Journal of Statistical Physics, Vol. 86, Nos. 1/2, 1997

Monte Carlo Study of the Θ-Point for Collapsing Trees

N. Madras[1] and E. J. Janse van Rensburg[1]

Received August 15, 1995; final March 19, 1996

We investigate the collapse transition of lattice trees with nearest neighbor attraction in two and three dimensions. Two methods are used: (1) A stochastic optimization process of the Robbins-Monro type, which is designed solely to locate the maximum value of the specific heat; and (2) umbrella sampling, which is designed to sample data over a wide temperature range, as well as to combat the quasiergodicity of Metropolis algorithms in the collapsed phase. We find good evidence that the transition is second order with a divergent specific heat, and that the divergence of the specific heat coincides with the metric collapse.

KEY WORDS: Lattice trees, collapse transition, scaling and crossover behavior; Robbins–Monro algorithm; umbrella sampling Monte Carlo.

1. INTRODUCTION

Polymers in solution are believed to undergo a collapse transition driven by solvent quality. The transition may be characterized by measuring a metric quantity associated with the polymers, such as the root mean square radius of gyration R_G. In the "good solvent" regime, or the "expanded" phase, one expects that $R_G \sim M^{\nu}$, where M is the molecular mass of the polymer, and ν is the *metric exponent* which describes the scaling of R_G with increasing mass. On the other hand, beyond the collapse transition, it is expected that $R_G \sim M^{1/d}$, where d is the dimension, so that the polymer is a compact or solid object. The change in scaling behavior occurs at the Θ-point, which is believed to be a tricritical point in the phase diagram of the collapsing polymer.[1-3]

[1] Department of Mathematics and Statistics, York University, North York, Ontario, M3J 1P3, Canada.

0022-4715/97/0100-0001$12.50/0

There is numerical evidence that branched polymers undergo a collapse transition of the kind described above.[4-12] Acyclic branched polymers are modeled as trees (weakly) embedded in a lattice with a nearest neighbor contact potential between "monomers" which are nearest neighbors in the lattice, but not in the tree (see Fig. 1). This is the so-called *t-model*[11,13] with partition function

$$\mathscr{Z}_n(\beta) = \sum_{c \geq 0} t_n(c)\, e^{\beta c} \tag{1.1}$$

where $t_n(c)$ is the number of trees with n vertices and c nearest neighbor contacts. The "contact chemical potential" is β. (Since β may be imagined to be proportional to the inverse of temperature, we will refer to it as the "inverse temperature.") The interaction between monomers is attractive if $\beta > 0$ and repulsive if $\beta < 0$. The free energy per monomer of the tree can be defined as[13]

$$F_n(\beta) = \frac{1}{n} \log \mathscr{Z}_n(\beta) \tag{1.2}$$

The specific heat per monomer is then

$$C_n(\beta) = \frac{d^2 F_n(\beta)}{d\beta^2} \tag{1.3}$$

and we will study it to determine if there is a thermodynamic transition which can be associated with a metric collapse in trees, as suggested, for

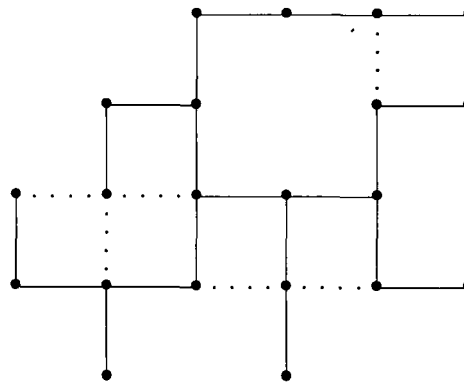

Fig. 1. A tree in the square lattice with 21 vertices (black dots), 20 edges (solid lines), and seven contacts (dotted lines).

example, in ref. 14 and 15. Available numerical data strongly support the association of a thermodynamic transition with the collapse; this includes the transfer matrix calculations by Derrida and Herrmann[4] for animals, and exact enumeration studies by Gaunt and Flesia[10, 11] for trees. These results were based on exact calculations, which limits the studies to small trees.

In this paper we revisit this model numerically, and in particular we seek to collect data on large trees in order to characterize the thermodynamic and metric behavior of lattice trees with nearest neighbor contacts. A cut-and-paste Monte Carlo algorithm for simulating trees in the hypercubic lattice[16] is adapted in two ways for this purpose. In the first instance, a stochastic optimization algorithm of the Robbins–Monro type is used with the cut-and-paste algorithm to determine the location and height of the peak in the specific heat with increasing n. The location of this peak, and its height can be used together with scaling arguments to determine the location of the critical point in the $n \to \infty$ limit. Second, we use umbrella sampling techniques[17] to study the dependence of the mean square radius of gyration as a function of β over a wide range of β; the collapse transition is characterized physically by a sudden change in the metric properties of the tree. Should the change in metric properties agree with the position of the peak in the heat capacity, which also increases with n, then we have strong evidence that the collapse transition is a second-order transition with a divergent specific heat characterized by both thermodynamic and metric properties. (In other words, the thermodynamic transition corresponds with the metric collapse, and is not driven by some other internal transition.)

This paper is organized as follows: In Section 2 we discuss some elementary results involving lattice trees. In Section 3 we discuss the Robbins–Monro algorithm and its implementation and we present numerical results obtained using this algorithm. In Section 4 we introduce umbrella sampling and study the collapse of lattice trees using this technique. We conclude the paper in Section 5 with a number of observations and conclusions.

2. ELEMENTARY CONSIDERATIONS

Equation (1.1) is the partition function of our model, with an associated reduced free energy per monomer defined by (1.2). Its $n \to \infty$ limit, the *limiting reduced free energy* (*per monomer*) $\mathscr{F}(\beta)$, is defined by

$$\mathscr{F}(\beta) = \lim_{n \to \infty} F_n(\beta) \tag{2.1}$$

Madras *et al.*[13] proved that the limit in (2.1) exists for $-\infty \leqslant \beta < \infty$, and is monotone nondecreasing, convex, and continuous for $-\infty < \beta < \infty$. Any nonanalyticity in $\mathscr{F}(\beta)$ at an inverse temperature β_c is defined to be a *thermodynamic phase transition*. Such a phase transition is *first order* if $\mathscr{F}(\beta)$ has a discontinuous first derivative at β_c, and is *second order* if it has a continuous first derivative at β_c but a discontinuous second derivative. At the present time, there is no rigorous proof of a nonanalyticity in $\mathscr{F}(\beta)$. (See refs. [18–20] for a discussion of the free energies of trees and animals.) Nonanalyticities in limiting reduced free energies have been proven only for directed walk models[21, 22] and a directed animal model.[23, 24] These results did not indicate the order of the internal transition(s) associated with the non-analyticities.

In the vicinity of a second-order transition one may apply nonrigorous (but very powerful) scaling arguments to define a set of scaling exponents which determines the universality class of the transition. In addition, the scaling assumptions will be very helpful in determining the location of the transition numerically. The following scaling arguments for collapse transitions are due to De'Bell and Whittington[25] (see also ref. 26). The starting point is the reduced free energy per monomer of the finite system given by equation (1.2). Then, from (2.1),

$$Z_n(\beta) = e^{\mathscr{F}(\beta)n + o(n)} \tag{2.2}$$

Let the critical value of β be denoted by β_c; and let $\tau = (\beta_c - \beta)/\beta_c$. For τ small (close to the transition), the singular part of the limiting free energy is expected to behave as

$$\mathscr{F}_{\text{Sing}}(\tau) = \tau^{2-\alpha}, \qquad \tau \geqslant 0 \tag{2.3}$$

where α is the "specific heat exponent.". If $\alpha > 0$, then this is associated with a divergence in the specific heat as the critical point is approached; if $\alpha < 0$, then it is associated with a cusp in the specific heat. Equation (2.3) applies only to the limiting free energy; for finite values of n we must modify it to take into account finite-size effects. A suitable scaling ansatz is

$$F_n(\beta) \sim \tau^{2-\alpha} f(n\tau^{1/\phi}) \tag{2.4}$$

where ϕ is a "crossover exponent" which relates the fugacity β to n, and $f(x)$ is a scaling function which relates $F_n(\beta)$ to the limiting free energy. Moreover, $f(x)$ is finite and $f(x) \to 1$ as $x \to \infty$. Other properties of $f(x)$ may be determined from those of $F_n(\beta)$ and $\mathscr{F}(\beta)$, defined in Eq. (2.1). Defining $g(x^\phi) = x^{(2-\alpha)\phi} f(x)$, in (2.4), we obtain

$$F_n(\beta) \sim n^{-\phi(2-\alpha)} g(n^\phi \tau) \tag{2.5}$$

The specific heat (per monomer) is defined by Eq. (1.3); taking the second derivative of (2.5) with respect to β for fixed n gives

$$C_n(\beta) \sim n^{\phi\alpha} g''(n^{\phi}\tau) \tag{2.6}$$

Let τ_n be that value of τ for which C_n is a maximum (that is, where $dg''/dx = 0$), and let H_n be the maximum value of C_n (with $\tau = \tau_n$). Then we conclude from (2.6) that

$$H_n \sim n^{\alpha\phi} \tag{2.7}$$

and

$$\tau_n \sim n^{-\phi} \tag{2.8}$$

Equations (2.7) and (2.8) have three unknown parameters, ϕ, α, and β_c, so that computing H_n and τ_n is not enough to solve for these unknowns. However, hyperscaling relates α and ϕ through $2 - \alpha = 1/\phi$, which together with Eqs. (2.7) and (2.8) can be used to estimate ϕ, α, and β_c.[4, 7, 11], [2]

However, even if we were to present convincing numerical evidence that the scaling law (2.4) holds (with suitable choices of ϕ, α, and β_c), and that H_n and τ_n scale as in (2.7) and (2.8), it still would not necessarily follow that this β_c corresponds to the collapse transition, or even that a collapse transition exists at all in this model. For it is possible that this β_c corresponds to an internal second-order transition that has nothing whatsoever to do with collapse of the polymer. The notion of collapse is a *metric* notion; it is observed physically by a change in the scaling exponents associated with a metric quantity, such as the root mean square radius of gyration. A necessary second step is thus to study a metric quantity and to corroborate the results from the thermodynamic data by illustrating a sudden change in the metric properties of the branched polymer at β_c. More formally, let R_n be any metric quantity; then it is expected that

$$R_n(\beta) \sim \begin{cases} n^{\nu} & \text{if } \beta < \beta_c \\ n^{\nu_c} & \text{if } \beta = \beta_c \\ n^{1/d} & \text{if } \beta > \beta_c \end{cases} \quad \text{as } n \to \infty \tag{2.9}$$

[2] It should be noted that the above only applies if H_n is divergent with increasing n (this occurs if $\alpha > 0$). If this does not happen (H_n is bounded as $n \to \infty$), then the maximum in the specific heat may occur at a analytic point of the reduced free energy. If this is the case, then the transition is signaled by a cusp in the specific heat, and not necessarily at a maximum. In the case of collapsing trees the numerical evidence will indicate that H_n diverges with n, and thus $\alpha > 0$.

where ν and ν_c are the metric exponents in the expanded phase and at the (tri)critical point, respectively.

The value of β_c can be obtained from the metric data by considering the ratios R_{kn}/R_n for k a fixed number (such that kn is an integer). It is seen from (2.9) that

$$\frac{R_{kn}}{R_n} \to \begin{cases} k^{\nu} & \text{if } \tau > 0 \\ k^{\nu_c} & \text{if } \tau = 0 \\ k^{1/d} & \text{if } \tau < 0 \end{cases} \quad \text{as } n \to \infty \tag{2.10}$$

for any fixed value of τ. If one computes $R_n(\beta)$ over a range of β and for a number of n-values and plots the ratios in (2.10) against β, a family of curves is found which intersect, in the large n limit, close to β_c. The $R_n(\beta)$ follows a scaling law similar to (2.6), namely

$$R_n(\beta) \sim n^{\nu_c} h(n^{\phi}\tau) \tag{2.11}$$

where h is a suitable scaling function [$h(x) \sim x^{(\nu - \nu_c)/\phi}$ as $x \to \infty$ and $h(x) \sim (-x)^{[(1/d) - \nu_c]/\phi}$ as $x \to -\infty$].

