Statistical models based on conditional probability distributions

1992 J. Phys. A: Math. Gen. 25 3549

(http://iopscience.iop.org/0305-4470/25/12/016)

J. Phys. A: Math. Gen. 25 (1992) 3549–3558. Printed in the UK

Statistical models based on conditional probability distributions

Rajamani S Narayanan
Theory Group, DESY, Notkestrasse 85, 2000 Hamburg 52, Federal Republic of Germany

Received 11 November 1991, in final form 4 April 1992

Abstract. We present a formulation of statistical mechanics models based on conditional probability distributions rather than a Hamiltonian. Closely linked with this formulation is a Monte Carlo algorithm, in which a configuration generated is guaranteed to be statistically independent of any other configuration for all values of the parameters, in particular near the critical point. The required internal symmetry and the lattice rotational symmetry are realized in a conventional manner, but the translational symmetry on the lattice is realized in an unconventional manner. By explicitly constructing a Z_2-invariant model in two dimensions, we show that it is possible to realize critical phenomena through this procedure. We also show that the specific heat exponent, α, and the susceptibility exponent, γ, are consistent with that of the Ising model in two dimensions.

1. Introduction

Theoretical studies of statistical mechanics models that exhibit critical behaviour play an important role in understanding various phenomena in physics. These models are usually defined on a lattice with the relevant degrees of freedom existing on the sites or links of the lattice. A model is specified by first giving the Hamiltonian, H, which defines the total energy of the system and is made up of some local interactions between the degrees of freedom. The strengths of the interactions become the parameters of the model. In the thermodynamic limit, where the number of degrees of freedom tend to infinity, interesting models exhibit collective phenomena in a region of the parameter space, or in other words, they behave critically. The central quantity in statistical mechanics is the Boltzmann distribution

$$P = e^{H} \tag{1.1}$$

which gives the probability distribution of the various configurations of the degrees of freedom. Measurable physical quantities are statistical averages of observables defined as some function of the degrees of freedom and are calculated with the weight given by P. An important insight gained from such an analysis of a statistical mechanics model is that, starting from a Hamiltonian with only local interactions one is able to observe long-range order, i.e. the statistical averages exhibit collective phenomena. Studies of various models seem to indicate that many critical properties remain the same for a wide choice of the interactions as long as the underlying symmetry of the Hamiltonian is preserved. This goes by the name of 'universality' [1].

0305-4470/92/123549+10$04.50

Physically relevant statistical mechanics models are usually not exactly solvable and one has to resort to various approximation techniques to extract the physical properties. Most difficult to extract are the properties describing the critical behaviour of the model. Analytical techniques, based on a high-temperature or a low-temperature expansion, break down in the vicinity of the critical temperature. Numerical techniques, referred to as Monte Carlo simulations, performed on a finite lattice are usually found to be a nice way of studying critical behaviour. Starting from some random initial configuration of the system, an update is made in a probabilistic manner to generate a new configuration. A sequence of configurations is generated and if the transition probability for the update is properly chosen, the sequence has the correct equilibrium Boltzmann distribution. If the terms in the sequence thus generated are pairwise-independent, then one gets a good estimate for the statistical average of an observable by just taking an average over the sequence. The longer the sequence, the better the estimate. Because the Hamiltonian has only local interactions, it is possible to perform a local update on the system to achieve this. Two such popular algorithms having a wide range of applicability are 'Metropolis' and 'heat bath'. Since only a local update is performed, the configuration changes slowly and several updates of the total lattice have to be made so that the final configuration obtained is independent of the starting configuration. As one gets close to the critical point, the number of updates of the whole lattice that have to be made before one gets an independent configuration becomes very large and the numerical study becomes impractical. This problem is referred to as 'critical slowing down'. The root of this problem is closely related to the insight stated in the previous paragraph. Although the Hamiltonian has only local interactions, there is long-range order present, implying that it is important to make large changes in the configuration to obtain an independent configuration. In recent years, using this very point as a basis, there have been various proposals to overcome this problem. They include multigrid techniques, cluster algorithms and Fourier acceleration [2]. All these algorithms attempt to perform a global update of the system rather than a local update, so that large changes are possible and an independent configuration can be obtained reasonably fast. These approaches have had successful applications, but they usually seem somewhat model-dependent.

