

Statistics of collapsed lattice animals: rigorous results and Monte Carlo simulations

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1996 J. Phys. A: Math. Gen. 29 7067

(<http://iopscience.iop.org/0305-4470/29/22/010>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.70

The article was downloaded on 02/06/2010 at 04:04

Please note that [terms and conditions apply](#).

Statistics of collapsed lattice animals: rigorous results and Monte Carlo simulations

L M Stratychuk† and C E Soteris

Department of Mathematics, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
S7N 5E6

Received 30 April 1995, in final form 2 August 1996

Abstract. In this paper a two-dimensional model of collapsed lattice animals with nearest-neighbour contact interactions is introduced and the limiting free energy for the model is proved to exist. We also establish that the partition function for collapsed lattice animals is related to the reliability polynomial for the $m \times m$ grid graph and to the generating function for spanning forests on the $m \times m$ grid graph. From these relationships and using results about directed animals we are able to obtain rigorous upper and lower bounds on the free energy for collapsed lattice animals. In addition we present Monte Carlo evidence that the heat capacity associated with this model remains bounded in the thermodynamic limit. In fact we do not find any evidence for a second-order phase transition for collapsed lattice animals.

1. Introduction

Understanding the statistical properties of lattice animals is important to the study of models of branched polymers and random-cluster models. Recently, information about the statistics of lattice animals has been used to make predictions about the collapse transition in branched polymers. In particular, Flesia *et al* (1992, 1994) have studied a general lattice animal model of branched polymer molecules in dilute solution which includes both monomer–monomer and monomer–solvent interactions. Based on the analysis of exact enumeration data they have predicted a curve of phase transitions corresponding to the ‘collapse’ of lattice animals from expanded to compact objects. This collapse can be driven by either strengthening the attraction between monomers or weakening the attraction between the monomer and solvent molecules. By making a connection with bond percolation, they proved that the collapse transition exists at least at one point along the predicted curve of phase transitions. In addition their results indicate that there are peaks in the heat capacity along a curve within the ‘collapsed’ regime suggesting the possibility of a collapse–collapse phase transition. If such a phase transition exists, it would be characterized by a transition from compact objects with few cycles to compact objects with many cycles. The goal of this paper is to investigate further the statistical properties of lattice animals in the collapsed regime, we introduce a model of collapsed lattice animals and study it for the square lattice by rigorous analysis of the limiting free energy of the model and by Monte Carlo simulation. The model developed here corresponds to the case $\beta_1 = -\infty$ (a nearest-neighbour infinite repulsion between monomer and solvent molecules) in the Flesia *et al* two-variable model and to $Q = 0$ in the random-cluster model for the hypercubic lattice Z^d . In particular, there is a

† Current address: Statistics Canada, Ottawa, Ontario, Canada K1A 0T6.

direct relationship between the partition function of the model and a reliability polynomial for the underlying lattice. We also show that the limiting free energy of our model is equal to the exponential growth rate of the generating function for spanning forests on certain subsets of the square lattice. We begin by defining the model and proving the existence of its limiting free energy. Next we make the appropriate connections to other models. We then obtain rigorous bounds on the limiting free energy and finally we discuss the results of the Monte Carlo simulation.

2. A model of collapsed lattice animals

A *lattice animal* is defined to be any finite connected subgraph of the d -dimensional hypercubic lattice Z^d . Consider a lattice animal with n vertices, cyclomatic index c , and e edges. These three quantities are related through Euler's relation

$$c = e - n + 1. \quad (2.1)$$

If a pair of vertices in the animal are near neighbours on the lattice and are not incident on a common edge of the animal, we call this pair of vertices a *contact*. Let the number of contacts in the animal be k . If an edge of the lattice is not an edge of the animal but is incident on exactly one vertex in the animal, we call this edge a *solvent contact* and write s for the number of solvent contacts associated with the animal. Counting edges gives the relation

$$2dn = 2e + 2k + s \quad (2.2)$$

and using (2.1),

$$s + 2k + 2c = 2(d - 1)n + 2. \quad (2.3)$$

Let $a_n(s, k)$ be the number (per lattice site) of animals with n vertices, k contacts and s solvent contacts. In Flesia *et al* (1992) the following two-variable partition function was defined

$$Z_n(\beta_1, \beta_2) = \sum_{s, k} a_n(s, k) e^{\beta_1 s + \beta_2 k} \quad (2.4)$$

and the limit

$$\lim_{n \rightarrow \infty} n^{-1} \log Z_n(\beta_1, \beta_2) = G(\beta_1, \beta_2) \quad (2.5)$$

was shown to exist and to be a convex function of β_1 and β_2 and continuous for finite β_1 and β_2 . These results hold for any hypercubic lattice Z^d . In the remainder of this paper we focus on the two-dimensional case, the square lattice.

In the 'collapsed' phase as β_1 goes to negative infinity, it is expected that the behaviour of $Z_n(\beta_1, \beta_2)$ will be influenced to the greatest extent by those animals with the minimum possible number of solvent contacts. Thus in order to study the collapsed region further we focus our model on those animals with the minimum possible number of solvent contacts and call these animals *collapsed*. Define $s_{min}(n) \equiv \min\{s \mid a_n(s, k) > 0\}$ and define the *partition function for collapsed animals* to be

$$Z_n(-\infty, \beta) \equiv \sum_k a_n(s_{min}(n), k) e^{\beta k}. \quad (2.6)$$

This notation is used for convenience and because of the definition above, $Z_n(-\infty, \beta_2) \neq \lim_{\beta_1 \rightarrow -\infty} Z_n(\beta_1, \beta_2) = 0$ but instead

$$Z_n(-\infty, \beta_2) = \lim_{\beta_1 \rightarrow -\infty} Z_n(\beta_1, \beta_2) e^{-\beta_1 s_{min}(n)} \geq a_n(s_{min}(n), 0). \quad (2.7)$$

The limiting free energy for this model is then defined to be

$$G(-\infty, \beta) \equiv \lim_{n \rightarrow \infty} n^{-1} \log Z_n(-\infty, \beta) \tag{2.8}$$

where the limit can be proved to exist and is finite for finite values of β . A sketch of the proof is given next and the full details are given in the appendix, sections A.1 and A.2. The existence proof relies first on using properties of lattice animals with minimum perimeter to show that for $0 \leq r \leq 2m$,

$$4m \leq s_{min}(m^2 + r) \leq 4m + 4 \tag{2.9}$$

and

$$a_{m^2}(s_{min}(m^2), k) \leq a_{m^2+r}(s_{min}(m^2 + r), k) \leq 2^{4m+o(m)} a_{M^2}(s_{min}(M^2), k) \tag{2.10}$$

where $M = m + 1 + \lfloor \sqrt{m} \rfloor$. Note that for any positive integer n there exists integers $m = \lfloor n^{1/2} \rfloor$ and r , $0 \leq r \leq 2m$, such that $n = m^2 + r$. Thus setting $m = \lfloor n^{1/2} \rfloor$, $r = n - m^2$, multiplying by $e^{\beta k}$ and summing over k in equation (2.10) gives that for any n

$$Z_{m^2}(-\infty, \beta) \leq Z_n(-\infty, \beta) \leq 2^{4m+o(m)} Z_{M^2}(-\infty, \beta). \tag{2.11}$$

The next step of the proof involves using an appropriate concatenation argument to prove that

$$\lim_{m \rightarrow \infty} m^{-2} \log Z_{m^2}(-\infty, \beta) \tag{2.12}$$

exists. The appropriate concatenation argument is similar to that used by Whittington and Guttmann (1990) for self-avoiding walks which cross a square. Equation (2.11), along with the existence of the limit in equation (2.12), guarantees the existence of the limit in equation (2.8).

We note that from the definitions in equation (2.4) and (2.6)

$$e^{\beta_1 s_{min}(n)} Z_n(-\infty, \beta_2) \leq Z_n(\beta_1, \beta_2). \tag{2.13}$$

Taking logarithms, dividing by n , letting n go to infinity in equation (2.13) and using the fact that $\lim_{n \rightarrow \infty} (s_{min}(n)/n) = 0$ (see equation (2.9)) gives

$$G(-\infty, \beta_2) \leq G(\beta_1, \beta_2) \tag{2.14}$$

for all β_1 and β_2 . We lack an appropriate upper bound to prove that

$$\lim_{\beta_1 \rightarrow -\infty} G(\beta_1, \beta_2) = G(-\infty, \beta_2) \tag{2.15}$$

however, it is expected that equation (2.15) is true.

3. Relationships with other models

In this section we establish relationships between the limiting free energy for collapsed lattice animals and functions associated with some other combinatorial problems on the square lattice. Establishing these connections allows us to obtain useful bounds on $G(-\infty, \beta)$ in the next section.

Consider the $m \times m$ grid graph, $S(m)$, to be the subgraph of the square lattice induced by the vertices $\{(i, j) \mid 0 \leq i \leq m - 1, 0 \leq j \leq m - 1\} \subset Z^2$. Note that $S(m)$ is the only (up to translation) collapsed animal with no contacts and with $n = m^2$ vertices (see the appendix, section A.1 for a justification of this). We next show that there is a direct connection between $Z_{m^2}(-\infty, \beta)$ and the reliability polynomial for $S(m)$.

