Journal of Statistical Physics, Vol. 82, Nos. 1/2, 1996

Monte Carlo Study of the Interacting Self-Avoiding Walk Model in Three Dimensions

M. C. Tesi,[1] E. J. Janse van Rensburg,[2] E. Orlandini,[3] and S. G. Whittington[1]

Received February 20, 1995

We consider self-avoiding walks on the simple cubic lattice in which neighboring pairs of vertices of the walk (not connected by an edge) have an associated pair-wise additive energy. If the associated force is attractive, then the walk can collapse from a coil to a compact ball. We describe two Monte Carlo algorithms which we used to investigate this collapse process, and the properties of the walk as a function of the energy or temperature. We report results about the thermodynamic and configurational properties of the walks and estimate the location of the collapse transition.

KEY WORDS: Self-avoiding walks; lattice models; Markov chains; Monte Carlo; phase transitions.

1. INTRODUCTION

Linear polymer molecules in dilute solution are believed to exist in three distinct states, depending on the quality of the solvent. In a good solvent (at high temperature) the polymer is an expanded random coil, while in a poor solvent (at low temperature) it is a compact ball. Between these two forms there is a state which exists only at a single temperature, the theta temperature, in which the exponents characterizing the dimensions of the polymer are different from those in either the good or poor solvent regimes. In this paper we describe two Monte Carlo algorithms for investigating a model of the collapse of a linear polymer from a coil to a ball. Although one of these algorithms (umbrella sampling) is wellknown in the statistical

[1] Department of Chemistry, University of Toronto, Toronto, Ontario, Canada M5S 1A1.
[2] Department of Mathematics, York University, North York, Ontario, Canada M3J 1P3.
[3] Theoretical Physics, University of Oxford, Oxford OX1 3NP, United Kingdom.

0022-4715/96/0100-0155$09.50/0

mechanics of liquids,[1] it has not previously been applied to the collapse of linear polymers. The other algorithm (which we call multiple Markov chains) is a variant of a method described by Geyer and Thompson[2] in the statistics literature. Both of these algorithms are designed to circumvent the quasi-ergodic problems associated with sampling in the collapsed phase.

The abrupt change in the dimensions of linear polymer molecules in dilute solution when the temperature (or solvent quality) is varied can be measured directly by light scattering[3,4] or indirectly by measuring the intrinsic viscosity of the polymer.[5] Although the collapse has been studied for many years, the detailed nature of the transition is still not well understood. The most frequently studied model is an interacting self-avoiding walk. We consider the simple cubic lattice Z^3, and write (x, y, z) for the integer coordinates of a typical point. An n-step self-avoiding walk is an ordered sequence of $(n+1)$ vertices of the lattice with coordinates (x_i, y_i, z_i), $i=0, 1,..., n$, such that (x_0, y_0, z_0) is the origin, pairs of vertices whose index differs by unity are unit distance apart, and every vertex is distinct. Self-avoiding walks have proved to be useful models of the dilute solution behavior of linear polymers in good solvents. To mimic the effect of solvent quality, we introduce a potential associated with the walk, $V=\sum_{i<j} v_{ij}$, where v_{ij} is infinity if the vertices i and j are coincident, ε if the vertices are unit distance apart and $|i-j| \neq 1$, and zero otherwise. We shall be mainly interested in the case $\varepsilon<0$, so that there is a short-range attractive force between pairs of vertices. We call each pair of vertices for which $v_{ij}=\varepsilon$ a *contact*. The canonical partition function can be written as

$$Z_n(\beta)=\sum_m c_n(m)\, e^{-m\varepsilon/k_B T}=\sum_m c_n(m)\, e^{m\beta} \tag{1.1}$$

where $c_n(m)$ is the number of self-avoiding walks with n edges and m contacts, $\beta=-\varepsilon/k_B T$, k_B is Boltzmann's constant, and T is the absolute temperature.

It is easy to prove that

$$\lim_{n\to\infty} n^{-1} \log Z_n(\beta) \equiv \mathscr{F}(\beta) \tag{1.2}$$

exists for $\beta \leqslant 0$, and that $\mathscr{F}(\beta)$ is monotone, convex, and continuous in this regime. However, for $\beta>0$ the existence of the limit has not been established. Nonetheless, the usual approach is to assume the existence of the limit, and that there is a single positive value of β at which $\mathscr{F}(\beta)$ is nonanalytic. Let us call this value β_o. For $\beta<\beta_o$ the walk is expanded and for $\beta>\beta_o$ it is compact. If we write $\langle S_n^2(\beta)\rangle$ for the mean square radius of gyration, then we expect that

$$\langle S_n^2(\beta)\rangle \sim n^{2\nu(\beta)} \tag{1.3}$$

where $\nu(\beta) = \nu(0) \equiv \nu$ for $\beta < \beta_o$, $\nu(\beta) = \nu_t$ for $\beta = \beta_o$, and $\nu(\beta) = 1/d$ for $\beta > \beta_o$, where d is the dimension of the lattice. In two dimensions there is little doubt about the value of ν,[6] and there is a prediction for the value of ν_t.[7] In three dimensions there are predictions based on renormalization group arguments.[8–12] In addition, the problem has been studied by transfer matrix methods[13] and by exact enumeration and series analysis techniques.[14–18]

This seems to be a natural problem for Monte Carlo methods and, indeed, there have been a number of attempts to attack the problem in this way.[19–22] One approach is to use a method in which the self-avoiding walk is constructed step by step,[23] and Mazur and McCrackin[19] and Meirovitch and Lim[22] both use methods of this general type. The other idea is to construct a Markov chain whose states are the n-step self-avoiding walks and to sample along a realization of this Markov chain. This is the approach followed by Webman *et al.*[21] Since the relative weights of the walks will be different and will depend on the value of β, one usually chooses the Markov chain to have the appropriate Boltzmann distribution for its limit distribution. This method will work well provided that β is not too large, but for larger (positive) values of β the "mobility" of the Markov chain can be very low. That is, there are different regions of configuration space which are important at a particular value of β but which are not easily accessible from one another in a realization of the Markov chain described above. The rate of convergence of the Markov chain to its limit distribution can be very slow and, worse still, this might not be apparent in a run of reasonable length, so that error estimates might be seriously underestimated. Clearly one needs to design a Markov chain which is more mobile and, in the next section, we describe two different methods for achieving this.

The primary aim of this paper is to describe the sampling schemes and their implementation. In addition, in Section 4 we report some numerical results about the β-dependence of thermodynamic and metric properties and use these to form estimates of the location of the Θ point.

2. MARKOV CHAIN SAMPLING

One is interested in estimating the expectation value of some property Q of the walks with respect to the Boltzmann distribution, i.e.,

$$\langle Q(\beta) \rangle = \frac{\sum_{k=1}^{c_n} Q(k)\, e^{m(k)\beta}}{\sum_{k=1}^{c_n} e^{m(k)\beta}} = \frac{1}{Z_n(\beta)} \sum_{k=1}^{c_n} Q(k)\, e^{m(k)\beta} \tag{2.1}$$

where the sums are over all the c_n n-step self-avoiding walks. $Q(k)$ is the value of the property Q for the kth walk and $m(k)$ is the number of

contacts in the kth walk. A natural idea is to define a Markov chain on the set of n-step walks, whose unique limit distribution $\{\pi_k(\beta)\}$ is the Boltzmann distribution $\{e^{\beta m(k)}/Z_n(\beta)\}$. For example, the Markov chain could be chosen such that the elements p_{ij} of the transition matrix from state i to state j are given by

$$p_{ij} = q_{ij} \min(1, \pi_j/\pi_i) \tag{2.2}$$

for $j \neq i$ and with p_{ii} chosen to make the row sums of the matrix equal to unity. Here q_{ij} are the elements of the transition matrix of an underlying symmetric Markov chain ($q_{ij} = q_{ji}$) which is ergodic. It is easy to see that, provided that all the π_j are positive, the Markov chain with transition probabilities given by (2.2) has unique limit distribution $\{\pi_i\}$.[23] In principle it is only necessary to set π_i to be the Boltzmann probability for the ith walk and to choose a suitable underlying symmetric Markov chain, such as that associated with the pivot algorithm.[24, 25] In this case the natural estimator of (2.1) is the sample mean

$$\bar{Q}(\beta) = \frac{1}{N} \sum_{t=1}^{N} Q(S(t)) \tag{2.3}$$

where $S(t)$ is the state of the Markov chain at time step t in the N-step realization.