In this paper, we make use of the notion that many different Hamiltonians, or equivalently many different Boltzmann distributions, could have the same critical properties and hence we arrive at a choice for the Boltzmann distribution that is easily accessible to numerical techniques, but one that still has the appropriate physical properties we require. We present our general idea in the next section and discuss the details of defining a model in section three. We apply this idea to a model in two dimensions with spin degrees of freedom and having a Z_2 symmetry. We find the critical exponents to be in agreement with the 2D Ising model. This forms the contents of section 4. A summary and some criticisms of the idea constitute the last section of this paper.

2. The idea

Consider a statistical mechanics model on a finite lattice with n sites. Let the degrees of freedom be defined on sites and be labelled $S_1, S_2, \ldots, S_n$. For simplicity, let there be only one parameter called β. The Boltzmann distribution, defined in (1.1), is explicitly $P(S_1, S_2, \ldots, S_n; \beta)$. What follows can be trivially extended to the situation where there is more than one coupling and where $n \to \infty$.

We start with the identity

$$\frac{\sum_{S_2,\ldots,S_n} P(S_1,S_2,\ldots,S_n;\beta)}{\sum_{S_1,S_2,\ldots,S_n} P(S_1,S_2,\ldots,S_n;\beta)} \frac{\sum_{S_3,\ldots,S_n} P(S_1,S_2,\ldots,S_n;\beta)}{\sum_{S_2,\ldots,S_n} P(S_1,S_2,\ldots,S_n;\beta)} \times \cdots \frac{P(S_1,S_2,\ldots,S_n;\beta)}{\sum_{S_n} P(S_1,S_2,\ldots,S_n;\beta)}$$
$$= \frac{P(S_1,S_2,\ldots,S_n;\beta)}{\sum_{S_1,S_2,\ldots,S_n} P(S_1,S_2,\ldots,S_n;\beta)} . \tag{2.1}$$

If the Boltzmann distribution is normalized, then the denominator on the right-hand side of (2.1) is unity. We interpret the left-hand side of (2.1) as follows. The first term is the probability distribution of S_1 ($p(S_1;\beta)$); the second term is the probability distribution of S_2 conditional on S_1 ($p(S_2|S_1;\beta)$). We sequentially proceed in this manner till we reach the last term, which is the probability distribution of S_n conditional on $S_1, S_2, \ldots, S_{n-1}$ ($p(S_n|S_1,S_2,\ldots,S_{n-1};\beta)$). The specifics of the ordering of the degrees of freedom present in the left-hand side of (2.1) will be addressed in this section.

It is usually very difficult, and in most models impossible, to obtain the conditional probability distribution starting from the Hamiltonian. Since the Boltzmann distribution, P, is the central quantity in statistical mechanics and not the Hamiltonian itself, one can think of defining a model by giving all the conditional probability distributions. The conditional probability distributions could be local or non-local, in the sense that it could be conditional on the degrees of freedom on some nearby sites or it could be conditional on all the degrees of freedom that have already been determined. In this paper, we assume it is local and address the question as to whether the statistical averages of observables computed using the resulting Boltzmann distribution exhibit critical behaviour at some value of the parameter β. Further, if some properties near the critical point of a model defined in this manner are the same as for another model defined via the Hamiltonian approach, this idea becomes interesting. This approach will then have an added attraction because associated with (2.1) is a natural procedure for a Monte Carlo simulation that will not suffer from critical slowing down. A configuration of the n degrees of freedom is formed by first picking S_1 according to $p(S_1;\beta)$; then picking S_2 according to $p(S_2|S_1;\beta)$ and so on, till we pick S_n according to $p(S_n|S_1,S_2,\ldots,S_{n-1};\beta)$. The terms in the sequence of configurations thus obtained are independent of each other. This is in contrast to the usual Monte Carlo algorithms, where the new configuration is obtained by a local updating of the previous configuration. Therefore by construction, the algorithm described here is not expected to suffer from critical slowing down.

