Long-Range Frustration in a Spin-Glass Model of the Vertex-Cover Problem

Haijun Zhou
Max-Planck-Institute of Colloids and Interfaces, 14424 Potsdam, Germany
Institute of Theoretical Physics, The Chinese Academy of Sciences, Beijing 100080, China
(Received 1 September 2004; published 2 June 2005)

In a spin-glass system on a random graph, some vertices have their spins changing among different configurations of a ground-state domain. Long-range frustrations may exist among these unfrozen vertices in the sense that certain combinations of spin values for these vertices may never appear in any configuration of this domain. We present a mean field theory to tackle such long-range frustrations and apply it to the NP-hard minimum vertex-cover (hard-core gas condensation) problem. Our analytical results on the ground-state energy density and on the fraction of frozen vertices are in good agreement with known numerical and mathematical results.

DOI: 10.1103/PhysRevLett.94.217203 PACS numbers: 75.10.Nr, 02.10.Ox, 05.20.-y, 89.75.-k

The energy landscape of a large spin-glass system is very complex. There may exist (exponentially) many ground-state and metastable domains in the configurational space; these domains are mutually separated from each other by (infinitely) high energy barriers. At low temperature, the system may get trapped in one of these configurational domains, and ergodicity is broken. In the cavity field formalism [1] of mean field theory of finite connectivity spin glasses [2–4], microscopic configurations of a system are therefore grouped into different macroscopic states (hereafter, a macroscopic state is simply referred to as a "state" and a microscopic configuration as a "configuration"). In a given state each vertex i feels a cavity field h_i that may be different for different vertices, and the fluctuation of this field among all the states is characterized by a probability distribution $P_i(h_i)$ that again may be different for different vertices.

The ground-state energy landscape of a spin-glass system can be studied by the zero temperature limit of the cavity field theory [5]. In this limit and in a given state α, the spin value σ_i of a vertex i either is positively frozen ($\sigma_i \equiv +1$ in all configurations) or is negatively frozen ($\sigma_i \equiv -1$) or is unfrozen (σ_i fluctuates over ± 1 among configurations of state α). A crucial assumption of the cavity field theory [1,5] is that, with probability unity, each of the 2^n combinations of spin values for n randomly chosen unfrozen vertices is realized in configurations of state α. However, this conventional cavity field theory leads to negative values of structural entropy Σ [5] when loops of spin-spin interactions become abundant (see, e.g., [6–10]) or even causes a certain type of divergence in the population dynamics [7,11]. To overcome these difficulties, a positive reweighting parameter y can be introduced, and its value can be determined self-consistently by requiring $\Sigma(y) = 0$ [5]. This procedure is, however, not quite satisfactory; in case of the vertex-cover problem, it predicts a ground-state energy that is systematically lower than the actual value [9].

Here we discuss the possibility of long-range correlations among spins of different *unfrozen* vertices. Both the spins of two unfrozen vertices i and j will certainly fluctuate among configurations of a state α. On the other hand, we find that with certain probability σ_i and σ_j may be prohibited to take a certain combination of values (e.g., $\sigma_i = \sigma_j = -1$) in *all* configurations of state α, even if i and j are far apart from each other in terms of the shortest path length. To detect such long-range frustrations among unfrozen vertices, our idea is to flip the spin of one unfrozen vertex and then check whether this perturbation propagates to other unfrozen ones. This Letter reports our calculations on a spin-glass model [9] of the NP-hard minimum vertex-cover problem [12–15], which is equivalent to the hard-core gas condensation of physics [16]. A long-range frustration order parameter R is defined. In this model the quenched randomness comes from the underlying random graph. Work on systems with additional quenched randomness of spin-spin interactions is reported in an accompanying paper [11].