3. LOCATING β_c AND ϕ

The scaling arguments in Section 2 produced the relations (2.7) and (2.8) involving ϕ, α, and β_c. Numerical estimates for these quantities can be found by computing H_n and τ_n. We do this by using a stochastic optimization algorithm, specifically of the Robbins–Monro type,[27] which we implement with the cut-and-paste Monte Carlo algorithm for lattice trees.[16]

3.1. Implementation

We follow the implementation due to Kersting[28] and to Glynn.[29] Let P_β be a family of probability distributions on the real line, indexed by a real parameter β. Define the function

$$M(\beta) = \sum_x x P_\beta(x) \tag{3.1}$$

[In our case, P_β will be the distribution of $(c - \langle c \rangle_{\beta,n})^2$, and $M(\beta) = C_n(\beta)$; see Section 3.2.] Assume that $M(\beta)$ has a unique maximum at $\beta = \beta_c$. If $M(\beta)$ is differentiable, then $M'(\beta_c) = 0$. To find a root of M', consider Newton's method:

$$\beta_{i+1} = \beta_i - \frac{M'(\beta_i)}{M''(\beta_i)} \tag{3.2}$$

and we expect that $\beta_i \to \beta_c$ as $i \to \infty$. Since $M(\beta)$ has a maximum, we may assume that $M''(\beta_c) < 0$. For large i, replace $M''(\beta_i) \approx M''(\beta_c)$ by $-\delta^{-1}$. Then (3.2) becomes

$$\beta_{i+1} = \beta_i + \delta M'(\beta_i) \tag{3.3}$$

Suppose that we cannot evaluate either $M(\beta)$ or $M'(\beta)$ in closed form, but we can get a Monte Carlo estimate Y_i such that

$$E\{Y_i | \beta_1, \beta_2, \ldots, \beta_i\} = M'(\beta_i) \tag{3.4}$$

where $E\{S|T\}$ is the conditional expectation of S given T. Then a reasonable value for β_{i+1} is $\beta_i + \delta Y_i$. However, Y_i is in general a very "noisy" estimate of $M'(\beta_i)$; it therefore makes more sense to weight β_i by its corresponding relative sample size i, and the new estimate of β_{i+1} by its corresponding relative sample size 1. This gives the recursion which we actually use:

$$\beta_{i+1} = \beta_i + \delta Y_i/(i+1) \tag{3.5}$$

If Y_i is an unbiased estimator of $M'(\beta_i)$, then (3.5) is the Robbins–Monro algorithm with one free parameter δ. In applications one tries to choose δ close to $|M''(\beta_c)|^{-1}$. The rate of convergence is $1/\sqrt{i}$,[29] except in rare cases (in other words, the Robbins–Monro algorithm attains the best possible Monte Carlo convergence rate). The convergence of the algorithm may be tracked by relying on Corollary 2 of [ref. 28]: Given a number of (fairly weak) conditions, then if $\delta\rho > 1/2$, where $\rho = |M''(\beta_c)| \neq 0$, then the random function

$$\frac{\sqrt{2\delta\rho - 1}}{\delta\sigma} t^{\delta\rho} \sqrt{i}(\beta_{[it]} - \beta_c) \tag{3.6}$$

converges weakly as $i \to \infty$ to the process $W(t^{2\delta\rho - 1})$ for $0 \leqslant t \leqslant 1$, where $W(t)$ is Brownian motion and where σ^2 is the variance of $Y_i(\beta_c)$. Thus, provided that $\delta\rho > 1/2$, β_i will be asymptotically normal about β_c, or from (3.6), if $t = k/i$,

$$\left(\frac{2\delta\rho - 1}{\delta^2\sigma^2}\right) k^{2\delta\rho} i^{1-2\delta\rho} \operatorname{Var}(\beta_k) \approx \frac{k^{2\delta\rho - 1}}{i^{2\delta\rho - 1}} \qquad (1 \leqslant k \leqslant i) \tag{3.7}$$

with the result that

$$\operatorname{Var}(\beta_k) \approx \frac{\delta^2\sigma^2}{k(2\delta\rho - 1)} \tag{3.8}$$

The recursion [Eq. (3.5)] implies that

$$E\{(\beta_k - \beta_{k-1})^2\} \approx \frac{\delta^2\sigma^2}{k^2} = \frac{2\delta\rho - 1}{k} \operatorname{Var}(\beta_k) \tag{3.9}$$

So we conclude that $\operatorname{Var}(\beta_k) \approx kE\{(\beta_k - \beta_{k-1})^2\}/(2\delta\rho - 1)$. During a simulation, δ is an input parameter, and we can estimate the value of $\rho = |M''(\beta_c)|$ by estimating $\langle |M''(\beta_c)| \rangle$ [if this is possible; in our case $M(\beta) = C_n(\beta)$, so that ρ is the absolute value of the fourth cumulant of the contact number, which we can approximate in our simulation by estimates of the first four central moments of the contact number]. The expectation of $k^2(\beta_k - \beta_{k-1})^2$ can similarly be estimated in order to approximate $\delta^2\sigma^2$, and we can therefore obtain confidence intervals on β_k.

3.2. Numerical Results

Let $\langle x \rangle_{\beta, n}$ be the canonical expectation of a measurable quantity $x(c)$ which depends on the number of contacts c in a tree. From Eq. (1.1) we have

$$\langle x(c) \rangle_{\beta, n} = \mathscr{Z}_n^{-1}(\beta) \sum_{c \geq 0} x(c)\, t_n(c)\, e^{\beta c} \tag{3.10}$$

By Eqs. (1.2) and (1.3), the specific heat per monomer (for finite n) is given by

$$C_n(\beta) = \frac{\langle c^2 \rangle_{\beta, n} - \langle c \rangle_{\beta, n}^2}{n} = \frac{\langle (c - \langle c \rangle_{\beta, n})^2 \rangle_{\beta, n}}{n} \tag{3.11}$$

That is, $C_n(\beta)$ is the second central moment of the number of nearest neighbor contacts. Every tree can be given an energy $-c$ and therefore an associated weight $e^{\beta c}$. Then the average of the contact energy is simply $\mathscr{E}_n(\beta) = -\langle c \rangle_{\beta, n}$.

The cut-and-paste algorithm for lattice trees[16] can be used to realize a symmetric Markov chain whose state space is the set of lattice trees with n vertices. If a Metropolis-style rejection of proposed states in the chain based on the relative weights of the states is added, then the algorithm will generate a sample of trees with partition function given by Eq. (1.1).[30, 31] The maximum in $C_n(\beta)$ can be obtained by the Robbins–Monro recursion by identifying $C_n(\beta)$ with $M(\beta)$ as defined in (3.1). Comparison with the recursion in (3.2) and (3.5) indicates that

$$\beta_{i+1} = \beta_i + \delta \frac{\overline{[(c - \bar{c})^3]}_{\beta_i, n}}{i+1} \tag{3.12}$$

where $\bar{\cdot}$ denotes the sample mean of simulated observations from the distribution $\langle \cdot \rangle_{\beta_i, n}$. Then the sequence β_i should converge to that value of β which maximizes the specific heat. There are two parameters and one initial guess in (3.12) which must be carefully chosen. These are δ, the number N of measurements done in estimating the third central moment, and the initial value β_0. Ideally, δ should be the inverse of the curvature of the specific heat at β_c, but we cannot know this *a priori*. Instead, we perform an initial run at a value of β which we believe to be close to β_c and we compute the curvature at this point. This estimate will be very uncertain, since it is the fourth cumulant of c; and practically, we found that we had to settle for a very rough estimate, or for any value which will give convergence. (The simulation was harder in two dimensions, and we had to settle for a value of δ which is probably far from the optimal choice.) Note that if a guess for δ is too small, then the convergence of the recursion will be slow, especially if β_0 is far from β_c. The choice of N and of the number M of recursions of (3.12) results in a total of NM iterations of the Monte Carlo algorithm.

If this total is fixed, then there is a tradeoff between N and M. Empirically, we found that good approximations to the third central moments in (3.12) are essential for convergence to take place. Therefore, in all our simulations, we fixed $M = 50$, with $N = 5 \times 10^7$, for a total of 2.5×10^9 Monte Carlo iterations. Initial guesses were made by performing short runs for small trees to find a rough idea of the location of the peak. Subsequent guesses for larger trees are made by "boot-strapping" from the results for small trees. To remove bias in the estimates for the third central moment, we relaxed the simulation for 10^6 iterations between recursions of (3.12). The uncertainty in the resulting estimate of β_c was computed using (3.9), where we approximate ρ by the Monte Carlo estimate of the fourth cumulant about the peak of the specific heat (this is also the curvature of the peak in the specific heat), and $\delta^2\sigma^2$ by the Monte Carlo estimate of the expectation $E\{k^2(\beta_k - \beta_{k-1})^2\}$. [We ignored the first 40 recursions of (3.12) before we took data on the fourth cumulant and this expectation. At this point, we are close to the peak in the specific heat, and the fourth cumulant at the peak can be approximated by the Monte Carlo estimates of $C_n''(\beta_i)$.] Once the algorithm is close to the peak, one can compute the height of the peak by taking the result of each iteration as an independent estimate (this is reasonable because the number of attempted MC moves in each iteration of the Robbins–Monro algorithm far exceeds the autocorrelation time of the underlying cut-and-paste algorithm). We took the last ten iterations in each case and computed an average peak height.

The simulations were performed for trees in two and in three dimensions. The evolution of β_i is plotted in Figs. 2 and 3, and in Tables I and II

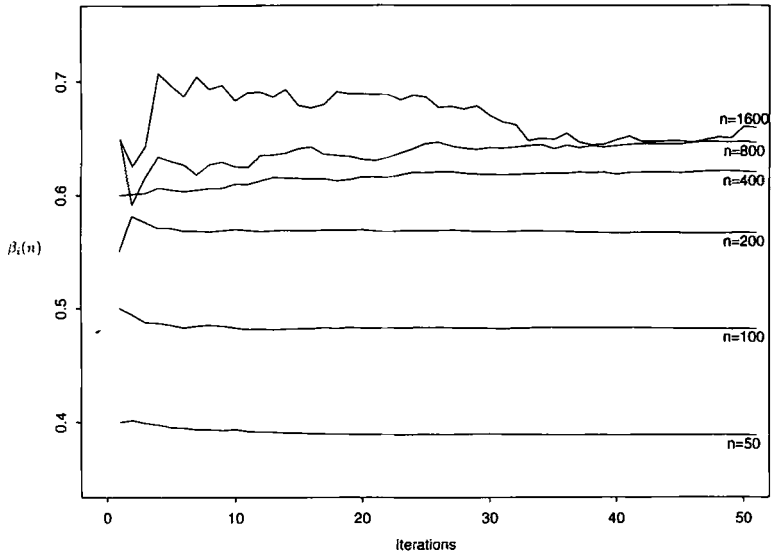

Fig. 2. The evolution of β_i in two dimensions under the Robbins–Monro iteration.

