Journal of Statistical Physics, Vol. 52, Nos. 3/4, 1988

A Note on the Cluster Variation Method

Guozhong An[1]

Received January 5, 1988; revision received April 11, 1988

Kikuchi's cluster variation method (CVM) is reformulated as the truncation of a Möbius inversion. An attempt is made to explicate and simplify the various approaches to the CVM. This formulation makes apparent the connection of the method with other types of cluster approximation. An illustration of the procedure is provided.

KEY WORDS: Cluster-variation method; Möbius inversion.

1. INTRODUCTION

The cluster variation method (CVM),(1) a closed-form approximation scheme often used in statistical physics, is powerful and accurate in determining phase diagrams (for reviews on applications of the CVM, see ref. 2 and 3). The method has been used successfully in computing phase diagrams involving both first-order and continuous transitions.(1-3) In addition, exact series expansion coefficients up to seventh order have been extracted using the CVM with moderate computation efforts.(4,5) Results obtained from the CVM compare favorably with those of Monte Carlo simulations.(6,7)

In Kikuchi's formulation of the CVM, one generates an approximate expression for the entropy through Boltzmann's relation $S = k_B \ln G$ by counting the thermodynamic weight G approximately(1) (k_B is the Boltzmann constant). For example, to derive an approximation for the entropy of a spin-1/2 system in the pair approximation, one introduces an ensemble of systems and counts the number of ways that new sites can be introduced into the ensemble without changing the equilibrium fraction of up and down spins and the fraction of pairs of all configurations.(1) This

---

[1] Department of Physics, University of Washington, Seattle, Washington 98195.

ensemble method becomes difficult to follow for large clusters because of the complexity of the counting involved. To overcome this difficulty, the CVM has been reformulated several times.$^{(8-10)}$ Barker$^{(8)}$ reformulated the CVM employing a generalized quasichemical argument and obtained a generally applicable approximation for the entropy. Barker's formulation has the virtue of relative computational simplicity, but its mathematical basis is not well founded. Morita reformulated$^{(9)}$ the CVM, using a generalized cumulant expansion for the entropy, which is conceptually more favorable than the ensemble or generalized quasichemical method. However, his method is rarely employed due to its mathematical complexity. Very recently, Schlijper, in his investigation of the mathematical aspects of the convergence$^{(11)}$ of CVM, noted without further elaboration that Morita's cumulant expansion is a Möbius inversion. As we shall see, the use of the Möbius inversion$^{(12)}$ enables a conceptually clear and mathematically simple formulation of the CVM. Though this paper reaches essentially the same results as previous works, I believe the present exposition to be clearer and therefore of more use. Definitions and notations are introduced in Section 2. I then proceed in Section 3 to the general formulation of the CVM. A specific application is given in Section 4.

2. DEFINITIONS

Consider a set P consisting of elements $\{\alpha, \beta, \gamma, ...\}$. There exists a binary relation R between certain pairs of elements, satisfying (1) reflexivity ($\alpha R \alpha$ for every $\alpha \in P$), (2) transitivity ($\alpha R \beta$ and $\beta R \gamma$ together imply $\alpha R \gamma$), and (3) antisymmetry ($\alpha R \beta$ and $\beta R \alpha$ together imply $\alpha = \beta$). A pair of elements is said to be *comparable* (*noncomparable*) if a relation R exists (does not exist) between them. The relation R is said to define a partial ordering on P, and is written as $\leqslant$. Furthermore, the set P itself is called a partially ordered set.

Define$^{(12)}$ a function ζ on $P \times P$ by

$$\zeta(\beta, \alpha) = \begin{cases} 1, & \beta \leqslant \alpha \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

The Möbius function μ of a partially ordered set P is defined through the following relation:

$$\sum_{\alpha \leqslant \beta \leqslant \gamma} \zeta(\alpha, \beta)\, \mu(\beta, \gamma) = \delta(\alpha, \gamma) \tag{2}$$

where $\delta(\alpha, \gamma)$ is the Kronecker delta. The Möbius function of a partially ordered set P is uniquely determined by its constituting elements and the binary relation R among them. For any pair of real-valued functions f, g defined on a partially ordered set P related by

$$f(\alpha) = \sum_{\beta \leqslant \alpha} g(\beta) \tag{3}$$

we have[12]

$$g(\alpha) = \sum_{\beta \leqslant \alpha} f(\beta)\, \mu(\beta, \alpha) \tag{3'}$$

where $(\alpha, \beta) \in P$.

