A new simulation of branched polymers

1987 J. Phys. A: Math. Gen. 20 6059

(http://iopscience.iop.org/0305-4470/20/17/039)

J. Phys. A: Math. Gen. **20** (1987) 6059-6073. Printed in the UK

A new simulation of branched polymers

Hagai Meirovitch

Polymer Research Department, Weizmann Institute of Science, Rehovot 76100, Israel

Received 14 April 1987, in final form 17 June 1987

Abstract. The scanning method for the simulation of linear chains is extended to general models of branched polymers without loops. A branched chain grows in 'time' (namely, a number of steps from the origin). Therefore (i) in contrast to other simulation techniques, which are of a relaxation type, the chains are statistically independent and the statistical error can reliably be estimated, (ii) the probability of a chain is known and hence the entropy and (iii) the scanning construction enables one to study geometrical properties which depend on time. For self-avoiding trees on a square lattice we obtain the relatively accurate estimates for the static critical exponents, $\nu = 0.640 \pm 0.004$ and $\theta = 1.003 \pm 0.02$, and for the connective constant $\mu = 5.1419 \pm 0.003$. We also obtain critical exponents $\gamma_t \simeq 1.26$ and $\nu_t \simeq 0.82$, which characterise the growth in time of the number of bonds and the gyration radius respectively. Application of the scanning method to more complex branched polymers is discussed.

1. Introduction

Various types of branched polymers are known. 'Stars', for example, have one branching centre of relatively long linear chains. Proteins and nucleic acids, on the other hand, consist of a long backbone which is 'decorated' by small side chains. Also, randomly branched polymers can be formed by combining polyfunctional units (Flory 1953). Dilute branched polymers of that type in a good solvent have been modelled by 'lattice animals' which are the ensemble of equally probable configurations of N connected self-avoiding bonds on a lattice. Lattice animals have loops; when loops are forbidden the structures are called self-avoiding lattice trees (SAT) or branched polymers without loops (Gaunt *et al* 1982). Theoretical and numerical studies suggest that lattice animals with and without loops and with different extent of branching all belong to the same universality class and have the upper critical dimensionality eight (Zimm and Stockmayer 1949, Dobson and Gordon 1965, de Gennes 1968, Stauffer 1978, Lubensky and Isaacson 1979, Gaunt 1980, Gaunt *et al* 1980, 1982, Duarte and Ruskin 1981, Derrida and De Seze 1982, Margolina *et al* 1984, Privman 1984, Caracciolo and Glaus 1985). Parisi and Sourlas (1981) have found a connection between this universality class in dimension d and the Lee-Yang edge singularity of the Ising model in $d-2$; this enables one to calculate the entropy exponent θ for $d=2$ and 3 and the exponent ν for $d=3$. The known simulation techniques for lattice animals with and without loops are of a relaxation type. Some of them are based on changing the location of the bonds on the chain during the simulation (Stauffer 1978, Peters *et al* 1979, Gould and Holl 1981, Seitz and Klein 1981, Dickman and Schieve 1984). With another technique bonds are added to the chain or removed from it with a certain

0305-4470/87/176059+15$02.50

stochastic process (Caracciolo and Glaus 1985, Glaus 1985). In this work we develop a new simulation technique for branched chains without loops (i.e. self-avoiding lattice trees), which is independent of the methods mentioned above. It is based on the concepts of the scanning method for generating linear chains, suggested by the author some time ago (Meirovitch 1982, 1985a, c). With the scanning method extended here, a lattice tree is started from the origin and grows shell-by-shell with the help of transition probabilities. The first shell ($t=1$) consists of bonds connecting the origin to some (or to all) of its nearest-neighbour sites; these sites constitute 'growth tips' for the next step $t=2$. At $t=2$ another shell of bonds is determined, which connect these tips to their nearest-neighbour sites and the process continues. Thus, t can be considered as time and the shells are therefore isotime shells (see Alexandrowicz 1985). The scanning construction provides the probability of a tree and hence the entropy is known. Another advantage is that the sampled configurations are statistically independent, which enables one to estimate the statistical error by employing standard statistical methods. Moreover, the growth of a SAT in time enables one to study its geometrical properties. (Other step-by-step simulation procedures have also been employed, but for models of branched polymers which differ from the (equally probable) lattice trees studied here (Redner 1979, Alexandrowicz 1985).) For trees, where the branching is random, it proves convenient to perform the scanning construction in the framework of the grand canonical ensemble. However, for star polymers, proteins, etc, where the branching is fixed it would be more natural to use the canonical ensemble. Finally, the trees studied in this work are on a square lattice which enables one to test the Parisi and Sourlas (1981) prediction $\theta=1$ and to estimate the exponent ν. We also calculate critical exponents which characterise the growth in time of quantities such as the gyration radius.

2. The model

We study a model of self-avoiding bond trees (SAT) without loops on a square lattice. This model has been treated by Seitz and Klein (1981), Gaunt *et al* (1982), Family (1983) and Caracciolo and Glaus (1985). Let us denote by C_N the total number of SAT of N bonds, which start from the origin of the coordinate system. The Boltzmann probability of SAT i is therefore

$$P_i^{\mathrm{B}}(N)=1/C_N \tag{1}$$

and the entropy is

$$S_N=-k_{\mathrm{B}}\sum_{\mathrm{SAT}}P_i^{\mathrm{B}}(N)\ln P_i^{\mathrm{B}}(N)=k_{\mathrm{B}}\ln C_N \tag{2}$$

where k_{B} is the Boltzmann constant. For large N the number of configurations per site, a_N, scales like

$$a_N=C_N/(N+1)\simeq A\mu^N N^{-\theta} \tag{3}$$

where the prefactor A and the connective constant μ are lattice dependent while θ is a critical exponent which is expected to be universal. The square gyration radius $\langle G_N^2\rangle$ scales with a critical exponent ν

$$\langle G_N^2\rangle\simeq\sum_{\mathrm{SAT}}P_i^{\mathrm{B}}(N)G_i^2\simeq N^{2\nu} \tag{4}$$

where G_i is the gyration radius of SAT i. The grand partition function of the model is

$$\Xi(\beta) = 1 + \sum_{N=1} C_N \beta^N \tag{5}$$

where β is the bond activity.