The problem with this method is that at large positive values of β the Markov chain converges only very slowly to its limit distribution. Typical realizations of the Markov chain spend long periods sampling relatively small regions of the configuration space and only rarely move to other regions. This is the classical *quasiergodic problem.*[23] One might hope to avoid this difficulty by sampling the state space at lower values of β. This approach depends on the following identity:

$$\begin{aligned} \langle Q(\beta) \rangle &= \frac{\sum_{k=1}^{c_n} Q(k)\, e^{\beta m(k)} \pi_k/\pi_k}{\sum_k e^{\beta m(k)} \pi_k/\pi_k} \frac{\sum_k \pi_k}{\sum_k \pi_k} \\ &= \frac{\sum_k Q(k)(e^{\beta m(k)}/\pi_k)(\pi_k/\sum_k \pi_k)}{\sum_k (e^{\beta m(k)}/\pi_k)(\pi_k/\sum_k \pi_k)} \\ &= \frac{\langle Q e^{\beta m}/\pi \rangle_\pi}{\langle e^{\beta m}/\pi \rangle_\pi} \end{aligned} \tag{2.4}$$

where the subscript π on the angular brackets denotes expectation with respect to the distribution $\{\pi_i\}$. This quantity can be estimated by the ratio estimate

$$\hat{Q}(\beta) = \frac{\sum_{t=1}^{N} Q(S(t))\, e^{\beta m(S(t))}/\pi_{S(t)}}{\sum_{t=1}^{N} e^{\beta m(S(t))}/\pi_{S(t)}} \tag{2.5}$$

By choosing $\pi_k = e^{\beta' m(k)}/Z_n(\beta')$ with $\beta' < \beta$ we can sample at β', where the mobility of the Markov chain should be higher, and reweight the data appropriately to obtain an estimate of the expectation at β.

In practice this is not very useful because if β' is small enough to give sufficient extra mobility, then the overlap between the distributions corresponding to β and β' will usually be too small to have efficient sampling over the parts of configuration space relevant to the distribution at β. However, several variants of this idea based on the identity (2.4) are useful, and some of them will be described in the following subsections.

2.1. Multiple Markov Chains

Geyer and Thompson[2] have described a method for increasing the mobility based on sampling along a set of Markov chains[4] run in parallel. We describe a minor variant of this approach, which we have found to be very useful in these problems. One wants to sample at some value β for which convergence is very slow, and one knows that convergence is fast at some other value β'. The idea is to select a set of values $\beta' = \beta_1 < \beta_2 < \beta_3 < \cdots < \beta_m = \beta$ to interpolate between β' and β so that there is considerable overlap in the distributions at β_l and β_{l+1}.

The Markov chains at $\beta_1, \beta_2, \ldots, \beta_m$ are evolved in parallel for a specified number of time steps. Then an adjacent pair, β_l, β_{l+1}, of β values is chosen uniformly from the $m-1$ adjacent pairs and, as a trial move, the configurations at these β values are swapped. Suppose that, when the swap is attempted, the state of the lth chain is S_l and the state of the $(l+1)$th chain is S_{l+1}. Writing $\pi_k(\beta_l)$ for the probability of the state k at β_l, the trial move is accepted with probability $r(l, l+1)$ given by

$$r(l, l+1) = \min\left(1, \frac{\pi_{S_{l+1}}(\beta_l)\, \pi_{S_l}(\beta_{l+1})}{\pi_{S_l}(\beta_l)\, \pi_{S_{l+1}}(\beta_{l+1})}\right) \tag{2.6}$$

This whole process is itself a Markov chain, which we call the *composite* Markov chain. Since each of the separate Markov chains is ergodic, so is the composite Markov chain, and the unique limit distribution of the composite Markov chain is the product distribution of the separate Markov chains at $\beta_1, \beta_2, \ldots, \beta_m$. This is immediate since, if $r(l, l+1) < 1$, then

[4] In this section we shall be primarily interested in Markov chains with Boltzmann limit distribution $\{\pi_k(\beta_i)\} = \{\exp[\beta_i m(k)]/Z_n(\beta_i)\}$.

$$\left[\prod_{k=1}^{m} \pi_{S_k}(\beta_k)\right] \frac{r(l, l+1)}{m-1} = \left[\prod_{k=1}^{m} \pi_{S_k}(\beta_k)\right] \frac{1}{m-1} \left(\frac{\pi_{S_{l+1}}(\beta_l)\, \pi_{S_l}(\beta_{l+1})}{\pi_{S_l}(\beta_l)\, \pi_{S_{l+1}}(\beta_{l+1})}\right)$$

$$= \left[\prod_{k \neq l, l+1} \pi_{S_k}(\beta_k)\right] \pi_{S_{l+1}}(\beta_l)\, \pi_{S_l}(\beta_{l+1}) \frac{1}{m-1} \tag{2.7}$$

with a similar result when $r(l, l+1) = 1$.

One can understand the advantage of this method by focusing on a particular β value. Every successful swap involving this and a neighboring chain corresponds to a large change in the configuration at that value of β, so the correlation time at each β is markedly reduced. Swaps move the system (at a particular β) to new regions of configuration space and therefore tend to mitigate quasi-ergodic problems. Of course, some β values should be chosen so that convergence at those values is rapid.

There are some practical considerations which require attention. First one has to decide on the set of β values. Clearly they should be sufficiently close together (and the value of m should be sufficiently large) that swapping occurs rather frequently, and β_1 should be small enough that convergence at that β value is rapid. We have always chosen $\beta_1 = 0$ to ensure this. It is also necessary to decide for how long to evolve the Markov chains between attempted swaps. One can collect information at every value of β at the same time, so that there is very little overhead compared with sampling at separate β values without the swapping. Hence the gain in convergence rate is obtained at only marginal cost in computer time. In Section 3 we make some further comments concerning our implementation of the method and we present numerical results.

2.2. Umbrella Sampling

The technique of umbrella sampling was invented by Torrie and Valleau[1] and has been extensively applied to problems in the theory of liquids[26] and to spin systems.[27] The method relies on the identity (2.4) and on the observation that this relation is true for *any* probability distribution $\{\pi_k\}$. There is no need for this distribution to be directly related to the *natural* probability distribution of the problem. The sampling from $\{\pi_k\}$ can be carried out using the usual Metropolis criterion with a suitable selection of trial moves. The original motivation was to allow efficient estimation of free energy differences, but the approach is also a very effective way to avoid quasi-ergodic problems. One can choose $\{\pi_k\}$, which we call the umbrella distribution, to overlap the Boltzmann distribution(s) at the temperature(s) of interest, and to extend to higher temperatures so that the mobility of the Markov chain is increased. This allows enormous flexibility.

However, this flexibility implies that there is no obvious criterion for the best choice of the umbrella distribution.[28] One possibility[29] is to take a weighted linear combination of the usual Boltzmann factors, at a number of temperatures, including the range of interest and also higher temperatures to increase mobility. We write the distribution π_k as follows:

$$\pi_k = \sum_j w(\beta_j)\, e^{\beta_j m(k)} \tag{2.8}$$

The problem is to determine appropriate values for the weighting factors $w(\beta_j)$. One might choose to make the sampling approximately uniform in the temperature variable.

To satisfy this requirement, we consider the quantity $\Pi = \sum_k \pi_k$, which may not be unity if the separate π_k are not normalized. We can write Π as

$$\begin{aligned}\Pi &= \sum_{k=1}^{c_n} \pi_k \\ &= \sum_k \sum_j w(\beta_j)\, e^{\beta_j m(k)} \\ &= \sum_j w(\beta_j) \sum_k e^{\beta_j m(k)} \\ &= \sum_j w(\beta_j)\, Z_n(\beta_j)\end{aligned} \tag{2.9}$$

The choice of the weights $w(\beta_j)$ in the above depends on the type of simulation that we wish to perform. We want to sample from an umbrella which covers a number of distributions at different β values and, moreover, we want to sample equally from each of those distributions. The contribution of the jth distribution to Π is $w(\beta_j)\, Z_n(\beta_j)/\Pi$, and if we set the weights such that these are all a constant, then we can expect our algorithm to sample equally from these. Hence $w(\beta_j)\, Z_n(\beta_j)$ must be constant. A convenient choice for the constant is $Z_n(0)$, the partition function at infinite temperature. With this choice, since the partition function is related to the finite-n free energy $F_n(\beta)$ by the relation $Z_n(\beta) = \exp[nF_n(\beta)]$, the weights are given by

$$w(\beta_j) = e^{-n(F_n(\beta_j) - F_n(0))} \tag{2.10}$$

Consequently, to employ this approach we need a reasonable approximation to the free energy differences for systems of the required size.

