TOPICAL REVIEW — Statistical physics and complex systems

Statistical physics of hard combinatorial optimization: Vertex cover problem*

Zhao Jin-Hua(赵金华) and Zhou Hai-Jun(周海军)†

State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, *China*

(Received 20 March 2014; published online 22 May 2014)

Typical-case computation complexity is a research topic at the boundary of computer science, applied mathematics, and statistical physics. In the last twenty years, the replica-symmetry-breaking mean field theory of spin glasses and the associated message-passing algorithms have greatly deepened our understanding of typical-case computation complexity. In this paper, we use the vertex cover problem, a basic nondeterministic-polynomial (NP)-complete combinatorial optimization problem of wide application, as an example to introduce the statistical physical methods and algorithms. We do not go into the technical details but emphasize mainly the intuitive physical meanings of the message-passing equations. A nonfamiliar reader shall be able to understand to a large extent the physics behind the mean field approaches and to adjust the mean field methods in solving other optimization problems.

Keywords: spin glass, energy minimization, replica symmetry breaking, belief propagation

PACS: 89.20.Ff, 75.10.Nr, 02.10.Ox, 05.10.–a **DOI:** 10.1088/1674-1056/23/7/078901

1. Introduction

The notion of computation complexity was introduced by Cook in 1971,[1–4] which distinguishes computation problems according to how the computing time scales with the size of the problem. Class P (polynomial) problems can be solved in polynomial time, namely the computing time t is bounded by a polynomial function of the problem's number N of variables, $t = O(N^c)$ with c being a finite constant. Class NP (nondeterministic polynomial) problems, however, may need an exponentially increasing time ($t \sim e^{cN}$) to solve in the worst case. The most difficult problems in the class NP are referred to as NP-complete problems, which are problems that can be mutually converted into each other by a polynomial algorithm. If one can solve all the instances of one NP-complete problem in polynomial time, she or he can simultaneously solve all the NP-complete problems in polynomial time. The existence or not of a polynomial-time algorithm for NP-complete problems is the famous and basic $P = NP$? problem of computation complexity.

Whether a problem belongs to the NP-complete class is judged by the worst-case computation difficulty. However, a typical problem instance of a NP-complete problem might actually be very easy to solve. In the last twenty years, typical-case computation complexity has been intensively studied as an emerging interdisciplinary research topic of mathematics, theoretical computer science and statistical physics.[5] Statistical physics concepts and methods have played a very significant role in understanding typical-case computation complexity.[6–10] Using the replica method[6,7] and the cavity method[11,12] of spin glass mean field theory, many interesting and fundamental optimization problems have been investigated in the statistical physics community, such as the travelling salesman problem,[13] the K-satisfiability problem,[14–18] the exclusive-or-satisfiability (XOR-SAT) problem,[19–21] the vertex cover problem (or independent set) problem and the hitting set problem,[22–32] the graph coloring problem,[33–35] the maximal matching problem,[36–38] and the feedback vertex set problem.[39] In this review, we take a single prototypic combinatorial optimization problem, the vertex cover problem, as an example to introduce the ideas behind the statistical physical methods and algorithms. Aimed at a reader outside the spin glass research field, we do not discuss the technical details but focus mainly on the intuitive physical picture behind the message-passing equations. Hopefully a motivated reader will easily grasp the essential ingredients of the mean field approaches and further adjust the methods to other optimization problems.

The layout of the paper is as follows. Section 2 introduces the vertex cover problem. Section 3 summarizes some mathematical results on the minimal vertex cover problem. Section 4 briefly mentions some local search algorithms. We introduce in Section 5 the concept of long range frustration and link it to computational difficulty. We then introduce a spin glass model in Section 6 and discuss two message-passing algorithms. Some additional discussions are made in Section 7.

*Project supported by the National Basic Research Program of China (Grant No. 2013CB932804), the Knowledge Innovation Program of Chinese Academy of Sciences (Grant No. KJCX2-EW-J02), and the National Natural Science Foundation of China (Grant Nos. 11121403 and 11225526).

†Corresponding author. E-mail: zhouhj@itp.ac.cn

© 2014 Chinese Physical Society and IOP Publishing Ltd

2. Vertex cover problem

A graph $G = (V, E)$ is composed of a set V of vertices and a set E of edges. There are N vertices in the graph, therefore $V = 1, 2, \ldots, N$. Each edge connects between two different vertices; for example, (i, j) denotes an edge between vertex i and vertex j. Here we consider sparse graphs such that the number of edges in graph G is of the same order as the number of vertices. A vertex cover (VC) $V_{\rm vc}$ of graph G is a subset of vertices of the graph, which contains at least one of the incident vertices of every edge in the set E. For example, if edge $(i, j) \in E$, then either $i \in V_{\rm vc}$ or $j \in V_{\rm vc}$ or both. Figure 1 shows three VCs for a small graph.