3. Details of model definition

In order to assign the conditional probability distributions, defined in (2.1), we first define the sequence of lattice sites necessary for the assignment. We give the specifics for a two-dimensional infinite square lattice (see figure 1). Each lattice site is labelled [3] by two indices, l and i. l denotes the level of the lattice site and ranges from 0 to ∞. i denotes the specific site on each level and ranges from 0 to $4l-1$. '00' is any site on the infinite lattice. The sites belonging to the level l are the ones that are removed

by l steps from '00', i.e. there are l links connecting '00' with a site on level l. The degree of freedom on the site 'li' is S_{li}. A configuration on the lattice is formed level by level in the following manner. The conditional probability distributions are assigned level by level, i.e. first the conditional probability distribution is given for the degree of freedom on the site at level 0, then for all the degrees of freedom on the sites in level 1 and so on. The algorithm that generates configurations will then be growing a whole configuration from a single seed and each configuration will be an independent growth process. The conditional probability distribution of a degree of freedom at some level is, in general, conditional upon all the degrees of freedom on the lower levels and all the degrees of freedom on the same level that have already been assigned. This statement implies that we should also specify the ordering of the sites within a level, but if we choose the conditional probability distribution of a site in one level to be dependent only on sites belonging to lower levels, this is not necessary.

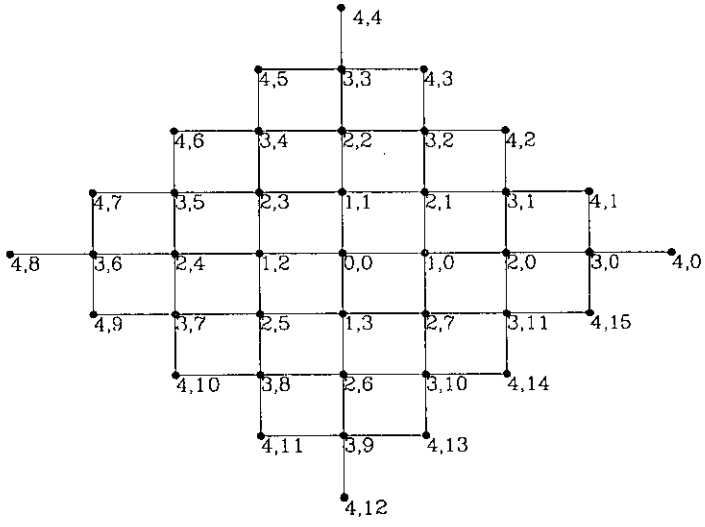

Figure 1. 2D square lattice showing sites up to level 4. Notation is l, s; l: level, s: site.

A typical example of only incorporating short-range interactions, would be to make the conditional probability distribution dependent on only a few degrees of freedom on the level that is one step lower. One criterion in choosing the conditional probability distribution, would be to have it invariant under certain changes of the degrees of freedom, so that the model has an underlying symmetry. An analysis of this model would then show whether it belongs in the same 'universality class' as another model, defined via the Hamiltonian having the same underlying symmetry. Before we completely specify a model, we have to make sure that the lattice symmetries are preserved by the procedure described above. To ensure the lattice symmetries, we can proceed in two different ways. One way is to choose the seed randomly anywhere on the lattice and to choose the conditional probability distributions to preserve the discrete rotational symmetry of the lattice. Another way is to fix the seed and choose the conditional probability distribution in such a way that it preserves all the lattice symmetries.