3. IMPLEMENTATION AND ANALYSIS OF THE METHODS

To implement any of these Markov chain sampling methods one must choose an underlying symmetric Markov chain, i.e., one must decide how a trial move is to be chosen. The important technical condition is that this underlying Markov chain must be ergodic. In practice this means that every state can be reached from every other state in a finite number of steps. It is known that Markov chains which use only local moves, and which work on walks with a fixed number of steps, are not ergodic,[30] so it is essential to incorporate some nonlocal moves into the algorithm.

The pivot algorithm[24,25] is the most efficient algorithm currently known for simulating self-avoiding walks of fixed length. With this algorithm one effectively independent sample can be produced in a computer time of order n (where n is the length of the walk), and this is the best possible order of magnitude, since it takes a time of order n merely to write down an n-step walk. A vertex of the walk is chosen uniformly and at random, the walk is disconnected into two subwalks at this vertex, one of the two subwalks is subjected to a randomly chosen lattice symmetry operation, and the two subwalks are reconnected at the chosen vertex. If the resulting walk is self-avoiding, the move is accepted, otherwise it is rejected. For a detailed description of the algorithm see ref. 31.

Unfortunately, the efficiency of the algorithm is dramatically affected by the introduction of a contact potential which favors compact configurations of the walk. For example, the probability for a pivot move to be accepted for a walk of $n = 200$ steps, sampled at $\beta = 0$, is approximately 0.32, whereas for the same walk sampled at the theta point, it is of the order of 0.11–0.12. This is because the pivot algorithm often proposes to move "large" pieces of the walk and, for more compact walks, these large-scale moves are likely to be rejected because of the self-avoidance constraint. In order to increase the probability of a move being accepted in these circumstances and to reduce the autocorrelation between samples, one needs to introduce other types of moves.

3.1. Local Moves

The number of contacts in the walk plays an important role in Markov chain sampling for the collapse problem, since this quantity turns out to be one of the slowest modes. This mode reflects a local property of the walk and it seems reasonable to try to speed up convergence by adding to the pivot moves some kind of *local move*.[31]

A local move is one that alters only a few consecutive vertices of the self-avoiding walk, leaving the other vertices unchanged. We used three

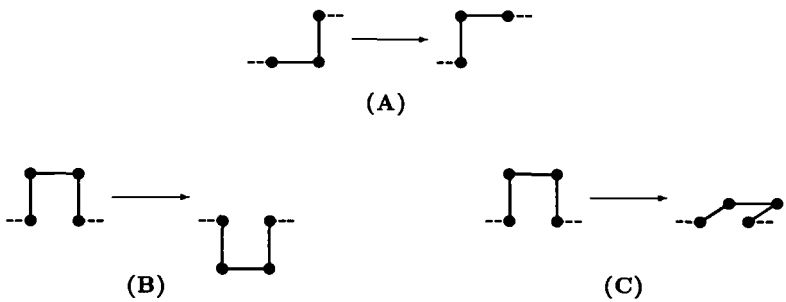

Fig. 1. Local moves used: (A) One-bead flip, (B) 180° crankshaft, and (C) 90° crankshaft.

kinds of moves typically known as "one-bead flip, 180° crankshaft, and 90° crankshaft" (see Fig. 1).

Although these moves affect only a small piece of the original self-avoiding walk, the inclusion of a sufficient number of them can produce a reasonable improvement in the autocorrelation time of the slow modes. This seems to be because the probability of such moves being accepted is of the order of 0.5, independent of the length of the walk and of the value of β, at least when β is not too large. Moreover, these local changes in conformation can, for compact configurations, be more efficient than pivots in changing the numbers of contacts.

From the point of view of the efficiency of the algorithm, a decision must be made about the relative numbers of local and pivot moves. Since the pivot moves involve pieces of the walk with size proportional to n (the length of the walk), it seems reasonable to combine each of them with a number of local moves which is proportional to n. After some initial trials we found it convenient to perform $n/4$ attempted local moves for every attempted pivot move.

In Tables I and II we compare estimates of the integrated autocorrelation times (given in units of sampling, i.e., per attempted pivot move or per attempted pivot $+n/4$ attempted local moves) of several observables, as a function of β, obtained both from the pivot algorithm and from the pivot + local algorithm with the usual Metropolis sampling scheme. In both cases (pivot and pivot + local) the integrated autocorrelation times for all the observables increase as β approaches the Θ point, as would be expected.[32] However, it is clear that the introduction of the local moves produces an appreciable reduction in the autocorrelation times of all the variables considered. This indicates that the addition of local moves makes the pivot algorithm more effective, especially around the transition region. From now on, the algorithm which we shall use and refer to will always be this combination of pivot and local moves.

Table I. Autocorrelation Times for $\langle S^2 \rangle$ and $\langle c \rangle$ as a Function of β for Walks of Length $n = 200$[a]

	$\tau(\langle S^2 \rangle)$				$\tau(\langle c \rangle)$			
β	Pivot	Pivot + local	MMC	Umbrella	Pivot	Pivot + local	MMC	Umbrella
0.000	0.77 ± 0.04	0.71 ± 0.03	0.62 ± 0.02	0.63 ± 0.01	3.9 ± 0.3	0.85 ± 0.03	0.79 ± 0.03	0.86 ± 0.01
0.089	0.90 ± 0.04	0.80 ± 0.03	0.60 ± 0.01	0.78 ± 0.01	5.3 ± 0.5	1.04 ± 0.04	0.78 ± 0.03	1.27 ± 0.03
0.178	1.4 ± 0.1	1.15 ± 0.05	0.80 ± 0.04	1.02 ± 0.02	8 ± 1	1.9 ± 0.2	1.05 ± 0.06	2.29 ± 0.08
0.267	5.4 ± 0.7	2.2 ± 0.1	1.4 ± 0.1	2.0 ± 0.1	22 ± 4	3.1 ± 0.2	2.2 ± 0.2	4.3 ± 0.2
0.356	19 ± 4	10 ± 2	4.3 ± 0.4	5.5 ± 0.4	95 ± 36	12 ± 2	6.1 ± 0.6	9.5 ± 0.6
0.444	78 ± 27	28 ± 6	14 ± 2	13 ± 2		30 ± 7	15 ± 3	28 ± 7
0.533		48 ± 10	20 ± 4	20 ± 3			11 ± 2	20 ± 3
0.622			21 ± 5	39 ± 5			8 ± 1	23 ± 3
0.711				53 ± 7			6 ± 1	22 ± 3

[a] For each group of four columns the first column refers to the pure pivot algorithm, the second column is for a mixture of pivot and local moves, the third column refers to the multiple Markov chain algorithm, and the fourth column refers to the umbrella algorithm. In the multiple Markov chain run, we used nine parallel Markov chains at the nine β values shown. For the umbrella sampling a much larger number (about 100) of contributing Boltzmann factors were used.

Table II. As in Table I, but for Walks of Length $n = 400$

	$\tau(\langle S^2 \rangle)$				$\tau(\langle c \rangle)$			
β	Pivot	Pivot + local	MMC	Umbrella	Pivot	Pivot + local	MMC	Umbrella
0.000	0.96 ± 0.06	0.80 ± 0.03	0.69 ± 0.03	0.72 ± 0.01	9 ± 1	1.3 ± 0.1	1.1 ± 0.1	1.04 ± 0.02
0.089	1.06 ± 0.05	0.94 ± 0.04	0.68 ± 0.02	0.89 ± 0.02	11 ± 1	1.6 ± 0.1	0.99 ± 0.04	2.4 ± 0.1
0.178	1.8 ± 0.1	1.6 ± 0.1	1.03 ± 0.05	1.18 ± 0.04	16 ± 2	2.9 ± 0.2	1.6 ± 0.1	4.9 ± 0.3
0.267	9 ± 2	6.2 ± 0.8	2.9 ± 0.3	3.3 ± 0.2	58 ± 16	8 ± 1	5.3 ± 0.6	10 ± 1
0.356		37 ± 9	34 ± 9	48 ± 12		38 ± 9	38 ± 10	92 ± 22
0.389		66 ± 20	24 ± 5	64 ± 17		68 ± 20	25 ± 5	89 ± 23
0.444		65 ± 16	25 ± 6	151 ± 46		53 ± 17	12 ± 2	73 ± 16
0.474			17 ± 3	168 ± 53			9 ± 2	
0.503			18 ± 4				9 ± 2	

