always possible to perform a type II move at vertex v_b if $v_b \neq N$. There are now various possibilities.

(a) $\{v_b, v_t\} = \{0, N\}$. We claim that for this case the walk is in either case (1) (if $v_b = 0$ and $v_t = N$) or case (2) (if $v_b = N$ and $v_t = 0$). For example, suppose $v_b = 0$ and $v_t = N$. Assume that ω has either a south or a west step. Suppose the first step that is not north or east is the step which leaves the ith vertex of the walk. $i > 0$, since otherwise v_b would not be 0. The step leaving vertex $i - 1$ is thus either a north or east step neither of which can be followed by a south step. Hence the step leaving vertex i is a west step and the step leaving vertex $i - 1$ is a north step. In order to get from the end of the west step (i.e. from vertex $i + 1$ with coordinates $(x_i - 1, y_i)$) to vertex $v_t = N$ (located in the region $\{(x, y)|y = y_i, x > x_i\} \cup \{(x, y)|y > y_i\}$) one must use a sequence of west and south steps followed by a sequence of east and north steps and avoid the first $0 - i$ vertices of the walk (located in the region $\{(x, y)|0 \leqslant y \leqslant y_i, 0 \leqslant x \leqslant x_i\}$). The only way to do this is to go south west of vertex 0 but then v_b would not be 0. It is therefore not possible to have either a south or west step in ω.

(b) $\{v_b, v_t\} \neq \{0, N\}$ and $v_t = N$. Since v_b is not in $\{0, N\}$, the step leaving v_b is in the east direction and the step leaving $v_b - 1$ is in the south direction. Perform a type II move at v_b. This results in a new walk which has its bottom vertex $v_b^1 = N$ and its top vertex $v_t^1 < v_b$. If $v_t^1 = 0$ then we are in case (2). Otherwise, the step leaving v_t^1 is in the west direction and the step leaving $v_t^1 - 1$ is in the north direction. Perform a type II move at v_t^1. This results in a new walk which has its top vertex $v_t^2 = N$ and its bottom vertex $v_b^2 < v_t^1 < v_b$. If $v_b^2 = 0$ then we are in case (1). Otherwise, continue the process until we end up with a walk which is in either case (1) or (2).

(c) $\{v_b, v_t\} \neq \{0, N\}$, $v_t > v_b$,$v_t \neq N$. Perform a type II move at v_t. This results in a new walk which has either its top vertex, v'_t, satisfying $v'_t = N$ or else its top vertex has y

coordinate at least one greater than the y-coordinate of v_t. Thus if we continue with this procedure it must eventually end with a walk having top vertex equal to N. Once we have $v_t = N$ then we are in case 3 (b).

(d) $\{v_b, v_t\} \neq \{0, N\}$ and $v_b = N$. Perform a type II move at vertex 0 and obtain a walk in case 3 (b).

(e) $\{v_b, v_t\} \neq \{0, N\}$, $v_b > v_t$,$v_b \neq N$. Perform a type II move at vertex 0 and obtain a walk in case 3 (c).

Since each move is its own inverse, the above argument is sufficient to show that one can get from any walk ω_1 in the appropriate class to any other walk ω_2 in the same class by a sequence of type I and type II moves. In particular, one uses the moves dictated by the argument above to take ω_1 to the appropriate straight walk. The argument above also dictates a sequence of moves which would take ω_2 to the straight walk. Reversing this sequence of moves takes the straight walk to ω_2. Thus irreducibility is proved.

We note that these moves can also be used to yield an irreducible Markov chain for the state space of all directed walks (rule (j) in GPO), however, an additional move is needed to generate the set of all spiral walks (rule (d) in GPO reflected through the line $y = x$). In particular, we need to introduce a type III move such that at a vertex i the step leaving it can be reflected through the line $y = -x$ and all other steps remain unchanged.

3. Monte Carlo simulations and results

The algorithm, realized by the implementation described in the previous section, was used to simulate rule (h) ASSAW for the various values of length N listed in table 1, from $N = 44$ up to $N = 12\,800$. Values were spread on a logarithmic scale for $N \geqslant 100$ using approximate factors of $2^{1/3}$ moving up from $N = 100$ to $N = 400$ and then $2^{1/2}$ up to 6400 with the largest N value chosen being 12 800. Extra values at small N were also chosen. In particular, to test the algorithm and method as described below we simulated $N = 44$—the largest length for which exact enumeration data was known.

For each value of N we first had to estimate the integrated autocorrelation (step) times, t_{auto}, for the quantities of interest. This time, t_{auto}, is essentially the minimum number of Monte Carlo steps after which effectively independent configurations are produced, with respect to measuring those quantities. In our case we calculated the radius of gyration, and other 'size measuring' quantities, only, so this time t_{auto} was more-or-less constant for all the quantities measured—we confirmed this. To find an estimate for t_{auto} we started at small lengths where very large samples could be chosen and made an initial guess of the autocorrelation time (apart from on the first occasion this was based on the previous value of N). We made an autocorrelation analysis of this using the statistical package Statistica to determine a revised estimate of the autocorrelation time. We then made a new simulation to test this hypothesis. We also used the method of blocking [17] to corroborate this. As N was increased we noted the dependence of t_{auto} on N, which was $o(N)$, and made guesses appropriately. Again we used the autocorrelation analysis and blocking analysis; although it was over smaller samples, all of which were larger than 10^5. For some quantities this analysis was carried out on the final data samples to check correlations (as we recorded the whole time series of some of the quantities).