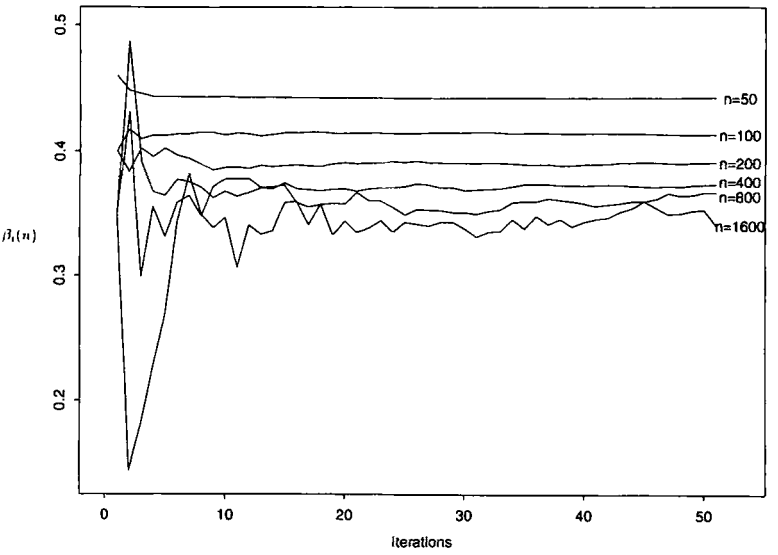

Fig. 3. The evolution of β_i in three dimensions under the Robbins–Monro iteration.

Table I. Robbins–Monro Results in Two Dimensions

n	$\beta_c(n)$	H_n	C_n''	δ	β_0
50	0.3885(2)	0.3020(24)	12.70(58)	0.333	0.400
100	0.4817(50)	0.3783(30)	41.3(15)	0.200	0.500
200	0.5672(48)	0.4527(60)	119(11)	0.100	0.550
400	0.6209(64)	0.5307(90)	334(38)	0.100	0.600
800	0.6468(88)	0.595(15)	839(130)	0.100	0.650
1600	0.660(15)	0.655(16)	2020(360)	0.100	0.650

we list the initial data and the results of the simulations. All the indicated error bars are 95% confidence intervals.[3]

3.3. Estimating φ, α, and β_c

The exponents ϕ and α and the critical value of β in the $n \to \infty$ limit can be extracted from the data in Tables I and II, assuming that the hyperscaling relation $2-\alpha = 1/\phi$ holds. Equation (2.7) suggests that a log–log plot of H_n against n should be linear, but significant curvature is present in a plot of the data in Tables I and II. Consequently, a significant correction to scaling must be present in the data.[4]

Two Dimensions. We first analyze our data under the assumption that $\alpha > 0$. In other words, we assume that the specific heat diverges as a power law of n. If $H_n = C_0 n^b$, then a linear least squares approximation, $\log H_n = \log C_0 + b \log n$, gives a very large χ^2-statistic and we conclude that the simple power-law assumption does not describe the data well. A modified ansatz $H_n = C_0 n^b (1 + Dn^{-1})$ (incorporating an analytic correction) can be approximated by the linear relation $\log H_n = \log C_0 + b \log n + Dn^{-1}$, and a linear least squares approximation gives $\chi_3^2 \approx 10.5$, which is not acceptable. If the data point at $n = 50$ is neglected, then $\chi_2^2 \approx 3.3$, which is a good fit. However, there is still a systematic error

[3] We will state all *statistical error bars* as 95% confidence intervals, unless we say explicitly otherwise.

[4] One can guess the significance of correction to scaling terms by keeping track of the χ^2-statistic when least squares analysis is done, assuming various scaling laws. We will accept a least squares fit if the χ^2-statistic is acceptable at the 95% level. Otherwise, we will attempt to improve the situation by neglecting datapoints at the smallest value of n (where correction terms are expected to be important), or, if that fails, by changing the model. At the 95% level, a χ^2-statistic with 2 degrees of freedom will have value at most 6.0; we indicate the number of degrees of freedom by a subscript: $\chi_2^2 \leqslant 6.0$ at the 95% level. Similarly, $\chi_3^2 \leqslant 7.8$ and $\chi_4^2 \leqslant 9.5$ at the 95% level.

Table II. Robbins–Monro Results in Three Dimensions

n	$\beta_c(n)$	H_n	C_n''	δ	β_0
50	0.4425(9)	0.6563(17)	58.9(11)	0.100	0.460
100	0.4133(32)	1.0103(60)	288(11)	0.0667	0.400
200	0.3910(54)	1.444(11)	1196(68)	0.0500	0.400
400	0.3734(84)	1.939(22)	4490(300)	0.0333	0.350
800	0.367(11)	2.507(50)	14700(1500)	0.0250	0.350
1600	0.341(2)	3.071(82)	44600(9600)	0.0200	0.350

present, and we attempt to estimate this by neglecting the data point at the smallest value of n.[5] The good fit above produced $b = 0.138 \pm 0.022$, and if we take the difference in the estimated regression coefficients in the last two fits as a systematic error, then we get $b = 0.138 \pm 0.022 \pm 0.033$, where the format is *best value* $\pm 95\%$ *confidence interval* $\pm$*systematic error*. From Eq. (2.7) and the hyperscaling assumption, $b = 2\phi - 1$; we get

$$\phi = 0.569 \pm 0.011 \pm 0.017 \tag{3.13}$$

By substituting this value of ϕ into (2.8), one can determine the value of β_c. However, a fit to $\beta_c(n) = \beta_c + Cn^{-\phi}$ [with $\phi = 0.569$ and where $\beta_c(n)$ is the value of β corresponding to the maximum in the specific heat] has a very large χ^2-statistic, and we modify the model by including a term to take into account corrections to scaling. If we assume that $\beta_c(n) = \beta_c + Cn^{-\phi} + Dn^{-1}$ (including an analytic correction), then $\chi_3^2 = 22.1$, and if we neglect the data point at 50, then $\chi_2^2 = 1.4$, which is a good fit. We neglect yet another datapoint to estimate a systematic error. The result is

$$\beta_c = 0.693 \pm 0.023 \pm 0.022 \tag{3.14}$$

As a last check, we compute β_c by taking values of ϕ at its confidence interval limits. If $\phi = 0.541$, then $\beta_c = 0.694 \pm 0.025$ ($\chi_2^2 = 1.5$) and if $\phi = 0.597$, then $\beta_c = 0.692 \pm 0.023$ ($\chi_2^2 = 1.4$), where the data point at $n = 50$ was ignored. These values are well within the confidence intervals stated in (3.14), and we conclude that the uncertainty in ϕ does not affect our estimate of β_c in any significant way.

Other Assumptions. We now consider the possibility that H_n diverges with $\log n$ (which corresponds to $\alpha = 0$), and the possibility that

[5] Systematic errors are due to correction to scaling which are unaccounted for by our model. If we neglect the data point where this error is likely to be most significant (the smallest value of n) and repeat the fit, then the absolute difference between the computed regression coefficients will be taken as a systematic error.

H_n is finite as $n \to \infty$ (this corresponds to the case that $\alpha < 0$). The value of b estimated above is close to zero, and we expect that an assumption that $H_n = C \log n + D$ will not be unreasonable. A linear least squares approximation gives $\chi_4^2 \approx 10.19$, which is not acceptable. If we ignore the data point at $n = 50$, then we obtain $\chi_3^2 \approx 6.7$, which is acceptable at the 95% level. Thus, we are unable to rule out this model from the data corresponding to the heights of the peaks in the specific heat. If we instead assume that $H_n = H_\infty + An^{-1/2}$ (corresponding to $\phi = 0.25$), then a least squares approximation gives $\chi_4^2 \approx 340$ and if the data point at $n = 50$ is ignored, $\chi_3^2 \approx 65$. Both these are ruled out at the 95% level. Smaller values of ϕ give increasingly worse fits, and the situation improves as ϕ increases to 0.5 (where $\alpha = 0$).

Three Dimensions. As in two dimensions, we first analyze our data under the assumption that H_n diverges as a power of n. A power law assumption $H_n = C_0 n^b$ has χ^2-statistic very large, and we conclude that corrections to scaling must be taken into account. If $\log H_n = \log C_0 + b \log n + Dn^{-1}$, then the χ^2-statistic is still too large, and we neglect the data point at $n = 50$ in the next fit. This gives $\chi_2^2 = 4.2$, which is a good fit. If we compute a systematic error by neglecting yet another term, then our estimate for b is $b = 0.307 \pm 0.020 \pm 0.030$. Thus

$$\phi = 0.654 \pm 0.010 \pm 0.015 \tag{3.15}$$

The critical value of β can be obtained as before; if $\phi = 0.654$, then $\beta_c = 0.333 \pm 0.012$, $\chi_3^2 \approx 4.8$, and if the data point at $n = 50$ is discarded, then $\beta_c = 0.326 \pm 0.020$, $\chi_2^2 \approx 3.9$. When ϕ takes values at its confidence limits, then $\beta_c = 0.334 \pm 0.012$, $\chi_3^2 \approx 4.9$, if $\phi = 0.679$, and $\beta_c = 0.332 \pm 0.012$, $\chi_3^2 \approx 4.7$, if $\phi = 0.629$. Estimating a systematic error as before, we find

$$\beta_c = 0.333 \pm 0.012 \pm 0.007 \tag{3.16}$$

Other Assumptions. An assumption that $H_n = C \log n + D$ gives very large χ^2-statistic, even if we ignore the data point at $n = 50$. If we add an analytic correction and assume that $H_n = C \log n + D + Fn^{-1}$, then $\chi_3^2 \approx 10.2$, which is not acceptable. Ignoring the data point at $n = 50$ gives $\chi_2^2 \approx 1.3$, which is acceptable. Thus, this ansatz is not inconsistent with our data. An assumption that $H_n = H_\infty + An^{-1/2}$ gives large χ^2 statistics, which gets larger with decreasing values of ϕ, but the situation improves as ϕ approaches 0.5.

Discussion. The curvature C_n'' of the specific heat peak may be computed by taking the second derivative of (2.6) at β_c. Hence, $C_n'' \sim n^{\alpha\phi + 2\phi}$, and by taking the fourth cumulant of the contact number, one may estimate

$\mu = \alpha\phi + 2\phi$. The measured values for the curvature are listed in Tables I and II. By hyperscaling, $2\phi - \alpha\phi = 1$, and one may compute ϕ in an alternative manner [$\phi = (\mu + 1)/4$]. A three-parameter assumption $\log C''_n = A + \mu \log n + B/n$ gives $\mu = 1.329 \pm 0.082$ with $\chi^2_3 \approx 1.1$, and if we ignore the data point at $n = 50$, then we get $\mu = 1.275 \pm 0.156$, $\chi^2_2 \approx 0.4$, in two dimensions. Taken together, our best estimate is $\mu = 1.329 \pm 0.082 \pm 0.054$ in two dimensions. In three dimensions we obtain $\mu = 1.719 \pm 0.066$, $\chi^2_3 \approx 3.7$, and if the data point at $n = 50$ is ignored, then $\mu = 1.628 \pm 0.132$, $\chi^2_2 \approx 1.1$. Thus $\mu = 1.719 \pm 0.066 \pm 0.092$ in three dimensions. Assuming hyperscaling, one gets $\phi = 0.582 \pm 0.021 \pm 0.014$ in two dimensions and $\phi = 0.680 \pm 0.017 \pm 0.023$ in three dimensions. These estimates are completely consistent with (3.13) and (3.15). On the other hand, by solving for ϕ and α from b and μ, one may obtain estimates for ϕ which do not rely on hyperscaling. Solving directly gives $\phi = (\mu - b)/2$. In two dimensions, $\phi = 0.596 \pm 0.10$, and in three dimensions, $\phi = 0.71 \pm 0.11$. The uncertainties in these estimates are large, but the results are consistent with (3.13) and (3.15), and there seems no reason to believe that hyperscaling should be in doubt (the large error bars in these estimates are due to the uncertainty in estimating μ from fourth cumulants). Therefore, we assume that (3.13) and (3.15) represent the best estimates for ϕ, modulo hyperscaling.