The second path will, in general, be technically difficult and as such we opt to pursue the first path here. We now discuss the symmetry issues pertaining to that way of defining the model. The conditional probabilities can be chosen to have the

required internal symmetry without much difficulty. It is the lattice symmetries that are hard to realize. In particular, translational invariance is difficult to handle. The reason for this is the presence of a certain specific ordering of the lattice sites in (2.1). One can choose the order and then suitably assign conditional probabilities so that the rotational symmetries on the lattice are realized. But due to the fact that the first site, namely S_1, in (2.1) is singled out, there is no translational invariance. To restore translational invariance, we do the following. We start with some lattice site as the seed and define a conditional probability distribution, associated with that site, that has the required internal symmetry and the discrete rotational symmetry. Picking two different sites on the lattice as the seed, we can associate conditional probability distributions with each of them that are simply related by a lattice translation connecting the two sites. This procedure can be repeated for every site on the lattice so that there is a conditional probability distribution associated with each site and they are all related to each other by lattice translations. If we now define the Boltzmann distribution for the model to be the average of all these conditional probability distributions, then it is evident that the resulting model has translational invariance. Since the resulting distribution is positive, an associated Hamiltonian can be obtained by formally taking the logarithm of the Boltzmann distribution. The Hamiltonian associated with the resulting Boltzmann distribution is expected to have complicated interactions because there is a sum over infinitely many conditional probability distributions. In fact, the Hamiltonian could even have non-local interactions and it is not *a priori* clear that one can invoke the concepts of universality in critical phenomena. But the Hamiltonian will have the required internal symmetry and the lattice symmetries. In the language of renormalization group [4], the resulting Hamiltonian will have many irrelevant terms. This situation is not much different from the one that arises out of Kadanoff's blocking scheme [1]. There again each blocking step adds extra interactions that are irrelevant. Of course, in Kadanoff's blocking scheme, the sequence of Hamiltonians obtained as a result of blocking, all lie in one basin of attraction and are therefore guaranteed to have the same critical behaviour. In our way of defining the model, we end up with a Hamiltonian with many irrelevant terms but we have no *a priori* notion of where it lies in the parameter space. In particular, we do not know if it lies in the same basin of attraction as another model defined via the standard Hamiltonian procedure and having the same internal symmetry. But this question can be answered by an explicit numerical study of a model defined by the procedure described here.

4. Application to 2D Z_2 model

The ideas discussed in the previous sections are applied to the case of a Z_2-symmetric model on a two-dimensional infinite square lattice. The aim is to see if there is a phase transition and to compare the behaviour near the critical point with the nearest-neighbour Ising model [5].

The degrees of freedom on each site are the usual Ising spins, $S_{li} = \pm 1$. The spin, S_{00}, is assigned ± 1 with equal probability. The conditional probability distribution of spins on level l will be chosen to depend on the state of spins on the level just below l. To this end, we classify the sites on the infinite square lattice into two categories:

(i) Includes all the sites that have one nearest neighbour in the previous level. There are exactly four such sites in each level (see figure 1).

(ii) Includes all the sites that have two nearest neighbours in the previous level. There are exactly $4(l-1)$ such sites in level l (see figure 1).

Consider a spin, S_{li}, belonging to category (i). Let the nearest neighbour in the level $(l-1)$ be $S_{(l-1)j}$. The conditional probability distribution of S_{li} is given by

$$p_a(S_{li}|S_{(l-1)j};b)=\frac{1+bS_{li}S_{(l-1)j}}{2} \tag{4.1}$$

where b is a parameter such that $0 \leqslant b \leqslant 1$. Now consider a spin, S_{li}, belonging to category (ii). Let the two nearest neighbours in the level $l-1$ be $S_{(l-1)j}$ and $S_{(l-1)k}$. For example, in figure 1, S_{42} is a spin belonging to category (ii) in level 4, and S_{31} and S_{32} are its two nearest neighbours in level 3. The conditional probability of S_{li} is given by

$$p_b(S_{li}|S_{(l-1)j},S_{(l-1)k};b)=\frac{(1+bS_{li}S_{(l-1)j})(1+bS_{li}S_{(l-1)k})}{\sum_{S_{li}=\pm 1}(1+bS_{li}S_{(l-1)j})(1+bS_{li}S_{(l-1)k})}. \tag{4.2}$$

This assignment is a simple one, where the choice of the state of the spin to be added is only influenced by nearest-neighbour sites. Choosing the above conditional probability distribution and summing over all sites on the lattice for the seed, ensures the lattice symmetries. From (4.1) and (4.2), it is clear that the model thus defined has an underlying Z_2 symmetry.