The grand canonical probability of configuration i is

$$P_i^{BG} = \beta^N / \Xi(\beta) \tag{6}$$

and the average number of bonds in the grand ensemble is

$$\langle N \rangle^G = \sum_{N=1} N C_N \beta^N / \Xi. \tag{7}$$

From (3) and (5) one obtains

$$\Xi \simeq 1 + \sum_{N=1} N^{-\theta}(N+1)(\beta\mu)^N \tag{8}$$

which is expected to converge for $\beta < \beta_c = \mu^{-1}$, where β_c is the critical activity; for such values of β a very good approximation would be to carry out the summation up to a cutoff value t_N (rather than ∞), where $t_N \gg \langle N \rangle^G$.

3. Exact scanning procedure

Let us first describe an *exact* scanning procedure for SAT in the grand canonical ensemble. A SAT is started from the origin and constructed step-by-step, where at step t polyfunctional units of 0-4 bonds are added to the growth tips determined at step $t-1$ (figure 1). Altogether, there are sixteen configurations of these five units, which

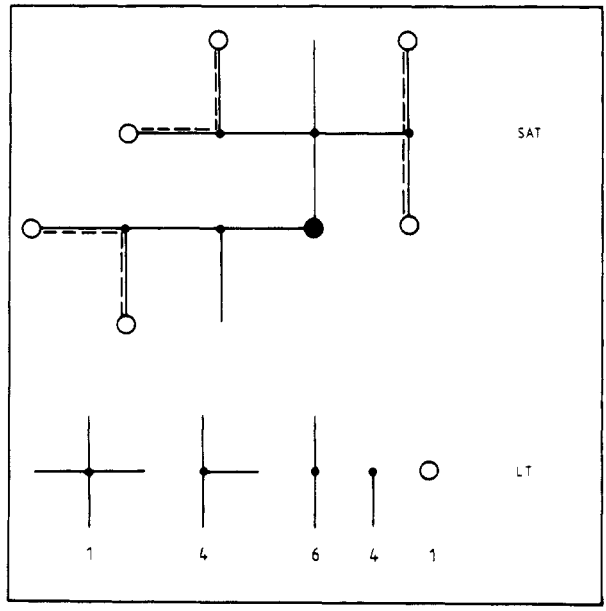

Figure 1. A partial SAT at step $t = 4$ of the scanning construction. Each of the $K_t = 6$ growth tips (circular) will be continued by adding a local tree (LT). The bonds of the shell configuration determined at $t = 3$ are denoted by broken lines. Tips which are not circled were terminated at $t < 3$ and will not be considered in future steps. Five LT and their degeneracies are presented in the lower part of the figure. The empty LT is denoted by a circle.

we call local trees (LT): one LT of four bonds (it can be added only at the first step), four LT of three bonds, six of two bonds, and four of one bond; the empty LT, which leads to the termination of a growth tip, is also considered (figure 1). For the first step ($t=1$) all sixteen LT has a non-zero transition probability and thus an LT of K_2 bonds is selected with a Monte Carlo lottery and placed at the origin. These bonds define K_2 growth tips for $t=2$ and the process continues; if the empty LT has been chosen at $t=1$ one obtains the 'vacuum' tree, which is added to the sample and the construction of a new tree is started. At step t a SAT of N_{t-1} bonds has already been generated at the $t-1$ previous steps and the K_t growth tips determined at step $t-1$ should be continued. Obviously, because of the excluded volume interaction between the LT and the bonds defining the growth tips, the maximum number of LT available for a tip is eight, rather than sixteen. Assume now that a set of K_t local trees is added to the K_t tips and that the excluded volume condition is not violated; we call such an arrangement of LT a shell configuration, denoting it by $B_k(t)$ and its number of bonds by $m_k(t)$ (figure 1). In general the number of shell configurations is smaller than 8^{K_t}. A shell configuration can be continued at the future steps in many ways. We call such a continuation a future SAT and denote its number of bonds by N_f. This enables one to define a future grand partition function for $B_k(t)$

$$\Xi_k(t,\beta)=\beta^{m_k(t)}\sum_{N_f=0} M_k(t,N_f)\beta^{N_f}=\beta^{m_k(t)}\Xi'_k(t,\beta) \tag{9}$$

where $M_k(t,N_f)$ is the number of future SAT of N_f bonds. The transition probability for $B_k(t)$ is thus

$$p_k(t,\beta)=\Xi_k(t,\beta)\left(\sum_k \Xi_k(t,\beta)\right)^{-1}. \tag{10}$$

Notice that $p_k(t,\beta)$ is a conditional probability which depends on the partial SAT generated at the $t-1$ past steps and on $B_k(t)$. A shell configuration $B_k(t)$ is selected with a Monte Carlo lottery according to the p_k and the process continues until the chain has been terminated at step $\bar{t}$ (because the empty LT has been chosen for each one of the tips). The SAT i of N_i bonds, thus constructed, has the probability P_i^{G}, which is the product of the $\bar{t}$ transition probabilities with which the shell configurations have been chosen:

$$P_i^{\mathrm{G}}=\prod_{t=1}^{\bar{t}} p_k^*(t,\beta)=\frac{\beta^{m_k(1)}\Xi'_{k(1)}}{\Xi}\,\frac{\beta^{m_k(2)}\Xi'_{k(2)}}{\Xi'_{k(1)}}\cdots\frac{1}{\Xi'_{k(\bar{t}-1)}}$$

$$\beta^{N_i}/\Xi=P_i^{\mathrm{BG}}. \tag{11}$$

Here $p_k^*(t,\beta)$ is the transition probability of the shell configuration selected at step t and for simplicity we have abbreviated the notation for the corresponding future grand partition functions (9) to read $\Xi'_{k(t)}$. The second and the third equalities, which are based on (9) and (10), prove that the construction is exact, i.e. the configurations are selected with the ensemble probability P_i^{BG} (6). It should be pointed out that a growth tip is considered only once during the process at a certain step t, where t is equal to the number of bonds on the chain (but not necessarily on the lattice) connecting the tip to the origin. Therefore t can be considered as time. However, this exact procedure is not feasible since already for the first step one has to calculate the grand partition function Ξ (5) for the whole system.