The *reliability polynomial* for an n vertex graph G (see Colbourn 1987, Welsh 1993b), $R(G; p)$, is defined as

$$R(G; p) = \sum_B p^{e(B)} (1-p)^{e(G)-e(B)} \quad (3.1)$$

where the sum is over all spanning connected subgraphs B of G and $e(B)$ is the number of edges in B . If one considers deleting each edge from G with probability $1-p$ or keeping it with probability p as in a bond percolation process, then $R(G; p)$ is the probability that the resulting graph is connected. A standard way, known as the F -form (Colbourn 1987), to express $R(G; p)$ is

$$R(G; p) = \sum_{i=0}^b F_i p^{b-i} (1-p)^i \quad (3.2)$$

where $b = e(G)$ and F_i is the number of spanning connected subgraphs of G composed of $b-i$ edges. For planar graphs, algorithms for calculating F_i are known (Liu and Chow 1983, Ramesh *et al* 1987, Myrvold 1992); however, there are no known polynomial algorithms for calculating $R(G; p)$.

Consider $G = S(m)$. In this case $e(S(m)) = 2m^2 - 2m$. If we define $\mathcal{B}_{m,k}$ to be the set of spanning connected subgraphs of $S(m)$ with exactly $2m^2 - 2m - k$ edges and define $\mathcal{B}_m = \cup_{k \geq 0} \mathcal{B}_{m,k}$, then the sum in equation (3.1) is over all elements of \mathcal{B}_m and $F_k = |\mathcal{B}_{m,k}|$ in equation (3.2). Hence we obtain

$$\begin{aligned} R_m(p) \equiv R(S(m); p) &= \sum_{k \geq 0} \sum_{B \in \mathcal{B}_{m,k}} p^{2m^2-2m-k} (1-p)^k \\ &= p^{2m^2-2m} \sum_{k=0}^{m^2-2m+1} |\mathcal{B}_{m,k}| \left(\frac{1-p}{p} \right)^k. \end{aligned} \quad (3.3)$$

However, $|\mathcal{B}_{m,k}| = a_{m^2}(4m, k)$ and hence

$$\begin{aligned} R_m(p) \equiv R(S(m); p) &= p^{2m^2-2m} \sum_{k=0}^{m^2-2m+1} a_{m^2}(4m, k) \left(\frac{1-p}{p} \right)^k \\ &= p^{2m^2-2m} Z_{m^2} \left(-\infty, \log \left(\frac{1-p}{p} \right) \right). \end{aligned} \quad (3.4)$$

If we define $p = 1/(1 + e^\beta)$ or equivalently $\beta = \log(1-p/p)$, then the above equation can be rewritten as

$$Z_{m^2}(-\infty, \beta) = (1 + e^\beta)^{2m^2-2m} R_m \left(\frac{1}{1 + e^\beta} \right). \quad (3.5)$$

Taking logarithms, dividing by m^2 , and letting m go to infinity in equation (3.5) gives that

$$G(-\infty, \beta) = 2 \log(1 + e^\beta) + \lim_{m \rightarrow \infty} m^{-2} \log R_m \left(\frac{1}{1 + e^\beta} \right). \quad (3.6)$$

Since $G(-\infty, \beta)$ is finite for $-\infty < \beta < \infty$ the limit

$$\mathcal{R}(p) \equiv \lim_{m \rightarrow \infty} m^{-2} \log R_m(p) \quad (3.7)$$

exists and is finite for $0 < p < 1$. Equation (3.6) yields

$$\mathcal{R}(p) = G \left(-\infty, \log \left(\frac{1-p}{p} \right) \right) + 2 \log p. \quad (3.8)$$

The reliability polynomial $R(G; p)$ of a graph G is directly related to the Tutte polynomial $T(G; 1, 1/(1 - p))$ (see Welsh 1993b and references therein). For $G = S(m)$ the two polynomials are related through the following equation:

$$R_m(p) = p^{m^2-1}(1 - p)^{m^2-2m+1}T(S(m); 1, 1/(1 - p)). \tag{3.9}$$

This implies that

$$G(-\infty, \beta) = \beta + \lim_{m \rightarrow \infty} m^{-2} \log T(S(m); 1, 1 + e^{-\beta}). \tag{3.10}$$

For any planar graph G , $T(G; 1, 2)$ counts the number of spanning forests of the dual of G (see Brylawski and Oxley 1992, p 131). A *spanning forest* of a graph G is any subgraph (not necessarily connected) of G which contains all the vertices of G and has no cycles. Thus $G(-\infty, 0)$ is the growth constant for the total number of spanning forests of the dual of $S(m)$ in addition to being the growth constant for the total number of spanning connected subgraphs of $S(m)$. In our case, because of the self-dual nature of the square lattice, it is possible to find a relationship between $R_m(p)$, $Z_{m^2}(-\infty, \beta)$ and the generating function for the k edge spanning forests of $S(m)$. Given the graph $S(m)$, its dual graph $S^*(m)$ is obtained by placing a vertex in the middle of each square of $S(m)$ and placing a vertex v_o outside $S(m)$. Two vertices in $S^*(m)$ are then joined by an edge if the corresponding squares in $S(m)$ share a common edge and each vertex on the boundary of $S^*(m)$ is joined by an edge to v_o . Thus the vertex set of $S^*(m)$ equals the vertex set of $S(m - 1)$ union v_o and the edge set equals the edge set of $S(m - 1)$ union the edges joining the boundary of $S(m - 1)$ to v_o . Define $f_{m,k}^*$ to be the number of spanning forests of $S^*(m)$ with k edges. Define the generating function for this sequence of numbers to be

$$\mathcal{F}_m^*(\beta) = \sum_k f_{m,k}^* e^{\beta k}. \tag{3.11}$$

In addition we define $f_{m,k}$ to be the number of spanning forests of $S(m)$ with k edges and the generating function for this sequence as

$$\mathcal{F}_m(\beta) = \sum_k f_{m,k} e^{\beta k}. \tag{3.12}$$

We claim that

$$G(-\infty, \beta) = \lim_{m \rightarrow \infty} m^{-2} \log \mathcal{F}_m^*(\beta) = \lim_{m \rightarrow \infty} m^{-2} \log \mathcal{F}_m(\beta). \tag{3.13}$$

The proof of this is given in the appendix, section A.3.

Finally we note that $R_m(p)$ corresponds to the partition function for the case $Q = 0$ of the random-cluster model. In particular, given a graph G the random-cluster model (see Welsh 1993a) for G consists of a random process on the edge set, $E(G)$, of G such that the probability of choosing the edge set $A \subseteq E(G)$ is given by:

$$\mu_A(p, Q) = \frac{p^{|A|}(1 - p)^{|E(G)-A|} Q^{k(A)}}{\sum_{A \subseteq E(G)} p^{|A|}(1 - p)^{|E(G)-A|} Q^{k(A)}} = \frac{p^{|A|}(1 - p)^{|E(G)-A|} Q^{k(A)-1}}{\mathcal{Z}} \tag{3.14}$$

where $k(A)$ is the number of connected components of the edge set A and where

$$\mathcal{Z} = \sum_{A \subseteq E(G)} p^{|A|}(1 - p)^{|E(G)-A|} Q^{k(A)-1}$$

is the partition function. If Q is a positive integer then the random-cluster model is related to the Q state Pott's model and in the special case $Q = 1$ the random-cluster model corresponds to bond percolation on the edge set $E(G)$. If one lets Q and p go to zero at the same rate (for example if $Q = \lambda p$) then, as pointed out by Fortuin and Kastelejan

(1972) and Seno and Vanderzande (1994), in the limit one obtains the spanning trees of G . Thus, in this limit with $G = S(m)$, one obtains a special class of the collapsed lattice animals, the spanning trees of $S(m)$, for which there are no cycles. On the other hand if one lets Q go to zero for fixed p then, when $Q = 0$, the terms which contribute to \mathcal{Z} are the terms with $k(A) = 1$, i.e. the set of all spanning connected subgraphs of G , so that for $Q = 0$ we have $\mathcal{Z} = R(G; p)$.

4. Bounds on $G(-\infty, \beta)$

In this section we use the connections made in the last section and results from the Flesia *et al* (1994) study of $G(\beta_1, \beta_2)$ to obtain bounds on $G(-\infty, \beta)$.