We study this model by the numerical algorithm described at the end of section 2. For this purpose, we will work on a finite lattice with periodic boundary conditions and then we will have some sites on the lattice that do not fit into category (i) or (ii), described above, because they will have three or four nearest-neighbour spins in the previous level. For those spins, we extend (4.2) to include all the nearest neighbours. The only control parameter in this model is b and we let

$$b=\tanh\beta. \tag{4.3}$$

Since $1/\beta$ is expected to be proportional to the physical temperature near the critical point of the model, this redefinition of the control parameter is useful.

We proceed to analyse the model defined above. Consider the two basic statistical observables,

$$M=\sum_{li}S_{li} \tag{4.4}$$

$$E=\sum_{[l_1i_1,l_2i_2]}S_{l_1i_1}S_{l_2i_2} \tag{4.5}$$

where the sum in (4.4) runs over all lattice sites, and the sum in (4.5) runs over all nearest neighbouring pairs in the lattice. Since we will be working with periodic boundary conditions and because the quantities M and E are summed over all sites on the lattice, it is sufficient in our Monte Carlo simulation to start from one fixed site on the lattice. M measures the total magnetization and E measures the nearest-neighbour correlations. Due to the Z_2 symmetry present in the model, $\langle M\rangle$ will be zero in any finite lattice. We study the histogram of M to show qualitatively the

critical behaviour present in the model. A study of the scaling behaviour will be done through the two averages

$$C = \frac{1}{2L^2}(\langle E^2\rangle - \langle E\rangle^2) \tag{4.6}$$

$$\chi = \frac{1}{L^2}(\langle |M|^2\rangle - \langle |M|\rangle^2) \tag{4.7}$$

where L is the length of the square lattice. The definition of χ chosen here uses $|M|$ instead of M since it is more suitable for finite-size scaling analysis [5]. By studying the data for C on a 32×32 lattice, we will show that all the measurements are independent, indicating that there is no critical slowing down in our simulation.

We now present the results of our simulation. Figure 2 shows the distribution of magnetization per site, $m = M/L^2$, on a 32×32 lattice for $\beta = 0, 1.4, 3$. A total of 5000 configurations were generated. $\beta = 0$ corresponds to infinite temperature and total disorder. This is evident from the sharp peak in the distribution at $m = 0$. On the other hand, the distribution at $\beta = 3$ shows two sharp peaks at $m = \pm 1$. This shows that a strong ferromagnetic ordering has set in at $\beta = 3$. One therefore expects a phase transition somewhere between $\beta = 0$ and $\beta = 3$, where the distribution of m is expected to be flat, indicating that spin clusters of varying sizes are equally preferred. In figure 2, we see this happening at $\beta = 1.4$.