4. Approximate scanning procedures

One can approximate the exact scanning procedure, described in the previous section, in many ways. We impose the following simplifications. (i) Not the whole future is taken into account, but only a small number b of future steps. (ii) This limited scanning is carried out for each tip separately, i.e. under the assumption that LT have not yet been added to the other tips. (iii) An adjustable parameter c replaces the activity β. The future grand partition function for tip l ($1 \leq l \leq K_t$) is thus (see (9))

$$\Xi^l_{k'}(t, c, b) = c^{m_{k'}(t)} \sum_{N_f} M_{k'}(N_f, t, b, l) c^{N_f} \tag{12}$$

where k' denotes an LT, which can be added to tip l without violating the excluded volume conditions; $m_{k'}(t)$ is the number of bonds of local tree k' ($m_{k'}(t) = 0$–3 for $t > 1$). $M_{k'}$ is the number of future SAT of N_f bonds, which can be added to k' in $b-1$ future time steps. At this stage two remarks should be made. (i) For the empty LT ($k' = 1$), which satisfies $\Xi^l_1 = 1$, we introduce another parameter e such that $\Xi^l_1 = e$. (ii) In our calculations we also define LT, which are based on two nearest-neighbour tips, employing $b = 1$; obviously, the maximal number of LT of such a pair is sixteen (of 0-4 bonds). However, for the sake of simplicity we omit the parameter e from the equations and only describe the process in which each tip is treated separately. The transition probability for the local tree k' of tip l is

$$p_{k'}(t, c, b, l) = \Xi^l_{k'}(t, c, b) \left(\sum_{k'} \Xi^l_{k'}(t, c, b) \right)^{-1}. \tag{13}$$

For each tip $l (1 \leq l \leq K_t)$ a set of $p_{k'}$ is calculated and an LT is selected by a lottery. These LT are then added to the tips and the excluded volume condition is checked; if it is violated the tree is discarded and a new tree is started. In the other case the set of LT is accepted and the process continues. The probability of the shell configuration $B^*_k(t)$ thus chosen is (see (10))

$$p^*_k(t, c, b) = \prod_{l=1}^{K_t} p^*_k(t, c, b, l) \tag{14}$$

and the probability of SAT i is (see (11))

$$P^G_i(c, b) = \prod_{t=1}^{\tilde{t}} p^*_k(t, c, b). \tag{15}$$

Obviously $P^G_i(c, b)$ is approximate (biased) and is not normalised over the ensemble of SAT but over a larger ensemble which also includes self-intersecting trees. However, one can define a normalised probability $\bar{P}^G_i(c, b)$ over the ensemble of SAT by

$$\bar{P}^G_i(c, b) = P^G_i(c, b)/A \tag{16}$$

where

$$A = \sum_{i \in \mathrm{SAT}} P^G_i(c, b) < 1. \tag{17}$$

A can easily be estimated from the attrition ratio

$$A \sim n_{\mathrm{succ}}/n_{\mathrm{start}} \tag{18}$$

where n_{start} and n_{succ} are the number of chains started and succeeded (i.e. without violating the excluded volume condition) respectively. In order to determine the

optimal value of c one defines a free energy functional $J(\beta, c, b)$ for the grand canonical ensemble, based on the approximate probability $\bar{P}_i^G$ (16)

$$J(\beta, c, b)/k_B T = \sum_{\text{SAT}} \bar{P}_i^G(c, b)(\ln \bar{P}_i^G(c, b) - N_i \ln \beta). \tag{19}$$

$J(\beta, c, b)$ is never smaller than the exact free energy defined with $P_i^{BG}(\beta)$ (6) and therefore c^*, the optimal value of c, can be obtained by minimising $J(\beta, c, b)$ with respect to c. One would expect that $c^* > \beta (e^* < 1)$ and $c^* \to \beta$ $(e^* \to 1)$ as b is increased, i.e. the approximation improves.

The approximate scanning procedure described thus far can conveniently be programmed for the computer since the growth tips are treated independently. However, sample attrition is expected to be significant for large N; it can be reduced by defining better scanning procedures in which b is increased and pairs or larger groups of tips are treated together. In the next section we show how the bias, introduced by the scanning method, can be removed.

5. Corrected averages for the canonical ensemble

In order to study the properties of SAT in the canonical ensemble one divides the grand canonical sample of n $(n = n_{\text{succ}})$ SAT into subsamples of size $n(N)$ for SAT of N bonds. Thus an approximate but normalised probability $P_i(N, c, b)$ for SAT i of N bonds in the canonical ensemble can be obtained from the grand canonical probability $P_i^G(c, b)$ (see (15)-(18)):

$$P_i(N, c, b) = P_i^G(c, b)/B(N) \tag{20}$$

where

$$B(N) \sim n(N)/n_{\text{start}}. \tag{21}$$

Since $P_i(N, c, b)$ is biased it would lead to incorrect statistical averages of the entropy and the gyration radius (see (2) and (4)):

$$S_N(c, b)/k_B = -\sum P_i(N, c, b) \ln P_i(N, c, b) \tag{22}$$

$$\langle G_N^2 \rangle_{c,b} = \sum P_i(N, c, b) G_i^2 \tag{23}$$

where the summations in (22) and (23) are carried out over the ensemble of SAT of N bonds. Obviously, $S_N(c, b)$ is never larger than the correct S_N (2). However, this bias can be removed by employing two different standard procedures, importance sampling (Hammersley and Handscomb 1964) and a generalised Monte Carlo procedure suggested by Schmidt (1983). With importance sampling the estimators for the exact values of G and S are

$$\bar{G}_N^2(\text{IS}) = \sum_{i=1}^{n(N)} G_i^2 P_i^{-1}(N, c, b) \left(\sum_{i=1}^{n(N)} P_i^{-1}(N, c, b) \right)^{-1} \tag{24}$$

$$\bar{S}_N(\text{IS})/k_B = -\ln \left(n_{\text{start}}^{-1} \sum_{i=1}^{n(N)} P_i^{-1}(N, c, b) \right) \tag{25}$$

where IS stands for importance sampling. One can also estimate with importance sampling averages in the grand ensemble. The average number of bonds (7) is

$$\bar{N}^G(\text{IS}) = \sum_{i=1}^{n} \beta^{N_i} N_i / P_i^G(c, b) \left(\sum_{i=1}^{n} \beta^{N_i} / P_i^G(c, b) \right)^{-1}. \tag{26}$$