3.2. Implementation of Multiple Markov Chain Sampling

The main problem to be addressed in the implementation of the multiple Markov chain algorithm concerns the determination of the number of elementary Markov chains (i.e., the number of β values) which define the composite Markov chain and the way in which these elementary chains should be distributed in the range of temperatures which are of interest. To be more precise, let $[\beta_{\min}, \beta_{\max}]$ be the interval of inverse temperatures over which we are interested in sampling. In addition, suppose that the swapping procedure is allowed only between Markov chains that are nearest neighbors in β. If $\{\beta_l\}_{l=1,m}$ is the set of β values in the interval $[\beta_{\min}, \beta_{\max}]$, the goal is to find a value of m and a set of $\delta\beta(l) = \beta_{l+1} - \beta_l$ such that there is a reasonable degree of overlap between the distributions $\pi_k(\beta_l)$ and $\pi_k(\beta_{l+1})$. In our specific case, since we are interested in locating the β value corresponding to the Θ transition, we need to have good sampling in the region close to the Θ point. Moreover, for the problems studied in this paper, it is convenient to consider a β range which depends on the length of the walk being studied, since the width of the heat capacity peak decreases as n increases.

We must also address the question of how frequently swapping should be attempted, bearing in mind that our goal is to have a high degree of mixing between the chains at the different β values.

A partial answer to these questions comes from the observation that the scheme we have described has two built-in diagnostics. Suppose we decide to attempt only to swap Markov chains at adjacent β values. If pairs of adjacent values are too far apart, very few proposed swaps will be accepted, and the realization of the Markov chain will have a low acceptance fraction, which will diagnose this problem. The second kind of diagnostic involves keeping track of where the swaps take the simulated realizations. Imagine that at the beginning of the run we attach a different color to the configuration associated with each elementary Markov chain and that the color stays with the configuration when it is swapped to a different β value. Ergodicity of the composite Markov chain implies that over a long run each color will spend an equal fraction of time at each β value. The sample can be considered well mixed when these fractions are nearly equal and when swapping is frequently successful.

By performing several test runs we found that reasonable results could be obtained by attempting a swap every five attempted pivots. The degree of uniformity in β reached with this choice can be seen in Figs. 2 and 3, in which we report the relative amounts of time spent by each "color" at each β value for walks of two different lengths. The degree of uniformity is quite good for $n = 100$ and reasonable for $n = 600$. The data for $n = 100$ are

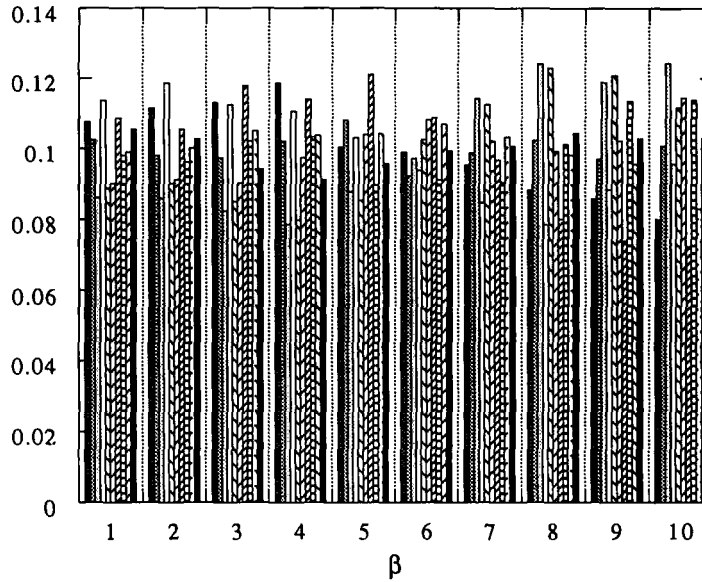

Fig. 2. Histogram showing, for a self-avoiding walk of $n = 100$ steps, the amount of time each configuration spent at each different β value. Different configurations are indicated by different symbols. The length of the run considered is 10^6 attempted pivots and the number of β values involved is 10.

obtained from a run of length 10^6 attempted pivots, while the data for $n = 600$ come from a run four times longer. Of course, in order to achieve reasonably uniform sampling, the length of run needed will increase as n increases.

In Tables I and II we report integrated autocorrelation times of different observables obtained using the multiple Markov chain algorithm for the same values of n considered previously. For all the observables, the value of τ increases as β increases up to a value around the critical point. For global observables such as the radius of gyration, this increase continues into the collapsed phase. However, for local observables such as the number of contacts, τ decreases as β is increased further. We note that $\tau(\langle c \rangle)$ is bounded below by a function which is proportional to the heat capacity[32] and, since the heat capacity decreases at larger β, the decrease which we observe in $\tau(\langle c \rangle)$ is consistent with this bound.

It is clear that swapping leads to a significant decrease in the integrated autocorrelation times for the larger values of β. This in turn leads to significantly smaller error bars in the resulting estimates, at very little cost in computer time.

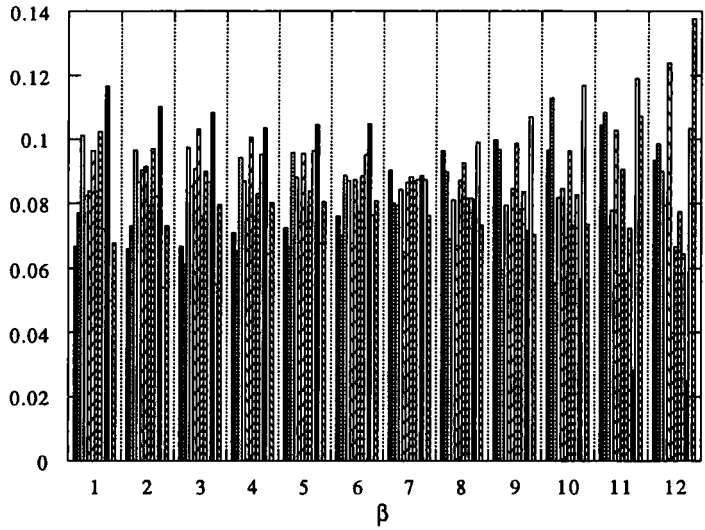

Fig. 3. As in Fig. 2, for a self-avoiding walk of $n = 600$ steps. The length of the run considered is 4×10^6 attempted pivots and the number of β values involved is 12.

The multiple Markov chain method gives estimates at a discrete set of β values, but these data can be reweighted to obtain estimates at intermediate values of β. However, as we shall see in the section which deals with the numerical results, the correlation between data sampled at different values of the temperature must be taken into account in the data analysis.

3.3. Implementation of Umbrella Sampling

The primary problem with umbrella sampling is the construction of a suitable umbrella distribution. Our approach has been to write this distribution as a linear combination of Boltzmann distributions, where the weights are given by (2.10). To use this method it is essential to have a good initial approximation to the free energy of the system of interest. One possibility, which has proven useful in various problems in the theory of liquids,[26, 29] is a bootstrap method. If we know the free energy for a walk of n steps, we can use this as a first approximation to the free energy for a walk of n' steps, with $n' > n$. Then we can use umbrella sampling to form a more accurate estimate of the free energy for the walk of n' steps and proceed until the range of n values of interest has been investigated. This method turned out to be completely inadequate for our problem. For instance, we exactly enumerated walks with up to 19 steps and from these

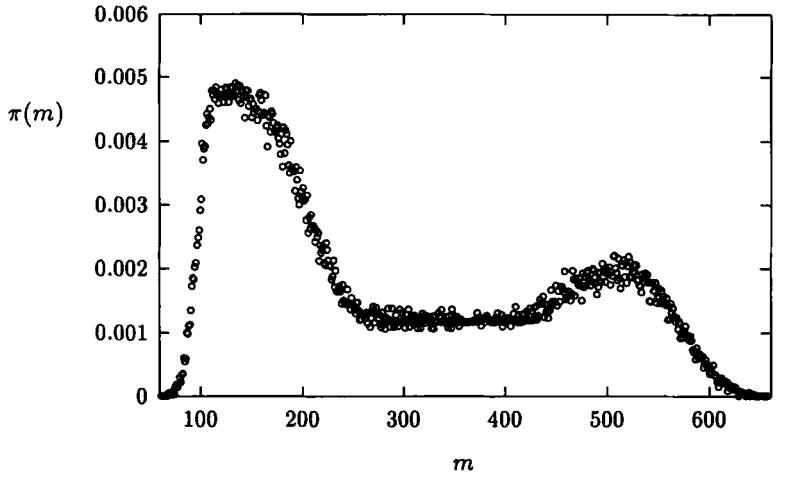

Fig. 4. Umbrella distribution for a self-avoiding walk of $n = 500$ steps.

data computed the exact values of the free energy for various values of β. We tried using the exact values of the free energy for $n = 19$ and also various extrapolations of these values to approximate the free energies for $n = 25$, but with very little success. It seems clear that more accurate estimates are needed to form a suitable umbrella distribution.