Now we specialize the consideration to a lattice with N sites. A cluster α is a subset of n_α lattice sites. There exists a natural ordering of all clusters, namely $\alpha \leqslant \beta$ if α is contained in β. This partial ordering defines a partially ordered set P with the largest cluster L containing all N lattice sites with its subclusters (subsets) as its elements. The Möbius function for this partially ordered set is[12]

$$\mu(\alpha, \beta) = (-1)^{n_\beta - n_\alpha} \tag{4}$$

We refer to site variables s_i as spins regardless of their nature as being discrete or continuous. The interaction among the spins in cluster α is written as $h_\alpha(s_i, i \in \alpha)$. It is understood that $h_\alpha(s_i, i \in \alpha)$ cannot be decomposed into interactions involving only subclusters of spins.

The Hamiltonian can be written as

$$\mathscr{H} = \sum_{\alpha \in \Gamma} h_\alpha(s_i, i \in \alpha) \tag{5}$$

where Γ is a collection of large clusters and their subclusters such that $h_\alpha = 0$ if $\alpha \notin \Gamma$. The free energy of the system can be obtained by minimizing the functional

$$F[\rho] = \operatorname{Tr} \rho \mathscr{H} + k_{\mathrm{B}} T \operatorname{Tr} \rho \ln \rho \tag{6}$$

with respect to ρ, subject to the constraint $\operatorname{Tr} \rho = 1$. The minimization is achieved at

$$\rho = e^{-\mathscr{H}/k_{\mathrm{B}}T} / \operatorname{Tr} e^{-\mathscr{H}/k_{\mathrm{B}}T} \tag{7}$$

which gives the density matrix of the system. If the minimization is carried out on a restricted class of functions, one is led to a particular closed-form approximation.

The reduced density matrix for a cluster α is defined in the usual manner,

$$\rho_\alpha = \operatorname*{Tr}_{L-\alpha} \rho \tag{8}$$

where the partial trace is taken over all variables in L, except those in α. It follows from the definition that

$$\begin{aligned} \rho_\alpha &= \operatorname*{Tr}_{\beta-\alpha} \rho_\beta, \qquad \beta \geqslant \alpha \\ \operatorname{Tr} \rho_\alpha &= 1 \end{aligned} \tag{9}$$

The cluster entropy associated with ρ_α is defined via

$$S_\alpha = -k_{\mathrm{B}} \operatorname{Tr} \rho_\alpha \ln \rho_\alpha \tag{10}$$

The entropy of the entire lattice S is then the cluster entropy S_L. A new set of functions is defined from the S_α using

$$\tilde{S}_\alpha = \sum_{\beta \leqslant \alpha} (-1)^{n_\alpha - n_\beta} S_\beta \tag{11}$$

The advantage of the $\tilde{S}_\alpha$ is that they are expected to die off quickly for clusters with linear size larger than the correlation length of the system$^{(5)}$, while the S_α obviously do not.

Since Eq. (11) is a summation over an interval in the partially ordered set P, we can invert it using Möbius inversion. Identifying $(-1)^{n_\alpha} \tilde{S}_\alpha$ and $(-1)^{-n_\beta} S_\beta$ with $f(\alpha)$ and $g(\beta)$ of Eq. (3), respectively, it follows from Eqs. (3′) and (4) that

$$S_\alpha = \sum_{\beta \leqslant \alpha} \tilde{S}_\beta \tag{12}$$

The $\tilde{S}_\alpha$ are simply the familar cumulants.$^{(9,13)}$

3. CLUSTER APPROXIMATION

From Eq. (12) we can write the entropy of the entire system as

$$S = \sum_{\alpha \leqslant L} \tilde{S}_\alpha = \sum_{\alpha \in P} \tilde{S}_\alpha \tag{13}$$