For small N the time between samples (τ_{sample}: being a number of Monte Carlo steps) was effectively chosen as several factors of the auto-correlation time t_{auto}. We confirmed that we did obtain independent samples using the autocorrelation analysis and blocking methods. Using these τ_{sample} the data collected was almost perfectly independent: figure 4 shows a blocking histogram and Statistica output using the appropriate chosen value of

Table 1. The lengths N simulated with the sample times τ_{sample} and total sample size S of independent configurations. For $N \geqslant 1130$ samples were not quite independent and this was accounted for in the calculation of errors.

N	τ_{sample}	S
44	440	16 777 216
50	500	16 777 216
60	600	16 777 216
70	700	16 777 216
80	800	16 777 216
100	1 000	16 777 216
130	1 200	16 777 216
165	1 400	16 777 216
200	1 700	16 777 216
250	2 000	16 777 216
320	2 400	16 777 216
400	3 000	8 388 608
560	4 000	4 194 304
800	5 000	1 048 576
1 130	3 000	1 048 576
1 600	3 500	1 048 576
2 260	4 500	524 288
3 200	6 000	524 288
4 520	8 000	131 072
6 400	10 000	131 072
12 800	20 000	131 072

sample time for that length. At larger N we chose to use a sampling time about the autocorrelation time t_{auto}. We corrected the underestimation of the size of the statistical errors in quantities by utilizing the blocking analysis. This allowed us to find the appropriate factor to multiply our raw statistical errors to take account of the fact that our samples were not quite independent. This is possible because the blocking analysis produces a revised estimate of the true statistical error. We tested this method thoroughly, and kept the values for every sample of some of the major quantities calculated (in addition to keeping running averages). The sample (step) times τ_{sample} are given in table 1. In general t_{auto} was roughly proportional to $N^{0.8-1.0}$ so our inhomogeneous pivot algorithm had a CPU time increase roughly proportional to $N^{1.8-2.0}$ for this problem. This is unfortunately not as good as the standard pivot algorithm for ordinary SAW [1].

For small lengths a virtual lattice was stored in memory so that memory requirements were increasing with N^2, while for larger lengths a hashing routine (as explained in section 2) was used to minimize memory usage which increased as a multiple of N. However, this meant that the algorithm was substantially slower at longer lengths (about a factor of 6). With the sample Monte Carlo times τ_{sample} quoted in table 1 it took 3.3 CPU milliseconds to produce an independent configuration of walk length $N = 100$, while it took about 100 CPU seconds between samples in the simulations of the walks of length 12 800: these times were estimated on a Dec AlphaStation 500/266. The simulations were performed on a number of machines including a Dec AlphaStation 500/266, a Dec AlphaStation 250/4 266 and a Sun Ultra 170 workstation.

For lengths $N < 320$ to we used very large total sample sizes of independent configurations, that is, $S = 2^{22} = 16\,777\,216$. As N was increased smaller sample sizes were used to accommodate the longer running times. For the longest lengths we used sample

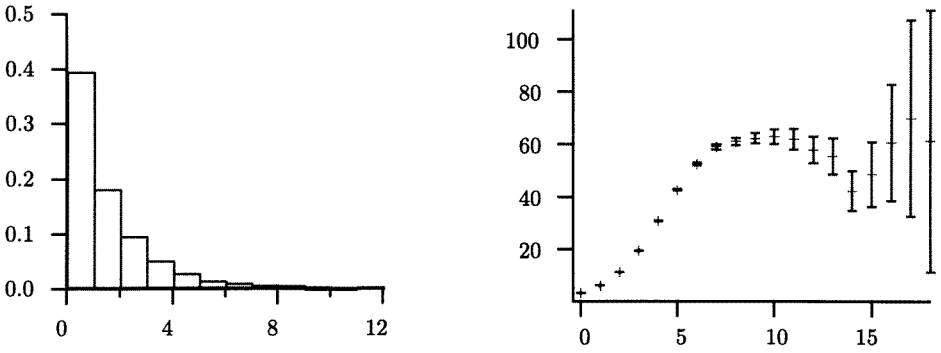

Figure 4. The two plots illustrate output from a statistical autocorrelation analysis (left) and via the renormalization blocking method (right). Data was taken from a run of length $N = 100$ walks sampling every 64 Monte Carlo steps. The autocorrelation histogram gives the average correlation between each data sample and the successive data samples one apart, two apart etc. The autocorrelation of sample 4 apart (= 256 steps) can be seen to be small so our sampling time of 1000 is a good over-estimate. The blocking output illustrates the calculated error on the parallel component of the end-to-end distance as a function of the blocking iteration. The blocking method course grains the sample at each iteration. Theoretically the flat (within error bars) section of the graph indicates that the uncorrelated regime has been reached. Here we use a sample run of walks of length $N = 100$ where we sampled every eight Monte Carlo steps. The x-axis indicates numbers of powers of two and so $1000 \approx 2^7 \times 8$ can be seen to be at the beginning of the flat regime and so represents an effectively uncorrelated sampling rate.

sizes of $S \geqslant 2^{17} = 131\,072$ (see table 1). The 131 072 samples of length $N = 12\,800$ took about 150 CPU days on the Dec AlphaStation 500/266 and so accounted for just under half of the total computer time of all runs.