For an infinite sample the estimators (24)-(26) always lead to the correct ensemble averages (2), (4) and (7) respectively; for a finite sample their *statistical* convergence depends on their standard deviations which are expected to decrease as the bias decreases. A practical criterion which enables one to verify that a statistical convergence has been attained is to check that the importance sampling results remain unchanged for different approximate scanning procedures (see Meirovitch 1985a, c). The Schmidt procedure enables one to extract from a sample selected with a biased probability $P_i(N, c, b)$ an effectively smaller sample in which SAT are weighted correctly with the Boltzmann probability $P_i^{\mathrm{B}}(N)$ (1). This process is performed as follows: The first SAT generated (i) is always accepted. The second one (j) is accepted with probability p_{ij}. If j has been accepted the sample contains the SAT i and j; in the other case i is accepted again with probability $1-p_{ij}$ and the sample contains SAT i twice. Thus the third SAT is tested with respect to the SAT accepted at the second step and the process continues until the whole biased sample of $n(N)$ SAT has been examined and an *unbiased* sample of $n(N)$ SAT has been obtained. p_{ij} is defined as follows:

$$p_{ij}=\begin{cases} P_j(N, c, b)/P_i(N, c, b) & \text{if } P_j(N, c, b)/P_i(N, c, b)<1 \\ 1 & \text{otherwise.} \end{cases} \tag{27}$$

One can define an acceptance rate AR

$$\mathrm{AR}=n(N)_{\mathrm{accep}}/n(N) \tag{28}$$

where $n(N)_{\mathrm{accep}}$ is the number of different SAT accepted to the unbiased sample. Note that for the Boltzmann probability $P_i^{\mathrm{B}}(N)$ (1), AR = 1 whereas for a biased probability $P_i(N, c, b)$ generally AR < 1. However, one would expect AR to increase as the approximation improves. One can estimate the various averages directly from the unbiased sample without compensation (as is necessary in the case of importance sampling):

$$\bar{G}_N^2(\mathrm{A})=\frac{1}{n(N)}\sum_i{}' G_i^2 \tag{29}$$

$$\bar{S}_N(\mathrm{A})=\frac{-k_{\mathrm{B}}}{n(N)}\sum_i{}' \ln P_i(N, c, b) \tag{30}$$

and Σ' means summation over the accepted sample where A stands for accepted. It should be pointed out that $\bar{S}_N(\mathrm{A})$, in contrast to $\bar{G}_N^2(\mathrm{A})$, is always approximate (since the $P_i(N, c, b)$ are biased) and is expected to overestimate the correct entropy (2) (Meirovitch 1985c, d). On the other hand, we have argued that $S_N(c, b)$ (22) underestimates the entropy; therefore if the deviations of these entropies from the correct value are approximately equal, one would expect their average $\bar{S}_N^{\mathrm{M}}$ to be a better approximation than each one of them individually:

$$\bar{S}_N^{\mathrm{M}}=\tfrac{1}{2}(\bar{S}_N(\mathrm{A})+\bar{S}_N(c, b)) \tag{31}$$

where $\bar{S}_N(c, b)$ is the estimator for $S_N(c, b)$ (22). Our interest in the properties of $\bar{S}_N^{\mathrm{M}}$ stems from the fact that this average can also be calculated from a sample of chains obtained with *any* computer simulation technique and thus enables one to estimate the entropy approximately (Meirovitch 1985b, d). It has already been found that for several linear chain models $\bar{S}_N^{\mathrm{M}}$ constitutes a very good approximation for the entropy and it is important to check its properties also for branched structures. The extent of convergence of the estimators (29) and (30) can be deduced from the behaviour of AR during the simulation. For a small sample AR is relatively large since the first SAT is always accepted. As the sample increases AR relaxes to its (smaller) constant value

and therefore the initial accepted sample, obtained before AR has been stabilised, should be discarded. Thus, the number of accepted SAT, $n(N)_{\text{accep}}$, in the remaining sample can be considered as an effective sample size for the unbiased sample, which means that the standard deviation of extensive properties is expected to decrease as $(N/n(N)_{\text{accep}})^{1/2}$.

6. Results and discussion

In this work we study SAT (without loops) on a square lattice as described in § 2, employing an approximate scanning procedure, which is slightly more complex than that developed in § 3 in that it is based on several types of local states. For the first step we define future grand partition functions $\Xi_{k'}(t=1, b=2, c_1, e_1)$ (12) for the various LT ($1 \leqslant k' \leqslant 16$) using the scanning parameter $b=2$, the activity parameter c_1 and the parameter e_1 for the empty LT($k'=1$). These define transition probabilities for $t=1$ (13). For $t>1$ three types of transition probabilities are employed. First the K_t growth tips are visited in a predefined order and pairs of tips, which are nearest neighbours on the lattice, are identified in the following manner. We start by searching for a nearest-neighbour for the first tip; if it is not found the tip will be treated later independently and the search is continued for the next tip among the K_t-2 remaining ones. In the other case the two nearest-neighbour tips define a pair and the search for the next tip is carried out over the group of K_t-3 tips, etc. At the end of the process two groups, of k_{pair} pairs and k_1 individual tips, are defined such that $2k_{\text{pair}}+k_1=K_t$. The transition probabilities for a pair are defined by calculating future grand partition functions $\Xi_{k'}$ of $b=1$ for the two tips together rather than for each one separately. Also, these transition probabilities depend on parameters which are denoted c_{pair} and e_{pair} (see § 3). The k_1 individual tips are treated separately one-by-one, using in most cases transition probabilities which are based on $b=2$ and the corresponding parameters c_2 and e_2; however, for k_2 of these tips the third future shell is partially occupied, which makes it easy to calculate transition probabilities which are based on $b=3$. The parameters for this better approximation are denoted c_3 and e_3. These transition probabilities are utilised for selecting a set of LT (a shell configuration) for the K_t tips which is accepted if the excluded volume condition is not violated, as described in § 2.