Following the derivation in Flesia *et al* (1994) of bounds for $G(\beta_1, \beta_2)$, it is possible to obtain bounds for $G(-\infty, \beta)$. In particular, the combinatorial inequality of Flesia *et al* (1994) (equation (4.1)) with $d = 2$ gives

$$\begin{aligned} \binom{n+1-k-s_{\min}(n)/2}{j} a_n(s_{\min}(n), k) &\leq \binom{k+j}{j} a_n(s_{\min}(n), k+j) \\ &\leq \binom{2n-k-s_{\min}(n)/2}{j} a_n(s_{\min}(n), k). \end{aligned} \quad (4.1)$$

Setting $k = 0$ in equation (4.1), multiplying through by $e^{\beta j}$ and summing over j gives

$$(1 + e^\beta)^{n+1-s_{\min}(n)/2} a_n(s_{\min}(n), 0) \leq Z_n(-\infty, \beta) \leq (1 + e^\beta)^{2n-s_{\min}(n)/2} a_n(s_{\min}(n), 0). \quad (4.2)$$

Since $a_{m^2}(4m, 0) = 1$, taking $m = \lfloor n^{1/2} \rfloor$, $r = n - m^2$ and $k = 0$ in equation (2.10) gives

$$1 \leq a_n(s_{\min}(n), 0) \leq 2^{4n^{1/2}+o(n^{1/2})} \quad (4.3)$$

so that

$$\lim_{n \rightarrow \infty} n^{-1} \log a_n(s_{\min}(n), 0) = 0. \quad (4.4)$$

Thus taking logarithms, dividing by n and letting n go to infinity in equation (4.2) gives the bounds

$$\log(1 + e^\beta) \leq G(-\infty, \beta) \leq 2 \log(1 + e^\beta). \quad (4.5)$$

Similarly, setting $k = n + 1 - j - s_{\min}(n)/2$ in equation (4.1), multiplying throughout by $e^{\beta[n+1-j-s_{\min}(n)/2]}$, summing over j , and then rearranging the order of the inequalities leads to

$$\begin{aligned} \left(\frac{2n}{n-1} \right)^{-1} (1 + e^\beta)^{n+1} a_n(s_{\min}(n), n+1-s_{\min}(n)/2) &\leq Z_n(-\infty, \beta) \\ &\leq (1 + e^\beta)^{n+1} a_n(s_{\min}(n), n+1-s_{\min}(n)/2). \end{aligned} \quad (4.6)$$

Lattice animals with n vertices and satisfying $s = s_{\min}(n)$, $k = n + 1 - s_{\min}(n)/2$ are the n vertex lattice animals with the maximum number of contacts. Thus from theorem 4.2 of Madras *et al* (1990) and equation (2.10)

$$\lim_{n \rightarrow \infty} n^{-1} \log a_n(s_{\min}(n), n+1-s_{\min}(n)/2) = 4\mathcal{C}/\pi \quad (4.7)$$

where \mathcal{C} is Catalan's constant. Taking logarithms, dividing by n and letting n go to infinity in equation (4.6) gives the bounds

$$\log(1 + e^\beta) - 2 \log 2 + 4\mathcal{C}/\pi \leq G(-\infty, \beta) \leq \log(1 + e^\beta) + 4\mathcal{C}/\pi. \quad (4.8)$$

In addition,

$$a_n(s_{min}(n), n + 1 - s_{min}(n)/2)e^{[n+1-s_{min}(n)/2]\beta} \leq Z_n(-\infty, \beta) \tag{4.9}$$

so that

$$\frac{4C}{\pi} + \beta \leq G(-\infty, \beta). \tag{4.10}$$

In summary,

$$\left. \begin{aligned} \beta < -\log(e^{4C/\pi} - 1) & \quad \log(1 + e^\beta) \\ \beta \geq -\log(e^{4C/\pi} - 1) & \quad \frac{4C}{\pi} + \beta \end{aligned} \right\} \leq G(-\infty, \beta) \tag{4.11}$$

$$\leq \begin{cases} 2 \log(1 + e^\beta) & \beta < \log(e^{4C/\pi} - 1) \\ \log(1 + e^\beta) + \frac{4C}{\pi} & \beta \geq \log(e^{4C/\pi} - 1). \end{cases} \tag{4.12}$$

Flesia *et al* (1994) found that the connection between bond percolation and the two-variable model leads to the upper bound

$$G(\beta_1, \beta_2) < 2 \log(1 + e^{\beta_2}) \tag{4.13}$$

for $\beta_2 > 0$ and for all $\beta_1 \leq \beta_2 - \frac{1}{2} \log(1 + e^{\beta_2})$ (see equation (4.14) of their paper). Equation (2.14) then implies that

$$G(-\infty, \beta) < 2 \log(1 + e^\beta) \tag{4.14}$$

for $\beta > 0$.

Using the connection to the reliability polynomial it is possible to extend the strict inequality in equation (4.14) to all finite values of β . To do this we need to introduce a few definitions. A *cut set* C for a connected graph G is a set of edges of G which when removed from G yield a disconnected spanning subgraph of G , $G - C$. A cut set C of G is *minimal* if adding in any edge of C to the graph $G - C$ results in a spanning connected subgraph of G . Let $\mathcal{C} = \{C_i \mid i = 1, \dots, s\}$ be a set of pairwise disjoint cut sets of G then the following is a standard result from reliability theory (Colbourn 1987)

$$R(G; p) \leq \prod_{i=1}^s [1 - (1 - p)^{|C_i|}] \tag{4.15}$$

where the product on the right is the probability, in the bond percolation process on G described in section 3, that at least one edge in each cut set is not deleted. A useful choice of \mathcal{C} for the case $G = S(m)$ is a set of minimal cut sets each composed of four edges and obtained as follows. The set of four edges incident on a vertex of degree 4 in $S(m)$ is a minimal cut set of $S(m)$. If one considers any set of degree 4 vertices of $S(m)$ such that no two vertices are joined by an edge, then the four-edge minimal cut sets associated with these vertices are pairwise disjoint. For m even, one can choose $(m - 2)^2/2$ degree 4 vertices of $S(m)$ so that no two vertices are joined by an edge and hence

$$R_m(p) \leq (1 - (1 - p)^4)^{(m-2)^2/2}. \tag{4.16}$$

For m odd, one can choose $((m - 1)^2/2) - m + 2$ degree 4 vertices of $S(m)$ so that no two vertices are joined by an edge and hence

$$R_m(p) \leq (1 - (1 - p)^4)^{((m-1)^2/2) - m + 2}. \tag{4.17}$$

Taking logarithms, dividing by m^2 and letting m go to infinity in equations (4.16) and (4.17) implies that

$$\mathcal{R}(p) \leq \frac{1}{2} \log(1 - (1 - p)^4) \quad (4.18)$$

where the term on the right-hand side is strictly less than zero for $p < 1$. Using this in equation (3.6) with $p = 1/1 + e^\beta$ gives

$$G(-\infty, \beta) \leq 2 \log(1 + e^\beta) + \frac{1}{2} \log \left[1 - \left(\frac{e^\beta}{1 + e^\beta} \right)^4 \right] \quad (4.19)$$

where the term on the right-hand side is strictly less than $2 \log(1 + e^\beta)$ for all finite values of β .

It is also possible to get an improved upper bound on $G(-\infty, \beta)$ for the region $\beta > 0$ using equation (3.13) and an upper bound on $\mathcal{F}_m(\beta)$ in terms of the generating function of the number of *rooted* spanning forests of $S(m)$. In particular, it is well known (see Stanley 1989) that the generating function for the number of rooted spanning forests with k edges of a graph can be obtained from the characteristic polynomial for the adjacency matrix of the graph. In a rooted spanning forest there is a distinguished vertex in each connected component of the forest. Each spanning forest with k edges in $S(m)$ has $m^2 - k$ components and thus is weighted by a factor $x_1 x_2 \dots x_{m^2 - k}$ in the generating function for rooted spanning forests of $S(m)$, where x_i is the number of vertices in the i th connected component of the forest. Thus clearly $\mathcal{F}_m(\beta)$ cannot be obtained easily from the rooted spanning forest generating function. However, the rooted spanning forest generating function is an upper bound for $\mathcal{F}_m(\beta)$ and this bound should be sharp for large β when $\mathcal{F}_m(\beta)$ is dominated by one-component spanning trees. In particular, it can be shown (see equation (18) of Kornilov and Priezzhev (1994) with $x = e^\beta$ and $z = 1$) that the limiting free energy, $\mathcal{F}^r(\beta)$, of the rooted spanning forest generating function for $S(m)$ satisfies

$$\mathcal{F}^r(\beta) = \beta + \pi^{-2} \int_0^\pi \int_0^\pi \log(4 + e^{-\beta} - 2 \cos \theta_1 - 2 \cos \theta_2) d\theta_1 d\theta_2. \quad (4.20)$$

We note that this free energy is analytic and hence there is no phase transition for rooted spanning forests (see Kornilov and Priezzhev (1994) for a more detailed discussion of the thermodynamics of this model). Furthermore, as expected, $\mathcal{F}^r(\beta)$ approaches $(4C/\pi) + \beta$ as β goes to infinity. Thus

$$G(-\infty, \beta) \leq \beta + \pi^{-2} \int_0^\pi \int_0^\pi \log(4 + e^{-\beta} - 2 \cos \theta_1 - 2 \cos \theta_2) d\theta_1 d\theta_2 \quad (4.21)$$

and this bound becomes sharper as β goes to infinity.

Finally, we can obtain some new bounds for both $G(\beta_1, \beta_2)$ and $G(-\infty, \beta)$ by considering the special case of directed lattice animals. The results presented here are based on the results of Dhar (1987) on the collapse of directed animals. A *directed animal* on the square lattice is a connected subgraph of n vertices (including the origin) such that a vertex with coordinates (i, j) belongs to the animal only if it is the origin or at least one of its predecessor vertices $(i - 1, j)$ and $(i, j - 1)$ belongs to the animal. Define n_1 to be the number of vertices of the animal with one predecessor in the animal and n_2 to be the number of vertices of the animal with two predecessors in the animal. For each animal we can also define s to be the number of solvent contacts and k to be the number of contacts of the animal where the meaning of these quantities is the same as for undirected animals. Let e be the number of edges in the directed animal ($e = N_1$ in Dhar's notation). Note that n_2 is the cyclomatic index of the animal (see equation (5b) of Dhar's paper) and hence

$n_2 = n + 1 - k - s/2$. Also note that for each vertex with two predecessors, either of the two predecessor edges (the edges between a vertex and its predecessor vertices) can be removed without disconnecting the animal. For any vertex with only one predecessor, removal of the predecessor edge automatically creates something that is not a directed animal.