For the vertex-cover model, we show that long-range frustration builds up ($R > 0$) when the mean vertex degree c of the graph exceeds $c = e = 2.7183$. Analytical predictions on the ground-state energy density and on the fraction of frozen vertices are both in very good agreement with known numerical and mathematical results. The calculations are carried out through the cavity approach. It remains open whether the same results are achievable by the replica method. Our approach is essentially replica symmetric in the sense that (a) we focus attention on just one of all possible macroscopic states, and (b) the statistical property of this state is specified by just three mean field parameters to be defined, R, q_+, and q_0. Competitions among multiple states will be included in the theory in future work.

We first introduce the random graph vertex-cover problem. A random graph $G(N, c)$ has N vertices, and between any two vertices an edge is present with probability $c/(N-1)$. The average number of edges incident to a vertex is c (the mean vertex degree). For large graph size N, a vertex's probability of having k edges is given by the Poisson distribution $P_c(k) = e^{-c}c^k/k!$. Denote $E(G)$ as the edge set of graph G. A vertex cover of G consists of a set of vertices $\Lambda = \{i_1, i_2, \ldots, i_m\}$ such that if edge $(i, j) \in E(G)$, then either $i \in \Lambda$ or $j \in \Lambda$ or both. The vertex-cover problem consists of finding a vertex cover Λ with size $|\Lambda| \leq n_0$, n_0 being a prescribed integer. This problem is mapped to a spin-glass model with energy functional

$$E[\{\sigma_i\}] = -\sum_{i=1}^{N} \sigma_i + \sum_{(i,j)\in E(G)} (1+\sigma_i)(1+\sigma_j). \quad (1)$$

$\sigma_i = -1$ if vertex $i \in \Lambda$ (covered) and $\sigma_i = +1$ otherwise.

The ground-state configurations of model (1) correspond to vertex-cover patterns with the global minimum size [9]. These configurations may be grouped into different states [5]. Two configurations in the same state are mutually reachable by flipping a finite number of spins in one configuration and then letting the system relax. (According to this definition of states, two configurations of the same state can have a Hamming distance scaling linearly with system size N.) Let us focus on one state, say, α. In state α, the spin value of a randomly chosen vertex i may be fixed to $\sigma_i \equiv +1$, or to $\sigma_i \equiv -1$, or fluctuate over ± 1. The fraction of positively frozen, negatively frozen, and unfrozen vertices in state α is q_+, q_-, and q_0, respectively. [By the way, we notice that in the minimum vertex-cover problem, the parameters (q_+, q_-, q_0) are the same for different ground-state states, due to the fact that the energy density is determined by Eq. (6).] The probability that, among k vertices that are randomly picked up from $G(N, c)$, k_0 are unfrozen, k_+ positively frozen, and k_- $(= k - k_0 - k_+)$ negatively frozen is $k!/(k_0!k_+!k_-!)q_0^{k_0} q_+^{k_+} q_-^{k_-}$ (in the large N limit).

Since the spin of an unfrozen vertex i fluctuates among different configurations of state α, the "correlation length" of this fluctuation is an important issue. We ask the following question: If σ_i is externally fixed to $\sigma_i = -1$, how many other unfrozen vertices must eventually fix their spins as a consequence?

For a random graph of size $N \to \infty$, the total number s of affected vertices may scale linearly with N. If this happens, vertex i is referred to as type-I unfrozen. The probability for this to happen is denoted as R (which defines our long-range frustration order parameter). The total number of affected vertices may also be finite. In this case, vertex i is type-II unfrozen. Based on insights gained from studies on random graphs [17], we know that the percolation clusters evoked by two type-I unfrozen vertices have a nonzero intersection (of size proportional to N). Therefore, the spin values of all the type-I unfrozen vertices must be strongly correlated. If we randomly choose two type-I unfrozen vertices i and j, then with probability *one-half* their spin values cannot be negative simultaneously: if $\sigma_i = -1$, then σ_j must be $+1$; if $\sigma_j = -1$, then σ_i must be $+1$. On the other hand, two randomly chosen type-II unfrozen vertices are mutually independent, since each vertex can influence only the spin values of $s \sim O(1)$ other unfrozen vertices while the shortest path length between two randomly chosen vertices of $G(N, c)$ scales as $\ln N$ and becomes divergent when $N \to \infty$ [17]. Denote $f(s)$ as the probability that a randomly chosen unfrozen vertex i, when flipped to $\sigma_i = -1$, will eventually fix the spin values of s unfrozen vertices with s being finite and therefore $\lim_{N\to\infty} s/N = 0$.