During the simulation running averages were kept of various quantities associated with the size of the configurations—for some quantities values were kept from each independent configuration to check this independence (see above). Let us denote the vector position of the $N + 1$ monomers (vertices) of a walk of N steps as $\{\boldsymbol{r}_j;\, j = 0, \ldots N\}$. Let $\boldsymbol{a}$ be a unit vector in the direction given by (1.3) and $\boldsymbol{b}$ be a unit vector in a direction orthogonal to $\boldsymbol{a}$ so $\boldsymbol{a} \cdot \boldsymbol{b} = 0$. The quantities calculated were the average square end-to-end distance projections along the axes given by the theoretical anisotropy direction and its orthogonal:

$$\langle R^2_{e,\parallel}\rangle(N) = \langle((\boldsymbol{r}_N - \boldsymbol{r}_0) \cdot \boldsymbol{a})^2\rangle \tag{3.1}$$

and

$$\langle R^2_{e,\perp}\rangle(N) = \langle((\boldsymbol{r}_N - \boldsymbol{r}_0) \cdot \boldsymbol{b})^2\rangle \tag{3.2}$$

where the average $\langle\cdot\rangle$ is over the configurations φ_s of the sample $\mathcal{S}$ so that $\{\varphi_s \in \mathcal{S};\, s = 1, \ldots, S\}$. The total average square end-to-end distance is simply then

$$\langle R^2_e\rangle(N) = \langle R^2_{e,\parallel}\rangle(N) + \langle R^2_{e,\perp}\rangle(N). \tag{3.3}$$

The variances of these quantities where also calculated so a statistical error was able to be found from twice the standard deviation (as samples were essentially independent—see above). The radius of gyration was calculated, that is,

$$\langle R^2_g\rangle(N) = \frac{1}{N+1}\sum_{i=0}^{N}\langle(\boldsymbol{r}_i - \mathcal{R}_c) \cdot (\boldsymbol{r}_i - \mathcal{R}_c)\rangle \tag{3.4}$$

where the centre-of-mass for a configuration φ is the vector

$$\mathcal{R}_c(\varphi) = \frac{1}{N+1}\sum_{j=0}^{N} \boldsymbol{r}_j. \tag{3.5}$$

The mean-square distance of a monomer from the endpoint $\boldsymbol{r}_0$ was calculated as

$$\langle R_m^2\rangle(N) = \frac{1}{N+1}\sum_{i=0}^{N}\langle(\boldsymbol{r}_i - \boldsymbol{r}_0)\cdot(\boldsymbol{r}_i - \boldsymbol{r}_0)\rangle. \tag{3.6}$$

We note that the mean square-distance of the centre-of-mass from the endpoint is given by

$$\langle R_c^2\rangle(N) \equiv \langle(\mathcal{R}_c - \boldsymbol{r}_0)\cdot(\mathcal{R}_c - \boldsymbol{r}_0)\rangle = \langle R_m^2\rangle(N) - \langle R_g^2\rangle(N). \tag{3.7}$$

A list of averages with error estimates against N are given for $\langle R_{e,\parallel}^2\rangle(N)$, $\langle R_{e,\perp}^2\rangle(N)$, $\langle R_g^2\rangle(N)$, and $\langle R_m^2\rangle(N)$ in table 2. We also calculated the full average moment of inertia tensor (without standard deviations) about various positions $\boldsymbol{p}$, which is given in general terms as the 2×2 matrix

$$\boldsymbol{T}_{\boldsymbol{p}} = \frac{1}{N+1}\sum_{i=0}^{N}\langle(\boldsymbol{r}_i - \boldsymbol{p})\cdot(\boldsymbol{r}_i - \boldsymbol{p})\boldsymbol{I} - (\boldsymbol{r}_i - \boldsymbol{p})(\boldsymbol{r}_i - \boldsymbol{p})\rangle. \tag{3.8}$$

Here we use dyadic notation (that is, $\boldsymbol{rr}$ gives a tensor/matrix), and $\boldsymbol{I}$ is the identity (tensor) matrix. In particular, we calculated the average moment-of-inertia tensor about the endpoint $\boldsymbol{p} = \boldsymbol{r}_0$ and about the centre of mass $\boldsymbol{p} = \mathcal{R}_c$. We calculated $\boldsymbol{T}$ in Cartesian coordinates and in coordinates rotated to the symmetry axis given by (1.3). Note that the trace of these matrices gives back the radius of gyration and mean-square distance of a monomer from the endpoint:

$$\langle R_m^2\rangle(N) = \mathrm{Tr}(\boldsymbol{T}_{\boldsymbol{r}_0}) \tag{3.9}$$

and

$$\langle R_g^2\rangle(N) = \mathrm{Tr}(\boldsymbol{T}_{\mathcal{R}_c}). \tag{3.10}$$

From the average moment-of-inertia tensor, $\boldsymbol{T}_{\boldsymbol{p}}$, calculated about the centre-of-mass in coordinates rotated to the theoretical symmetry axis we calculated a quantity that allowed us to determine the veracity of the axis conjecture. To do this we diagonalized the resulting matrix and considered the eigenvector associated to the largest eigenvalue: this vector has components $\cos(\theta_d)$ along the theoretical symmetry axis and $\sin(\theta_d)$ perpendicular to it, where θ_d is the angle of deviation of the moment-of-inertia from the theoretical axis. After simulation we were able to extract an estimate for

$$\eta_d(N) = \sin(\theta_d(N)) \tag{3.11}$$

as a function of N. A list of these $\eta_d(N)$ values is given in table 3. The diagonalization of $\boldsymbol{T}_{\mathcal{R}_c}$ also allowed us to calculate the components of $\langle R_g^2\rangle(N)$ in the moment of inertia axes (that is, the finite N estimates of the major and minor axes, rather than along the theoretical axes). The components $\langle R_{g,\mathrm{major}}^2\rangle(N)$ and $\langle R_{g,\mathrm{minor}}^2\rangle(N)$ are given as the eigenvalues of $\boldsymbol{T}_{\mathcal{R}_c}$. These are tabulated in table 4.