Define $A_{n,s,k}$ to be the number of directed animals with n vertices, s solvent contacts and k contacts. The above discussion tells us that starting with a directed animal with no contacts (i.e. $n_2 = n - s/2 + 1$) we can create an animal with k contacts by removing one predecessor edge from any k vertices with two predecessor edges. There are two choices of an edge for each vertex with two predecessor edges, thus

$$A_{n,s,k} = \binom{n_2}{k} 2^k A_{n,s,0} = \binom{n - s/2 + 1}{k} 2^k A_{n,s,0}. \tag{4.22}$$

Consider pure directed bond percolation on the lattice. Each bond is considered open with probability p_B and a vertex is open only if it is connected by an open bond to at least one of its open predecessors. The probability that the origin is in a connected open cluster consisting of exactly n open vertices is then

$$\text{Prob}(|C| = n) = \sum_{s,k} A_{n,s,k} (p_B)^{2n-s/2-k} (1 - p_B)^{k+s/2} \tag{4.23}$$

which can be obtained from equation (4) of Dhar's paper by taking $p_S = 1$. In this case only half the solvent contacts are required to be closed. The remaining solvent contacts are predecessor edges which go from a vertex not in the animal to a vertex in the animal and such edges, open or closed, cannot be edges in the animal.

Using equation (4.22) in equation (4.23) yields

$$\begin{aligned} P_n(p_B) \equiv \text{Prob}(|C| = n) &= p_B^{2n} \sum_s A_{n,s,0} \left[\frac{1 - p_B}{p_B} \right]^{s/2} \\ &\times \sum_{k=0}^{(n-s/2+1)} \binom{n - s/2 + 1}{k} 2^k \left[\frac{1 - p_B}{p_B} \right]^k. \end{aligned} \tag{4.24}$$

Thus

$$P_n(p_B) = p_B^{2n} \sum_s A_{n,s,0} \left[\frac{1 - p_B}{p_B} \right]^{s/2} \left[\frac{2 - p_B}{p_B} \right]^{n-s/2+1} \tag{4.25}$$

and

$$P_n(p_B) = p_B^{n-1} (2 - p_B)^{n+1} \sum_s A_{n,s,0} \left[\frac{1 - p_B}{2 - p_B} \right]^{s/2}. \tag{4.26}$$

Define the solvent model partition function for directed site animals ($k = 0$) by

$$D_n(\beta) = \sum_s A_{n,s,0} e^{\beta s} \tag{4.27}$$

and its limiting free energy by

$$\mathcal{D}(\beta) = \lim_{n \rightarrow \infty} n^{-1} \log D_n(\beta). \tag{4.28}$$

Similarly define the two-variable partition function for directed animals by

$$Z_n(\beta_1, \beta_2) = \sum_{s,k} A_{n,s,k} e^{\beta_1 s + \beta_2 k} \tag{4.29}$$

and its limiting free energy by

$$G(\beta_1, \beta_2) = \lim_{n \rightarrow \infty} n^{-1} \log Z_n(\beta_1, \beta_2). \tag{4.30}$$

Then setting $\beta = \frac{1}{2} \log\left(\frac{1-p_B}{2-p_B}\right)$ (i.e. $p_B = \frac{1-2e^{2\beta}}{1-e^{2\beta}}$) in equation (4.28) for $\beta < -\frac{1}{2} \log 2$ and using equation (4.26) we obtain:

$$\mathcal{D}(\beta) = 2 \log(1 - e^{2\beta}) - \log(1 - 2e^{2\beta}) + \lim_{n \rightarrow \infty} n^{-1} \log P_n \left(\frac{1 - 2e^{2\beta}}{1 - e^{2\beta}} \right). \quad (4.31)$$

In addition, equations (4.22), (4.29) and (4.30) imply

$$\mathbf{G}(\beta_1, \beta_2) = \log(1 + 2e^{\beta_2}) + \mathcal{D}(\beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2})). \quad (4.32)$$

From directed percolation it is known that

$$\begin{aligned} \lim_{n \rightarrow \infty} n^{-1} \log P_n(p_B) &= 0 & \text{for } p_B \geq p_{BC} \approx 0.6447 \\ \lim_{n \rightarrow \infty} n^{-1} \log P_n(p_B) &< 0 & \text{for } p_B < p_{BC}. \end{aligned} \quad (4.33)$$

Thus

$$\mathcal{D}(\beta) \begin{cases} = 2 \log(1 - e^{2\beta}) - \log(1 - 2e^{2\beta}) & \beta \leq \frac{1}{2} \log\left(\frac{1 - p_{BC}}{2 - p_{BC}}\right) \approx -0.6694 \\ < 2 \log(1 - e^{2\beta}) - \log(1 - 2e^{2\beta}) & -\frac{1}{2} \log 2 > \beta > \frac{1}{2} \log\left(\frac{1 - p_{BC}}{2 - p_{BC}}\right). \end{cases} \quad (4.34)$$

Hence for $\beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2}) < -\frac{1}{2} \log 2$,

$$\mathbf{G}(\beta_1, \beta_2) \begin{cases} = 2 \log(1 + 2e^{\beta_2} - e^{2\beta_1}) - \log(1 + 2e^{\beta_2} - 2e^{2\beta_1}) \\ \beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2}) \leq \frac{1}{2} \log\left(\frac{1 - p_{BC}}{2 - p_{BC}}\right) \\ < 2 \log(1 + 2e^{\beta_2} - e^{2\beta_1}) - \log(1 + 2e^{\beta_2} - 2e^{2\beta_1}) \\ \beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2}) > \frac{1}{2} \log\left(\frac{1 - p_{BC}}{2 - p_{BC}}\right) \end{cases} \quad (4.35)$$

and thus $\mathbf{G}(\beta_1, \beta_2)$ is analytic everywhere in the region $\frac{1}{2} \log\left(\frac{1-p_{BC}}{2-p_{BC}}\right) \geq \beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2})$ and there is a curve of collapse-phase transitions in the (β_1, β_2) plane given by the equation

$$\frac{1}{2} \log\left(\frac{1 - p_{BC}}{2 - p_{BC}}\right) = \beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2}). \quad (4.36)$$

(A similar conclusion was obtained recently and independently by Henkel and Seno (1996) for the grand canonical two-variable partition function of directed animals.) The full curve in figure 1 is a graph of the collapse-phase boundary given by equation (4.36). Also note that, in the region to the left of this phase boundary where the limiting free energy is known exactly and is analytic, one can calculate the heat capacity defined by

$$\mathbf{C}(\beta_1, \beta_2) \equiv \frac{\partial^2 \mathbf{G}(\beta_1, \beta_2)}{\partial \beta_2^2}. \quad (4.37)$$

It is found that for a range of fixed values of β_1 , $\mathbf{C}(\beta_1, \beta_2)$ can have a local maximum at a β_2 value, β_2^* , such that $\beta_2^* > \log\left[\left(\frac{1-p_{BC}}{2-p_{BC}}\right)e^{2\beta_1} - 1\right] - \log 2$ (or equivalently the point (β_1, β_2^*) is to the left of the phase boundary given in equation (4.36)). The broken curve on figure 1 is a graph of β_2^* as a function of β_1 . Therefore, in the directed animal two-variable model one has a curve of heat capacity maximums which do not correspond to a phase transition.

The above results lead to new bounds on $\mathbf{G}(\beta_1, \beta_2)$ and $\mathbf{G}(-\infty, \beta)$. Clearly,

$$A_{n,s,k} \leq a_n(s, k). \quad (4.38)$$

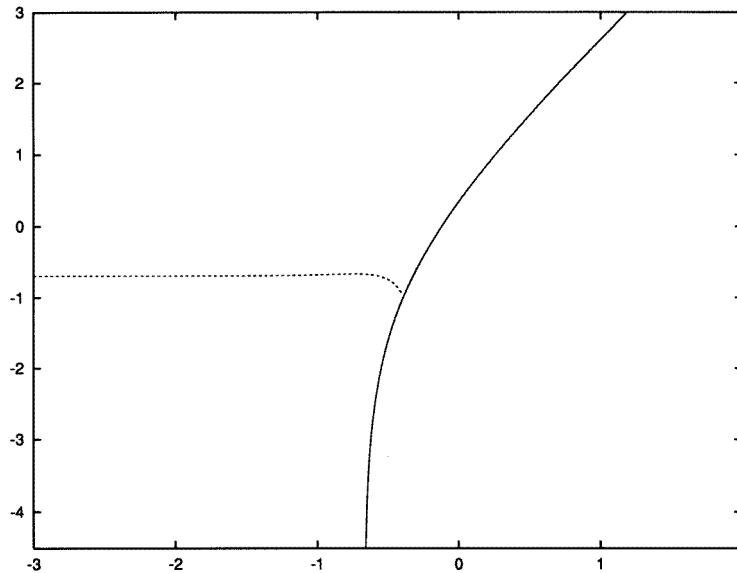


Figure 1. A β_2 versus β_1 plot of the directed animal collapse transition phase boundary (equation (4.36)) shown here as a full curve. The broken curve indicates the locations (β_2^*) of local maxima for fixed β_1 of the function $C(\beta_1, \beta_2)$; this line does not correspond to a phase boundary.