We calculate the parameters q_0, q_+, q_- by the cavity field method [1,5]: First a random graph $G(N, c)$ is generated, then a new vertex i is connected to a set V_i of k randomly chosen vertices in $G(N, c)$, k following the distribution $P_c(k)$; the unfrozen or frozen probabilities $\{q_0(i), q_+(i), q_-(i)\}$ of vertex i in the enlarged graph (denoted as G') are then calculated. We assume the following convergence condition: $\lim_{N\to\infty}\{q_0(i), q_+(i), q_-(i)\} = \{q_0, q_+, q_-\}$. This enables us to write down a set of self-consistent equations in the large N limit.

If the new vertex i is positively frozen, then none of the vertices in V_i are positively frozen in graph G. Furthermore, there are two possible situations: (i) no vertices in V_i are type-I unfrozen in G, or (ii) some of the vertices in V_i are type-I unfrozen. In case (ii), all these type-I unfrozen vertices will take spin value -1 simultaneously in some configurations of state α, so that vertex i will have $\sigma_i \equiv +1$ as it is added into the system. With this analysis, we get a self-consistent equation for q_+:

$$\begin{aligned} q_+ &= \sum_{k=1}^{\infty} P_c(k) \sum_{l=1}^{k} C_k^l (q_0 R)^l [q_0(1-R) + q_-]^{k-l} 2^{1-l} \\ &\quad + \sum_{k=0}^{\infty} P_c(k)[q_0(1-R) + q_-]^k \\ &= 2e^{-cq_+ - (1/2)cq_0 R} - e^{-cq_+ - cq_0 R}, \end{aligned} \quad (2)$$

where $C_k^l = k!/(l!(k-l)!)$. Equations (2)–(4) and (7) are derived elsewhere [18].

If the new vertex i is unfrozen, there are also two possibilities concerning the spin values of vertices in V_i: (iii) none of them is positively frozen in G, or (iv) one of them is positively frozen in G. To ensure vertex i will be unfrozen, in situation (iii) two or more of the vertices in V_i must be type-I unfrozen in G, among which one is in conflict with all the others, and in situation (iv) some of the vertices in V_i may be type-I unfrozen in G, but they must be capable of taking spin value -1 simultaneously. Therefore, we get a self-consistent equation for q_0 [18]:

$$\begin{aligned} q_0 &= (2cq_+ + cq_0 R)e^{-cq_+ - (1/2)cq_0 R} - [cq_+ + cq_0 R \\ &\quad + (cq_0 R)^2/4]e^{-cq_+ - cq_0 R}. \end{aligned} \quad (3)$$

If the spin of an unfrozen vertex i is flipped to $\sigma_i = -1$, it may not affect any other vertices ($s = 0$), provided its local environment is described by the above-mentioned situation (iii). This happens with probability

$p_1 = 1 - cq_+^2/q_0$. With probability $1 - p_1$, the unfrozen vertex i encounters a local environment of type (iv); that is, one of its nearest neighbors vertex j is positively frozen in graph G. This vertex j must face the local environment of type (i) in graph G if vertex i is type-II unfrozen. [If vertex j has the local environment of type (ii), flipping the spin value of vertex i to $\sigma_i = -1$ would cause a percolation cluster of size proportional to N.] With these preparations, we obtain the following self-consistent equation for the distribution $f(s)$ [18]:

$$f(s) = p_1 \delta_s^0 + (1 - p_1) \sum_{l=0} P_{c'}(l) \sum_{\{s_m\}} \prod_m^l f(s_m) \delta_{s-1}^{s_1 + \cdots + s_l}. \quad (4)$$