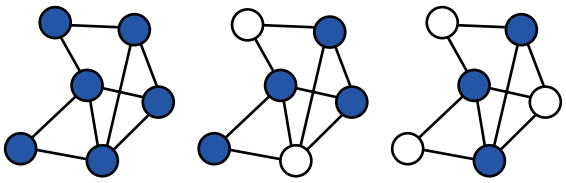


Fig. 1. Three vertex covers for a small graph of $N = 6$ vertices and $M = 9$ edges. A vertex is represented by a filled circle if it belongs to the VC, otherwise it is represented by an open circle.

With respect to a given vertex cover $V_{\rm vc}$, a vertex i is referred to as being covered if it belongs to $V_{\rm vc}$, otherwise the vertex is referred to as being uncovered.

The vertex cover problem is one of the first 21 problems shown to be NP-complete; it has fundamental importance in the field of computation complexity.[2] This problem also has wide practical applications, for example internet traffic monitoring,[40] prevention of denial-of-service attacks,[41] immunization strategies in networks,[42] and network source location problem.[43]

The vertex cover problem can be expressed either as a decision problem or as an optimization problem. As a decision problem, we ask whether there exists a vertex cover $V_{\rm vc}$ with cardinality $|V_{\rm vc}|$ less than a certain given value. As an optimization problem, we need to construct a vertex cover whose cardinality is the global minimum over all possible VCs for a given graph G. In this review, we focus on the optimization problem for random graphs. In our following discussions, we refer to the relative size (with respect to the vertex number N) of a vertex cover $V_{\rm vc}$ as its energy density and denote it by x, namely

$$x \equiv \frac{|V_{\rm vc}|}{N} . \tag{1}$$

The global minimal energy density for a given graph is denoted as x_0. If the cardinality of a VC for a given graph is the global minimum among all the VCs, it is referred to as an optimal VC.

3. Mathematical bounds and asymptotics

Here we list some established results on rigorous bounds and asymptotic behaviors concerning the global minimum of VCs.

On a general graph G, Harant's upper bound[44] on the minimal energy density x_0, obtained by generalizing the earlier work of references,[45,46] is expressed as

$$x_{\rm c}(G) \leq 1 - \frac{1}{N} \frac{(\sum_{i\in V} \frac{1}{d_i+1})^2}{\sum_{i\in V} \frac{1}{d_i+1} - \sum_{(i,j)\in E} \frac{(d_i-d_j)^2}{(d_i+1)(d_j+1)}} , \tag{2}$$

where d_i is the degree (the number of attached edges) of vertex i. More refined upper bounds of x_0, not in the form of explicit expressions, can be found in Refs. [47]–[49].

A random graph of mean vertex degree c is obtained by setting up $M = (c/2)N$ edges completely at random starting from an empty graph of N vertices.[50,51] For such random graphs, the work of Gazmuri[52] predicts that the minimal energy density x_0 almost surely are bounded by $x_l < x_0 < 1 - \ln c/c$, where x_l is the root of

$$x \ln x + (1-x)\ln(1-x) + (c/2)(1-x)^2 = 0 . \tag{3}$$

These bounds are further improved by using the method of weighted second moment.[53] In the case of $c \to \infty$, Frieze obtained the following asymptotic expression for the minimal energy density[54]

$$x_0 = 1 - (2/c)(\ln c - \ln\ln c + 1 - \ln 2) + o(1/c) . \tag{4}$$

4. Some heuristic local algorithms

An algorithm that is guaranteed to find a VC of global minimal cardinality is branch-and-bound (see Ref. [25] for a detailed description). This algorithm performs an optimized search over all the VCs of a given graph to determine the global minimum of VC cardinality. Since the search space increases exponentially with vertex number N, this algorithm works only for small graphs.

There exist also many heuristic algorithms, which construct VCs based on some local rules. One very simple heuristic algorithm is maximum degree decimation, which recursively adds a vertex of the largest degree into the vertex cover set and then reduces the graph by deleting this vertex and its connected edges. We can improve the performance of this greedy algorithm by combining it with a leaf-removal process.[55] A vertex i is considered as a leaf vertex if this vertex is attached by only one edge, say (i, j). For such a leaf vertex i, it is always an optimal choice to add the neighboring vertex j instead of vertex i into a VC. The combined heuristic algorithm then works as follows: as long as there is a leaf vertex, add the neighboring vertex of this leaf vertex to the VC

and simplify the graph G, otherwise add a vertex of the largest degree into the VC and simplify G. The performance of such an algorithm on random graphs is shown in Fig. 2. When the mean vertex degree $c < e = 2.718\cdots$, this algorithm has a high probability of constructing a VC of global minimal cardinality, but for $c > e$, the energy density of the constructed VC is higher than the minimal value x_0.