This is an exact expression. We now introduce the approximation by truncating (13) and keeping only terms associated with the subclusters of a chosen set of clusters $\gamma_1, \gamma_2, \ldots, \gamma_k$. The γ's are pairwise noncomparable, i.e., no cluster γ_i is a subcluster of another γ_j for all i and j. We denote the set consisting of $\gamma_1, \gamma_2, \ldots, \gamma_k$ and all of their subclusters by P'. The truncation is motivated by the expected rapid convergence of $\tilde{S}_\alpha$ with increasing α, except when the system is near criticality, where there are very long-ranged correlations. After the truncation, we have

$$S \approx \sum_{\alpha \in P'} \tilde{S}_\alpha \tag{14}$$

where $P' = \{\alpha \leqslant \gamma_i;\ i = 1, \text{ or } 2, \ldots, \text{ or } k\}$. This relation can be rewritten in terms of the S_α, hence the ρ_α, which will be our variational parameters. Substituting Eq. (11) in Eq. (14), we have

$$\begin{aligned} S \approx \sum_{\alpha \in P'} \tilde{S}_\alpha &= \sum_{\alpha \in P'} \sum_{\beta \leqslant \alpha} (-1)^{n_\alpha - n_\beta} S_\beta \\ &= \sum_{\alpha \in P'} \sum_{\beta} (-1)^{n_\alpha - n_\beta} \zeta(\beta, \alpha) S_\beta \end{aligned} \tag{14'}$$

Interchanging the order of summation of α and β in Eq. (14′), we have

$$S \approx \sum_{\beta \in P'} \sum_{\alpha \in P'} (-1)^{n_\alpha - n_\beta} \zeta(\beta, \alpha) S_\beta \equiv \sum_{\beta \in P'} a_\beta S_\beta \tag{15}$$

This relation was first obtained by Barker$^{(8)}$ (see also ref. 10) by a heuristic argument without obtaining an explicit expression for a_β. We now have

$$a_\beta = \sum_{\alpha \in P'} (-1)^{n_\alpha - n_\beta} \zeta(\beta, \alpha) = \begin{cases} N_e(\beta) - N_o(\beta), & n_\beta = \text{even} \\ N_o(\beta) - N_e(\beta), & n_\beta = \text{odd} \end{cases} \tag{16}$$

where $N_e(\alpha)$ is the number of clusters in P' larger than or equal to α with even numbers of sites, and $N_o(\alpha)$ is similarly defined for the odd clusters. It is clear that a_β vanishes for those clusters having only one cluster larger than itself $(1 - 1 = 0)$. It follows from Eqs. (2) and (16) that

$$\begin{aligned} \sum_{\substack{\beta \geqslant \alpha \\ (\alpha, \beta) \in P'}} a_\beta &= \sum_{\gamma \in P'} \sum_{\alpha \leqslant \beta \leqslant \gamma} \zeta(\alpha, \beta) \mu(\beta, \gamma) \\ &= \sum_{\gamma \in P'} \delta(\alpha, \gamma) = 1 \end{aligned} \tag{16'}$$

In practice, it is often easier to obtain the a_β from Eq. (16′) than from Eq. (16), since Eq. (16′) avoids the inclusion of the clusters that have vanishing a_α.

For the energy part of the free energy (6), we have

$$\operatorname{Tr} \rho \mathscr{H} = \sum_{\alpha \in \Gamma} \operatorname{Tr} \rho h_\alpha = \sum_{\alpha \in \Gamma} \operatorname{Tr} \rho_\alpha h_\alpha \tag{17}$$

without approximation. Thus, a restricted functional is derived for the free energy,

$$F[\rho_\alpha, \alpha \in P'] = \sum_{\alpha \in \Gamma} \operatorname{Tr} \rho_\alpha h_\alpha + k_B T \sum_{\alpha \in P'} a_\alpha \operatorname{Tr} \rho_\alpha \ln \rho_\alpha \tag{18}$$

It is usual to choose P' such that all clusters of Γ, which define the range of interactions, are contained in P'. However, in the usual mean field approximation one takes P' as the collection of all single sites regardless of Γ, and $\Gamma \nsubseteq P'$. Another exception is found in ref. 16. Minimizing F subject to the constrains

$$\begin{aligned} &\operatorname{Tr} \rho_\alpha = 1 && \alpha \in \{\text{maximum clusters in } P'\} \\ &\rho_\alpha = \operatorname*{Tr}_{\beta - \alpha} \rho_\beta && \beta \geqslant \alpha \end{aligned} \tag{19}$$

we obtain a system of nonlinear algebraic equations which can be solved numerically by Kikuchi's natural iteration algorithm(12) or by other numerical methods.