In Eq. (4), δ is the Kronecker symbol, and $c' = cq_0(1 - R)$ is the mean number of type-II unfrozen vertices adjacent to a positively frozen vertex. Since $R = 1 - \sum_{s=0}^{\infty} f(s)$, we establish that the long-range frustration order parameter R is determined by the following equation:

$$R = (cq_+^2/q_0)(1 - e^{-cq_0 R(1-R)}). \quad (5)$$

A positive R signifies the appearance of a percolation cluster of unfrozen vertices whose spin values are strongly correlated.

Figure 1 shows the value of R as a function of mean vertex degree c. $R \equiv 0$ when $c \leq e$; this is consistent with Ref. [19] that, a minimal vertex-cover pattern can be found by a polynomial leaf-removal algorithm. When $c > e$, a finite fraction of the unfrozen vertices are long-range frustrated; the leaf-removal algorithm outputs a looped subgraph [19]. At mean vertex degree $c \simeq 40$, the order parameter R reaches a maximal value; then it gradually decays as c is further increased.

The fraction of vertices that are covered in a minimal vertex cover is [18]

$$X_{\min} = 1 - q_+ - q_0/2. \quad (6)$$

Figure 2 shows the relationship between $X_{\min}$ and mean vertex degree c. At large c values, Eq. (6) is in agreement with a rigorous asymptotic expression given by Frieze [20]; at low values of c, it is in agreement with the exact enumeration results of Weigt and Hartmann [13]. These excellent agreements are quite encouraging, in view of the fact that all previous efforts failed [9,13,14]. It has already been established that when $c > e$ the replica symmetric solution of the vertex-cover problem becomes unstable [13,14], but earlier replica symmetry breaking solutions either resulted in negative structural entropy or predicted a minimal vertex-cover size noticeably lower than the actual value [9].

So far we have focused on only one ground-state state of the vertex-cover problem. When $c > e$, it is believed that there are many such states (replica symmetry breaking). This is consistent with our observation that, when $c > e$, the fraction of frozen vertices ($= q_+ + q_-$, dashed lines in Fig. 3) in one state is much higher than the actual fraction of frozen vertices estimated numerically (symbols in Fig. 3) [14]. This is easy to understand: A frozen vertex in one state may be unfrozen or be frozen to the opposite spin value in another state. At the moment we are unable to construct a theory to include the competitions among different states. As a first attempt, we make the following conjectures: (a) if a vertex is positively frozen in one state, it is positively frozen in all states, and (b) a vertex is negatively frozen in all states only if it is adjacent to two or more positively frozen vertices. Then an expression on the fraction of frozen vertices is obtained [18]:

$$\Gamma = q_+ + 1 - e^{-cq_+} - cq_+ e^{-cq_+}. \quad (7)$$

The agreement of Eq. (7) with the numerical data of Ref. [14] is quite good (Fig. 3). This is an issue to be understood more deeply.

To summarize, we have studied long-range frustrations among *unfrozen* vertices in a macroscopic state of a spin-glass system. We found that, with certain probability, the

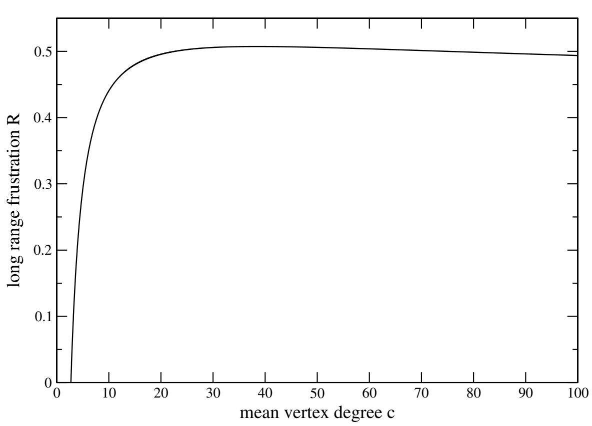

FIG. 1. The long-range frustration order parameter R.