The simulation has been carried out at $\beta=0.192$ ($\beta^{-1}=5.2$), which is lower than the expected critical activity $\beta_c=\mu^{-1}=0.195$. First we have carried out short runs in order to minimise the free-energy functional J (19) with respect to the various parameters; the optimal values are presented in table 1. As one would expect the values of c are larger than β and they monotonically decrease as the approximation improves, from $c_{\text{pair}}=0.47(b=1)$ to $c_3=0.23(b=3)$, which is only slightly larger than $\beta=0.192$. An opposite trend is observed for the parameters e (besides e_1) where $e_{\text{pair}}=0.56$ increases to $e_3=0.88$. Notice that a smaller value of the parameter e causes a decrease in the probability of the empty LT and thus increases the probability to generate larger SAT. Also, since e is a parameter of only one LT, J is expected to be less sensitive to changes in e than in c; this indeed is borne out. In the table, results are also presented for a worse scanning approximation based on $b=1$ where all the tips are treated separately. As expected, the grand canonical free energy $|J|=0.17$ and the average number of bonds $\bar{N}^{\text{G}}(\text{IS})\sim 40$ (26) are smaller than the corresponding values 0.194 and 60 obtained for the better approximation.

Table 1. Results for the optimal parameters c^* and e^*, the minimum value of the grand canonical free-energy functional J (19) and the average number of bonds $\bar{N}^G$ (26). The construction of SAT has been carried out up to a cutoff shell $t_N = 150$. Results are also presented for the worse approximation based on $b = 1$.

	c^*	e^*	$J/150$	$\bar{N}^G$
Approximation $b = 1, 2, 3$				
Pairs ($t > 1$)	0.47	0.56		
$b = 2 (t = 1)$	0.3	0.4	−0.0194	~60
$b = 2 (t > 1)$	0.3	0.8		
$b = 3 (t > 1)$	0.23	0.88		
Approximation $b = 1$				
$b = 1 (t \geq 1)$	0.54	0.93	−0.0170	~40

Our major calculation is based on two computer runs (each required 30 h of CPU time on an IBM 3081) carried out with different random number sequences. This enables one to estimate the statistical error. The two grand canonical samples have then been unified into one larger sample from which subsamples of SAT of given N, $10 \leq N \leq 140$ have been defined. The bias of these samples has been corrected by both importance sampling and the Schmidt procedure. For the latter we have included in the samples only the SAT accepted after AR (28) has been stabilised. Also, we have carried out calculations with the worse scanning approximation (based on $b = 1$) for $N \leq 50$ and verified that the importance sampling results are equal (within the statistical error) to those obtained with the better approximation. This gives confidence that convergence has been attained and thus the effective sample size is n_{accep}, as discussed at the end of § 5. Some results for the entropy, the gyration radius, the sample size and the acceptance rate are presented in table 2. This table reveals that, as expected, the results for the approximate entropy $S_N(c, b)$ (22) are always smaller than the importance sampling values, $\bar{S}_N(\text{IS})$ (25), which are considered to be correct; the deviation increases with increasing N, i.e. as the bias increases. The accuracy of $\bar{S}_N(\text{IS})$ can be checked only for $N \leq 11$ where exact enumeration values are available (Gaunt *et al* 1982). For $N = 10$ the exact value is $a_N = 672\,390$ (see (3)) which is equal, within the statistical error, to our value $672\,001 \pm 518$. It should be pointed out that $\bar{S}_N^M(c, b)$ (31) is very close to $\bar{S}_N(\text{IS})$, which means that the results for $S_N(b, c)$ (22) and $\bar{S}_N(\text{A})$ (30) deviate almost equally from the correct entropy. As has been already pointed out

Table 2. Results for the entropies $S_N(c, b)$ (22), $\bar{S}_N(\text{IS})$ (25) and $S_N^M(c, b)$ (31) (expressed in units of $k_B(N+1)$), the gyration radius $\bar{G}_N^2(\text{IS})$ (24), the sample size $n(N)$, the number of accepted trees $n(N)_{\text{accep}}$ and the acceptance rate AR (28). N is the number of bonds. The statistical error is denoted by parentheses; for example, $1.616(8) = 1.616 \pm 0.008$.

N	$S_N(c, b)$	$\bar{S}_N(\text{IS})$	$\bar{S}_N^M$	$G_N^2(\text{IS})/N$	$n(N)$	$n(N)_{\text{accep}}$	AR
10	1.422 37 (4)	1.437 81 (7)	1.437 92 (9)	0.293 7 (3)	1197 720	811 260	0.68
30	1.544 08 (7)	1.565 2 (8)	1.565 2 (4)	0.357 3 (6)	310 990	130 260	0.42
50	1.569 5 (2)	1.593 6 (5)	1.593 2 (7)	0.400 (3)	103 850	27 950	0.27
70	1.579 8 (4)	1.605 6 (9)	1.604 4 (9)	0.437 (6)	37 150	7 070	0.19
90	1.586 1 (7)	1.612 (2)	1.611 (8)	0.461 (9)	13 680	1 900	0.14
110	1.589 (2)	1.616 (8)	1.612 (8)	0.47 (2)	4 990	550	0.11

this behaviour is important in the context of a method which enables one to extract the entropy from a sample of chain conformations (Meirovitch 1985b, d). It should be emphasised that for the calculation of the critical exponent θ we use the results of $\bar{S}_N$(IS) (25), rather than those of $\bar{S}_N^M(c, b)$ (31). It should also be noticed that the values of the approximate gyration radius $\langle G_N^2\rangle_{c,b}$ (not appearing in table 2) are always smaller than the values of $\bar{G}_N^2$(IS) (considered as correct), 0.270 being smaller than 0.290 for $N = 10$, and 0.378 smaller than 0.472 for $N = 110$. This means that the SAT generated with the scanning method are too compact, as has also been observed for self-avoiding walks (Meirovitch 1985a). Finally, we emphasise the dramatic decrease of $n(N)$ and $n(N)_{\text{accep}}$ as N increases. This is followed by a decrease in the acceptance rate AR, which constitutes, in addition to $S_N(c, b)$ (22), a measure of the extent of bias. Obviously, the sample can be enriched with larger SAT by utilising better approximations at activity values closer to β_c.