Hence for $\beta_1 - \frac{1}{2} \log(1 + 2e^{\beta_2}) \leq \frac{1}{2} \log(\frac{1-p_{BC}}{2-p_{BC}}) \approx -0.6694$

$$2 \log(1 + 2e^{\beta_2} - e^{2\beta_1}) - \log(1 + 2e^{\beta_2} - 2e^{2\beta_1}) \leq G(\beta_1, \beta_2) \tag{4.39}$$

and for all β

$$\log(1 + 2e^\beta) \leq G(-\infty, \beta). \tag{4.40}$$

Resummarizing the bounds for $G(-\infty, \beta)$ gives

$$\left. \begin{aligned} \beta < -\log(e^{4C/\pi} - 2) & \quad \log(1 + 2e^\beta) \\ \beta \geq -\log(e^{4C/\pi} - 2) & \quad \frac{4C}{\pi} + \beta \end{aligned} \right\} \leq G(-\infty, \beta) \tag{4.41}$$

$$G(-\infty, \beta) \leq \min \left\{ 2 \log(1 + e^\beta) + \frac{1}{2} \log \left[1 - \left(\frac{e^\beta}{1 + e^\beta} \right)^4 \right], \mathcal{F}^r(\beta), \log(1 + e^\beta) + \frac{4C}{\pi} \right\} \tag{4.42}$$

where at least for $-2 \leq \beta \leq 2$, $\mathcal{F}^r(\beta) < \log(1 + e^\beta) + 4C/r\pi$. In figure 2 we plot the lower and upper bounds given by equations (4.41) and (4.42) for the range $-2 \leq \beta \leq 2$.

5. Monte Carlo simulation

As indicated by equations (2.10) and (2.11), the behaviour of $G(-\infty, \beta)$ is determined by the sequence of collapsed animals with m^2 vertices. We thus focus our Monte Carlo algorithm on such animals and we study the set of spanning connected subgraphs of $S(m)$, \mathcal{B}_m , directly as opposed to studying the reliability or spanning forest problems. To do this we have developed a Metropolis-style dynamic Monte Carlo algorithm which generates a

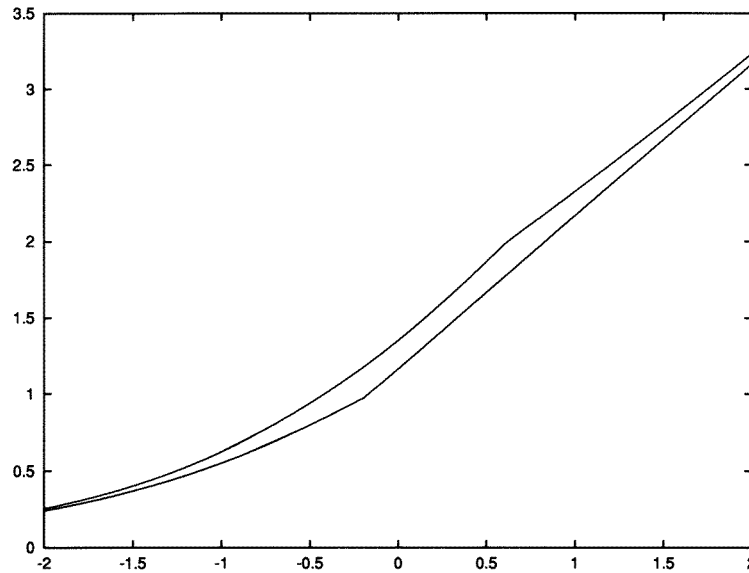


Figure 2. The bounds in equations (4.41) and (4.42) are plotted versus β as full curves.

Markov chain on the state space \mathcal{B}_m . For a fixed value of β , the generated Markov chain is such that its unique equilibrium distribution satisfies

$$\mu_\beta(A) = \frac{e^{\beta k}}{Z_{m^2}(-\infty, \beta)} \quad (5.1)$$

where for any $A \in \mathcal{B}_m$, $\mu_\beta(A)$ is the probability that the chain is in state A and where k is the number of contacts in the spanning connected subgraph A .

The algorithm is as follows. We start at time $t = 0$ with any spanning connected subgraph A_0 of $S(m)$. Suppose at time $t \geq 0$ the state of the Markov chain is $X_t = A \in \mathcal{B}_{m,k}$. Then, an edge e is chosen uniformly at random from the entire set of $2m^2 - 2m$ edges in $S(m)$. If e is an edge (a contact) of the animal A , we attempt to remove (add) this edge from (to) A to create a new animal A' . If this move is not accepted then $X_{t+1} = A$. The transition probabilities for these moves are determined by the Boltzmann weights associated with A and A' and whether or not the graph resulting from the move is connected. In particular, if e is a contact of the animal A then we can always add this edge to A without affecting the connectedness of the animal and create an animal $A' \in \mathcal{B}_{m,k+1}$; in this case we accept the move with probability $\min\{1, e^\beta\}$ and thus

$$\text{Prob}(X_{t+1} = A' | X_t = A) = \frac{1}{2m^2 - 2m} \min\{1, e^\beta\}. \quad (5.2)$$

If e is an edge of the animal A then we determine if removal of e will disconnect the animal. If removal of e disconnects the animal then we do not make any move and $X_{t+1} = A$. On the other hand if removal of e leaves the animal connected then we can remove e and create an animal $A' \in \mathcal{B}_{m,k-1}$; in this case we accept the move with probability $\min\{1, e^{-\beta}\}$, so that

$$\text{Prob}(X_{t+1} = A' | X_t = A) = \frac{1}{2m^2 - 2m} \min\{1, e^{-\beta}\}. \quad (5.3)$$

Of course the efficiency of this algorithm is limited by the efficiency of the connectedness check step. However, we take advantage of the fact that at most one edge changes in the

graph at each step and are able to minimize the loss of efficiency due to the connectedness check (see Paulhus 1994). Note that the Markov chain $\{X_t\}$ satisfies the detailed balance condition

$$\text{Prob}(X_{t+1} = A' | X_t = A) \mu_\beta(A) = \text{Prob}(X_{t+1} = A | X_t = A') \mu_\beta(A') \quad (5.4)$$

and thus the chain $\{X_t\}$ is reversible. The chain is also irreducible because it is always possible to get from one spanning connected subgraph A' of $S(m)$ to another spanning connected subgraph A'' of $S(m)$ by a sequence of edge additions and edge deletions. For example, first add to A' all the edges corresponding to contacts in A' to obtain $S(m)$ and then delete the edges in $S(m)$ corresponding to the contacts in A'' . Finally, because the chain is irreducible and because of the rejection techniques used in the algorithm (which ensure there exists a state A such that $\text{Prob}(X_{t+1} = A | X_t = A) > 0$), the chain is aperiodic. These properties (equation (5.4), irreducibility, and aperiodicity) guarantee that the Markov chain $\{X_t\}$ is ergodic and has the unique equilibrium distribution given by equation (5.1).

Due to limitations in storage space we chose the method of block averages (see Law and Kelton 1982) for collecting data and estimating the autocorrelation time τ_{int} (see Sokal 1989). We assumed that sampling the Markov chain after an initial equilibration time of at least $10\tau_{int}$ would be sampling from the distribution μ_β . We also assumed that for time steps greater than $10\tau_{int}$ estimates of an average quantity from two distinct blocks each of length at least $2\tau_{int}$ would be essentially independent estimates. For a total number of vertices $n = m^2$, we observed that τ_{int} grows faster than linearly but less than quadratically with n . Using estimates of τ_{int} and the above assumptions, we obtained 40 essentially independent observations of each variable for each n value. For example, for $n = 9$, 40 blocks each of size 5×10^4 were used, while for $n = 196$, 40 blocks each of size 3.5×10^6 were used.

In order to investigate the possible singular behaviour of $G(-\infty, \beta)$, we define the heat capacity

$$C_n(\beta) \equiv n^{-1} \frac{d^2 \log Z_n(-\infty, \beta)}{d\beta^2} = \frac{\text{var}(k)}{n} \quad (5.5)$$

where the variance of k is calculated over the probability distribution μ_β defined in equation (5.1). We can therefore estimate $C_n(\beta)$ using the Monte Carlo algorithm described above. Estimates of $C_{m^2}(\beta)$ for $m = 2, 3, 4$ from our Monte Carlo simulation compare well with exact values which can be obtained from the exact enumeration data reported in Madras *et al* (1990). In figure 3 we show our estimates of $C_{m^2}(\beta)$ for $m = 8, 10, 12, 14, 20, 30, 40$ and for a range of β values. The error bars at each β give a 95% confidence interval for the average over 40 values of $C_n(\beta)$. For a given value of m , the graph of $C_{m^2}(\beta)$ versus β attains a maximum value on the interval $\beta \in (-1.5, 0)$. As m increases the maximum values, h_{m^2} , are increasing with m but at a rate which is decreasing as m gets larger. The location of the maximum, $\hat{\beta}_{m^2}$, becomes less negative as m increases. In order to quantify this behaviour, we have estimated h_{m^2} , the peak height, and $\hat{\beta}_{m^2}$, the peak location, for $m = 3, \dots, 15, 20, 30, 40, 50, 60$ by fitting a polynomial to the estimated values of $C_{m^2}(\beta)$. We next obtained, by a jack-knife procedure, estimates of the variances of the peak height and peak location. Figure 4 shows a plot of the estimates of h_{m^2} versus $1/m$. Figure 5 shows a plot of the estimates of $\hat{\beta}_{m^2}$ versus $1/m$. In both cases the peak height and peak location appear to be approaching, approximately linearly with $1/m$, a constant as m goes to infinity. To make this more precise we assume that h_n and $\hat{\beta}_n$ satisfy the following scaling forms as n goes to infinity,