Table 2. Our estimates for $\langle R^2_{e,\parallel}\rangle(N)$, $\langle R^2_{e,\perp}\rangle(N)$, $\langle R^2_g\rangle(N)$, and $\langle R^2_m\rangle(N)$ with error estimates. For $N = 44$ our Monte Carlo results can be compared with the exact enumeration data of $\langle R^2_{e,\parallel}\rangle(N) = 452.398\,55$ and $\langle R^2_{e,\perp}\rangle(N) = 7.702\,78$.

N	$\langle R^2_{e,\parallel}\rangle(N)$	$\langle R^2_{e,\perp}\rangle(N)$	$\langle R^2_g\rangle(N)$	$\langle R^2_m\rangle(N)$
44	452.24(20)	7.701(5)	65.340(14)	200.06(6)
50	563.53(26)	8.559 7(58)	82.387(18)	250.765(83)
60	770.8(4)	9.955 9(67)	114.810(26)	346.24(12)
70	1 004.3(5)	11.321 4(77)	152.162(35)	455.01(16)
80	1 261.8(6)	12.654 1(87)	194.292(45)	576.36(21)
100	1 846.7(9)	15.247(11)	292.633(69)	855.75(31)
130	2 890.6(1.5)	19.028(14)	474.40(12)	1 363.27(51)
165	4 340.4(2.3)	23.275(17)	736.55(19)	2 081.31(79)
200	6 029.4(3.3)	27.415(20)	1 051.21(27)	2 929.8(1.2)
250	8 823.3(4.9)	33.170(24)	1 588.53(40)	4 355.8(1.7)
320	13 457.7(7.6)	40.980(29)	2 510.06(65)	6 758.8(2.7)
400	19 721(16)	49.682(50)	3 796.4(1.4)	10 055.7(5.8)
560	35 146(41)	66.416(94)	7 091.8(3.8)	18 329(15)
800	65 008(155)	90.18(26)	13 767(15)	34 649(59)
1 130	117 854(311)	122.16(36)	26 124(29)	64 200(120)
1 600	215 919(580)	166.19(49)	49 976(62)	119 950(230)
2 260	393 343(1 465)	224.46(96)	94 860(170)	222 780(600)
3 200	724 120(2 960)	304.5(1.2)	181 520(360)	417 500(1 300)
4 520	1328 030(11 750)	416.4(3.7)	346 000(1500)	776 600(4 500)
6 400	2469 170(21 300)	566.8(4.9)	663 200(2 900)	1466 400(9900)
12 800	8406 600(68 000)	1049.0(9.3)	2418 600(9 900)	5152 300(33 000)

4. Analysis and discussion

In the previous exact enumeration work [7, 3] the mean square end-to-end distance $\langle R^2_e\rangle(N)$ and its components along the theoretical symmetry axis, $\langle R^2_{e,\parallel}\rangle(N)$ and $\langle R^2_{e,\perp}\rangle(N)$, up to $N = 44$ had been analysed assuming various possible asymptotic forms. The basic assumption was a form of the type

$$\langle R^2\rangle(N) \sim AN^{2\nu} \qquad \text{as } N \to \infty \tag{4.1}$$

for $\langle R^2_{e,\parallel}\rangle$, $\langle R^2_e\rangle$ and also for $\langle R^2_{e,\perp}\rangle^2$ (due to the anisotropy conjecture). Here we also consider $\langle R^2_g\rangle(N)$, and $\langle R^2_m\rangle(N)$ and extend our information up to $N = 12\,800$. Implicit in some of the series analysis via differential approximants [3] are various corrections to scaling assumptions. On the other hand, direct assumptions were made on the possible corrections to scaling in [7]. They obtained a similar answer to the differential approximant analysis. In [7] local exponent values were obtained from

$$\hat{\nu}(N) = \frac{\log(\langle R^2\rangle(N)) - \log(\langle R^2\rangle(N-2))}{\log(N) - \log(N-2)} \tag{4.2}$$

so as to test whether

$$\langle R^2\rangle(N) \sim AN^{2\nu}(1 + c/N^{\Delta}) \qquad \text{as } \to \infty \tag{4.3}$$

or

$$\langle R^2\rangle(N) \sim AN^{2\nu}(\log(N))^{\alpha} \qquad \text{as } N \to \infty \tag{4.4}$$

produced better fits to the data. The first, (4.3), implies

$$2\hat{\nu}(N) \sim 2\nu + C/N^{\Delta} \qquad \text{as } N \to \infty \tag{4.5}$$

Table 3. The eigenvalues $\langle R^2_{g,\mathrm{major}}\rangle(N)$ and $\langle R^2_{g,\mathrm{minor}}\rangle(N)$ of the moment of inertia tensor evaluate around the centre of mass, and the minor component of the eigenvector associated with $\langle R^2_{g,\mathrm{major}}\rangle(N)$.