Note that in the above discussion no specific information about the function to be approximated has been used except the requirement that the cumulants approach zero for large clusters. Thus, in place of the cluster entropies S_α, one can consider H_α or F_α defined by

$$\rho_\alpha = e^{-H_\alpha/k_B T} / \operatorname{Tr} e^{-H_\alpha/k_B T}, \qquad F_\alpha = -k_B T \ln \operatorname{Tr} e^{-H_\alpha/k_B T} \tag{20}$$

Up to this point, the ρ_α have been considered as functions which are varied to minimize the approximate free energy. Alternatively, one can treat them as the true reduced density matrices of the system, which therefore satisfy Eq. (9). Then the Hamiltonian and the free energy can be approximated in a manner analogous to that in which entropy was in Eq. (15),

$$\mathscr{H} \approx \sum_{\alpha \in P'} a_\alpha H_\alpha, \qquad F \approx \sum_{\alpha \in P'} a_\alpha F_\alpha \tag{20a}$$

Equations (19), (20), and (20a) together form a set of self-consistent equations for the H_α.(10) It has been shown that these procedures are, in fact, equivalent.(10) Note that H_α is in general quite different from the sum of bare interaction within the cluster α, namely,

$$\mathscr{H}_\alpha \equiv \sum_{\beta \leqslant \alpha} h_\beta(s_i, i \in \beta) \tag{20b}$$

In some cases it may be profitable to define the cluster Hamiltonian in a mixed way[5]; namely, if $\alpha \in P^-$, H_α is given by Eq. (20), otherwise, by $\mathscr{H}_\alpha$ of Eq. (20b), where $P^- \subset P'$. We then again have a set of self-consistent equations for $\{H_\alpha, \alpha \in P^-\}$ given by Eqs. (19), (20, (20a), and (20b).

4. PARTICULAR APPROXIMATIONS

In actual application, simple symmetry considerations often reduce the sums in Eq. (18) to a very small number of terms,

$$F/N = \sum_{\alpha \in \Gamma} \operatorname{Tr} h_\alpha \rho_\alpha / N - T \sum_{\alpha \in P'} a_\alpha S_\alpha / N$$

$$= \sum_{\alpha_0} [(N_{\alpha_0}/N) \operatorname{Tr} h_{\alpha_0} \rho_{\alpha_0} - T a_{\alpha_0} (N_{\alpha_0}/N) S_{\alpha_0}] \tag{21}$$

where $N \to \infty$ is the total number of sites of the lattice, and N_{α_0} the total number of clusters of the α_0 kind. As an example, consider a lattice gas model on an fcc lattice with first-neighbor exclusion and second- and third-neighbor interaction. The Hamiltonian is

$$\mathscr{H}_{HC} = \sum_{\langle 2\text{nd} \rangle} n_i n_j \varepsilon_2 + \sum_{\langle 3\text{rd} \rangle} n_i n_j \varepsilon_3 - \mu \sum_i n_i \tag{22}$$

where the first sum is over all second-neighbor pairs, the second term is over all third-neighbor pairs, μ is the chemical potential, and $n_i = 0$, 1 for empty or occupied sites, respectively.

The clusters P' are taken to be all tetrahedrons consisting of four nearest neighbor sites, all second-neighbor pairs, all third-neighbor pairs, and all their subclusters, (first neighbor pairs and points). An fcc lattice can be divided into four interpenetrating simple cubic lattices. They are labeled as sublattices 1–4 respectively. The first and third neighbors of a site are located on a different sublattice than that of the site itself, and are labeled as (AB), where A stands for the sublattice of one site and B for the other. Second neighbors are located on the same sublattice.