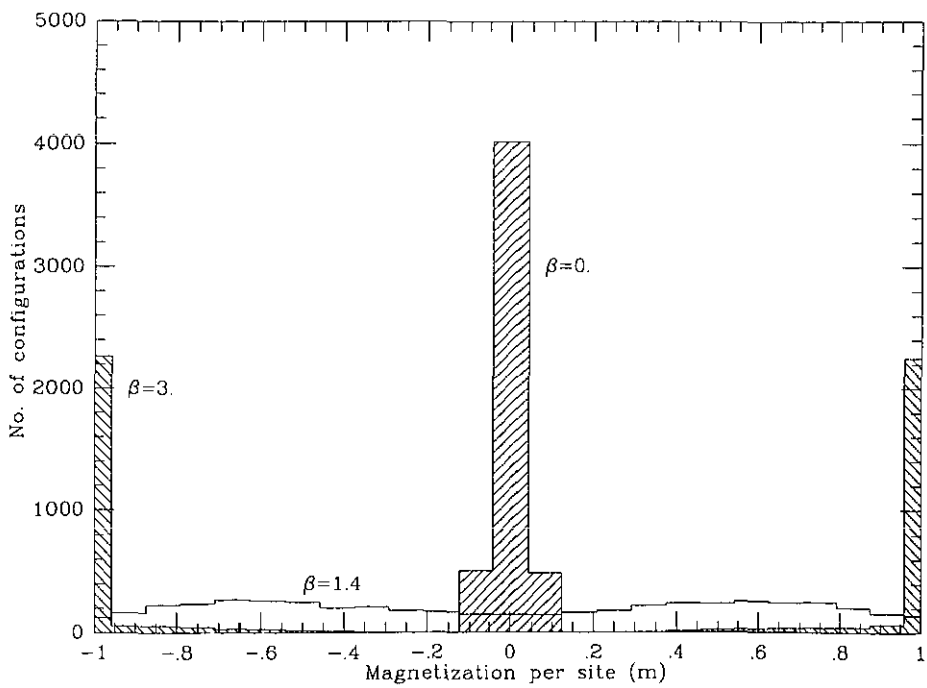

Figure 2. Distribution of magnetization per site.

Before we analysed C and χ for their critical behaviour, we verified that the configurations generated by this algorithm are independent by analysing the data for C on a 32×32 lattice at $\beta = 1.6$. We generated a total of 5000 configurations. These were then put into bins of sizes varying from 1 to 20. An average value was

computed for each bin. The bin averages were then assumed to be independent and an overall average and standard deviation were computed. We found the standard deviation to be 0.08, independent of the bin size, confirming that the original set of 5000 measurements are truly independent.

In order to perform a quantitative analysis in the scaling region, we perform a finite-size scaling analysis of C and χ. We study periodic lattices of several lengths ranging from $L = 4$ to $L = 32$ in steps of 4. By varying β on all these lattices, we obtain the maximum values of C and χ, namely, $C_{\max}(L)$ and $\chi_{\max}(L)$. Finite-size scaling [6, 7] means that

$$C_{\max}(L) \sim L^{\alpha/\nu} \tag{4.8}$$

$$\chi_{\max}(L) \sim L^{\gamma/\nu} \tag{4.9}$$

where α is the exponent characterizing the divergence of C, γ is the exponent characterizing the divergence of χ and ν is the exponent characterizing the divergence of the correlation length in the infinite lattice. If either α or γ is zero, as is the case in the 2D Ising model ($\alpha = 0$), then the corresponding quantity on the finite lattice is expected to scale as $\ln L$. Our finite-size analysis for C and χ are shown in figures 3 and 4 respectively. In figure 3, we plot $C_{\max}(L)$ versus $\ln L$. A straight line fits the data quite well asymptotically indicating that $\alpha = 0$. In figure 4, we plot $\ln \chi_{\max}(L)$ versus $\ln L$. A simple-minded least-square fit of the data to a straight line gives a slope of 1.88 ± 0.28, giving an estimate for γ/ν. This estimate is consistent with the exact result of $\gamma/\nu = 7/4$ for the Ising model [5]. This shows that the analysis performed here is consistent with the statement that the model investigated is in the same universality class as the 2D Ising model.