The approach which we used to form these preliminary estimates of the free energies was as follows. We carried out preliminary runs using the multiple Markov chain method to estimate the *energy* on a discrete set of values of β, fitted these data to a polynomial of appropriate order in β, and integrated this fitted energy to obtain the (relative) free energy $F_n(\beta) - F_n(0)$. This procedure turned out to be successful over a wide range of β and n values.

With these estimates of the free energies we can use an umbrella distribution constructed as a linear combination of Boltzmann distributions, with the weights given by (2.10). The resulting distribution should cover a suitable range of β values and we show the umbrella distribution which we obtained for $n = 500$ in Fig. 4. We found that such umbrella distributions gave quite satisfactory sampling over a wide range of β values, even though the distribution is not "flat" over the complete range of interest.

4. NUMERICAL RESULTS

In this section we report data on the β dependence of the heat capacity and, in particular, on how the location of the heat capacity peak depends on n. We use these results to estimate the location of the Θ point in the

$n \to \infty$ limit. We also estimate the location of the Θ point using metric quantities by the flatness criterion[20, 22] and this estimate turns out to be in very good agreement with the one obtained from the thermal quantities. Finally we estimate the value of the exponent ν at the Θ point.

4.1. Multiple Markov Chains

Each multiple Markov chain run gives estimates, at a discrete set of values of β, of quantities such as the mean number of contacts $\langle m \rangle$ and its variance $C = \langle m^2 \rangle - \langle m \rangle^2$, which are equal to the (reduced) energy and (reduced) heat capacity of the system. These estimates and their associated errors are obtained from averages over each of the elementary Markov chains at $\beta_1, \beta_2, \ldots, \beta_m$. In addition, one can obtain estimates at intermediate values of β by reweighting the data obtained from the elementary Markov chain at β_i, using (2.5) with $\pi_k = e^{\beta_i m(k)}/Z_n(\beta_i)$. Rewriting (2.5) as

$$\hat{Q}(\beta) = \frac{\sum_{t=1}^{N} X(t)}{\sum_{t=1}^{N} Y(t)} \tag{4.1}$$

we find that the variance of the ratio estimate $\hat{Q}$ is given approximately[33, 34] by

$$\mathrm{var}(\hat{Q}) = \left(\frac{x_N}{y_N}\right)^2 \left(\frac{\mathrm{var}(x_N)}{x_N^2} + \frac{\mathrm{var}(y_N)}{y_N^2} - 2\frac{\mathrm{cov}(x_N, y_N)}{x_N y_N}\right) \tag{4.2}$$

where $x_N = \sum_{t=1}^{N} X(t)$ and $y_N = \sum_{t=1}^{N} Y(t)$. Similarly, the estimates at β obtained from the reweighted data at each β_i can be combined by forming a linear combination with weights proportional to the inverses of their respective variances. In practice, only the data from the two or three β_i values closest to β contribute significantly and we have followed this route in forming our final estimates.

An alternative way to estimate the variance of a ratio is based on considering the time series

$$\{Z(t)\} = \left\{\frac{X(t)}{x_N} - \frac{Y(t)}{y_N}\right\}_{t=1,N} \frac{x_N}{y_N} \tag{4.3}$$

The variance of Z is precisely the same as the variance of $\hat{Q}$ as given in (4.2). For most calculations we constructed this time series and computed the variance in this way. The autocorrelation time can conveniently be computed from this time series (see appendix).

In Fig. 5 we show the β dependence of the heat capacity $C(\beta)$ for $n = 200$ and for $n = 600$. Both curves are smooth and show a peak, which sharpens as n increases.

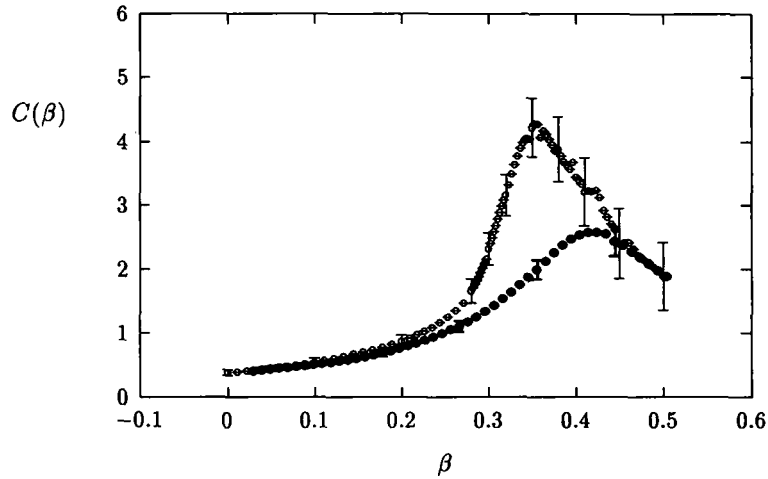

Fig. 5. Heat capacity for a self-avoiding walk of $n = 200$ and 600 steps obtained by reweighting the data obtained from a multiple Markov chain run.

By reweighting the data (as explained above), it is possible to obtain $C(\beta)$ at intermediate values of β, and hence it is possible to estimate the position of the maximum for each curve. The location of the maximum (in the large-n limit) is the Θ point. In Table III we give our estimates of the peak positions for values of n between 50 and 1600. In the next section we compare these results with those obtained by umbrella sampling.

Table III. Peak Positions of the Heat Capacity Estimated by Multiple Markov Chains and by Umbrella Sampling[a]

n	MMC	Umbrella
50	0.520 ± 0.030	0.532 ± 0.050
100	0.470 ± 0.020	0.468 ± 0.040
200	0.420 ± 0.020	0.414 ± 0.020
300	0.400 ± 0.020	0.397 ± 0.010
400	0.380 ± 0.010	0.383 ± 0.012
500	0.370 ± 0.010	0.375 ± 0.012
600	0.360 ± 0.010	0.366 ± 0.011
800	0.370 ± 0.015	0.350 ± 0.010
1200	0.340 ± 0.010	0.334 ± 0.010
1600	0.329 ± 0.010	

[a] The error bars are one standard deviation.

4.2. Umbrella Sampling

Umbrella sampling requires a reweighting of the data from the umbrella distribution to the Boltzmann distribution at each β value of interest. This is based on the identity (2.4) and uses the ratio estimate (2.5). Since we are using a ratio estimate, the estimates of the statistical uncertainties can be obtained by block average methods (see appendix) with variance given by (4.2).

The time series (4.3) contains information about the observable of interest at a given fixed β value, and from it it is also possible to estimate the autocorrelation times for the observables at that β value.[35] Such times are reported in Tables I and II for the same values of β and n previously considered. For both multiple Markov chains and umbrella sampling, the autocorrelation times increase as the transition is approached from small β. This reflects the large fluctuations in the energy around the transition.[32] At larger values of β the behavior is different for different variables. For small values of β the autocorrelation times are similar for multiple Markov chains and for umbrella sampling. At larger values of β the autocorrelation times for umbrella sampling are larger. This is because multiple Markov chains have large jumps in β whenever a swap occurs (leading to smaller values of τ), but umbrella sampling leads to smaller moves, typically, in the β coordinate. Umbrella sampling gives estimates at all values of β, but this gives rise to bigger values of τ.

We have estimated the energy, heat capacity, and free energy at a range of values of β for values of n from $n = 19$ to $n = 1200$. Since we are interested in the behavior around the Θ point, we have constructed umbrella distributions which are designed to sample in the range $0 \leqslant \beta \leqslant \beta_{\max}$ where we have chosen smaller values of $\beta_{\max}$ at larger values of n. In Fig. 6 we show the dependence of $C(\beta)$ on β for $n = 200$, 600, and 1200.