N	$\langle R^2_{g,\mathrm{major}}\rangle(N)$	$\langle R^2_{g,\mathrm{minor}}\rangle(N)$	$\eta_d(N)$
50	80.3548	2.0326	0.012 841
60	112.386	2.4244	0.010 813
70	149.345	2.8175	0.009 425
80	191.083	3.2097	0.008 340
100	288.639	3.9936	0.006 817
130	469.229	5.1683	0.005 346
165	730.012	6.5357	0.004 307
200	1 043.31	7.9007	0.003 587
250	1 578.68	9.8512	0.002 936
320	2 497.48	12.573	0.002 341
400	3 780.77	15.681	0.001 894
560	7 069.95	21.872	0.001 415
800	13 736.2	31.053	0.000 994
1 130	26 080.4	43.773	0.000 7613
1 600	49 913.8	61.726	0.000 5453
2 260	94 774.7	86.740	0.000 3963
3 200	181 397	122.21	0.000 3157
4 520	345 841	172.16	0.000 1476
6 400	662 970	241.87	0.000 1955
12 800	2418 146	477.59	0.000 1001

while the second, (4.4), implies

$$2\hat{\nu}(N) \sim 2\nu + \alpha/\log(N) \qquad \text{as } N \to \infty. \tag{4.6}$$

However, it was found that in the case of the form (4.4) that for two-choice-spiral walks $\hat{\nu}(N)$ approached the apparent limiting value from above. Hence this implied that $\alpha > 0$. Although, correspondingly analysis for three-choice-spiral walks seemed to imply $\alpha < 0$ so no consistent confluent logarithmic correction could be implied for both models. The series analysis then tended to support the form (4.3) with exponent $2\nu = 1.69(1)$.

Before embarking on the description of our analyses let us first describe the possible asymptotic forms with corrections to scaling that we have considered. First is the form (4.3) above, which we shall refer to as 'scenario 1':

$$\langle R^2\rangle(N) \sim AN^{2\nu}(1 + c/N^{\Delta}) \qquad \text{as } N \to \infty \qquad \text{(scenario 1)}. \tag{4.7}$$

This would be the first choice for someone analysing a SAW problem. The possibility of multiplicative logarithms would imply that

$$\langle R^2\rangle(N) \sim AN^{2\nu}(\log(N))^{\alpha}(1 + c/N^{\Delta}) \qquad \text{as } N \to \infty \qquad \text{(scenario 2)} \tag{4.8}$$

is the asymptotic form. Additive logarithmic corrections instead of power law ones give two further possibilities:

$$\langle R^2\rangle(N) \sim AN^{2\nu}(1 + d/(\log(N))^{\beta}) \qquad \text{as } N \to \infty \qquad \text{(scenario 3)} \tag{4.9}$$

and

$$\langle R^2\rangle(N) \sim AN^{2\nu}(\log(N))^{\alpha}(1 + d/(\log(N))^{\beta}) \qquad \text{as } N \to \infty \qquad \text{(scenario 4)}. \tag{4.10}$$

Let us now revisit the type of analysis described above for the previous series analysis. We note the implied asymptotic forms of the local exponent estimates $\hat{\nu}(N)$ of our scenarios are: for scenario 1,

$$2\hat{\nu}(N) \sim 2\nu + C/N^{\Delta} \qquad \text{as } N \to \infty \tag{4.11}$$

for scenario 2,

$$2\hat{\nu}(N) \sim 2\nu + \alpha/\log(N) + C/N^{\Delta} \qquad \text{as } N \to \infty \tag{4.12}$$

for scenario 3,

$$2\hat{\nu}(N) \sim 2\nu + D/(\log(N))^{\beta} \qquad \text{as } N \to \infty \tag{4.13}$$

while scenario 4 gives

$$2\hat{\nu}(N) \sim 2\nu + D/(\log(N))^{\beta} + \alpha/\log(N) \qquad \text{as } N \to \infty. \tag{4.14}$$

As one can readily see, if the additive logarithmic correction exponent β is close to or smaller than 1 it will be virtually impossible for any analysis to differentiate it from a multiplicative logarithmic correction. However, one can try to do two things. First one can test scenario 1 by examining the corrections to scaling to see if they are so strong that they must be logarithmic. Secondly one can test scenario 2 by demonstrating whether a consistent value of α can be associated with the data. We shall do this and in fact conclude that the asymptotic form is not likely to be of either of the forms, (4.7) or (4.8). We hence conclude that the data must follow a form of either of the type scenario 3, (4.9), or scenario 4, (4.10), that is, with additive logarithmic corrections.