$$\hat{\beta}_n \sim \hat{\beta}_\infty - b_1 n^{-\psi_1} \quad (5.6)$$

$$h_n \sim h_\infty - c_1 n^{-\psi_2}. \quad (5.7)$$

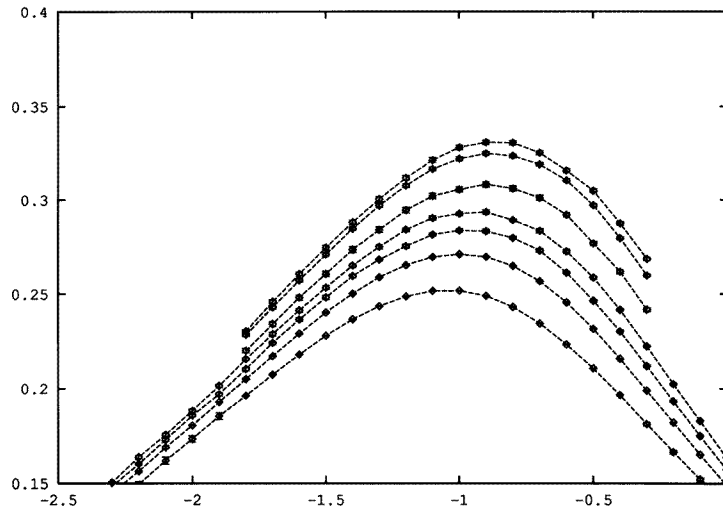


Figure 3. The heat capacity, C_n^2 , versus β for $m = 8, 10, 12, 14, 20, 30, 40$. The broke curves are guides to the eye. The error bars are 95% confidence intervals.

In order to estimate $\hat{\beta}_\infty$ we assume that for n large enough

$$\hat{\beta}_n = \hat{\beta}_\infty - bn^{-\psi}. \quad (5.8)$$

For a fixed value of ψ , and $n \geq N_{min}$ we perform a two-parameter weighted linear least squares fit to find $\hat{\beta}_\infty$ and b in equation (5.8) and then vary ψ in order to minimize the χ^2 statistic of the fit. Taking $N_{min} = 49$, we obtain our best estimate at $\psi = 0.508$ and estimate $\hat{\beta}_\infty$ to be

$$-0.82 \pm 0.005 \pm 0.07. \quad (5.9)$$

The first error term gives a 95% confidence interval for $\hat{\beta}_\infty$ based on the weighted linear least squares fit. The second error term is a systematic error estimated by comparing our estimate to another estimate obtained by taking $N_{min} = 121$. Next, we assume that $\hat{\beta}_\infty = -0.82$ in equation (5.6) and obtain an estimate for ψ_1 . That is, we perform a two-parameter weighted linear least squares fit to find $\log b_1$ and $-\psi_1$ in the equation

$$\log(\hat{\beta}_\infty - \hat{\beta}_n) = -\psi_1 \log n + \log b_1. \quad (5.10)$$

From this procedure we obtain our best estimate for ψ_1 to be

$$0.50 \pm 0.03 \pm 0.3. \quad (5.11)$$

We then follow the same procedure to get estimates of h_∞ and ψ_2 . We find our best estimate when $\psi = 0.459$ and estimate h_∞ to be

$$0.3553 \pm 0.0006 \pm 0.02 \quad (5.12)$$

and our best estimate of ψ_2 is

$$0.459 \pm 0.005 \pm 0.2. \quad (5.13)$$

We followed a similar procedure and attempted to fit h_n to scaling forms other than that given in equation (5.7), for example $h_n \sim bn^\psi$ and $h_n \sim b(\log n)^\psi$; however, the scaling form of equation (5.7) yielded the best fit. We therefore conclude that the peak heights are going to a constant h_∞ as estimated above. From looking at the plots of $C_n(\beta)$ for wide

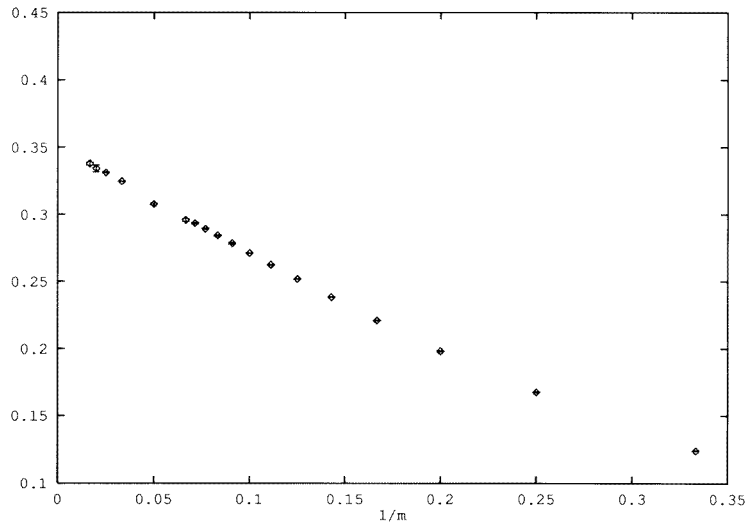


Figure 4. The estimated value of h_{m^2} versus $1/m$ for $m = 3, \dots, 15, 20, 30, 40, 50, 60$. The error bars are estimates of 95% confidence intervals.

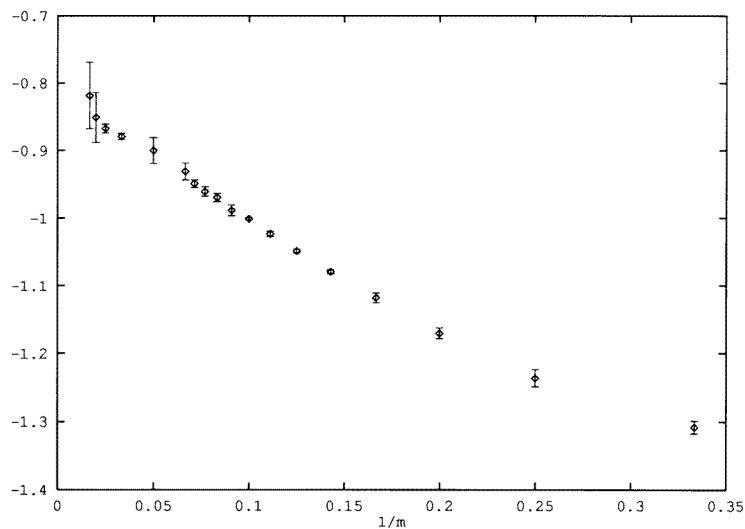


Figure 5. The estimated value of $\hat{\beta}_{m^2}$ versus $1/m$ for $m = 3, \dots, 15, 20, 30, 40, 50, 60$. The error bars are estimates of 95% confidence intervals.

ranges of β , it appears that the functions $C_n(\beta)$ have a bounded continuous limit as n goes to infinity and that hence the second derivative of $G(-\infty, \beta)$ exists for all β . Thus we find no evidence for a second-order collapse-collapse phase transition. This does not rule out the possibility of a higher-order phase transition or even a first-order phase transition. We investigated higher-order derivatives of $Z_n(-\infty, \beta)$; however, the fluctuations in the data were too large to draw any conclusions about even the third or fourth derivatives. Furthermore, from the results obtained for ψ_2 , we cannot rule out the possibility that the

correct scaling form for h_n is

$$h_n \sim c_0 + \frac{c_1}{\sqrt{n}} + \frac{c_2}{n} + \dots \quad (5.14)$$

which would be the scaling form for h_n if there was no collapse-collapse phase transition and where the $1/\sqrt{n}$ term is a finite size surface correction term (Barber 1983).

6. Discussion

In summary, we have introduced a model for studying collapsed lattice animals and proved that the limiting free energy exists for this model. We are also able to establish that the partition function for collapsed lattice animals is directly related to the reliability polynomial for the $m \times m$ grid graph and the limiting free energy is related to the generating function for spanning forests of the $m \times m$ grid graph. In addition we obtain rigorous bounds for the free energy of collapsed lattice animals and a new lower bound for the free energy of the two-variable lattice animal model. Although we do not rigorously establish the existence or non-existence of a collapse-collapse transition, our results provide strong evidence for the fact that if a collapse-collapse phase transition exists for this model then it is not a second-order phase transition.