In order to estimate the critical exponents ν and θ and the connective constant μ ((3) and (4)) we employ a procedure suggested recently (Berretti and Sokal 1985), which takes into account corrections to scaling. This procedure is essential for obtaining reliable results. Since the data are too 'noisy', only the analytic correction $1/N$ is considered and we fit the data to the functions

$$\langle G_N^2\rangle \sim A_G(N+g)^{2\nu} \tag{32}$$

$$\exp(S/k_B)/(N+1) \sim A_S\mu^N N^{-\theta}(1+s/N) \tag{33}$$

where g and s are parameters which should be optimised (notice that (32) has the same functional form as (33) to order $1/N$; we use it in order to be consistent with Caracciolo and Glaus (1985)). The main idea is to calculate for a given value of g, for example, several estimates for ν ($\nu(N_{\min})$) based on values of G between $N_{\min}$ and $N_{\max}(N_{\min} \ll N_{\max})$ where $N_{\min}$ increases while $N_{\max}$ remains fixed. The best parameter g is the one which leads to the 'flattest' graph of $\nu(N_{\min})$. In order to estimate θ and μ reliably one needs highly accurate results for the entropy; therefore we use $N_{\max} \leqslant 60$ for which the statistical error is relatively low. Results for μ and θ, for several values of the parameter s (using $N_{\max} = 60$), are summarised in table 3. For both quantities the flattest graphs are observed for $s = -0.3$. However, our estimates for μ and θ are obtained by averaging their values for $s = -0.2$ to -0.5; the error of this average is defined as twice the standard deviation and it constitutes only a part of the systematic error, which also depends on $N_{\max}$, for example. The statistical error has been obtained by performing a similar analysis for the two smaller samples (discussed previously), which are based on different random number sequences. This last error, however, is relatively large for $N_{\max} = 60$ and has been found to increase significantly for $N_{\max} \geqslant 65$. On the other hand, for $N_{\max} = 55$ the statistical error is much lower while the partial systematic error is only slightly changed. The results with these errors combined are:

$N = 60$	$\mu = 5.1428 \pm 0.0012$	$\theta = 1.009 \pm 0.046$
$N = 55$	$\mu = 5.1410 \pm 0.0018$	$\theta = 0.996 \pm 0.015$
best	$\mu = 5.1419 \pm 0.003$	$\theta = 1.003 \pm 0.02.$

Our best estimates (the third line) have been obtained by averaging the values of the first two lines and adopting the errors of $N = 55$. Our results for θ agree with $\theta = 1.00 \pm 0.02$ obtained by exact enumeration of SAT of $N \leqslant 11$ (Gaunt *et al* 1982) and are slightly better than $\theta = 1.001 \pm 0.08$ obtained from a Monte Carlo simulation

Table 3. Sets of results for μ and θ as functions of N_{min} for several values of s for $N_{max} = 60$. They were obtained by a least-squares procedure based on equation (34). The flattest sets are observed for $s = -0.3$.

	N_{min}				
s	10	15	20	25	30
Values of μ					
−0.6	5.1454	5.1444	5.1428	5.1444	5.1433
−0.5	5.1444	4.1438	5.1428	5.1444	5.1433
−0.4	5.1438	5.1433	5.1423	5.1438	5.1428
−0.3	5.1428	5.1428	5.1418	5.1433	5.1428
−0.2	5.1423	5.1423	5.1413	5.1433	5.1423
−0.1	5.1413	5.1418	5.1408	5.1428	5.1423
0	5.1048	5.1413	5.1402	5.1423	5.1418
Values of θ					
−0.6	1.033	1.026	1.014	1.027	1.018
−0.5	1.023	1.019	1.011	1.024	1.015
−0.4	1.016	1.013	1.004	1.017	1.008
−0.3	1.006	1.006	0.998	1.011	1.006
−0.2	0.999	0.999	0.991	1.008	0.999
−0.1	0.989	0.992	0.984	1.001	0.997
0	0.982	0.986	0.978	0.995	0.990

(Caracciolo and Glaus 1985). Also,exact enumeration study of lattice animals has led to $\theta = 1.00 \pm 0.02$ (Guttmann 1982); these estimates suggest that the Parisi and Sourlas (1981) prediction $\theta = 1$ is correct. Also our value for μ agrees with $\mu = 5.1434 \pm 0.0013 \pm 0.057$ obtained by Caracciolo and Glaus (1985) and $\mu = 5.14 \pm 0.01 \pm 0.26\Delta\theta$ by Gaunt *et al* (1982); our error, however, is smaller. All these estimates differ significantly from the value $\mu = 3.88$ obtained with a real space renormalisation group technique (Family 1983). In order to estimate ν we have plotted $\nu(N_{min})$ as a function of $1/N_{min}$ for various parameters g, as shown in figure 2. Such figures have been drawn for eight values of N_{max} (55, 60, 65 ... 90). For each figure

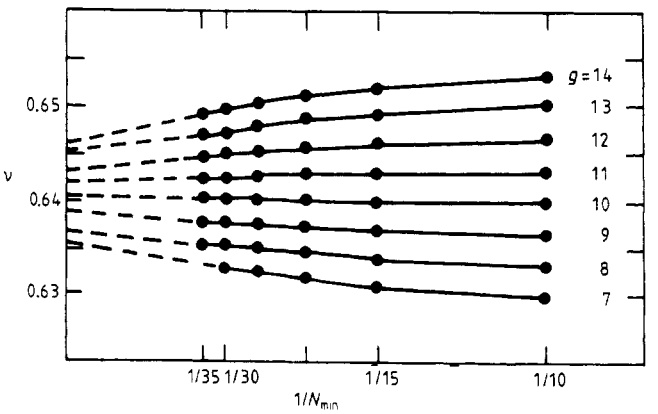

Figure 2. Plots of ν as a function of $1/N_{min}$, obtained for $N_{max} = 75$ for several values of the parameter g (full lines). These lines are extrapolated to $1/N_{min} = 0$ (broken lines). The flattest graph is obtained for $g = 10$ and its extrapolation leads to $\nu_{extru} = 0.641$.