To accommodate the Monte Carlo data we widen our definition of the local exponent by finding local estimates $\hat{\nu}(N)$ from linear regression analyses of $\log(\langle R^2\rangle(N))$ against $\log(N)$ over m consecutive values of N given in table 2, using the various choices for $\langle R^2\rangle(N)$ given there. We considered various values of m including 3, 4, 5 and 6, and satisfied ourselves that the conclusions of the analysis below were robust to this change. The analysis included the exact enumeration data where possible. Figure 5 illustrates our local exponent estimates $2\hat{\nu}(N)$, using $m = 4$, for $\langle R^2_{e,\parallel}\rangle(N)$, against $1/\log(N)$. Here, the values of $\log(\langle R^2_{e,\parallel}\rangle(N))$ on which the regressions were performed were groups of four ($m = 4$) data points that successively overlap by two data points. More precisely, the four points of largest N were considered (regression was performed to give a local exponent estimate) then the two with largest N were dropped and another two of smaller N added and another regression performed, and so on until it was not possible to continue. Each regression gave a local exponent estimate. Each estimate was plotted against the inverse of the average of $\log(N)$ over the four points in question, we denote this scale as $1/\overline{\log(N)}$. The error bars are simple statistical errors (two standard deviations) from the regression analysis. The small N data can be seen to fall on a reasonably straight line on this scale: this being the exact enumeration data for $N \leqslant 44$. Importantly however, the exponent estimates from Monte Carlo do not continue to follow this trend and have a turning point about $\overline{\log(N)} \approx \log(150)$. While the error bars become large it is clear that most extrapolations will yield a value of 2ν substantially larger than 1.7. So our first conclusion is that previous series analysis is not accurate since there is a turning point in the effective exponent and that 2ν is substantially larger than 1.7.

We now consider similar analyses of $\langle R^2_{e,\perp}\rangle^2(N)$, $\langle R^2_g\rangle(N)$, $\langle R^2_m\rangle(N)$, $\langle R^2_{g,\text{major}}\rangle(N)$ and $\langle R^2_{g,\text{minor}}\rangle^2(N)$ in comparison with that of $\langle R^2_{e,\parallel}\rangle(N)$. Figure 6 shows local estimates $2\hat{\nu}(N)$ of 2ν obtained from the four sources, $\langle R^2_{e,\parallel}\rangle(N)$, $\langle R^2_m\rangle(N)$, $\langle R^2_g\rangle(N)$, and $\langle R^2_{g,\text{minor}}\rangle^2(N)$ plotted against $1/\overline{\log(N)}$ in the same manner as described above. This scale was chosen

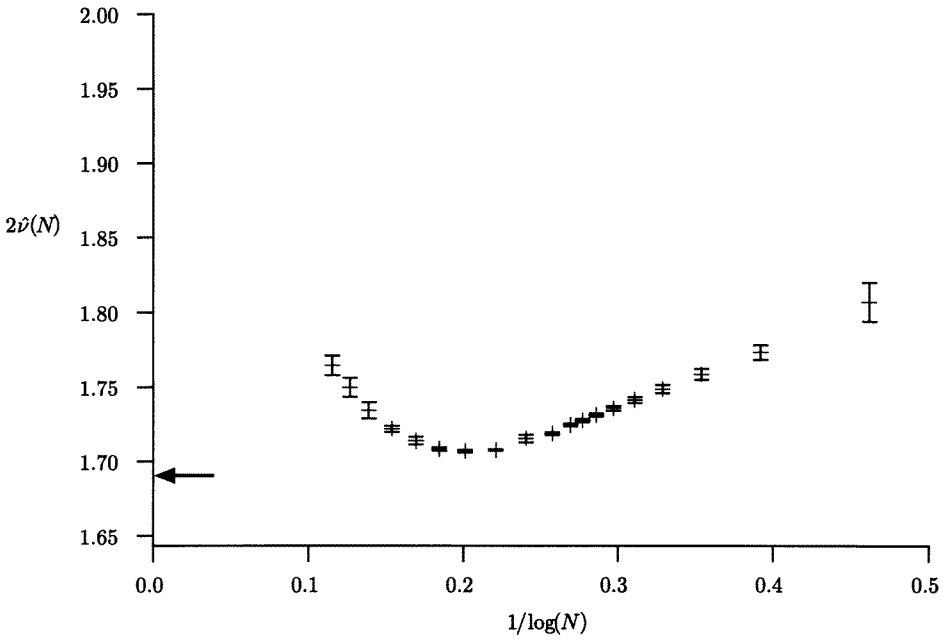

Figure 5. Local exponent estimate $2\hat{\nu}$ from the exact enumeration data and Monte Carlo data of $\langle R^2_{e,\parallel}\rangle(N)$. They were obtained from local regressions over four points moving over the whole range two at a time. The arrow approximately indicates the exact enumeration extrapolation.

with hindsight as the most appropriate to allow a consistent final estimate of 2ν from each of the data sets. More precisely, if we extrapolate over $1/\log(N)^{1/2}$, $1/\log(N)^2$, $1/N^{1/3}$, or $1/N$ say, the four extrapolations give widely different answers. Extrapolating linearly on the scale, $1/\overline{\log(N)}$, over the last three points of each graph (which were obtained from the top eight points of $\langle R^2\rangle(N)$ data: $N \geqslant 800$) gives four estimates of 2ν. From these, and leaving some allowance for systematic error by using different scales and different regression analyses, gives $2\nu = 1.91 \pm 0.03$. Remember that while this scale is more appropriate than $1/N$, or $1/\sqrt{N}$ say, it is only empirical and we cannot say that it is the best scale, of course. We note that the total radius of gyration seems the best converged quantity. Hence, the likely logarithmic correction rules out scenario 1. Also, since the sign of the correction is not consistent it is unlikely that scenario 2 can hold unless there exist further turning points for some of the quantities local exponent estimates (which may of course occur since we have already seen one!). Hence, we might conclude that scenarios 3 or 4 are likely candidates *given* that we have reached the asymptotic regime.