The results for directed animals indicate that a curve of heat capacity peaks occurs without any corresponding phase transition. The same may be the case for general lattice animals and hence there may be no collapse-collapse transition for the general two-variable lattice animal model. However, one important feature for the lattice animal problem which is missing from the directed animal problem is the following. There is evidence that there are two universality classes for the collapse of lattice animals (Flesia *et al* 1992, Seno and Vanderzande 1994) and that the line of heat capacity peaks in the collapsed region (for increasing values of β_1) intersects the line of collapse transition points at the dividing point for the two universality classes (thought to be the point corresponding to bond percolation). Hence the heat capacity peaks in the collapsed region may be of interest even if there is no collapse-collapse transition. From our Monte Carlo results we are able to estimate the peak location and height for $\beta_1 = -\infty$. We also point out that even if one proves that there is no collapse-collapse transition at $\beta_1 = -\infty$ this does not necessarily imply that there is no collapse-collapse transition for $\beta_1 > -\infty$ and in particular it is still an open problem to show that

$$\lim_{\beta_1 \rightarrow -\infty} G(\beta_1, \beta_2) = G(-\infty, \beta_2). \quad (6.1)$$

Acknowledgments

Both authors wish to acknowledge helpful conversations with M Paulhus and M Bickis. CES wishes to acknowledge helpful conversations with M Bousquet-Mélou, R Brak, E Janse van Rensburg, N Madras, A Owczarek, and S Whittington. Part of this paper was completed while CES was on sabbatical thus CES would like to thank R Brak and the Department of Mathematics, University of Melbourne, S Whittington and the Chemical Physics Theory Group, University of Toronto, and the Institute for Mathematics and its Applications, University of Minnesota, for their hospitality and support. LMS would like to acknowledge financial assistance from the Challenge '93, '94 and Best Matches '94 programmes. This research was funded by NSERC of Canada.

Appendix

A.1. Proof of equations (2.8) and (2.9)

We consider first the set of animals with n vertices, $s_{min}(n)$ solvent contacts and $k = 0$ contacts and call these the *collapsed site animals*. Due to the minimum solvent contact requirement such animals will have no holes in their interior and their boundary will form a convex polygon. (A *convex polygon* is a self-avoiding polygon with the property that any line on the dual lattice intersects the polygon in either zero or two edges of the polygon (see Enting and Guttmann 1990).) Furthermore, given a collapsed site animal we can create its *dual animal* by placing a square cell around each vertex of the original animal. The dual animal then has exactly $s_{min}(n)$ edges in its polygonal boundary and n interior cells (i.e. area n). Hence $a_n(s_{min}(n), 0) = p_{n,s_{min}(n)}$ where $p_{n,m}$ is the number of self-avoiding polygons on the square lattice with m edges enclosing an area of n unit cells (see Enting and Guttmann 1990). For fixed area n , the polygons with the least number of edges (i.e. smallest perimeter) are convex so $a_n(s_{min}(n), 0) = \hat{p}_{n,s_{min}(n)}$ where $\hat{p}_{n,m}$ is the number of convex polygons with m edges enclosing an area of n unit cells. Next note that $s_{min}(n)$ has two possible forms, either $s_{min}(n) = 4m$ for some m or $s_{min}(n) = 4m + 2$ for some m . In particular, if we set $s = 4m$ and ask for site animals with the largest number of vertices, then an appropriate isoperimetric inequality guarantees that there is exactly one such lattice animal which forms a square with m^2 vertices. That is $a_{m^2}(4m, 0) = 1$. Similarly, if we set $s = 4m + 2$ and ask for site animals with the largest possible number of vertices then there are two such animals both of which are m by $m + 1$ rectangles so that $a_{m(m+1)}(4m + 2, 0) = 2$. Since for any positive integer n there exists integers $m = \lfloor n^{1/2} \rfloor$ and r , $0 \leq r \leq 2m$, such that $n = m^2 + r$, then $s_{min}(n)$ satisfies

$$4\lfloor n^{1/2} \rfloor = 4m \leq s_{min}(n) \leq 4m + 4 = 4\lfloor n^{1/2} \rfloor + 4 \tag{A.1}$$

and furthermore for $0 < r \leq m$, $s_{min}(n) = 4m + 2$ and for $m < r \leq 2m + 1$, $s_{min}(n) = 4m + 4$.

We can obtain a lower bound on $a_n(s_{min}(n), 0)$ by noting first that if one adds a single edge to the outside of the $n = m^2$ square site animal one increases the number of vertices by one and the number of solvent contacts by two. If one adds two parallel edges onto neighbouring solvent contacts on one face of the square animal then one increases the number of vertices by two and the number of solvent contacts by two. Similarly, if one replaces r neighbouring solvents on one side of the square animal by edges ($1 \leq r < m$) then one increases the number of vertices by r and the number of solvent contacts by two. In this process we replace any contacts that are created by edges so that the resulting animal is a site animal. Thus for $1 \leq r \leq m$,

$$m - r + 1 \leq a_{m^2+r}(4m + 2, 0) = a_{m^2+r}(s_{min}(m^2 + r), 0). \tag{A.2}$$

The same argument indicates that each animal with m^2 vertices, $4m$ solvent contacts and k contacts is contained in at least one animal with $m^2 + r$ vertices, $4m + 2$ solvent contacts and k contacts. Thus for $1 \leq r \leq m$

$$(m - r + 1)a_{m^2}(4m, k) \leq a_{m^2+r}(4m + 2, k) = a_{m^2+r}(s_{min}(m^2 + r), k). \tag{A.3}$$

A similar argument involving adding edges to a long side of an m by $m + 1$ rectangle yields the lower bound

$$(m - r + 2)a_{m^2+m}(4m + 2, k) \leq a_{m^2+m+r}(4m + 4, k) = a_{m^2+m+r}(s_{min}(m^2 + m + r), k) \tag{A.4}$$

where $1 \leq r \leq m + 1$. We note also that for $1 \leq r \leq m$

$$a_{m^2+r}(4m + 2, 0) = \hat{p}_{m^2+r, 4m+2} \leq \hat{p}_{4m+2} \leq 2^{4m+2+o(m)} \quad (\text{A.5})$$

and for $1 \leq r \leq m + 1$

$$a_{m^2+m+r}(4m + 4, 0) = \hat{p}_{m^2+m+r, 4m+4} \leq \hat{p}_{4m+4} \leq 2^{4m+4+o(m)} \quad (\text{A.6})$$

where \hat{p}_m is the total number of convex polygons with m edges. The last inequality comes from the fact that $\lim_{n \rightarrow \infty} (2n)^{-1} \log \hat{p}_{2n} = 2$ (Delest and Viennot 1984).

We next show that each collapsed site animal is a subgraph of a square site animal. For a collapsed site animal with $n = m^2 + r$, $1 \leq r \leq m$, vertices and $s = 4m + 2$ solvent contacts, its dual animal (with a convex polygon as its boundary) is bounded by a rectangle with perimeter $4m + 2$ (see for example Bousquet-Mélou and Fédou 1995) and area a such that $m^2 + r \leq a \leq m^2 + m$ (since $m^2 + m$ is the maximum area bounded by a polygon with $4m + 2$ edges). All such bounding rectangles have dimensions $m - l$ by $m + 1 + l$ for some $l \geq 0$ such that $(m - l)(m + 1 + l) \geq m^2 + 1$ and hence $l \leq \lfloor \sqrt{m - \frac{3}{4}} - \frac{1}{2} \rfloor$. Similarly, for a collapsed site animal with $n = m^2 + m + r$, $1 \leq r \leq m + 1$, vertices and $s = 4m + 4$ solvent contacts, its dual animal is bounded by a rectangle with perimeter $4m + 4$ and area a such that $m^2 + m + r \leq a \leq (m + 1)^2$. Such bounding rectangles have dimensions $m + 1 - l$ by $m + 1 + l$ for some $l \geq 0$ such that $(m + 1)^2 - l^2 \geq m^2 + m + 1$ and hence $l \leq \lfloor \sqrt{m} \rfloor$. Therefore in either case discussed above the length of the longest side of the bounding rectangle is at most $m + 1 + \lfloor m^{1/2} \rfloor$.

To determine what the above discussion about the dual animal implies for the original collapsed site animal we note the following. Consider a rectangular site animal with n vertices and s solvent contacts. Its dual animal is a rectangle with perimeter s and area n . If there are m_1 vertices on a side of the original rectangular site animal then the corresponding side of the dual rectangle has m_1 edges. Hence the discussion in the previous paragraph implies that a collapsed site animal with $n = m^2 + r$, $1 \leq r \leq 2m + 1$, vertices is a subgraph of a rectangular site animal with its longest side having at most $m + 1 + \lfloor m^{1/2} \rfloor$ vertices and thus it is a subgraph of a square site animal with $(m + 1 + \lfloor m^{1/2} \rfloor)^2$ vertices. This gives the following bound

$$a_{m^2+r}(s_{\min}(m^2 + r), k) \leq [a_{M^2}(4M, k)][a_{m^2+r}(s_{\min}(m^2 + r), 0)] \quad (\text{A.7})$$

for $1 \leq r \leq 2m$ and $M = m + 1 + \lfloor m^{1/2} \rfloor$. Equations (A.3)–(A.7) give that

$$a_{m^2}(4m, k) \leq a_{m^2+r}(s_{\min}(m^2 + r), k) \leq a_{M^2}(4M, k) \hat{p}_{4m+4} \leq a_{M^2}(4M, k) 2^{4m+o(m)} \quad (\text{A.8})$$

and this gives equation (2.10).