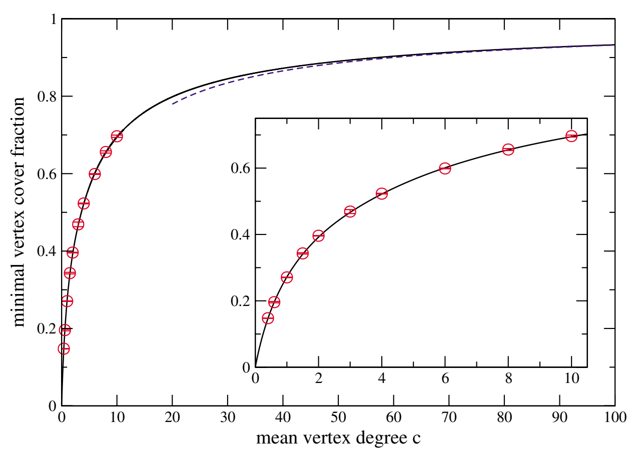

FIG. 2 (color online). The minimal vertex-cover fraction $X_{\min}$ [Eq. (6), solid line] and its comparison with the asymptotic formula of Ref. [20] (dashed line) and the numerical results of Refs. [13] (symbols).

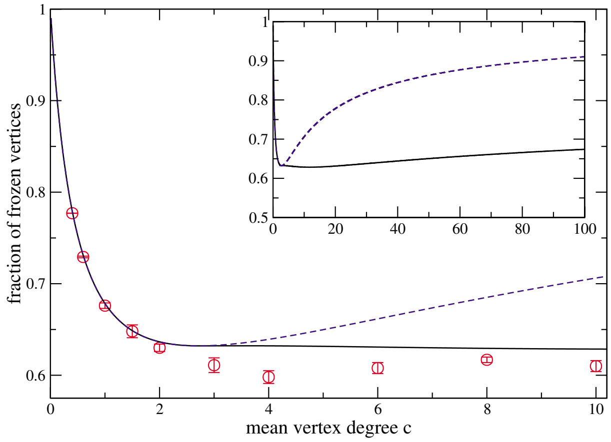

FIG. 3 (color online). Fraction of frozen vertices in *all* macroscopic states [Eq. (7), solid lines] and its comparison with the numerical results of Ref. [14] (symbols) and the fraction of frozen vertices in *one* macroscopic state (dashed lines).

fluctuations of the spin values of two or more distantly separated unfrozen vertices are highly correlated. A long-range frustration order parameter R was calculated to quantify this strong correlation. When applying our method to the NP-hard minimum vertex-cover (hard-core gas condensation) problem, the analytical predictions concerning the ground-state energy density and the fraction of frozen vertices are in good agreement with known numerical and rigorous results. The basic idea behind this Letter is also applicable to other spin-glass systems [11].

We emphasize that the appearance of many macroscopic states in the energy landscape of a spin-glass system does not necessarily mean the existence of long-range frustrations among unfrozen vertices in a single macroscopic state. As a counterexample, in the maximum matching problem [21] there is no long-range frustrations ($R \equiv 0$), but there exist an exponential number of macroscopic states. It is interesting to notice that the maximum matching problem can be solved by polynomial algorithms. It appears that the proliferation of macroscopic states is not the real reason of the computational complexity in finding a ground-state configuration for a disordered system. As another example, there are many macroscopic states in a typical random three-satisfiability formula when $3.921 < \alpha < 4.267$ (here α is the clauses-to-variables ratio), but the survey propagation algorithm is able to find a solution efficiently [6,7].