Estimates for the specific heat exponent α may be obtained from b and ϕ. We found $\alpha = 0.24 \pm 0.11$ in dimensions (excluding 0 by two error bars), and $\alpha = 0.47 \pm 0.10$ in three dimensions (excluding 0 by at least four error bars).[6] It is important to note that one arrives at these values by assuming that the transition is second order with a divergent specific heat, and that we then get answers consistent with our assumption. While there is some circularity in this process (if we assume that $\alpha > 0$, then $\alpha = 0.24 > 0$ in two dimensions, and a similar result in three dimensions), we should note that the assumption that $\alpha = 0$ ($H_n \sim \log n$) also gives a consistent (albeit circular) analysis. However, the belief that the transition is second order with a divergent specific heat is supported to a significant extent by the fact that the estimated values of α exclude 0 by such a wide margin.

4. UMBRELLA SAMPLING

The Robbins–Monro procedure is designed for the sole purpose of pinpointing the maximum of the specific heat curve $C_n(\beta)$. Apart from gaining information on the specific heat close to the critical value of β, we

[6] Note that these error bars contain contributions from the systematic errors estimated in ϕ and b. The 95% confidence intervals (which are due to the statistical errors in ϕ and b) on α are smaller than the stated error bars here.

do not get any other information which will describe the true nature of the collapse transition; for that, we need to graph $C_n(\beta)$ for a wide range of β, and in particular in the general vicinity of the scaling region. In addition, we do not have information about the metric behavior of trees as β changes through the critical—value this characterizes the collapse transition (the specific heat only gives thermodynamic information). Thus, we would also like to obtain information about the behaviour of the radius of gyration $R_n(\beta)$ as β changes through its critical value.

4.1. Background and Implementation

A naive Monte Carlo approach would be to perform a set of independent simulations at several values of β, followed by an interpolation on the data. Despite the simplicity of this approach, there are several methodological difficulties, including the poor behavior of the Monte Carlo algorithm in the collapsed regime, characterized by very long autocorrelation times. This is due to "quasiergodicity": the tree falls into conformations which are essentially "frozen" for many iterations due to the strength of the attractive interactions.

An alternative Monte Carlo solution is *umbrella sampling*.[17, 32] Let T_n denote the set of n-site trees (modulo translation). For a tree $\omega \in T_n$, let $c(\omega)$ be the number of nearest neighbor contacts of ω. Let

$$p_{\beta, n}(\omega) = \frac{e^{\beta c(\omega)}}{\mathscr{Z}_n(\beta)} \tag{4.1}$$

be the probability distribution on T_n for given β. Next, let π be some positive function (measure) on T_n [which assigns weights $\pi(\omega)$ to each tree ω] and let $\mathscr{Z}_\pi := \sum_{\psi \in T_n} \pi(\psi)$ be the associated normalizing constant [or partition function as in Eq. (1.1)]. In practice, we choose π to be easily computable, but it will be too hard to compute $\mathscr{Z}_\pi$ exactly.

Let $X_1, X_2, \ldots$ be the Markov chain obtained by implementing the Metropolis algorithm with respect to the weights $\pi(\cdot)$, using the cut-and-paste tree algorithm as the underlying symmetric Markov chain for the proposed moves. Then the equilibrium probability distribution for X_i is $\pi/\mathscr{Z}_\pi$. (We assume that a sufficiently long initial section of the chain has been discarded to ensure that the remaining observations are in equilibrium.) For any function f on T_n, let

$$S_k[f; \beta] := \frac{1}{k} \sum_{i=1}^{k} \frac{f(X_i)}{\pi(X_i)} e^{\beta c(X_i)} \tag{4.2}$$

Observe that

$$\langle S_k[f;\beta]\rangle_\pi = \left\langle \frac{f(X_1)}{\pi(X_1)} e^{\beta c(X_1)} \right\rangle_\pi$$
$$= \sum_{\omega \in T_n} \frac{f(\omega)}{\pi(\omega)} e^{\beta c(\omega)} \frac{\pi(\omega)}{\mathscr{Z}_\pi} = \langle f\rangle_{\beta,n} \frac{\mathscr{Z}_n(\beta)}{\mathscr{Z}_\pi} \tag{4.3}$$

The ergodic theorem further implies that

$$\lim_{k\to\infty} S_k[f;\beta] = \langle f\rangle_{\beta,n} \frac{\mathscr{Z}_n(\beta)}{\mathscr{Z}_\pi} \tag{4.4}$$

with probability one. Now define the ratio estimator

$$R_k[f;\beta] := \frac{S_k[f;\beta]}{S_k[1;\beta]} = \frac{\sum_{i=1}^k [f(X_i)/\pi(X_i)]\, e^{\beta c(X_i)}}{\sum_{i=1}^k [1/\pi(X_i)]\, e^{\beta c(X_i)}} \tag{4.5}$$

Then (4.4) implies that

$$\lim_{k\to\infty} R_k[f;\beta] = \langle f\rangle_{\beta,n} \tag{4.6}$$

with probability one. Thus, for any observable f, we can (in principle) estimate $\langle f\rangle_{\beta,n}$ for *every* β using data from a single simulation run. Of course, this is only really true in the $n\to\infty$ limit; one cannot expect these estimates to be uniformly good for all β. But one can hope that for a judicious choice of π, one can get good estimates simultaneously for all β in some reasonably large interval. Thus, we intend to choose π such that a wide range of interesting distributions are covered; π is therefore called an "umbrella distribution."

The basic approach, then, is to sample from an artificial distribution and reweight the data [as in (4.2)] to get a valid estimator that is (hopefully) more efficient than the natural estimator (i.e., sampling directly from the distribution of interest). This idea, known as *importance sampling*, is a classical method of Monte Carlo variance reduction.[29] Umbrella sampling in statistical mechanics was pioneered by Torrie and Valleau.[17] Our implementation has also been called "temperature scaling" (since we reweight by distributions corresponding to a common Hamiltonian with different temperatures on a single configuration space), as opposed to "density scaling," for example.[32]

A special case of umbrella sampling is when $\pi(\omega) = \exp[\beta_0 c(\omega)]$ for some fixed value β_0. This approach has been implemented for Ising and

Potts models by Ferrenberg and Swendsen[33, 34] and it has also recently been employed to estimate likelihood functions in statistics.[35] Another special case, somewhat more general, is when π is a convex combination of several fixed-temperature distributions:

$$\frac{\pi(\omega)}{\mathscr{Z}_\pi} = \sum_{i=1}^{r} \lambda_i p_{\beta^{(i)},n}(\omega) \tag{4.7}$$

where $\lambda_1, \ldots, \lambda_r$ are positive numbers that add up to 1. This choice has an intimate connection with the method of "simulated tempering"[36] that is explored in ref. 37. Simulated tempering was developed independently by Geyer and Thompson[38] in the context of statistical inference. The main idea of (4.7) is that if some $\beta^{(i)}$ are in the collapsed regime and others are in the expanded regime, then the system will not get stuck in a collapsed configuration (typical of large β) because a lot of the sampling is performed as if the system had $\beta < \beta_c$. Thus the Markov chain is expected to have manageable autocorrelation times. Also, since a significant fraction of the sampling is performed as if the system had $\beta = \beta^{(i)}$, one expects that the data would give good estimates for $\langle f \rangle_{\beta^{(i)},n}$ for each i. Some rigorous results that help to support this intuition are presented in ref. 37.

In the simulated tempering scheme, one is always constrained to the form (4.7). However, in the general framework of umbrella sampling, there is no such restriction. One can choose an arbitrary distribution π and use short Monte Carlo runs to refine π so as to improve the "mixing rate" of the Markov chain. [We note that simulated tempering also requires some short initial runs to get crude estimates on ratios of the $\mathscr{Z}_n(\beta^{(i)})$.] As is natural, we restricted our attention to weights $\pi(\omega)$ that depend on ω only through $c(\omega)$ (i.e., through the energy of ω); formally, there is a function π^c from $\{0, 1, \ldots, (d-1)n\}$ to $[0, +\infty)$ such that

$$\pi(\omega) = \pi^c(c(\omega)) \tag{4.8}$$

[Note that $(d-l)n$ is an upper bound for the maximum number of contacts in an n-site tree in $\mathbf{Z}^d$.]

One useful guideline in choosing an umbrella distribution is to aim for a π which makes the distribution of energy (number of contacts, in our case) *uniform* over a wide range of values. This should accomplish two things: (i) It makes for a rapidly mixing Markov chain by removing the (physical) energy barriers that create regions of metastable configurations in the physical system (also called quasi-ergodic barriers); and (ii) it gives good estimates for any distribution $p_{\beta,n}$ that puts most of its weight in the energy range "covered" by the umbrella π.[32, 37, 39] In the present project,

we used three different methods for getting initial umbrella distributions for different values of n:

(I) *Monte Carlo runs at different β values.* Fix $\beta^{(1)},\ldots,\beta^{(r)}$, and do r short, independent runs (one for each $p_{\beta^{(i)},n}$). For each pair $i, j \in \{1,\ldots,r\}$, obtain an estimate $\hat{Z}_{i,j}$ of the ratios $\mathscr{Z}_n(\beta^{(j)})/\mathscr{Z}_n(\beta^{(i)})$. (There are several ways to do this. One crude way is by our estimator $S_k[1;\beta^{(j)}]$ with $\pi = \exp(\beta^{(i)}c)$; see also refs. 33 and 40 for more thoughtful approaches.) Then, for some choice of $j \in \{1,\ldots,r\}$, set

$$\pi(\omega) = \sum_{i=1}^{r} \hat{Z}_{i,j} e^{\beta^{(i)} c(\omega)} \tag{4.9}$$

If the short runs are not too short, then this should be a good approximation of the mixture (4.7) with $\lambda_i \equiv 1/r$ [and $\mathscr{Z}_\pi \approx r\mathscr{Z}_n(\beta^{(j)})$]. Note that for implementation, we only have to compute and store the numbers

$$\pi^c(k) = \sum_{i=1}^{r} \hat{Z}_{i,j} e^{\beta^{(i)} k} \tag{4.10}$$

for $k = 0, 1,\ldots, (d-1)n$. This procedure worked well up to $n = 400$ in $d = 2$, but not at $n = 800$.

(II) *Series extrapolation.* As in (I), fix $\beta^{(1)},\ldots,\beta^{(r)}$, and aim for an umbrella of the form (4.7) with $\lambda_i \equiv 1/r$. But instead of doing preliminary Monte Carlo work, we use exact enumeration for small values of n and the resulting extrapolation of the limiting reduced free energy $\mathscr{F}(\beta^{(i)})$. Then set

$$\pi^c(k) = \sum_{i=1}^{r} e^{\beta^{(i)} k - n\mathscr{F}(\beta^{(i)})} \tag{4.11}$$

for $k = 0, 1,\ldots, (d-1)n$. Fortunately, the enumeration and extrapolation have already been done by Gaunt and Flesia.[11] (They only do the extrapolation for a discrete set of β's, so we interpolated the results to get values at $\beta^{(i)}$ that were not in this set.) This procedure worked well at $n = 800$ in $d = 2$, but not at $n = 1600$.

(III) *Histogram uniformization.* At $n = 1600$ in $d = 2$, methods (I) and (II) both gave poor results. We followed instead a method of Valleau[32] which aims to create a distribution which will be uniform over a certain range of energies. This is done by dynamically adjusting π during a short initial run. A long simulation at $n = 1600$ with an umbrella π_0 obtained by method (II) did not give good estimates; in. particular, the energy

histogram $HIST(k)$ (=number of observed conformations with exactly k contacts) was decidedly uneven. This could be for one of two reasons: either the run was too short and $HIST(k)$ still contained large fluctuations, or the umbrella π_0 had been chosen poorly. Since our run was long, we believed that the second reason was the true cause; it is quite plausible that the extrapolations from (II) were just not good enough for $n = 1600$. So how can we improve an umbrella? Observe that if the simulation is not too short, then we expect that $HIST(k)/N \approx Ct_n(k)\,\pi_0^c(k)$ (where N is the length of the run) for all k which are not too far into the tails of the distribution. With this reasoning, it is natural to try to improve on π_0 by taking

$$\pi_1^c(k) := \frac{\pi_0^c(k)}{HIST(k)}$$

in the hope that the result would be uniform over all k that were not too far out in the tails. [If $HIST(k)$ is 0, then we set $\pi_1^c(k)$ equal to $\pi_0^c(k)$.]