A.2. Proof of the existence of the limit in equation (2.11)

Using arguments similar to those used in Whittington and Guttmann (1990) for self-avoiding walks which cross a square, we next show that the limit

$$\lim_{m^2 \rightarrow \infty} m^{-2} \log Z_{m^2}(-\infty, \beta) \quad (\text{A.9})$$

exists. Consider an $M \times M$ square grid of M rows of M vertices, $S(M)$, we can create an animal which spans this grid by covering it with animals which span a smaller $m \times m$ square, $S(m)$. Note that $S(m)$ has $2m^2 - 2m$ edges and m^2 vertices. Define $\mathcal{B}_{m,k}$ to be the set of spanning connected subgraphs of $S(m)$ with exactly $2m^2 - 2m - k$ edges or, equivalently, the set of lattice animals with $n = m^2$ vertices, $s = s_{\min}(n) = 4m$ solvent contacts and k contacts. Thus $|\mathcal{B}_{m,k}| = a_{m^2}(4m, k)$. Define $\mathcal{B}_m \equiv \cup_{k \geq 0} \mathcal{B}_{m,k}$. Fix any m . Now for any integer $M > 0$ there exists integers $p > 0$ and $q > 0$ such that $M = pm + q$

where $0 \leq q < m$. Fix such p and q . Let $B_{k_1}, \dots, B_{k_{p^2}}$ be a set of p^2 animals in \mathcal{B}_m with k_1, \dots, k_{p^2} contacts respectively. We can create an animal which spans the $M \times M$ square grid by first building up a $pm \times pm$ square using the p^2 animals $B_{k_1}, \dots, B_{k_{p^2}}$. In the resulting $pm \times pm$ square, we join each pair of neighbouring vertices which belong to different animals by an edge and thus obtain an animal in \mathcal{B}_{pm} with $\sum_{i=1}^{p^2} k_i$ contacts. Finally, put the bottom left corner of this animal in the corresponding corner of an $M \times M$ square grid and add to the animal all the edges of the $M \times M$ grid that are not already either edges or contacts of the animal. We thus obtain an animal in \mathcal{B}_M with $\sum_{i=1}^{p^2} k_i$ contacts. This concatenation argument implies the following

$$\sum_{\{k_i, i=1, \dots, p^2 \mid \sum_{i=1}^{p^2} k_i = k\}} \prod_{i=1}^{p^2} a_{m^2}(4m, k_i) \leq a_{M^2}(4M, k). \tag{A.10}$$

Multiplying both sides by $e^{\beta k}$ and then summing over k in (A.10) yields

$$[Z_{m^2}(-\infty, \beta)]^{p^2} \leq Z_{M^2}(-\infty, \beta). \tag{A.11}$$

The number of connected spanning subgraphs of $S(M)$ is certainly less than the total number of spanning subgraphs of $S(M)$ so that

$$a_{M^2}(4M, k) \leq \binom{2M^2 - 2M}{k} \tag{A.12}$$

and thus

$$Z_{M^2}(-\infty, \beta) \leq (1 + e^\beta)^{2M^2 - 2M}. \tag{A.13}$$

Equation (A.13) implies that

$$[Z_{M^2}(-\infty, \beta)]^{1/M^2} \leq (1 + e^\beta)^2. \tag{A.14}$$

Whittington and Guttmann (1990) (see also Madras 1995) proved that the limit $\lim_{L \rightarrow \infty} L^{-2} \log W_{L^2}$ exists and is finite for a sequence $(W_{L^2})_{L=1}^\infty$ if for any m, p and $q < m$

$$[W_{m^2}]^{p^2} \leq W_{M^2} \tag{A.15}$$

where $M = pm + q$, and if there exists a finite number $B > 0$ such that for all L

$$[W_{L^2}]^{1/L^2} \leq B. \tag{A.16}$$

Thus from equations (A.11) and (A.14), we can conclude that, for fixed finite β , the limit in equation (2.12) exists and is finite. This combined with equation (2.11) proves the existence of the limiting free energy $G(-\infty, \beta)$.

A.3. Proof of equation (3.13)

In this section we prove equation (3.13). We begin by showing that

$$Z_{m^2}(-\infty, \beta) = \mathcal{F}_m^*(\beta) \tag{A.17}$$

and hence obtain the first equality in equation (3.13). To show this we define a bijection between the set $\mathcal{B}_{m,k}$ of spanning connected subgraphs of $S(m)$ with k contacts and the set $\mathcal{F}_{m,k}^*$ of spanning forests of $S^*(m)$ with k edges. We note that the definition of the dual provides a bijection between the edges of $S(m)$ and the edges of $S^*(m)$. Under this bijection an edge in $S^*(m)$ corresponds to the edge it crosses in $S(m)$. Given a spanning subgraph A of $S(m)$ with k contacts we can construct a spanning subgraph A^* of $S^*(m)$

which consists of all the vertices in $S^*(m)$ and only the edges in $S^*(m)$ corresponding to the k contact edges of A in $S(m)$. We note that a subgraph A of $S(m)$ is disconnected if and only if there exists a set of vertices of A whose ‘boundary’ (here this means edges in $S(m)$ but not in A which are incident on exactly one vertex of A) in $S(m)$ consists entirely of contacts. Furthermore there exists a set of vertices of A whose boundary in $S(m)$ consists entirely of contacts if and only if A^* contains a cycle. Thus A is a spanning connected subgraph of $S(m)$ if and only if A^* is a spanning forest of $S^*(m)$. Thus $a_{m^2}(4m, k) = f_{m,k}^*$ and equation (A.17) follows.

To show that the next equality in equation (3.13) holds, let $T_{m-1,k}$ be a spanning forest of $S(m-1)$ with k edges. Because the vertex set $V(S^*(m)) = V(S(m-1)) \cup \{v_o\}$, $T_{m-1,k}$ plus the vertex v_o forms a spanning forest of $S^*(m)$ with k edges. Thus

$$f_{m-1,k} \leq f_{m,k}^*. \quad (\text{A.18})$$

Next let $T_{m,k}^*$ be a spanning forest of $S^*(m)$ and suppose the degree of v_o in this forest is j , $0 \leq j \leq 4m-4$. Then removing the j edges adjacent to v_o yields a spanning forest of $S(m-1)$ with $k-j$ edges. The resulting spanning forest of $S(m-1)$ could have been obtained from at most $\binom{4m-4}{j}$ different spanning forests of $S^*(m)$. Thus

$$f_{m,k}^* \leq \sum_{j=0}^{\min\{k, 4m-4\}} \binom{4m-4}{j} f_{m-1, k-j}. \quad (\text{A.19})$$

Using equations (A.18) and (A.19) we obtain

$$\sum_k f_{m-1,k} e^{\beta k} \leq \sum_k f_{m,k}^* e^{\beta k} \leq (1 + e^\beta)^{4m-4} \sum_k f_{m-1,k} e^{\beta k}. \quad (\text{A.20})$$

Taking logarithms, dividing by m^2 and letting m go to infinity in equation (A.20) gives equation (3.13).

References

- Barber M N 1983 *Phase Transitions and Critical Phenomena* vol 8, ed C Domb and J L Lebowitz p 145
 Bousquet-Mélou M and Féduou J 1995 *Discrete Math.* **137** 53
 Brylawski T H and Oxley J G 1992 *Matroid Applications* ed N White (Cambridge: Cambridge University Press) p 123
 Colbourn C J 1987 *The Combinatorics of Network Reliability* (Oxford: Oxford University Press)
 Delest M and Viennot G 1984 *Theor. Comput. Sci.* **34** 169
 Dhar D 1987 *J. Phys. A: Math. Gen.* **20** L847
 Enting I G and Guttmann A J 1990 *J. Stat. Phys.* **58** 475
 Flesia S, Gaunt D S, Soteris C E and Whittington S G 1992 *J. Phys. A: Math. Gen.* **25** L1169
 ——— 1994 *J. Phys. A: Math. Gen.* **27** 5831
 Fortuin C M and Kasteleyn P W 1972 *Physica* **57** 536
 Henkel M and Seno F 1996 *Phys. Rev. E* **53** 3662
 Kornilov E I and Priezhev V B 1994 *Theor. Math. Phys.* **98** 61
 Law A M and Kelton W D 1982 *Simulation and Modelling Analysis* (New York: McGraw-Hill)
 Liu C I and Chow Y 1983 *Acta Math. Hung.* **41** 27
 Madras N 1995 *J. Phys. A: Math. Gen.* **28** 1535
 Madras N, Soteris C E, Whittington S G, Martin J L, Sykes M F, Flesia S and Gaunt D S 1990 *J. Phys. A: Math. Gen.* **23** 5327
 Myrvold W 1992 *Networks* **22** 647
 Paulhus M 1994 A study of computer simulations of combinatorial structures with applications to lattice animal models of branched polymers *MSc Thesis* University of Saskatchewan, Saskatoon, Canada
 Ramesh A, Ball M O and Colbourn C J 1987 *Ann. Discrete Math.* **33** 261
 Seno F and Vanderzande C 1994 *J. Phys. A: Math. Gen.* **27** 5813 (Corrigendum 1994 *J. Phys. A: Math. Gen.* **27** 7937)

- Sokal A D 1989 *Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms* Cours de Troisième Cycle de la Physique en Suisse Romande (Lausanne, June 1989)
- Stanley R P 1989 *Ann. NY Acad. Sci.* **576** 500
- Welsh D J A 1993a *J. Phys. A: Math. Gen.* **26** 2471
- 1993b *Complexity: Knots, Colourings and Counting* (*London Mathematical Society Lecture Note Series 186*) (Cambridge: Cambridge University Press)
- Whittington S G and Guttmann A J 1990 *J. Phys. A: Math. Gen.* **23** 5601