We have also tackled the question as to which scenario is likely to be correct by examining ratios of pairs of the size-measuring quantities $\langle R^2\rangle(N)$ such as $\langle R^2_m\rangle(N)/\langle R^2_{e,\parallel}\rangle(N)$. Such a quantity should converge to some constant with the same corrections as the $\langle R^2\rangle(N)$ quantities (unless some fortuitous cancellation occurs). For example one possibility is

$$\langle R^2_m\rangle(N)/\langle R^2_{e,\parallel}\rangle(N) \sim C\,(1 + g/(\log(N))^{\beta}) \qquad \text{as } N \to \infty \tag{4.15}$$

while another is

$$\langle R^2_m\rangle(N)/\langle R^2_{e,\parallel}\rangle(N) \sim C\,(1 + h/N^{\Delta}) \qquad \text{as } N \to \infty. \tag{4.16}$$

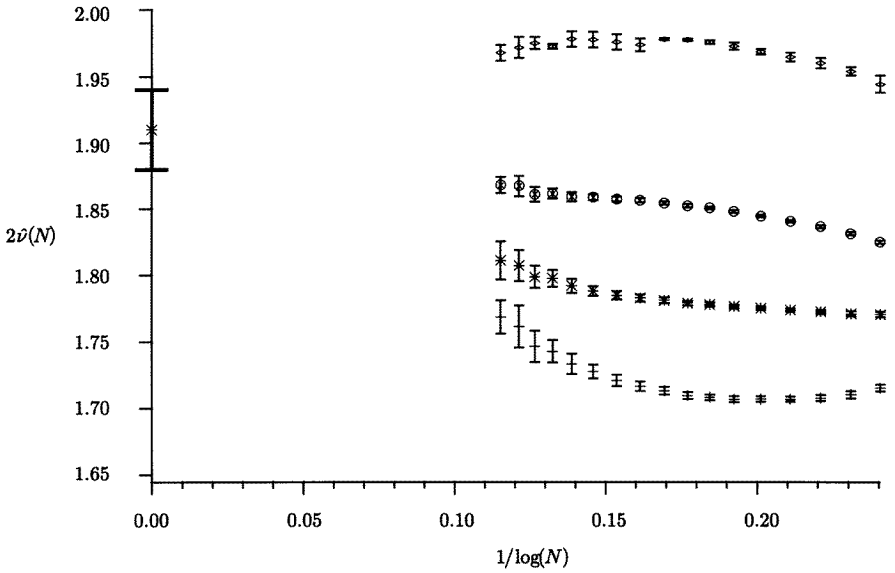

Figure 6. Local exponent estimate $2\hat{\nu}$ from the exact enumeration data and Monte Carlo data. They were obtained from local weighted regressions over four points moving over the whole range one at a time. From top to bottom they were obtained from $\langle R^2_{g,\mathrm{minor}}\rangle^2(N)$, $\langle R^2_g\rangle(N)$, $\langle R^2_m\rangle(N)$, and $\langle R^2_{e,\|}\rangle(N)$ respectively. On the axis the extrapolated estimate of 2ν is indicated with error.

In this way we eliminate one of the free parameters in our fitting forms relative to considering the $\langle R^2\rangle(N)$ quantities separately. We have compared (4.15) with $\beta = 1$ to (4.16) with $\Delta = \frac{1}{2}$ and $\Delta = 1$. In fact we have examined both linear and quadratic fits in those correction scales (for $N \geqslant 800$). Intriguingly the best scale is (4.16) with $\Delta = \frac{1}{2}$ which gives amplitude ratio estimates to two decimal places consistently. That is, we extrapolated $\langle R^2_m\rangle/\langle R^2_{e,\|}\rangle$, $\langle R^2_{e,\|}\rangle/\langle R^2_m\rangle$, $\langle R^2_m\rangle/\langle R^2_g\rangle$, and $\langle R^2_g\rangle/\langle R^2_{e,\|}\rangle$, and found constants C that were consistent with each other. While (4.15) with $\beta = 1$ was not so well behaved we cannot rule out the consistency of a quadratic fit, albeit with large error bars. Given that a quadratic form of (4.16) with $\Delta = \frac{1}{2}$ (that is, with $N^{-1/2}$ and N^{-1} terms) is close to the correct asymptotic form of the corrections (that is, power law corrections) the only way to *reconcile this with the previous evidence on extrapolating* $2\hat{\nu}$ *is to concede the asymptotic regime has not been reached*. It is therefore likely that combinations of logarithmic and power law (with small exponent) corrections combine in this problem to make asymptotic analysis difficult. Reflecting on our estimate of 2ν then gives us pause for concern, and implies that we should increase the possible size of systematic error. We therefore conclude that $2\nu = 1.91(4)$.

Finally, in figure 7 we have a plot of $\eta_d(N)$ against $1/N^{0.9}$. The scale $1/N^{0.9}$ was chosen to give a good linear fit over the data: note that there is no turning point feature in this data and no evidence of logarithmic corrections to scaling. We note that 0.9 is close to ν and one might expect corrections of $1/N^{\nu}$ to the major-axis-angle scaling. The line on the figure is the line of linear best fit. It is clearly extrapolating to 0 within statistical error (we find $-0.5 \times 10^{-5} \pm 2.0 \times 10^{-5}$ over the last 13 points, which are of the order 10^{-3}) and so we conclude that the theoretical angle conjectured in [3] does indeed hold.