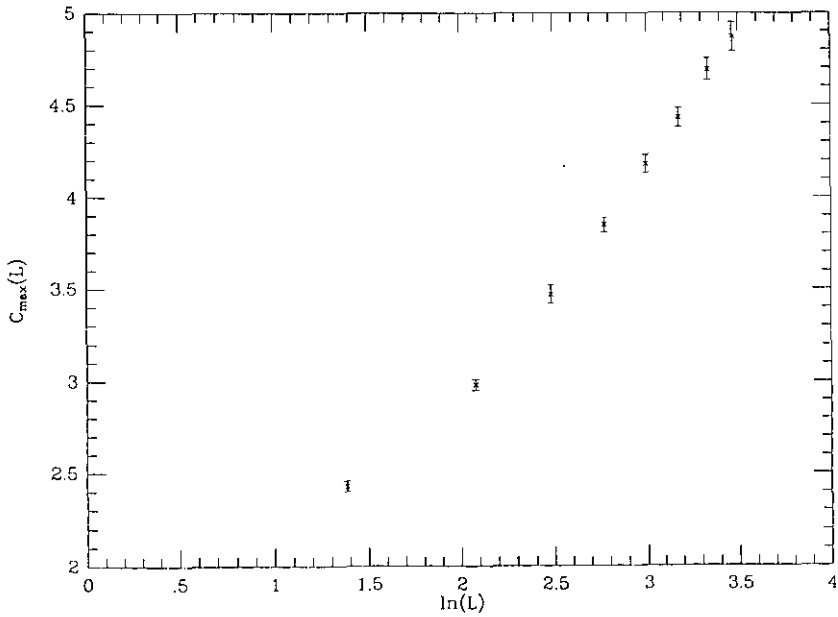

Figure 3. Finite-size scaling of specific heat.

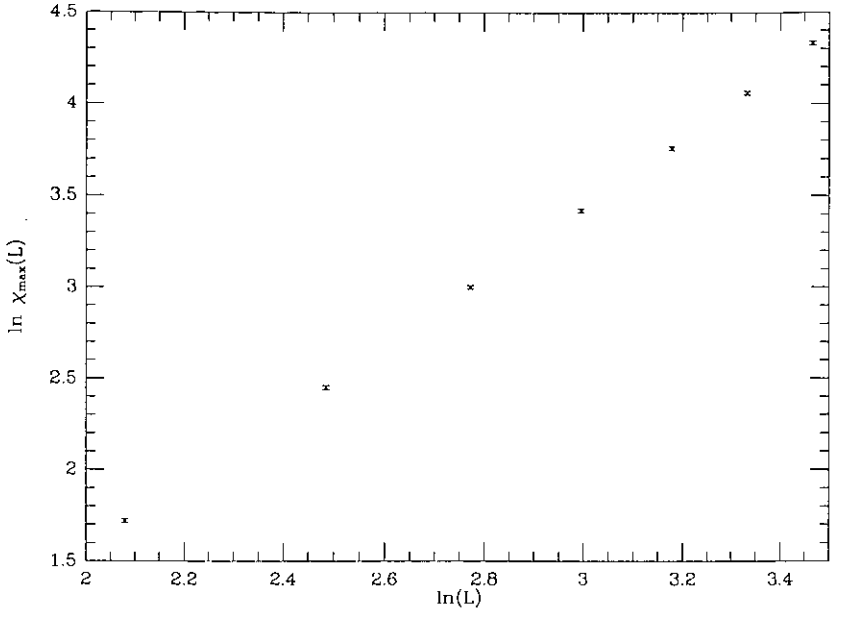

Figure 4. Finite-size scaling of susceptibility.

5. Summary and criticisms

In this paper, we investigated, from a slightly different viewpoint, the possibility of realizing long-range order in a statistical model starting from short-range interactions. Whereas in conventional methods one starts from a Hamiltonian with short-range interactions, here we started with short-range interactions in the conditional probability distributions of the degrees of freedom. This is based on an intuitive picture, where a configuration of the statistical mechanical system is formed by a growth process using a single degree of freedom as the seed. The growth is similar to the growth of a crystal. The state of the new degree of freedom that is added, should be dependent on the collective state of the degrees of freedom already present. Modelling this statistically, is achieved by specifying the conditional probability distribution. By locality or short-range interactions, in this context, it is meant that the probability distribution is conditional on only a few nearby degrees of freedom. This way of formulating a statistical mechanics model leads to a different Monte Carlo algorithm, where each configuration generated is guaranteed to be statistically independent. This statement is true for all values of the parameters and further, the effort put into generating a configuration is the same for all values of the parameters. Therefore the simulation does not suffer from critical slowing down.