4.2. Numerical Results

We generated data for the mean square radius of gyration $R_n^2(\beta)$ and the specific heat $C_n(\beta)$ over large intervals of β which included our estimates of the critical values from Section 3 for n equal to 50, 100, 200, 400, 800, and 1600. The performance of the algorithm worsened as n increased, especially in two dimensions.

We attempted to keep the amount of computer time spent on the umbrella sampling on the same order of magnitude as had been devoted to Robbins–Monro. The main run for each umbrella distribution was either 5×10^8 or 8×10^8 Monte Carlo iterations, taking data every 1000 or 2000 iterations, respectively. The time required for such a run varied from about 10 to 80 hr on our workstation, depending upon the size of the tree. We estimate very roughly that about 3 months of CPU time on an HP-730 workstation were spent on the umbrella sampling simulations, including initialization and runs with poor umbrellas that were eventually discarded. It is apparent that better results could be obtained with longer runs.

As a check, we compared our estimates of radius of gyration at $\beta = 0$ with those of Janse van Rensburg and Madras.[16] All estimates were completely consistent within error bars (in fact the error bars in the present study are usually smaller). Our methods for computing error bars are described in the Appendix.

Two Dimensions. For $n = 50, 100, 200, 400$, two large umbrellas easily covered a wide range of β: from $\beta = 0$ up to $\beta = 0.8$, which was the

largest value of β that we tried to analyze. For example, at $n = 200$, the first umbrella was based on nine regularly spaced $\beta^{(i)}$ values from 0 to 0.48; the second was based on nine regularly spaced values from 0.3 to 0.48. The overlap region could be used to doublecheck the validity of the simulations. We usually found that agreement was excellent on the overlap; in fact, the agreement frequently continued well beyond the overlap (especially for smaller n). We needed three umbrellas for $n = 800$ and four for $n = 1600$. The agreement at overlaps was not always good for these larger n values, and we had to repeat the histogram uniformization process (III) more than once before a reasonably uniform histogram was obtained (see Section 4.1). In addition to examining the uniformity of the histograms, we also monitored the behavior of each simulation by checking the frequency with which large cut-and-paste moves were accepted, and by looking at the autocorrelations in the output (see Appendix).

We plot the specific heat of trees in two dimensions in Fig. 4 as a function of β. The data exhibit a clear, albeit broad, peak, which increases in height, and whose location moves to larger values of β as n increases. One standard deviation about the data (68% confidence intervals) is also plotted. The locations and heights of the peaks in the data are listed in Table III. The results are entirely consistent with the results obtained from the Robbins–Monro method (Table I), although the error bars for β_c are

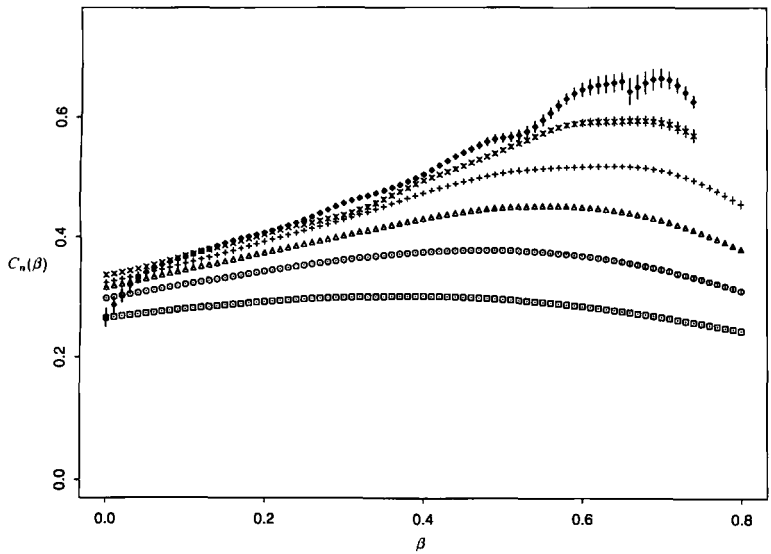

Fig. 4. The specific heat in two dimensions as a function of β. The data correspond to $n = 50$ ($\square$), 100 ($\bigcirc$), 200 ($\triangle$), 400 ($+$), 800 ($\times$), and 1600 ($\diamond$).

Table III. Umbrella Sampling Results in Two Dimensions

n	$\beta_c(n)$	H_n
50	0.378(13)	0.30106(83)
100	0.534(233)	0.384(12)
200	0.5611(30)	0.4516(18)
400	0.635(130)	0.5181(47)
800	0.669(215)	0.5936(99)
1600	0.699(62)	0.663(20)

considerably larger here. This is hardly a surprise, since only some of the umbrella sampling is spent in parts of configuration space that are weighted heavily near β_c. [Some of the error bars are disturbingly large in Table III; this is a combined effect of very broad peaks and the need to estimate fourth cumulants of c. We make no claim about error bars for our error bars for $\beta_c(n)$ in Table III, but we believe that they have the right order of magnitude.] The error bars for the peak heights H_n however, are more comparable with the Robbins–Monro error bars.

We can analyze the data in Table III in exactly the same manner as was done with the Robbins–Monro data in Tables I and II to obtain independent estimates of the crossover exponent ϕ and the critical value of β. If we assume the form $\log H_n = \log C_0 + b \log n + Dn^{-1}$, with C_0 and D constants and $b = 2\phi - 1$ by hyperscaling, then a least squares fit gives $b = 0.160 \pm 0.012$, $\chi_3^2 \approx 1.7$. We estimate a systematic error by ignoring the data point at $n = 50$, obtaining $b = 0.170 \pm 0.026$, $\chi_2^2 \approx 0.9$. We take the difference in these estimates as a systematic error, and find for ϕ

$$\phi = 0.580 \pm 0.006 \pm 0.005 \tag{4.12}$$

in good agreement with (3.13). The critical value of β can be computed exactly as before, but here a two-parameter fit proves adequate due to the large error bars; assuming that $\beta_c(n) = \beta_c + Cn^{-\phi}$, we find (by taking $\phi = 0.585$), $\beta_c = 0.709 \pm 0.012$, $\chi_4^2 \approx 1.6$, and if we neglect the data point at $n = 50$, then $\beta_c = 0.754 \pm 0.082$, $\chi_3^2 \approx 0.33$. At the confidence limits of ϕ we obtain $\beta_c = 0.700 \pm 0.011$, $\chi_4^2 \approx 1.9$ ($\phi = 0.609$), and $\beta_c = 0.718 \pm 0.013$, $\chi_4^2 \approx 1.3$ ($\phi = 0.561$). Thus, our estimate is

$$\beta_c = 0.709 \pm 0.012 \pm 0.045 \tag{4.13}$$

This result is close to the value obtained by the Robbins–Monro simulation.

The mean square radius of gyration data are best used to compute the amplitude ratios in Eq. (2.10). Here we expect that

$$\frac{R^2_{2n}(\beta)}{R^2_n(\beta)} \to \begin{cases} 2^{2\nu} & \text{if } \beta < \beta_c \\ 2^{2\nu_c} & \text{if } \beta = \beta_c \\ 2^{2/d} & \text{if } \beta > \beta_c \end{cases} \tag{4.14}$$

We plot the amplitude ratios in Fig. 5, which results in a family of curves which intersect one another close to the critical point and may approach the step function in (4.14) as n increases. The family of curves approaches a number close to 2.4 as β decreases, assuming a scaling form with a confluent correction $R^2_n = An^{2\nu}(1 + Bn^{-\Delta})$; this suggests that the convergence in (4.14) should be as $2^{2\nu} + Cn^{-\Delta}$ if $\beta < \beta_c$. The confluent correction for lattice trees was estimated from data describing the span in ref. 16; using that value (0.915) and executing a least squares fit to our data at $\beta = 0$ produces (we neglect the data point at $n = 50$ to guess a systematic error)

$$\nu = 0.6370 \pm 0.0022 \pm 0.0032 \tag{4.15}$$

with $\chi^2_3 \approx 5.5$. We have not taken into account the uncertainty in the value Δ in estimating (4.15), but a fit with $\Delta = 0.5$ gives $\nu = 0.6386$ and with $\Delta = 1$

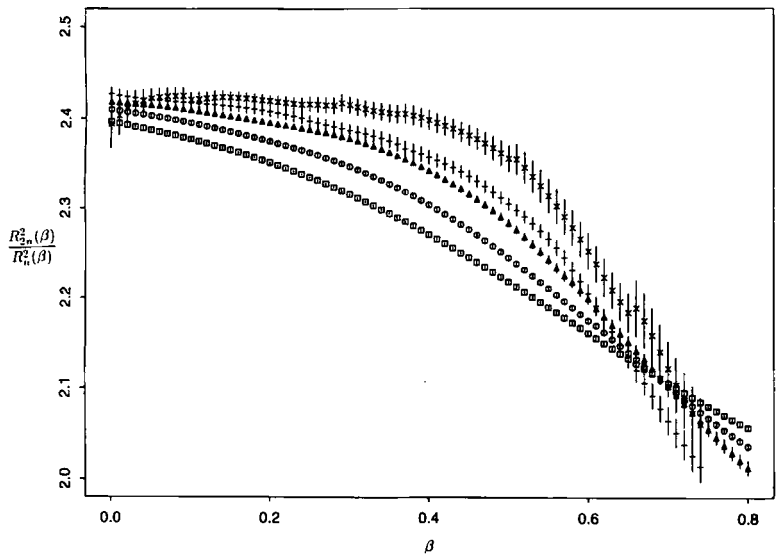

Fig. 5. The amplitude ratio in two dimensions. The data correspond to $n = 50$ (□), 100 (○), 200 (△), 400 (+), and 800 (×).

we get $\nu = 0.6368$, both with acceptable χ^2-statistics. These values are within the confidence intervals of (4.15), and we conclude that uncertainties in the value of Δ gives an unimportant contribution to the systematic error in (4.15).[7] The best estimate for ν is remarkably close to the values of ν for lattice trees obtained elsewhere (see ref. 16 and references therein).

The intersection in the family of curves is harder to determine. The three curves corresponding to $n = 50$, 100, and 200 intersect almost exactly in one point; if $\beta = 0.69$, then the values of the amplitude ratio are 2.1101, 2.1086, and 2.1116, respectively, for $n = 50$, 100, and 200. Consequently, we take $\beta_c = 0.69$ and the value of the amplitude ratio 2.110. We estimate confidence intervals by comparing the data from $n = 400$ and 800 to these values and taking the largest differences. This gives confidence intervals of size 0.03 in β_c and 0.03 in the exponent, from which we calculate ν_c:

$$\beta_c = 0.69 \pm 0.03 \tag{4.16}$$

$$\nu_c = 0.54 \pm 0.03 \tag{4.17}$$

which includes the previously estimated values of β_c [Eqs. (3.14) and (4.13)] within the confidence intervals.

Three Dimensions. With the two-dimensional work done, we had the benefit of hindsight and used method (II) of Section 4.1 in preference to method (I) for our three-dimensional work. For $n = 50$, 100, 200, this worked fine, and two umbrellas were always enough to get good coverage from $\beta = 0$ up to $\beta = 0.65$ (note that β_c is much smaller in three dimensions than in two, so we actually got farther into the collapsed regime in three dimensions than in two). For $n \geqslant 400$, we used method (III) as well. Two umbrellas eventually sufficed for $n = 400$ and for $n = 800$. However, for $n = 1600$, we were unable to get credible results for $\beta > 0.43$; in the end, two umbrellas were used for the reduced interval of β's.