The agreement between the estimates obtained from umbrella sampling and from the multiple Markov chain method is usually excellent. In particular, the values at $\beta_1, \beta_2, \ldots, \beta_m$, corresponding to the elementary Markov chains in the multiple Markov chain approach, agree essentially exactly with the values from the umbrella sampling. Between successive values there are some minor disagreements resulting from difficulties with the reweighting of the multiple Markov chain data. The estimates of the locations of the peak maxima are shown in Table III and agree well with the estimates from multiple Markov chain sampling.

In Fig. 7 we show the locations of the peak maxima plotted against $1/\sqrt{n}$. We have exactly enumerated walks with $n \leqslant 19$ and used the heat capacities calculated from these results for these values of n and the values

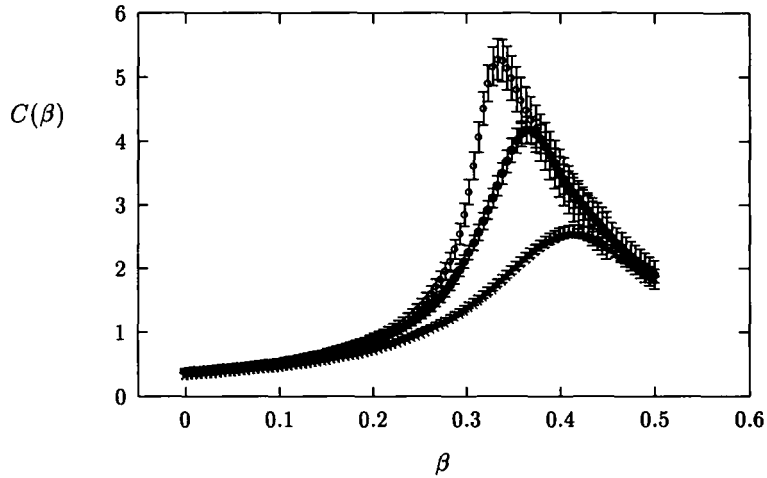

Fig. 6. Heat capacity for a self-avoiding walk of $n = 200$, 600, and 1200 steps from umbrella sampling.

obtained from umbrella sampling for n ranging from 28 to 1200. The value corresponding to $n = 1600$ has been obtained from a multiple Markov chain run.

The graph shows considerable curvature at small values of n, followed by a linear regime. We have computed least square fits to the data, with weights proportional to the reciprocals of the variances, including various

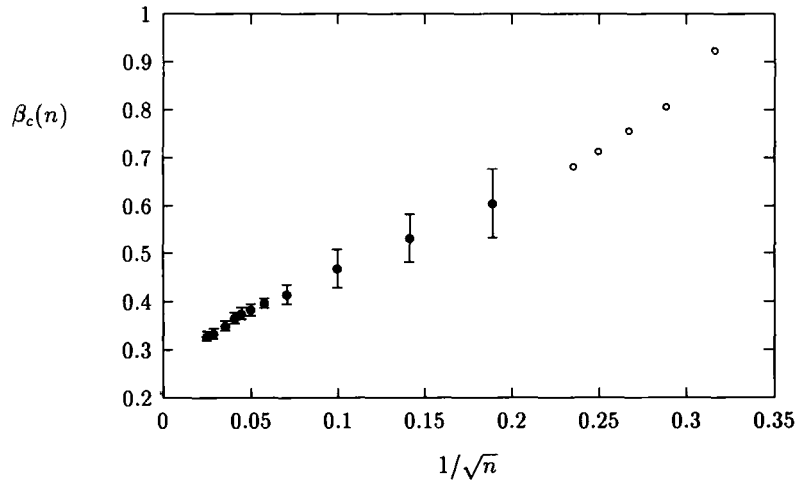

Fig. 7. Positions of the maxima of the heat capacity curves, using data from exact enumeration, umbrella sampling, and multiple Markov chains.

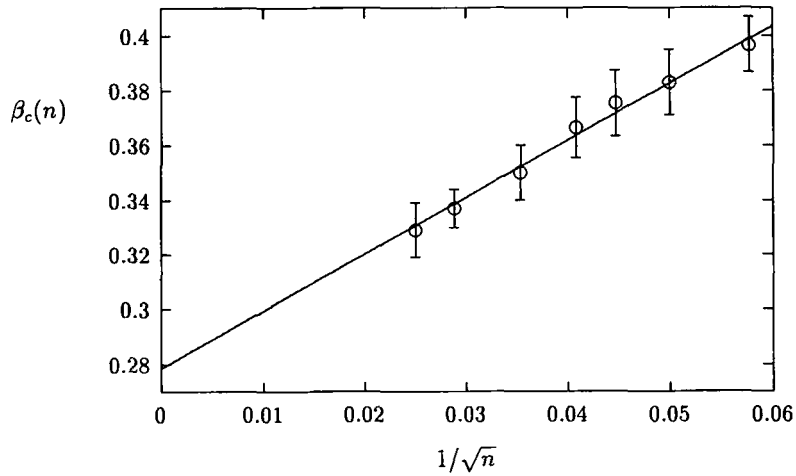

Fig. 8. The locations of the maxima of the heat capacity curves for $n \geqslant 300$ and the corresponding least squares fit.

numbers of points. For $n = 1200$ we combined the estimates from multiple Markov chains and umbrella sampling. The fit including the last seven points ($300 \leqslant n \leqslant 1600$) gives $\beta_\Theta = 0.276 \pm 0.004$ and is shown in Fig. 8. Considering the additional errors due to residual curvature (estimated by including one or two additional points), we give $\beta_\Theta = 0.276 \pm 0.006$ as our final estimate. (We note that the large-n data are essential in arriving at this estimate. For instance, using the data only for $n = 28$, 50, 100, 200, and 300 gives an estimate of 0.303 ± 0.003.)

To determine the location of the theta temperature it is also possible to use geometric quantities, such as the mean radius of gyration $\langle S_n^2 \rangle$ or the mean end-to-end distance $\langle R_n^2 \rangle$. The generalized scaling behavior for these quantities for self-avoiding walks at tricriticality in given by[8, 9]

$$\langle Q_n^2 \rangle^{1/2} \sim n^{\nu_\Theta} f_\pm(n^{\phi_\Theta} \tau) \tag{4.4}$$

where ϕ_Θ is a crossover exponent, $\tau = |T - \Theta|/\Theta$, and $\langle Q_n^2 \rangle$ can be the mean squared radius of gyration or the mean squared end-to-end distance. Let $x = n^{\phi_\Theta} \tau$. For small τ and large x (corresponding to large n values), $f_\pm(x)$ must have the form

$$f_\pm \sim x^\mu \begin{cases} \mu = \mu^+ = (\nu_{\mathrm{SAW}} - \nu_\Theta)/\phi_\Theta & \text{for } T > \Theta \\ \mu = 0 & \text{for } T = \Theta \\ \mu = \mu^- = (\nu_{\mathrm{comp}} - \nu_\Theta)/\phi_\Theta & \text{for } T < \Theta \end{cases} \tag{4.5}$$

The behavior reported above is expected to be valid in the large-n regime, with $\nu_{SAW} \simeq 0.588$[36] and $\nu_{comp} = 1/d = 1/3$. For smaller values of n, at values of T slightly above Θ the self-avoiding walk will expand with an effective value of $\nu < \nu_{SAW}$, which is expected to approach ν_{SAW} monotonically from below as n increases. On the other hand, for $T < \Theta$, the self-avoiding walk grows with an effective exponent $\nu > 1/3$, which approaches the limiting behavior monotonically from above. Thus at $T = \Theta$ the effective exponent ν is expected to depend only weakly on n, and this gives a useful method for locating the Θ point. A plot of $\langle S_n^2 \rangle / n$ against $\log n$ should show little n dependence for data obtained at the Θ point.

This criterion has been successfully used in several previous studies.[20, 22] However, it is important to remember that in $d = 3$, $f_{\pm}$ is also affected by logarithmic corrections proportional to $1/\ln n$,[10] which may contribute significantly for small n values. For this reason we used values of $n > 200$ in applying the flatness criterion.

As can be seen from Figs. 9 and 10, in which $\langle S_n^2 \rangle / n$ and $\langle R_n^2 \rangle / n$ are plotted against $\ln n$, these ratios are essentially independent of n for $\beta = 0.275 \pm 0.008$. In forming this estimate we have relied primarily on the radius-of-gyration data, but have included the less accurate end-to-end length data in determining the error bar. This estimate is in very good agreement with the location of the Θ point obtained using thermal quantities.