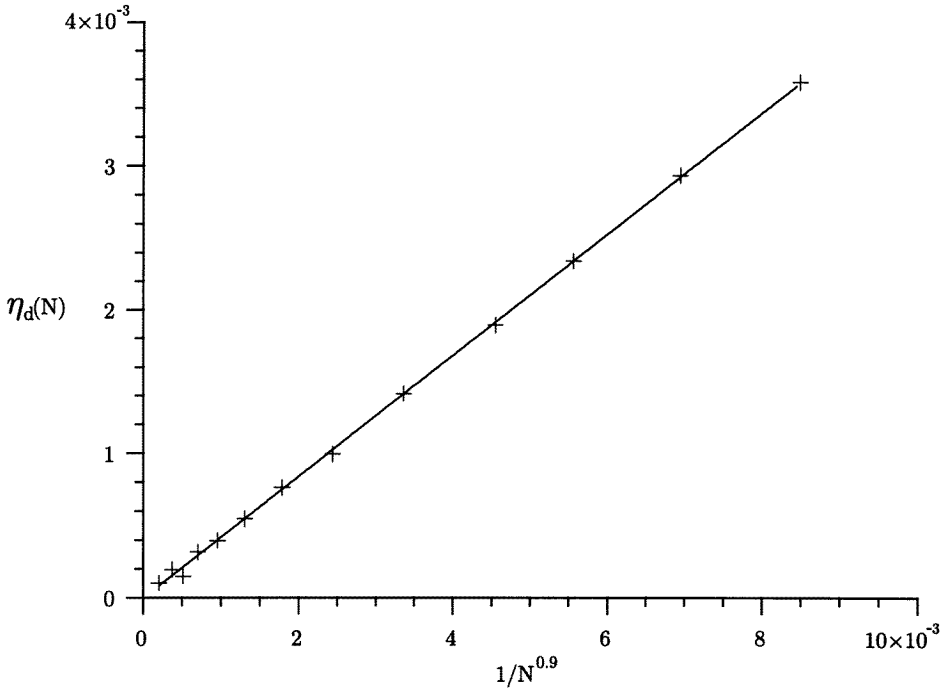

Figure 7. A plot of the minor component $\eta_d(N)$ of the eigenvector associated with the maximum eigenvalue of the inertia tensor against $N^{-0.9}$. This component $\eta_d(N)$ is the sine of the angle of deviation from the theoretical axis. The straight line is the line of best fit.

5. Summary

We have made extensive Monte Carlo simulations of so-called two-choice-spiral walks, also known as ASSAW, so as to ascertain the scaling of the 'size' of the model polymer with length. These results, for walks up to length $N = 12\,800$, show that indeed the ASSAW models form a distinct universality class, though with scaling exponents substantially different to those previously estimated from the series analysis of exact enumeration data. The major axis scaling exponent we estimate as $\nu_{\|} = 0.955 \pm 0.02$. This is opposed to previous estimates that differ by about 0.10. We explain the inaccuracy of the series estimates by noting that there is a turning point in the local exponent estimates as N increases, in the same way as there is for pure spiral walks (SSAW). This we believe is the result of a combination of strong power law and logarithmic corrections to scaling.

Acknowledgments

Financial support from the Australian Research Council is gratefully acknowledged by RB and ALO while financial support from NSERC of Canada is gratefully acknowledged by CS. Some of this work was carried out while CS was visiting the Department of Mathematics and Statistics of the University of Melbourne and CS is grateful to the members of that department for their kind and generous hospitality. Also, some of this work was carried out while RB and ALO were visiting the Department of Mathematics and Statistics of the University of Saskatchewan, and they are grateful to the members of that department for

their kind and generous hospitality. ALO is grateful to T Benyi for discussions on hashing algorithms.

References

[1] Madras N and Slade G 1993 *The Self-avoiding Walk* (Boston, MA: Birkhauser)
[2] Privman V and Švrakić N M 1989 Directed models of polymers, interfaces, and clusters: scaling and finite–size properties *Lecture Notes in Physics* vol 338 (Berlin: Springer)
[3] Guttmann A J, Prellberg T and Owczarek A L 1993 *J. Phys. A: Math. Gen.* **26** 6615
[4] Nienhuis B 1982 *Phys. Rev. Lett.* **49** 1062
[5] Blöte H W J and Hilhorst H J 1984 *J. Phys. A: Math. Gen.* **17** L111
[6] Manna S S 1984 *J. Phys. A: Math. Gen.* **17** L899
[7] Guttmann A J and Wallace K J 1985 *J. Phys. A: Math. Gen.* **18** L1049
[8] Whittington S G 1985 *J. Phys. A: Math. Gen.* **18** L67
[9] Jensen I and Guttmann A J 1995 *J. Phys. A: Math. Gen.* **28** 4813
[10] Lal M 1969 *Mol. Phys.* **17** 57
[11] L'Ecuyer P and Cote S 1991 *ACM Trans. Math. Softw.* **17** 98
[12] L'Ecuyer P 1988 *Commun. ACM* **31** 742
[13] Coddington P 1993 *Preprint* SSCS 526 Syracuse University
[14] Marsaglia G and Zaman A 1991 *Ann. Appl. Probab.* **1** 462
[15] van Rensburg E J J and Madras N 1992 *J. Phys. A: Math. Gen.* **25** 303
[16] Soteros C E and Paulhus M M 1997 *IMA Volumes in Mathemtics and Statistics* vol 102, ed S G Whittington p 121 in press
[17] Flyvbjerg H and Petersen H G 1989 *J. Chem. Phys.* **91** 461