The specific heat data for trees in three dimensions are plotted in Fig. 6. The peaks are narrower compared to two dimensions, and the location of the maximum decreases as n increases, in contrast to the results in two dimensions. We list the locations of the peak maximum and the heights in Table IV. Again, the results are consistent with the Robbins–Monro results of Table II (with larger error bars here), except for the H_{1600}, which is larger here than in Table II by about two error bars.

[7] A fit with $\Delta = 0.333$ gives $\nu = 0.640$ and with $\Delta = 3.000$ gives $\nu = 0.636$; both these values are within the stated error bars. In this study we sampled on large trees, and we expect that corrections to scaling will be unimportant at larger values of n, unless Δ is very close to 0. Since the available evidence indicates that Δ is not close to 0 (see ref. 16 and references therein), we do not expect Δ to play a significant role in the analysis.

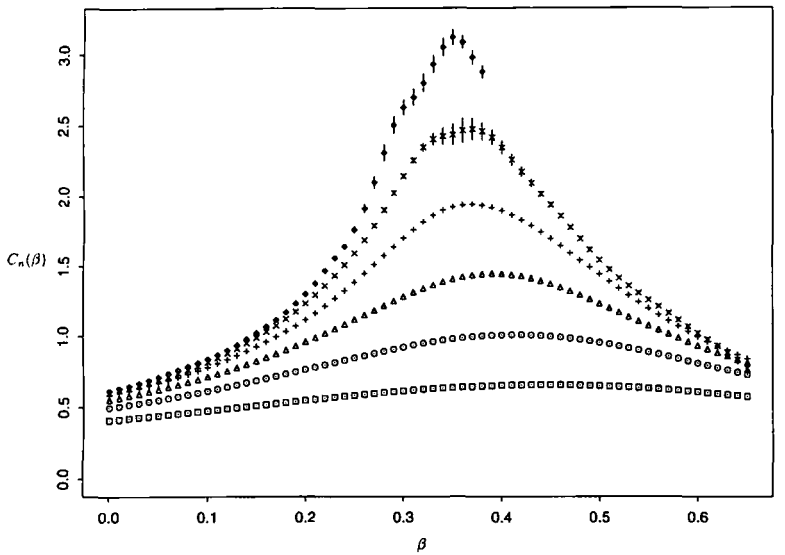

Fig. 6. The specific heat in two dimensions as a function of β. The data correspond to $n = 50$ ($\square$), 100 ($\bigcirc$), 200 ($\triangle$), 400 ($+$), 800 ($\times$), and 1600 ($\diamond$).

If we assume again that $\log H_n = \log C_0 + b \log n + Dn^{-1}$, then a least squares analysis of the heights in the peaks of the specific heat gives $b = 0.36641 \pm 0.0084$, but χ_3^2 is large, and this is not a good fit. We attempt a second fit with the data point at $n = 50$ left out to obtain $b = 0.312 \pm 0.015$, $\chi_2^2 \approx 2.5$, which is a good fit. A third fit with a second data point ($n = 100$) ignored gives $b = 0.286 \pm 0.038$, and we conclude that (using the hyperscaling relation)

$$\phi = 0.656 \pm 0.008 \pm 0.013 \tag{4.18}$$

Table IV. Umbrella Sampling Results in Three Dimensions

n	$\beta_c(n)$	H_n
50	0.4435(47)	0.6555(16)
100	0.4124(45)	1.0084(24)
200	0.3903(74)	1.4454(52)
400	0.366(22)	1.949(27)
800	0.369(63)	2.475(95)
1600	0.352(19)	3.125(71)

which includes our previous estimate in Eq. (3.15) within its confidence intervals. We can similarly estimate β_c: Assume that $\beta_c(n) = \beta_c + Cn^{-\phi}$; then a least square fit gives (with $\phi = 0.656$) $\beta_c = 0.352 \pm 0.009$, $\chi_3^2 \approx 2.8$, and at the limits of the confidence interval of ϕ we obtain $\beta_c = 0.353 \pm 0.009$, $\chi_3^2 \approx 2.9$ ($\phi = 0.662$) and $\beta_c = 0.352 \pm 0.009$, $\chi_3^2 \approx 2.7$ ($\phi = 0.650$). We neglected the data point at $n = 50$ to obtain $\beta_c = 0.345 \pm 0.015$, $\chi_2^2 \approx 1.0$. Thus, our best estimate for β_c is

$$\beta_c = 0.352 \pm 0.009 \pm 0.007 \tag{4.19}$$

which is barely inside the the confidence interval of (3.16). We note that the estimate 0.333 from (3.16) lies outside the confidence interval of (4.19).

The mean square radius of gyration data are plotted in Fig. 7 as a function of β. We observe a sharp decrease in the size of the trees, for larger n values, as β is increased beyond a certain value. The amplitude ratios are plotted in Fig. 8 and provide strong evidence for a collapse transition. The family of curves all cross at $\beta_c = 0.29 \pm 0.01$ where they have value 1.74 ± 0.01, estimated as in two 2 dimensions. The family of curves converges at $\beta = 0$ as discussed in the case for two dimensions, and we can estimate ν by extrapolating the data, using the value for the confluent correction obtained from span data in ref. 16 ($\Delta = 0.736$). We were not able to obtain a good fit, even by neglecting data points at $n = 50$ and $n = 100$. An examination of the residuals indicated that the data point at $n = 800$ gives by far the largest contribution to the χ^2-statistic, which suggests that this data point is an outlier. Neglecting it gives $\chi_2^2 \approx 4.4$, which is a good fit, giving $\nu = 0.4967 \pm 0.0011$. However, the value of Δ is very uncertain. If we repeat the fit with Δ taken to be equal to 0.5, then $\nu = 0.5001 \pm 0.0014$, and putting $\Delta = 1$ gives $\chi_2^2 = 8.4$, which is not a good fit.[8] Taken together, we have our best estimate as

$$\nu = 0.4967 \pm 0.0011 \pm 0.0034 \tag{4.20}$$

This is consistent with the generally accepted value of 0.5 for ν.[41]

At the point of intersection in Fig. 8 we obtain

$$\beta_c = 0.29 \pm 0.01 \tag{4.21}$$

$$\nu_c = 0.400 \pm 0.005 \tag{4.22}$$

The value of β_c is considerably lower than the values obtained in the Robbins–Monro simulation or in (4.19). However, we do not wish to

[8] A fit with $\Delta = 0.333$ gives $\nu = 0.506$ and with $\Delta = 3.000$ gives $\nu = 0.493$; both values are within the stated error bars.

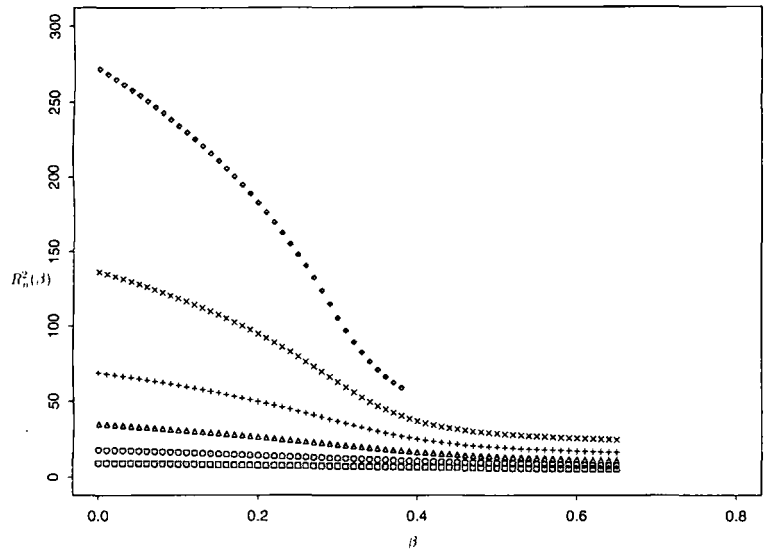

Fig. 7. The mean square radius of gyration in three dimensions as a function of β. The data correspond to $n = 50$ ($\square$), 100 ($\bigcirc$), 200 ($\triangle$), 400 ($+$), 800 ($\times$), and 1600 ($\diamond$).

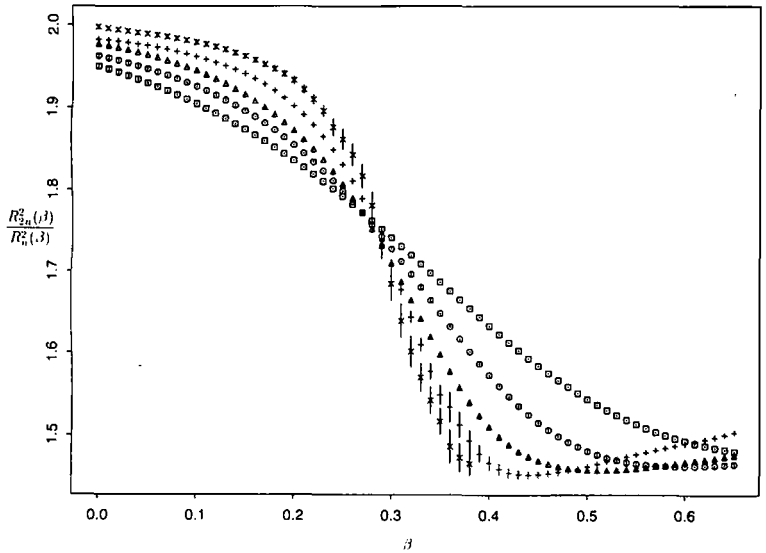

Fig. 8. The amplitude ratio in three dimensions as a function of β. The data correspond to $n = 50$ ($\square$), 100 ($\bigcirc$), 200 ($\triangle$), 400 ($+$), and 800 ($\times$).

overemphasize the significance of this. It may well be the case that simulations at even larger values of n will produce estimates of β_c that are closer together. Thus, the metric collapse is not far enough from the peak in the specific heat to be considered strong evidence against the transition being of second order.

Metric Scaling. The scaling assumption for the mean square radius of gyration is $R_n^2(\beta) \sim n^{2\nu_c} h_1(n^\phi \tau)$, where $h_1(x)$ is a suitable scaling function [see Eq. (2.11)]. It is therefore not unreasonable to expect that a plot of $R_n^2(\beta)$ against $n^\phi \tau$ will reveal the shape of the scaling function $h_1(x)$. We found, however, that correction to scaling tends to obscure $h_1(x)$, and it proved not practical to demonstrate the scaling of the mean square radius of gyration in this manner. Thus, we attempted to "cancel" the corrections to scaling by plotting the ratio $R_{2n}^2(\beta)/R_n^2(\beta)$ against $n^\phi \tau$; since

$$\frac{R_{2n}^2(\beta)}{R_n^2(\beta)} \sim 2^{2\nu_c} \frac{h_1(2^\phi n^\phi \tau)}{h_1(n^\phi \tau)} \tag{4.23}$$

such a plot will reveal the ratio $h_1(2^\phi x)/h_1(x)$. These plots can give additional support for the analysis in the preceding paragraphs, since we can choose ϕ and β_c independently in order to have the best coincidence (as measured by an "eyeball" test) of the curves for $n = 50$, 100, 200, 400, and 800 in (4.23). Our best effort in two dimensions produces Fig. 9, with the choices

$$\beta_c = 0.69, \qquad \phi = 0.58 \tag{4.24}$$

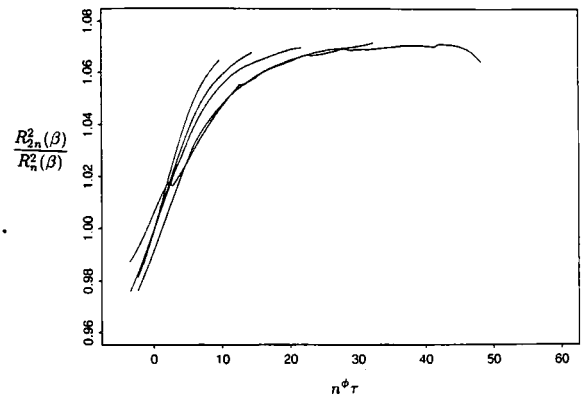

Fig. 9. Plot of $R_{2n}^2(\beta)/R_n^2(\beta)$ against $n^\phi \tau$ with $\beta_c = 0.69$ and $\phi = 0.58$ in two dimensions.