On the other hand, we believe the existence of long-range frustrations among unfrozen vertices will make it intrinsically difficult for a search algorithm to find a ground-state configuration. Because of these long-range effects, it is difficult (a) to determine whether a vertex is frozen or unfrozen in a macroscopic state and (b) to trace the percolation cluster associated with a given unfrozen vertex. Recently, some NP-hard combinatorial optimization problems in computer science were studied by the zero temperature cavity field method [6–9]. We hope the present work, besides improving our understanding of finite connectivity spin glasses, will stimulate further efforts in finding more efficient algorithms. We are presently implementing the physical picture of this Letter into an algorithm for the vertex-cover problem.

H. Z. is grateful to R. Lipowsky, Z.-C. Ou-Yang, and L. Yu for their support, to A. K. Hartmann and M. Weigt for sharing their numerical results, and to L. Yu and R. Lipowsky for helpful discussions.

[1] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).
[2] L. Viana and A. J. Bray, J. Phys. C **18**, 3037 (1985).
[3] I. Kanter and H. Sompolinsky, Phys. Rev. Lett. **58**, 164 (1987).
[4] R. Monasson, J. Phys. A **31**, 513 (1998).
[5] M. Mézard and G. Parisi, J. Stat. Phys. **111**, 1 (2003).
[6] M. Mézard, G. Parisi, and R. Zecchina, Science **297**, 812 (2002).
[7] M. Mézard and R. Zecchina, Phys. Rev. E **66**, 056126 (2002).
[8] R. Mulet, A. Pagnani, M. Weigt, and R. Zecchina, Phys. Rev. Lett. **89**, 268701 (2002).
[9] H. Zhou, Eur. Phys. J. B **32**, 265 (2003).
[10] O. Rivoire, J. Stat. Phys. **117**, 453 (2004).
[11] H. Zhou, New J. Phys. **7**, 123 (2005).
[12] M. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, San Francisco, 1979).
[13] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000).
[14] M. Weigt and A. K. Hartmann, Phys. Rev. E **63**, 056127 (2001).
[15] A. K. Hartmann and M. Weigt, J. Phys. A **36**, 11 069 (2003).
[16] G. Biroli and M. Mézard, Phys. Rev. Lett. **88**, 025501 (2002).
[17] B. Bollobás, *Random Graphs* (Academic Press, London, 1985).
[18] See EPAPS Document No. E-PRLTAO-94-030524 for technical details of the calculation. A direct link to this document may be found in the online article's HTML reference section. The document may also be reached via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html) or ftp.aip.org in the directory /epaps/. See the EPAPS homepage for more information.
[19] M. Bauer and O. Golinelli, Eur. Phys. J. B **24**, 339 (2001).
[20] A. M. Frieze, Discrete Mathematics **81**, 171 (1990).
[21] H. Zhou and Z.-C. Ou-Yang, cond-mat/0309348.

Erratum: Long-Range Frustration in a Spin-Glass Model of the Vertex-Cover Problem [Phys. Rev. Lett. 94, 217203 (2005)]

Haijun Zhou

(Received 9 October 2012; published 9 November 2012)

DOI: 10.1103/PhysRevLett.109.199901 PACS numbers: 75.10.Nr, 02.10.Ox, 05.20.−y, 89.75.−k, 99.10.Cd

We found several mistakes in this Letter, which significantly affect the quantitative theoretical results.

The expression [Eq. (5)] for the long-range frustration order parameter R was wrong. For an unfrozen vertex i to be type-I unfrozen in graph G, one of its nearest neighbors (say j) must be positively frozen and facing the local environment of type (ii) in G' (the graph obtained by removing i and its edges from G). The correct formula for R is

$$R = \frac{cq_+^2}{q_0}\left(1 - \frac{1}{q_+} e^{-cq_+ - cq_0 R}\right).$$

R as a function of the mean vertex degree c is shown in the corrected Fig. 1. It is positive for $c > e = 2.718\ldots$ and its maximum is reached at $c \simeq 14.85$.