In sections 2 and 3, we developed the details of the construction of a model, through the aforementioned procedure. That such an approach could lead to physically interesting models is demonstrated by constructing a Z_2-symmetric model in two dimensions with spin degrees of freedom and showing that the model exhibits a ferromagnetic order at low temperature. Analysis of the critical exponents of this model via finite-size scaling shows that they are the same as those for the two-dimensional Ising model. This analysis of the model is, of course, far from complete. It is important to study the two-point spin correlation function. A careful analysis will enable us to study the correlation length near the critical point and also enable the extraction

of the correlation length exponent, ν. A study of the two-point function will also help us understand the low-lying spectrum. Also, we need to verify whether scaling laws are obeyed in this model near the critical point. Since translational invariance is realized in an unconventional manner, it is important to study the clustering property of four-point functions. Some of these issues are currently under study.

If we take the viewpoint that the agreement of the critical exponents for standard thermodynamic quantities between two different models puts them in the same 'universality class', then we could conclude that it is possible to construct models via the procedure developed here and be in the same universality class as the standard models constructed via the Hamiltonian approach. But there could be a criticism to this statement, particularly if one is interested in defining a quantum field theory at the critical point of the statistical mechanics model. In this paper, we have not dealt with the existence of a Hamiltonian operator. Therefore, questions such as unitarity and locality of the model are not discussed here. In the context of quantum field theory, one also has to answer such questions as to whether the particle spectrum and strength of interactions are consistent with a physically relevant model. Although it is hard to extract the exact Hamiltonian associated with this procedure, it is reasonably clear that the interactions are rather complicated. As a consequence it is not certain that this approach will succeed in constructing physically useful quantum field theories. On the other hand, it could be useful in analysing statistical mechanics models that exhibit critical behaviour. In particular, since the Monte Carlo algorithm does not suffer from critical slowing down, the procedure described here could be of use in extracting critical exponents of interesting models to a good accuracy.

Acknowledgments

The author would like to thank Martin Lüscher for many helpful discussions and comments. Thanks are also due to the author's wife, Vijaya, for help with the manuscript.

References

[1] Kadanoff L P 1976 *Phase Transitions and Critical Phenomena* vol 5A ed C Domb and M S Green (New York: Academic)

[2] Parisi G 1984 *Progress in Gauge Field Theory (Cargese 1983)* ed G t'Hooft *et al* (New York: Plenum)
Goodman J and Sokal A D 1986 *Phys. Rev. Lett.* **56** 1015
Swendsen R H and Wang J S 1987 *Phys. Rev. Lett.* **58** 86
Wolff U 1989 *Phys. Rev. Lett.* **62** 361
Adler S L and Bhanot G 1991 *Phys. Rev. Lett.* **66** 1807
Adler S L *et al* IASSNS-HEP-91/39 preprint

[3] The idea of breaking a lattice into layers in the context of cluster algorithms is discussed by M Lüscher in an unpublished note titled 'Cluster Algorithm for the O(N) Non-Linear Sigma-Model', August 1989

[4] Wilson K G and Kogut J 1974 *Phys. Rep.* C **12** 75

[5] The Ising model in two dimensions has critical exponents, $\alpha = 0$, $\gamma = 7/4$ and $\nu = 1$. For a general reference see:
Amit D J 1984 *Field Theory, the Renormalization Group and Critical Phenomena* (Singapore: World Scientific)

[6] Binder K 1984 *Applications of the Monte Carlo Method in statistical physics* ed K Binder (New York: Springer)

[7] Barber M N 1983 *Phase Transitions and Critical Phenomena* vol 8 ed C Domb and J L Lebowitz (New York: Academic)