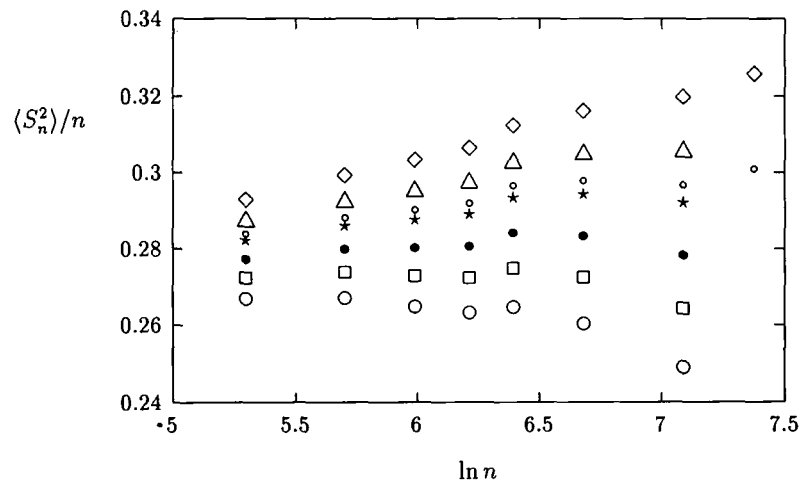

Fig. 9. Plot of $\langle S_n^2 \rangle / n$ versus $\ln n$ for different values of β: 0.262 (◇), 0.268 (△), 0.271 (◦), 0.273 (⋆), 0.278 (•), 0.283 (□), and 0.288 (○). The n values considered are $n = 200$, 300, 400, 500, 600, 800, 1200.

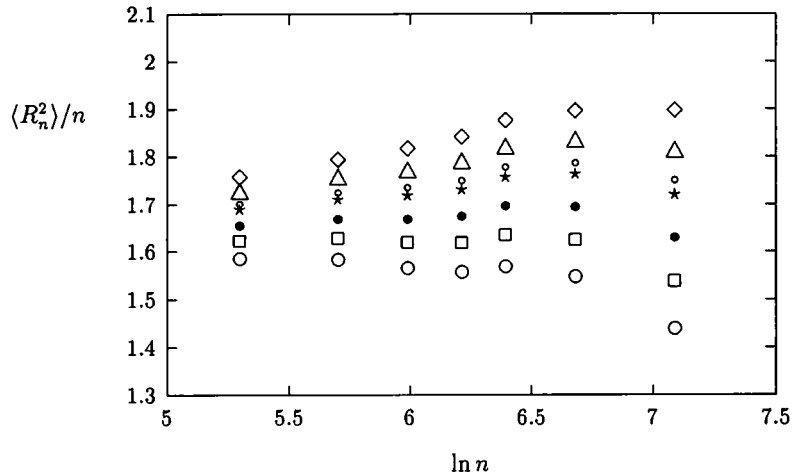

Fig. 10. Plot of $\langle R_n^2\rangle/n$ versus $\ln n$ for different values of β: 0.262 ($\diamond$), 0.268 ($\triangle$), 0.271 ($\circ$), 0.273 ($*$), 0.278 ($\bullet$), 0.283 ($\square$), and 2.88 ($\bigcirc$). The n values considered are $n = 200$, 300, 400, 500, 600, 800, 1200.

At this value of β we estimate ν_Θ from both the mean square radius of gyration $\langle S_n^2\rangle$ and the end-to-end distance $\langle R_n^2\rangle$, from log–log plots, as shown in Fig. 11. The estimates have been obtained considering data from runs with $n \geqslant 100$, by fitting $\ln(\langle S_n^2\rangle)$ and $\ln(\langle R_n^2\rangle)$ as functions of $x = \ln n$. We have used a function of the form $h(x) = A + Bx - C/x$, where A and B ($=2\nu$) are the parameters to be determined, and $C \simeq 0.102$[10]

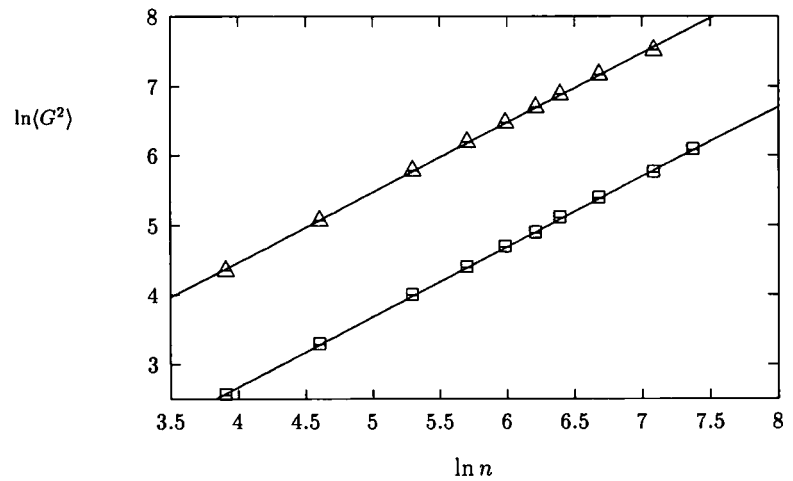

Fig. 11. Plots of $\ln(\langle S_n^2\rangle)$ and $\ln(\langle R_n^2\rangle)$ against $\ln n$ to obtain $2\nu_\Theta$.

takes into account the logarithmic corrections showing up at the Θ point. The values obtained are $2\nu_\Theta[\langle S_n^2\rangle] = 1.004 \pm 0.006$, $2\nu_\Theta[\langle R_n^2\rangle] = 0.99 \pm 0.01$. We estimate the amplitude ratio $\langle R_n^2\rangle/\langle S_n^2\rangle$ at the Θ point to be 6.75 ± 0.5.

5. SUMMARY AND DISCUSSION

We have described two methods for handling quasiergodic problems in Markov chain sampling Monte Carlo methods, and have reported our experience with these methods for the interacting self-avoiding walk model on the simple cubic lattice. We have shown that the performance of the pivot algorithm can be improved, when interactions are included in the model, by incorporating some local moves which are effective in improving the autocorrelation times of the slow modes in the problem. However, local moves (combined with pivots) are not sufficient to give good sampling in the region where the interactions are very strong, and we have shown that both multiple Markov chain methods and umbrella sampling can lead to dramatic improvements in the sampling efficiency. We have been able to estimate both thermodynamic and metric quantities for values of n up to 1600 for values of β up to 0.45, which is well inside the collapsed regime. For $n \leqslant 400$, properties were estimated successfully up to $\beta = 0.8$.

As a test of these methods, we have estimated the positions of the maxima in the heat capacities at different values of n and have extrapolated these to form estimates of the Θ point. We have made corresponding estimates using metric data and a flatness criterion. Our estimates of the location of the Θ point and of the exponent $\nu(\Theta)$ are similar to those obtained by others[22, 27] and agree within the combined error bars. In a future publication we shall report results extending to larger values of n and β, and comparing the behavior of walks and polygons.

The methods described here should be applicable to a wide range of problems with strong interactions, where standard Metropolis sampling methods fail because of slow convergence of the Markov chain. In addition, they can be used to give information not only at the Θ point, but over a wide range of values of the interaction parameter.

APPENDIX. ERROR ESTIMATES AND AUTOCORRELATION TIMES

Effective error estimates are essential in assessing the efficiency of a Markov chain Monte Carlo algorithm and in particular the way in which it handles quasiergodic problems. A realization of a Markov chain will produce a series of random states $S(1)$, $S(2)$,..., $S(N)$ that are usually highly correlated, and this correlation must be taken into account in estimating

the variances of estimates obtained by sampling along such realizations. These correlations will usually be larger when the Markov chain is less mobile, so a good assessment of the errors can give information about the efficiency of the Markov chain sampling procedure.

To be more precise, let the observed states of the Markov chain be represented by $S(t)$ and let Q be an arbitrary observable over the set of n-step self-avoiding walks. If the distribution of the initial state of the Markov chain is the unique limit distribution, then $\{Q_t\} = \{Q(S(t))\}$ is a stationary stochastic process with average

$$\langle Q_t \rangle = \sum_{k \in S} \pi_k Q(k) \tag{A.1}$$

A natural estimator is the sample mean (2.3) and, if the $S(t)$ are independent, the estimate of the corresponding variance is given by the usual formula

$$\mathrm{Var}(\bar{Q}) = \frac{1}{N(N-1)} \sum_{t=0}^{N-1} (Q_t - \bar{Q})^2 \tag{A.2}$$

When the random states $S(t)$ are not independent, $\bar{Q}$, given by (2.3), is still an unbiased estimate of $\langle Q_t \rangle$, provided that the initial states of the realization are discarded, to allow the system to relax to the limit distribution. However, the variance of this estimate is not given by Eq. (A.2), but is typically larger. We considered two different methods for estimating the variance of the observable Q.