The self-consistent equation (4) for the size distribution $f(s)$ was wrong. The quantity p_1 of this expression should be replaced by $p_1' = p_1/(1 - R)$, which is the conditional probability that a vertex i faces the environment of type (iii) given that i is a type-II unfrozen vertex in G. The function $f(s)$ satisfies $\sum_{s=0}^{\infty} f(s) = 1$.

Equation (6) for the fraction of covered vertices $X_{\min}$ was also wrong. According to the analysis in Ref. [1], the correct formula should be

$$X_{\min} = \frac{1}{c}\int_0^c [1 - q_+(c')]dc',$$

where $q_+(c)$ is the fraction of positively frozen vertices at mean vertex degree c. $X_{\min}$ as a function of c is shown in the updated Fig. 2. The theoretical prediction is in agreement with simulation results obtained on single large graphs [2]; it is slightly lower than the enumeration result obtained on single small graphs [3]. When the mean vertex degree c is large, the value of $X_{\min}$ is slightly lower than the asymptotic result of Ref. [4].

Financial support from the Chinese Academy of Sciences (KJCX2-EW-J02) and NSFC (Grants No. 10834014 and 11121403) is acknowledged.

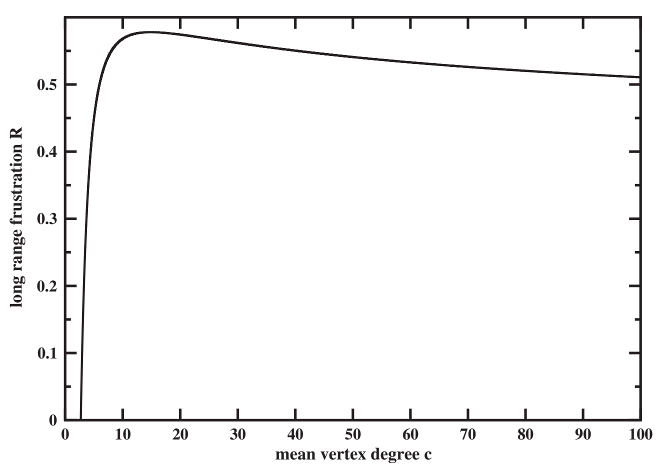

FIG. 1. The long-range frustration order parameter R.

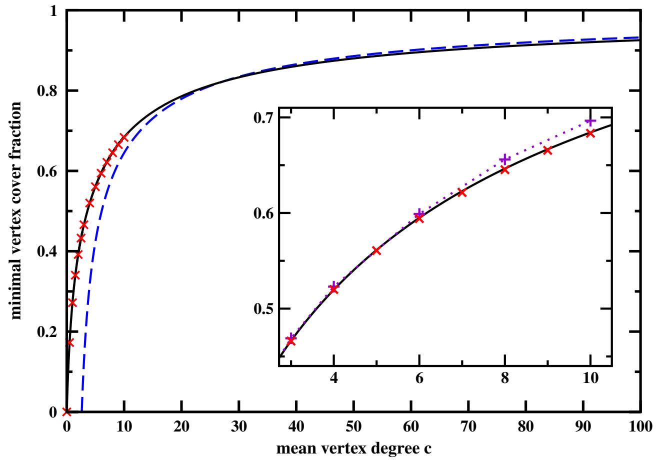

FIG. 2 (color online). The minimal vertex-cover fraction $X_{\min}$ (solid line) and its comparison with the asymptotic formula of Ref. [4] (dashed line), the numerical results of Ref. [3] (+ symbols), and Ref. [2] (× symbols).

[1] H. Zhou, New J. Phys. **7**, 123 (2005).
[2] M. Weigt and H. Zhou, Phys. Rev. E **74**, 046110 (2006).
[3] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000).
[4] A. M. Frieze, Discrete Math. **81**, 171 (1990).