the flattest graph has been extrapolated to $1/N_{\min}=0$ and the corresponding values ν_{extra} have been determined. We have taken the mean value of the eight results for ν_{extra} as our central estimate for ν and twice their standard deviation as the error,

$$\nu = 0.640 \pm 0.004.$$

This value equals (within the errors) the Monte Carlo result of $\nu = 0.640 \pm 0.004 \pm 0.004$ (Caracciolo and Glaus 1985) and Family's value of $\nu = 0.637$. Another Monte Carlo study for SAT (Seitz and Klein 1981) has led to $\nu = 0.612$, but without taking into account correction to scaling. Values of $\nu = 0.57$ and $\nu = 0.62$ have been obtained by Redner (1979) and Alexandrowicz (1985) respectively but, as has already been pointed out, their models of branched polymers differ from SAT; also, they have not carried out a correction to scaling analysis. Our result for ν is very close to the estimates $\nu = 0.64$–0.66 obtained for lattice animals by various techniques (Stauffer 1978, Gould and Holl 1981, Derrida and De Seze 1982, Family 1983, Margolina *et al* 1984, Privman 1984); in particular, within the errors our value equals the Derrida–De Seze result of $\nu = 0.6408 \pm 0.0003$ which is probably the best estimate for animals. This again suggests that animals with and without loops belong to the same universality class.

Another aspect of the scanning procedure is the fact that it describes a growth mechanism of SAT under equilibrium conditions, i.e. the generated SAT are approximately equally probable (1). It should be pointed out, however, that this mechanism does not seem to be physical mainly due to the fact that only the accepted SAT, which are determined *a posteriori*, are taken into account. However, for SAT of a given N one can study averages of the gyration radius G_t and the number of bonds N_t as functions of time. Alexandrowicz (1980, 1985) has defined exponents γ_t and ν_t for N_t and G_t, respectively,

$$N_t \sim t^{\gamma_t} \tag{34}$$

$$G_t \sim t^{\nu_t}. \tag{35}$$

This means that $G_t \sim N^{\nu_t/\gamma_t}$ or that

$$\nu = \nu_t/\gamma_t. \tag{36}$$

Similar properties have also been studied by Havlin and Nossal (1984) who refer to the time steps as chemical shells. In our procedures time steps (shells) are well defined and therefore γ_t and ν_t can be calculated. It should be pointed out that γ_t characterises the extent of branching. For a linear chain $\gamma_t = 1$; for an ideal tree (without excluded volume) $\gamma_t = 2$ (since $\nu = \frac{1}{4}$ and $\nu_t = \frac{1}{2}$) (Zimm and Stockmayer 1949, de Gennes 1968). Therefore, for a SAT, $1 < \gamma_t < 2$. The significance of ν_t can be explained as follows: for each growth tip at step t a linear chain of t bonds can be defined, which connects the tip to the origin. The average end-to-end distance of such a chain is $\sim t^{1/2}$ for an ideal tree and t^{ν_t} for a SAT, where ν_t is expected to differ from $\frac{3}{4}$, the value for self-avoiding walks (SAW) (Flory 1953). Therefore, ν_t measures the compactness of the branched chain. We have calculated ν_t and γ_t for $t = 3$–21 for seven chains of $N = 75$, 80, 85–100 and 110, using importance sampling, and have averaged the results for the different N. We have found that for values of t for which N_t is not too close to N, G_t and N_t are independent of N. Log–log plots of N_t and G_t as functions of t are presented in figure 3. They lead to

$$\nu_t \simeq 0.82 \qquad \gamma_t \simeq 1.26 \qquad \nu_t/\gamma_t \simeq 0.65.$$

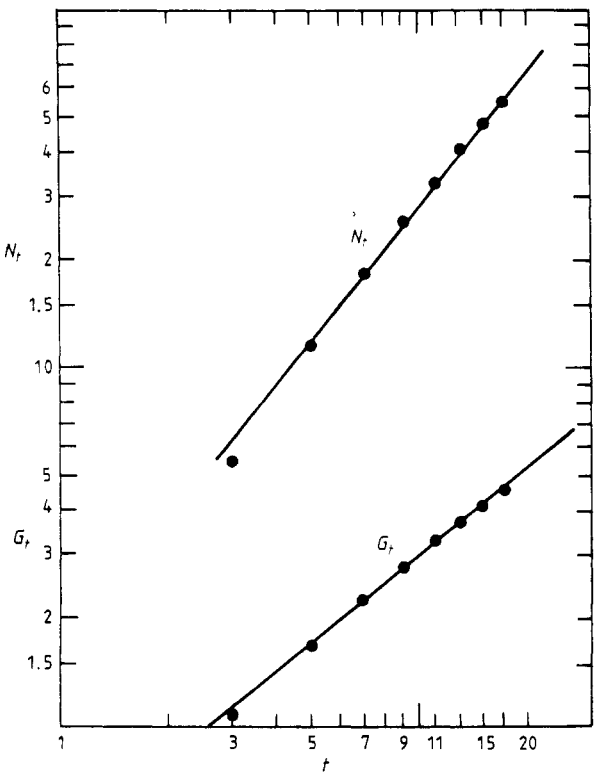

Figure 3. Log-log plots of N_t and G_t as functions of t, having slopes of $\gamma_t \simeq 1.26$ and $\nu_t \simeq 0.82$, respectively.

The fact that $\nu_t > \frac{3}{4}$ means that the branches are more open than SAW, in contrast to the behaviour in a melt where real chains become ideal. Also ν_t/γ_t is close to $\nu = 0.64$. It should be pointed out that Alexandrowicz (1985) obtained for his model T of branched polymers $\nu_t = 0.84$ and $\gamma_t = 1.35$, which means that his chains are even more branched and stretched than the present ones. We recall, however, that his chains are not equally probable.