A1. Batching or Block Average Method

The block average method is probably the simplest method of obtaining a reliable estimate of the variance of $\bar{Q}$. The N sample points ar divided into r blocks, each consisting of $m = N/r$ consecutive sample points. Estimates of $\langle Q_t \rangle$ are found using the data in each block, with the estimate from block i being

$$\bar{Q}_i = \frac{1}{m} \sum_{1+(i-1)m}^{im} Q_t \tag{A.3}$$

The estimate of $\langle Q_t \rangle$ based on the entire data set [Eq. (2.3)] can be expressed as the average of the block estimates:

$$\bar{Q} = \frac{1}{r} \sum_{i=1}^{r} \bar{Q}_i \tag{A.4}$$

If the block estimates $\bar{Q}_i$ are indeed independent, then the variance of $\bar{Q}$ can be estimated as in (A.2):

$$\mathrm{Var}(\bar{Q}) = \frac{1}{r(r-1)} \sum_{i=1}^{r} (\bar{Q}_i - \bar{Q})^2 \tag{A.5}$$

It is clear that the quality of this estimate depends crucially on the choice of the block size m. If m is too small, so that the $\bar{Q}_i$ are significantly correlated, then the variance estimate is likely to be too small, leading to an underestimate of the errors associated with the estimate of $\langle Q_t \rangle$. On the other hand, if m is too large, the smaller value for r that results leads to a larger variance in the estimate of the variance itself. For small r this will cause significant doubts concerning whether $\bar{Q}$ is accurate even if the estimated variance from (A.5) is small.

One way to cope with this problem of choosing an appropriate block size is to plot the estimated variance of $\bar{Q}$ as a function of the block size m. If there is were no noise in the variance estimates, these estimates would approach the true value in the limit as m increased. From the plot, one may be able to pick out an approximation to this limiting value despite the noise.

A2. Autocorrelation Function

Another way to estimate the accurancy of $\bar{Q}$ is to use methods from time series analysis. The idea is to introduce an unnormalized autocorrelation function of the stationary stochastic process $\{Q_t\} \equiv \{Q(S(t))\}$

$$C_{QQ}(s) = \langle Q_t Q_{t+s} \rangle - \langle Q_t \rangle^2 \tag{A.6}$$

The normalized autocorrelation function is defined by $\rho_{QQ}(t) = C_{QQ}(t)/C_{QQ}(0)$. Once the Markov chain is in equilibrium, then the *integrated autocorrelation time* is given by

$$\tau_{\mathrm{int}}(Q) = \tfrac{1}{2} \sum_{-\infty}^{\infty} \rho_{QQ}(t) \tag{A.7}$$

The integrated autocorrelation time controls the statistical error in MC estimates of the average $\langle Q \rangle$ of the observable Q. Indeed, the variance in the sample mean $\bar{Q}$ is given asymptotically by

$$\mathrm{Var}(Q) \sim \frac{1}{N} (2\tau_{\mathrm{int}}(Q))\, C_{QQ}(0) \tag{A.8}$$

In other words, the effective number of independent observations is $N/(2\tau_{\text{int}})$.[25] In practice, we calculated the integrated autocorrelation time by using a procedure suggested by Madras and Sokal.[25] A "window" $\lambda(t)$, which has $\lambda(t)=1$ if $|t|<M$ and zero otherwise, is chosen, and Eq. (A.7) is approximated by

$$\bar{\tau}_{\text{int}}(Q)=\frac{1}{2}\sum_{t=-M}^{M}\left[C_{QQ}^{-1}(0)\frac{1}{M-|t|}\sum_{i=1}^{M-|t|}(Q_i-\bar{Q})(Q_{i+|t|}-\bar{Q})\right] \quad \text{(A.9)}$$

where M, the size of the window, is determined by the "automatic windowing algorithm." For details and improvements of this algorithm, see refs. 25 and 36.

ACKNOWLEDGMENTS

We are pleased to acknowledge financial support from the NSERC of Canada. We would like to thank Neal Madras, Alan Sokal, and John Valleau for many pleasant and fruitful discussions.

REFERENCES

1. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**:187 (1977).
2. C. J. Geyer and E. A. Thompson, Preprint, University of Minnesota (1994).
3. S. T. Sun, I. Nishio, G. Swislow, and T. Tanaka, *J. Chem. Phys.* **73**:5971 (1980).
4. I. H. Park, J. H. Kim, and T. Chang, *Macromolecules* **25**:7300 (1992).
5. S. F. Sun, *J. Chem. Phys.* **93**:7508 (1990).
6. B. Nienhuis, *J. Stat. Phys.* **34**:731 (1984).
7. A. Coniglio, N. Jan, I. Majid, and H. E. Stanley, *Phys. Rev. B* **35**:3617 (1987); B. Duplantier and H. Saleur, *Phys. Rev. Lett.* **59**:539 (1987); F. Seno and A. L. Stella, *J. Phys.* (Paris) **49**:739 (1988).
8. P. G. de Gennes, *J. Phys. Lett.* (Paris) **36**:L55 (1975).
9. P. G. de Gennes, *J. Phys. Lett.* (Paris) **39**:L299 (1978).
10. B. Duplantier, *Europhys. Lett.* **1**:491 (1986); *J. Chem. Phys.* **86**:4233 (1987).
11. A. L. Kholodenko and K. F. Freed, *J. Chem. Phys.* **80**:900 (1984); *J. Phys. A* **17**:L191 (1984).
12. A. Maritan, F. Seno, and A. L. Stella, *Physica A* **156**:679 (1989).
13. H. Saleur, *J. Stat. Phys.* **45**:419 (1986).
14. R. Finsy, M. Janssens, and A. Bellemans, *J. Phys. A* **8**:L106 (1975).
15. D. C. Rapaport, *J. Phys. A* **9**:1521 (1976).
16. T. Ishinabe, *J. Phys. A* **20**:6435 (1985).
17. V. Privman, *J. Phys. A* **19**:3287 (1987).
18. V. Privman and D. A. Kurtze, *Macromolecules* **19**:2377 (1986).
19. J. Mazur and F. L. McCrackin, *J. Chem. Phys.* **49**:648 (1968).
20. K. Kremer, A. Baumgartner, and K. Binder, *J. Phys. A* **15**:2879 (1981).
21. I. Webman, J. L. Lebowitz, and M. H. Kalos, *Macromolecules* **14**:1495 (1981).
22. H. Meirovitch and H. A. Lim, *J. Chem. Phys.* **92**:5144 (1990).

23. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Methuen, London, 1964).
24. M. Lal, *Mol. Phys.* **17**:57 (1969).
25. N. Madras and A. D. Sokal, *J. Stat. Phys.* **56**:109 (1988).
26. J. P. Valleau, In *Proceedings of the International Symposium on Ludwig Boltzmann*, G. Battimelli, M. G. Ianniello, and O. Kresten, eds. (1993).
27. B. A. Berg and T. Neuhaus, *Phys. Rev. Lett.* **68**:9 (1992).
28. N. Madras and M. Piccioni, *Importance sampling for families of distributions*, Bernoulli (1994), submitted.
29. J. P. Valleau, *J. Chem. Phys.* **99**:4718 (1993).
30. N. Madras and A. D. Sokal, *J. Stat. Phys.* **47**:573 (1987).
31. A. D. Sokal, *Monte Carlo methods for the self-avoiding walk*, preprint (1994).
32. S. Caracciolo, A. Pelissetto, and A. D. Sokal, *Phys. Rev. Lett.* **72**:179 (1994).
33. F. Yates, *Sampling Methods for Censuses and Surveys* (Griffin, 1960).
34. W. E. Deming, *Some Theory of Sampling* (Dover, New York, 1966).
35. S. Caracciolo, A. Pelissetto, and A. D. Sokal, *Monte Carlo methods using reweighting: some warnings*, preprint (1994).
36. B. Li, N. Madras, and A. D. Sokal. *Critical exponents, hyperscaling and universal amplitude ratios for two- and three-dimensional self-avoiding walks*, preprint (1994).
37. P. Grassberger and R. Hegger, *Simulations of 3-dimensional Θ-Polymer*, preprint (1994).