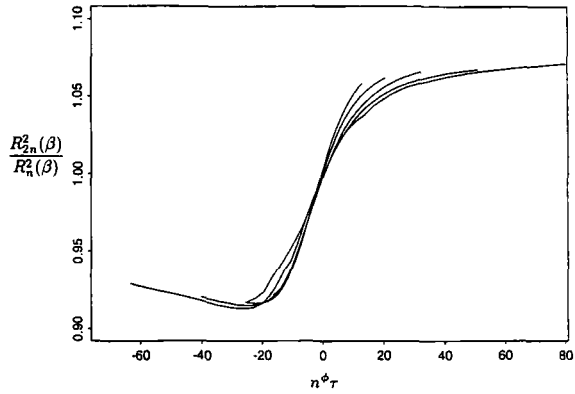

Fig. 10. Plot of $R^2_{2n}(\beta)/R^2_n(\beta)$ against $n^{\phi}\tau$ with $\beta_c = 0.29$ and $\phi = 0.66$ in three dimensions.

In three dimensions a similar analysis gives (Fig. 10)

$$\beta_c = 0.29, \qquad \phi = 0.66 \tag{4.25}$$

These results strongly support the assumption (2.11) and the analysis based on it. One can graph similar plots for the specific heat, but we found that corrections to scaling are so strong in that case that we cannot make independent estimates of ϕ and β_c from those data in a sensible way.

5. CONCLUSIONS

In this paper we have studied the collapse transition of lattice trees with the assumption that it is a second-order transition characterized by a divergent specific heat. The totality of our results strongly supports this view, but the exponent α is small enough that we cannot completely rule out the possibility that the specific heat may be divergent as a logarithm of the mass of the tree, or that the specific heat remains bounded. The evidence supporting our assumption includes two independent studies using two different methods (Robbins–Monro and umbrella sampling), and with the exception of the disparity in the location of the critical point in three dimensions, with completely consistent results.

There are several other studies reported in the literature which are devoted to the collapse transition of lattice trees or lattice animals. In two dimensions, the crossover exponent has been found to be $\phi = 0.657 \pm 0.025$ for lattice animals (transfer matrix[4]) and $\phi = 0.60 \pm 0.03$ for lattice trees (exact enumeration[11]). It was also found that $\beta_c = 0.5 \pm 0.1$ for lattice

trees.[11] (The uncertainties in these estimates were extrapolated from finite values of n, and are not 95% confidence intervals). The value of ϕ as estimated in ref. 11 is consistent with our estimates, which are 0.569 ± 0.028 (Robbins–Monro simulation) and 0.580 ± 0.011 (umbrella sampling). The locations of the critical point are also consistent; we obtained 0.693 ± 0.045 (Robbins–Monro) and 0.709 ± 0.057 (umbrella sampling). We can take a weighted average of these estimates to obtain our best values, shown in Table V.

The estimation of the metric exponent at $\beta = 0$ in two dimensions, using data from the umbrella simulation, gives a surprisingly good result: $\nu = 0.637 \pm 0.006$, which compares with the best current MC estimates for this exponent.[16] We can also estimate β_c from the metric data by estimating the intersection of the family of amplitude ratios in Fig. 8, which we note is close the the values estimated by the Robbins–Monro and umbrella sampling simulations. This intersection of amplitude ratios indicates a change in the metric properties of the trees, and the fact that it is so closely correlated with the peak in the specific heat data can be taken as strong support for a second-order transition. We estimate the metric exponent at the critical point to be $\nu_c \approx 0.54$, while we know that it is 0.5 in the collapsed phase. These results support the notion that the critical point is a tricritical point with coexistence of three phases: an *expanded phase* with $\nu \approx 0.64$, a *Θ-phase* with $\nu_c \approx 0.54$, and a *collapsed phase* with $\nu = 0.5$.

In two dimensions, the identification of collapsed animals with vesicles[42] gives $\phi = 2/3 \approx 0.66$. This value agrees closely with the Derrida and Herrmann estimate[4] ($\phi = 0.657 \pm 0.025$) for collapsing animals. In contrast to this, the collapsing tree crossover exponent is estimated to be about 0.58 [ref. 11 and our estimate (3.13)]. An important difference in these models is the nature of the collapsed phase; the animal models which are given a cycle fugacity generally collapse to section graphs (rich in

Table V. Best Estimates

Two dimensions	
β_c	0.699 ± 0.052
ϕ	0.579 ± 0.022
α	0.24 ± 0.11
ν_c	0.54 ± 0.03
Three dimensions	
β_c	0.346 ± 0.017
ϕ	0.655 ± 0.024
α	0.47 ± 0.10
ν_c	0.400 ± 0.005

cycles), while the tree models have a contact fugacity and collapse to spanning trees of compact regions of the lattice (see Refs. 19, 20, and 43 for a discussion of differences in a collapse to spanning trees and to section graphs (or "vesicles")]. In view of these results, and our own, we conclude that the value of the crossover exponent measured in any simulation will be determined by the nature of the collapsed phase. If the collapse is to spanning trees, then we expect $\phi \approx 0.58$, and if it is to section graphs, then $\phi \approx 0.66$. This strongly suggests the existence of two collapsed phases, namely a spanning tree phase and a section graph (or vesicle) phase. Coexistence of these phases with the expanded phase of lattice animals (with exponents identical to those of lattice trees) occurs at a multicritical point, which is believed to be the percolation point.[43,44] The mapping of a model of interacting branch polymers to the Potts model[45] has produced some predictions for various exponents. In particular, the crossover exponent for collapse to spanning trees (this occurs in the so-called contact model) is predicted to be determined by the Ising critical point, which gives $\phi = 8/15 \approx 0.53$. This is outside the confidence interval in our best estimate in Table V. On the other hand, the prediction for the metric exponent of the Θ-phase is $\nu_c = 8/15 \approx 0.53$, which is close to our result.

There is a wide range of estimates for the crossover exponent in three dimensions. Lam[8] estimated that $\phi \approx 0.814$ by a Monte Carlo simulation, while Chang and Shapir[9] estimated $\phi \approx 1$, by exact enumeration. The exact enumeration study of Gaunt and Flesia[11] estimated $\phi = 0.82 \pm 0.03$ and, less precisely, $\beta_c = 0.35 \pm 0.30$. (As above, these uncertainties are extrapolated from finite n data, and are not 95% confidence intervals.) A study which related collapsing animals to surfaces gives $\phi = 1$.[46] Our results from the Robbins–Monro and the umbrella sampling data agree within their confidence intervals, so we combine them to arrive at our best estimates (see Table V). Our value $\phi = 0.655 \pm 0.024$ for the crossover exponent is considerably lower than the results stated above, and the agreement between our estimate $\beta_c = 0.346 \pm 0.017$ and the estimate in ref. 11 should be taken as a coincidence, since the latter has such a large confidence interval. It is less clear from these results that there are two values of the crossover exponent, determined by the nature of the collapsed phase (as we conjectured for two dimensions). However, these data suggests $\phi \approx 1$ for a collapse to the section graphs. The results are less sure for a collapse to the spanning tree phase. We obtained $\phi \approx 0.66$, but the studies by Lam[8] and Gaunt and Flesia[11] suggests a value close to 0.80.

We were again able to compute an accurate value for the metric exponent of the expanded phase, $\nu = 0.4967 \pm 0.0045$. This is in good agreement with the accepted exact value of 0.5.[41] The amplitude ratios intersect each other in Fig. 8 at $\beta_c \approx 0.29$, which is slightly less than the best estimate we

have for the critical point from thermodynamic data in Table V. This discrepancy is not big, and with increasing n it is quite possible that the point of intersection may increase slightly to coincide with the value in Table V. On the other hand, we favor the results obtained from the metric data, since those signal the transition so clearly. There is a strong correlation, then, between the metric collapse and the divergence of the specific heat, exactly as we observed in two dimensions, and we interpret this as evidence for a second order transition. The value of the metric exponent at the critical point is $\nu_c = 0.400 \pm 0.005$, and in the collapsed phase we expect $\nu = 1/3$. Our best estimate for ν_c excludes the value obtained by Stella *et al.*[46] where the association of surfaces to lattice animals is conjectured to imply that $\nu_c = 0.5$.

APPENDIX. STATISTICAL PROCEDURES

The statistical procedures of Section 4 for estimating quantities and their error bars can be derived from the following standard situation. Suppose that $Y_1, Y_2, \ldots, Y_k$ are observations from a stationary (time-homogeneous) random process. In our case, these are functionals $Y_i = g(X_i)$ of a Markov chain $X_1, X_2, \ldots$ that has reached equilibrium. We estimate the mean $\langle Y_1 \rangle$ by the sample mean

$$\bar{Y}_k = \frac{1}{k} \sum_{i=1}^{k} Y_i \tag{A.1}$$

whose variance is given (asymptotically as $k \to \infty$) by

$$\mathrm{Var}(\bar{Y}_k) \sim \frac{1}{k} \left(\mathrm{Var}(Y_1) + 2 \sum_{i=1}^{\infty} \mathrm{Cov}(Y_1, Y_{i+1}) \right) \tag{A.2}$$

In our case, $\bar{Y}_k$ will always be asymptotically normal (since Y_i is always be a functional of an ergodic Markov chain), so our 95% confidence interval for $\langle Y \rangle$ will be $\bar{Y}_k \pm 2\hat{\sigma}(\bar{Y}_k)$, where $\hat{\sigma}(\bar{Y}_k)$ is the square root of an estimate for the variance (A.2), which we discuss in the next paragraph.

There are several ways to estimate (A.2). The simplest way is by "batching": divide the run into several long pieces. and assume (i.e., hope) that each piece is independent. Indeed, trying different batch sizes is one way to check for long autocorrelation times. Alternatively, one can use the methods of time series analysis to estimate each term in the sum on the right side of (A.2) for i up to some cutoff M, where M is chosen to be "sufficiently large" (see Appendix C of ref. 47 for discussion). We used slightly different methods at different times in the present study, but we regularly

checked our error bars by running two methods on the same data (usually the batching was done as a check on the time series). Moreover, we looked at the behaviour of the estimates of $\mathrm{Cov}(Y_l, Y_{i+l})$ for $i \geqslant 1$ until these estimates became indistinguishable from 0 due to noise (one sign of this is a balanced number of positive and negative estimates); this helped convince us that our choice of M was indeed large enough. The covariances often decayed very slowly, which could be due either to slow mixing or to a reweighting of the umbrella sampling data that put most of the weight on a small fraction of what was observed. If the decay was too slow, then we would try another umbrella distribution, as described in Section 4.