7. Conclusions

We have extended the concepts of the scanning method to branched polymers without loops and obtained for self-avoiding trees on a square lattice relatively accurate results for θ, μ and ν, which agree with values obtained before. We have also obtained estimates for the exponents ν_t and γ_t. Our results for θ support the Parisi and Sourlas (1981) prediction $\theta = 1$; this and our value $\nu \simeq 0.640$ constitute additional evidence that lattice animals with and without loops belong to the same universality class. The advantage of the scanning method for trees should be emphasised. (i) The sampled configurations are statistically independent and therefore the statistical error can be estimated by the Schmidt procedure using standard statistical methods. The other simulation techniques are of a relaxation type and therefore the generated configurations are correlated. This correlation can safely be removed with the grand canonical

(GC) method (Berretti and Sokal 1985, Caracciolo and Glaus 1985) where bonds are added to the chain or removed from it during the simulation. However, this is not the case with the canonical ensemble (CE) type techniques which only allow bonds to change places (Stauffer 1978, Seitz and Klein 1981, Dickman and Schieve 1984). (ii) With the scanning method the chain probability is always known and hence the entropy can be calculated. In general, it is very difficult to estimate the entropy with both the GC and CE methods. However, for the special case in which the scaling relation (3) holds the entropy can also be obtained with the GC method using a nice procedure suggested by Berretti and Sokal (1985). (iii) The scanning method enables one to study geometrical properties which depend on time. It is of interest to compare the efficiency of the scanning method to that of the GC and CE techniques. The average number of bonds of the SAT in our grand canonical sample is $\langle N\rangle^{G}\sim 60$ (see (7) and (26)) as compared to 40 obtained by Caracciolo and Glaus (1985). This means that we have generated larger trees. We have used $N\leqslant 60$ for the calculation of the entropy, $N\leqslant 90$ for the gyration radius and $N\leqslant 110$ for estimating γ_t and ν_t. Lattice animals of comparable size, $N\sim 100$ bonds, have been generated by CE techniques for $d=2$ and 3 (Stauffer 1978, Peters *et al* 1979, Dickman and Schieve 1984) and slightly larger clusters of $N\sim 150$ have been simulated by Gould and Holl (1981). Seitz and Klein (1981), on the other hand, generated significantly larger SAT of $N=600$. However, it is not clear whether their sample is uncorrelated. The efficiency of our method for generating larger SAT can further be improved by developing algorithms which enable one to calculate transition probabilities for larger groups of tips and a larger scanning parameter b. Finally it should be noticed that the scanning procedure developed here can also be applied to SAT in larger dimensions ($d>2$), to branched polymers with finite interactions (Derrida and Herrmann 1983, Dickman and Schieve 1986) and to random surfaces. On the other hand, in order to study branched polymers with a fixed number of branches, such as stars or proteins, it would be more convenient to work in the canonical ensemble, using a modified version of the present scanning procedure. Such a procedure would also be suitable for treating systems of many-linear chains.

Acknowledgment

I thank Professor Zeev Alexandrowicz for valuable discussions.

References

Alexandrowicz Z 1980 *Phys. Lett.* **80A** 284
—— 1985 *Phys. Rev. Lett.* **54** 1420
Berretti A and Sokal A D 1985 *J. Stat. Phys.* **40** 483
Caracciolo S and Glaus U 1985 *J. Stat. Phys.* **41** 95
De Gennes P G 1968 *Biopolymers* **6** 715
Derrida B and De Seze L 1982 *J. Physique* **43** 475
Derrida B and Herrmann H J 1983 *J. Physique* **44** 1365
Dickman R and Schieve W C 1984 *J. Physique* **45** 1727
—— 1986 *J. Stat. Phys.* **44** 465
Dobson G R and Gordon N 1965 *J. Chem. Phys.* **43** 705
Duarte J A M and Ruskin H J 1981 *J. Physique* **42** 1585
Family F 1983 *J. Phys. A: Math. Gen.* **16** L97
Flory P J 1953 *Principles of Polymer Chemistry* (Ithaca, NY: Cornell University Press)

Gaunt D S 1980 *J. Phys. A: Math. Gen.* **13** L97
Gaunt D S, Martin J L, Ord G, Torrie G M and Whittington S G 1980 *J. Phys. A: Math. Gen.* **13** 1791
Gaunt D S, Sykes M F, Torrie G M and Whittington S G 1982 *J. Phys. A: Math. Gen.* **15** 3209
Glaus U 1985 *J. Phys. A: Math. Gen.* **18** L609
Gould H and Holl K 1981 *J. Phys. A: Math. Gen.* **14** L443
Guttmann A J 1982 *J. Phys. A: Math. Gen.* **15** 1987
Hammersley J M and Handscomb D C 1964 *Monte Carlo Methods* (London: Methuen)
Havlin S and Nossal R J 1984 *Phys. A: Math. Gen.* **17** L427
Lubensky T C and Isaacson J 1979 *Phys. Rev.* A **20** 2130
Margolina A, Family F and Privman V 1984 *Z. Phys.* B **54** 321
Meirovitch H 1982 *J. Phys. A: Math. Gen.* **15** L735
—— 1985a *Macromol.* **18** 563
—— 1985b *Macromol.* **18** 569
—— 1985c *Phys. Rev.* A **32** 3699
—— 1985d *Phys. Rev.* A **32** 3709
Parisi G and Sourlas N 1981 *Phys. Rev. Lett.* **46** 871
Peters H P, Stauffer D, Hölters H P and Loewenich K 1979 *Z. Phys.* B **34** 399
Privman V 1984 *Physica* **123A** 428
Redner S 1979 *J. Phys. A: Math. Gen.* **12** L239
Schmidt K E 1983 *Phys. Rev. Lett.* **51** 2175
Seitz W A and Klein D J 1981 *J. Chem. Phys.* **75** 5190
Stauffer D 1978 *Phys. Rev. Lett.* **41** 1333
Zimm B H and Stockmayer W H 1949 *J. Chem. Phys.* **17** 1301