Tetrahedrons belong to the largest clusters of P'; it follows from Eq. (16) that $a_T = 1$. All tetrahedrons are equivalent, and their number per site is 2. A first-neighbor pair is contained in two tetrahedrons and four triangles. Thus we have $a_{1\text{st}(AB)} = 1 + 2 - 4 = -1$ from Eq. (16), or $a_{1\text{st}(AB)} = 1 - 2a_T = -1$, from Eq. (16′) (note that $a_{\text{triangle}} = 1 - 1 = 0$). The number of first-neighbor pairs of (AB) kind per site is 1. Similarly, we have the results listed in Table I. Where S^{α_0} is a simple sum, e.g.,

$$S^{3\text{rd}(AB)} = -k_B \sum_i L(\rho_i^{3\text{rd}(AB)}) \tag{23}$$

with $L(x) \equiv x \ln x$, and i labels the configurations of the third-neighbor

Table I

α_0	N_{α_0}/N	a_{α_0}	$S^{\alpha_0}\ (\equiv S_{\alpha_0})$
Third-neighbor pairs (AB)	2	1	$S^{3\mathrm{rd}(AB)}$
Second-neighbor pairs (AA)	3/4	1	$S^{2\mathrm{nd}(AA)}$
Tetrahedron	2	1	S^{T}
First-neighbor pairs (AB)	1	-1	$S^{1\mathrm{st}(AB)}$
Point (A)	1/4	-25	$S^{p(A)}$

pairs. The triangle cluster does not enter at all in the above, for it has no contributions either to the entropy or to the energy in the above approximation. One can now minimize the free energy (21) with the constraints expressed by (19). For details of the minimization, see ref, 14 and 15.

ACKNOWLEDGMENTS

The author is indebted to Prof. F. Y. Wu for encouragement and suggestions. He is grateful to Prof. M. Schick for stimulating discussion and a critical reading of the manuscript. He also thanks Prof. R. Kikuchi for useful conversations. This work was supported in part by National Science Foundation Grant DMR-8613598.

REFERENCES

1. R. Kikuchi, *Phys. Rev.* **81**:988 (1951).
2. D. M. Burley, in *Phase Transitions and Critical Phenomena*, Vol. 2, C. Domb and M. S. Green, eds. (Academic, New York, 1972).
3. D. de Fontaine, in *Solid State Physics*, Vol. 34, H. Ehrenreich, F. Seitz and D. Turnbull, eds. (Academic Press, New York, 1979).
4. S. K. Aggarwal and T. Tanaka, *Phys. Rev. B* **16**:3963 (1977).
5. S. J. Singer and J. D. Weeks, *Phys. Rev. B* **36**:2228 (1987).
6. K. Binder, *Phys. Rev. Lett.* **45**:811 (1980); K. Binder, J. L. Lebowitz, M. K. Phani, and M. H. Kalos, *Acta Metall.* **29**:1655 (1981); B. Dünwig and K. Binder, *Phys. Rev. B* **36**:6935 (1987).
7. R. Kikuchi, *Prog. Theor. Phys. Suppl.* **87**:69 (1986).
8. J. A. Barker, *Proc. R. Soc. A* **216**:45 (1953).
9. T. Morita, *J. Phys. Soc. Jpn.* **12**:753 (1957); *J. Math. Phys.* **13**:115 (1972).
10. J. M. Sanchez, F. Ducastelle, and D. Gratias, *Physica A* **128**:334 (1984).
11. A. G. Schlijper, *Phys. Rev. B* **27**:6841 (1983).
12. G.-C. Rota, *Z. Wahrsch.* **2**:340 (1964).
13. R. Kubo, *J. Phys. Soc. Jpn.* **17**:1100 (1962).
14. R. Kikuchi, *J. Chem. Phys.* **60**:1071 (1974).
15. R. Kikuchi and D. de Fontaine, in *Application of Phase Diagrams in Metallurgy and Ceramics*, C. G. Carter, ed. (NBS Special Publication No. 496, 1978).
16. A. G. Schlijper and J. Westerhof, *Phys. Rev. B* **36**:5458 (1987).