Observe that for any function f on the set of trees T_n and any β the estimator $S_k[f;\beta]$ defined in Eq. (4.2) is of the form (A.1) with

$$Y_i = \frac{f(X_i)}{\pi(X_i)} e^{\beta c(X_i)} \tag{A.3}$$

Moreover, let

$$\varepsilon[f;\beta] = S_k[f;\beta] - \langle f \rangle \frac{\mathscr{Z}_n(\beta)}{\mathscr{Z}_\pi} \tag{A.4}$$

then $\varepsilon[f;\beta]$ is asymptotically normal, with mean 0, and has magnitude of order $1/\sqrt{k}$. Therefore, we can approximate $R_k[f;\beta]$ [see Eq. (4.5)] using a Taylor approximation and neglecting terms of order $1/k$, as follows:

$$\begin{aligned} R_k[f;\beta] &= \frac{\langle f \rangle_{\beta,n} (\mathscr{Z}_n(\beta)/\mathscr{Z}_\pi) + \varepsilon_k[f;\beta]}{\mathscr{Z}_n(\beta)/\mathscr{Z}_\pi + \varepsilon[1;\beta]} \\ &\approx \langle f \rangle_{\beta,n} + \frac{\mathscr{Z}_\pi}{\mathscr{Z}_n(\beta)} \varepsilon_k[f;\beta] - \frac{\mathscr{Z}_\pi}{\mathscr{Z}_n(\beta)} \langle f \rangle_{\beta,n} \varepsilon_k[1;\beta] \\ &= \langle f \rangle_{\beta,n} + \frac{\mathscr{Z}_\pi}{\mathscr{Z}_n(\beta)} \varepsilon_k[f - \langle f \rangle_{\beta,n};\beta] \end{aligned} \tag{A.5}$$

Therefore, for error bars on our estimate of $\langle f \rangle_{\beta,n}$, we obtain $\hat{\sigma}(R_k[f;\beta])$ by applying the above time series methods to the second term in the last line of (A.5), i.e.,to $\bar{Y}_k$ with

$$Y_i = \frac{\mathscr{Z}_\pi}{\mathscr{Z}_n(\beta)} (f(X_i) - \langle f \rangle_{\beta,n}) \frac{e^{\beta c(X_i)}}{\pi(X_i)} \tag{A.6}$$

but in practice we cannot evaluate this exactly, so we instead perform the time series analysis on its approximation:

$$Y_i = \frac{1}{S_k[1;\beta]} (f(X_i) - R_k[f;\beta]) \frac{e^{\beta c(X_i)}}{\pi(X_i)} \tag{A.7}$$

Our error bars on the specific heat estimates in Section 4 are obtained this way, with $f(X) = (c(X) - R_k[c;\beta])^2 \approx (c(X) - \langle c \rangle_{\beta,n})^2$.

The error bars on our estimates of $\beta_c(n)$ in Tables III and IV in Section 4 also come from a linearization, but the derivation is a bit more subtle. We estimate $\beta_c(n)$ by $\hat{\beta}_k$, which is the value of β that maximizes our estimate of the specific heat; i.e., $\hat{\beta}_k$ maximizes $R_k[f;\beta]$, where

$$f(X) = f(X,\beta) = (c(X) - R_k[c;\beta])^2 \tag{A.8}$$

Let

$$U_k[f;\beta] = \frac{d}{d\beta} R_k[f;\beta] = R_k[cf + f';\beta] - R_k[f;\beta]\, R_k[c;\beta] \tag{A.9}$$

where $f' = df/d\beta$ (to obtain (A.9), we use $dS_k[f;\beta]/d\beta = S_k[cf + f';\beta]$). The definitions now imply that

$$0 = U_k[f;\hat{\beta}_k] \tag{A.10}$$

where f is given by Equation (A.8) with $\beta = \hat{\beta}_k$. To derive approximate asymptotics for the error $\hat{\beta}_k - \beta_c(n)$, we linearize Eq. (A.10) at $\beta = \beta_c(n)$:

$$0 \approx U_k[f;\beta_c(n)] + \frac{d}{d\beta} U_k[f;\beta_c(n)](\hat{\beta}_k - \beta_c(n)) \tag{A.11}$$

which leads to

$$\hat{\beta}_k - \beta_c(n) \approx -\frac{U_k[f;\beta_c(n)]}{(d/d\beta)\, U_k[f;\beta_c(n)]} \tag{A.12}$$

For general f, one can compute

$$\frac{d}{d\beta} U_k[f;\beta_c(n)] = R_k[c^2 f + 2cf' + f'';\beta] - 2R_k[cf + f';\beta]\, R_k[c;\beta] + 2R_k[f;\beta]\, R_k[c;\beta]^2 - R_k[f;\beta]\, R_k[c^2;\beta] \tag{A.13}$$

For our choice of f in Eq. (A.8), one can routinely check that

$$f' = 2(c - R_k[c;\beta])(R_k[c;\beta]^2 - R_k[c^2;\beta])$$

$$R_k[f';\beta] = 0 \qquad \text{for every } \beta \tag{A.14}$$

and

$$R_k[f'';\beta] = 2(R_k[c^2;\beta] - R_k[c;\beta]^2)^2 \tag{A.15}$$

Therefore

$$\frac{d}{d\beta} U_k[f;\beta] = R_k[(c - R_k[c;\beta])^4;\beta] - 3R_k[(c - R_k[c;\beta])^2;\beta]^2 \tag{A.16}$$

and as $k \to \infty$ this converges (with probability one) to

$$\langle (c - \langle c \rangle_{\beta,n})^4 \rangle_{\beta,n} - 3\langle (c - \langle c \rangle_{\beta,n})^2 \rangle_{\beta,n}^2 = \frac{d^4}{d\beta^4} \log \mathscr{Z}_n(\beta) \tag{A.17}$$

Therefore, assuming that the expression in (A.17) is nonzero, we can estimate the denominator of (A.12) by the expression of (A.16) evaluated at $\beta = \hat{\beta}_k$. Now, the numerator of (A.12) is 0 at $\beta = \hat{\beta}_k$ by definition [recall (A.10)], but we can still proceed. Equations (A.9) and (A.14) imply that for our choice of f,

$$U_k[f;\beta] = R_k[(c - R_k[c;\beta])^3;\beta] \tag{A.18}$$

so $U_k[f;\beta]$ is of the basic form $\bar{Y}_k$ with

$$Y_i = (c(X_i) - R_k[c;\beta])^3 \frac{e^{\beta c(X_i)}}{\pi(X_i)} \tag{A.19}$$

So the usual methods can now be employed to obtain $\hat{\sigma}(U_k[f;\beta])$. Finally, then, the approximate 95% confidence interval for $\beta_c(n)$ is

$$\hat{\beta}_k \pm 2 \frac{\hat{\sigma}(U_k[f;\hat{\beta}_k])}{(d/d\beta)\, U_k[f;\hat{\beta}_k]} \tag{A.20}$$

where the denominator is evaluated using the right-hand side of (A.16) at $\beta = \hat{\beta}_k$.

Remark. The second half of this Appendix is an adaptation of the standard proof of the symptotic normality of the "maximum likelihood estimator" in statistics [see, for example, Section 9.2(iii) of Cox and Hinkley[48]]; an application in a very similar context is in Theorem 7 of Geyer in ref. 49. We are grateful to Charlie Geyer for telling us about this.

ACKNOWLEDGMENTS

We are grateful to Richard Brak, Charlie Geyer, Peter Glynn, Enzo Orlandini, Andrea Pellissetto, Mauro Piccioni, Alan Sokal, Carla Tesi, John Valleau, and Stu Whittington for helpful conversations, comments, and suggestions. Our research was supported in part by Research Grants from the Natural Sciences and Engineering Research Council of Canada. The computations reported in this paper were all done on an HP730 purchased with an Equipment Grant to the authors from NSERC Canada.

REFERENCES

1. P. J. Flory, *Principles of Polymer Chemistry* (Cornell University Press, Ithaca, 1953).
2. J. Mazur and F. L. McCrackin, *J. Chem. Phys.* **49**:648 (1968).
3. J. Mazur and D. McIntyre, *Macromolecules* **8**:464 (1975).
4. B. Derrida and H. J. Herrmann, *J. Phys. (Paris)* **44**:1365 (1983).
5. R. Dickman and W. C. Schieve, *J. Phys.* **45**:1727 (1984).
6. R. Dickman and W. C. Schieve, *J. Stat. Phys.* **44**:465 (1986).
7. P. M. Lam, *Phys. Rev. B* **36**:6988 (1987).
8. P. M. Lam, *Phys. Rev. B* **38**:2813 (1988).
9. I. S. Chang and Y. Shapir, *Phys. Rev. B* **38**:6736 (1988).
10. D. S. Gaunt and S. Flesia, *Physica A* **168**:602 (1990).
11. D. S. Gaunt and S. Flesia, *J. Phys. A: Math. Gen.* **24**:3655 (1991).
12. S. Flesia and D. S. Gaunt, *J. Phys. A: Math. Gen.* **25**:2127 (1992).
13. N. Madras, C. E. Soteros, S. G. Whittington, J. L. Martin, M. F. Sykes, S. Flesia, and D. S. Gaunt, *J. Phys. A: Math. Gen.* **23**:5327 (1990).
14. D. C. Rapaport, *J. Phys. A: Math. Gen.* **10**:637 (1977).
15. H. A. Lim, A. Guha, and Y. Shapir, *Phys. Rev. A* **38**:371 (1988).
16. E. J. Janse van Rensburg and N. Madras, *J. Phys. A: Math. Gen.* **25**:303 (1992).
17. G. M. Torrie and J. P. Valleau, *J. Comput. Phys.* **23**:187 (1977).
18. S. G. Whittington, C. E. Soteros, and N. Madras, *J. Math. Chem.* **7**:87 (1991).
19. S. Flesia, D. S. Gaunt, C. E. Soteros, and S. G. Whittington, *J. Phys. A: Math. Gen.* **25**:3515 (1992).
20. S. Flesia, D. S. Gaunt, C. E. Soteros, and S. G. Whittington, *J. Phys. A: Math. Gen.* **26**:L993 (1992).
21. R. Brak, A. Guttmann, and S. G. Whittington, *J. Phys. A: Math. Gen.* **23**:4581 (1992).
22. A. L. Owczarek, *J. Stat. Phys.* **72**:737 (1993).
23. D. Dhar, *Phys. Rev. Lett.* **51**:853 (1983).
24. D. Dhar, *Physica A* **140**:210 (1986).
25. K. De'Bell and S. G. Whittington, Unpublished (1992).

26. R. Brak, Unpublished (1994).
27. H. Robbins and S. Monro, *Ann. Math. Stat.* **22**:400 (1951).
28. G. Kersting, *Ann. Prob.* **5**:954 (1977).
29. P. W. Glynn, In *Proceedings of the 1986 Winter Simulation Conference*, J. Wilson, J. Henriksen, and S. Roberts, eds., p. 52.
30. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, *J. Chem. Phys.* **21**:1087 (1953).
31. J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Chapman and Hall, London, 1964).
32. J. P. Valleau, *J. Comput. Phys.* **96**:193 (1991).
33. A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **61**:2635 (1988).
34. A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.* **63**:1195 (1989).
35. C. J. Geyer and E. A. Thompson, *J. Royal Stat. Soc. B* **54**:657 (1992).
36. E. Marinari and G. Parisi, *Europhys. Lett.* **19**:451 (1992).
37. N. Madras and M. Piccioni, Preprint (1994).
38. C. J. Geyer and E. A. Thompson, *J. Am. Stat. Assoc.* **90**:909 (1995).
39. B. A. Berg and T. Neuhaus, *Phys. Lett. B* **267**:249 (1991).
40. C. J. Geyer, Preprint (1993).
41. G. Parisi and N. Sourlas, *Phys. Rev. Lett.* **46**:871 (1981).
42. C. Vanderzande, *Phys. Rev. Lett.* **70**:3595 (1993).
43. S. Flesia, D. S. Gaunt, C. E. Soteros, and S. G. Whittington, *J. Phys. A: Math. Gen.* **27**:5831 (1994).
44. S. G. Whittington, Private communication (1994).
45. F. Seno and C. Vanderzande, Preprint (1995).
46. A. L. Stella, E. Orlandini, I. Beichl, F. Sullivan, M. C. Tesi, and T. L. Einstein, *Phys. Rev. Lett.* **69**:3650 (1992).
47. N. Madras and A. D. Sokal, *J. Stat. Phys.* **50**:109 (1988).
48. D. R. Cox and D. V. Hinkley, *Theoretical Statistics* (Chapman and Hall, London, 1974).
49. C. J. Geyer, *J. R. Stat. Soc. B* **56**:261 (1994).