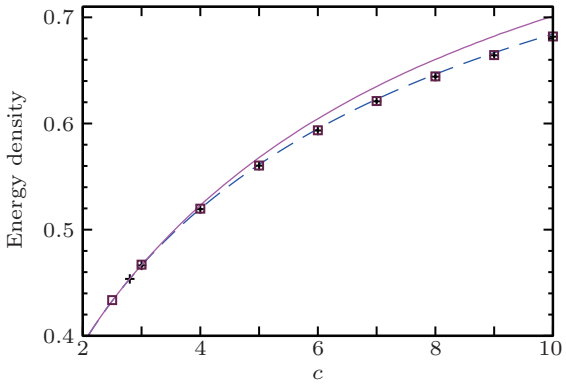


Fig. 2. The solid line is the energy density $x(c)$ obtained by the hybrid algorithm of leaf-removal and maximum degree decimation on a single random graph of vertex number $N = 10^6$ and mean vertex degree c. The dashed line is the energy density minimum $x_0(c)$ predicted by the long range frustration theory at $N = \infty$. Each square symbol is the mean energy density obtained by the SPD algorithm on 16 random graph instances of vertex number $N = 10^5$, while each plus symbol shows the energy density minimum $x_0(c)$ predicted by the first-step replica-symmetry-breaking (1RSB) mean field theory at $N = \infty$.

The vertex cover problem can also be solved by Monte Carlo optimization methods.[56,57] For example, reference [58] adopts the parallel tempering technique[59,60] and obtains near-optimal VCs for relatively large single random graphs.

5. Long range frustrations

Theoretical analysis revealed that the structure of a random graph has a continuous phase transition at mean connectivity $c = e$, characterized by the emergence of a core of macroscopic size.[55,61–63] A random graph has no core if its mean vertex degree $c < e$, therefore, the leaf-removal process can delete all the edges of the graph and construct an optimal VC.

On the other hand, the existence of a core at $c > e$ leads to very complicated long range frustrations among the covering states of different variables, making it impossible for a local algorithm to find an optimal VC.[61,62] Let us focus on the set Γ_0 of optimal VCs. If a vertex i is always being covered or always being uncovered in all the VCs of set Γ_0, then it is regarded as a frozen vertex with respect to Γ_0. Otherwise vertex i is an unfrozen vertex with respect to Γ_0, meaning that it is being covered in some (but not all) of the VCs of set Γ_0. To explain the picture of long range frustrations, let us randomly pick up two of such unfrozen vertices, say j and k. These two vertices might be far apart in terms of the shortest-distance path length. There are four possible joint states for these two vertices. If all these four joint states can be observed in at least one VC of set Γ_0, then vertices j and k are regarded as being unfrustrated, and fixing one vertex to the covered or uncovered state will not affect the other vertex. However, if at least one joint state (say vertex j being covered and vertex k being uncovered) of these two vertices is not observed in any VC of set Γ_0, then these two vertices are said to be (long-range) frustrated. In such a case, fixing vertex j to the covered state will cause vertex k to also be covered, even though j and k might be extremely separated in the graph.

Such complicated long range frustration effects can be quantitatively considered by a mean field theory. For a random graph of mean vertex degree c, the long range frustration theory of Refs. [61] and [62] predicts the minimal energy density x_0 to be

$$x_0 = 1 - \frac{1}{c}\int_0^c r_0(\tilde{c})\mathrm{d}\tilde{c}\,. \tag{5}$$

In this expression, r_0 is the fraction of vertices that do not belong to any optimal VC, whose value is determined by solving the following three equations involving r_0 and two other quantities r_* and R (for details, see Refs. [61], [62], and [64]):

$$r_0 = 2\,\mathrm{e}^{-cr_0-cr_*R/2} - \mathrm{e}^{-cr_0-cr_*R}\,, \tag{6}$$

$$r_* = (2cr_0 + cr_*R)\,\mathrm{e}^{-cr_0-cr_*R/2} - \left(cr_0 + cr_*R + (cr_*R)^2/4\right)\mathrm{e}^{-cr_0-cr_*R}, \tag{7}$$

$$R = \frac{cr_0^2}{r_*}\left(1 - \frac{1}{r_0}\,\mathrm{e}^{-cr_0-cr_*R}\right). \tag{8}$$

The predicted values of $x_0(c)$ by this long range frustration theory are in good agreement with the empirical results obtained by the survey propagation-guided decimation (SPD) algorithm[27] (see Section 6) and with the theoretical results obtained by the first-step replica-symmetry-breaking (1RSB) mean field theory,[26,64] see Fig. 2. This indicates that the physical picture behind the long range frustration theory is one of the main reasons for the difficulty of obtaining optimal VCs for random graphs with mean vertex degree $c > e$. An extension of the long range frustration theory was made in Ref. [65], which discussed the backbone structure of optimal VCs.

6. Message-passing algorithms

For random graphs with mean vertex degree $c > e$, because of the existence of long range frustrations among the vertices, the minimal vertex cover problem is in a spin glass phase. In such a phase, the optimal VCs are distributed into many clusters. Each VC cluster contains a number of highly similar VCs, while the VCs of different clusters are much less similar with each other. Besides optimal VCs, the system also

has an enormous number of local minimal VCs, which also form many clusters. The number of local minimal VCs exponentially exceeds that of optimal VCs. Therefore, a local search algorithm will be trapped into one of the local minimal VCs with certainty.

In the last twenty years, the mean field theoretical methods of spin glasses, namely the replica method[6,7,9,66] and the cavity method,[8,11] have been applied on the vertex cover problem[22,24,26,27,64] to describe its complex energy landscape. From the physical point of view, the cavity method is based on the Bethe–Peierls approximation of statistical physics[67–69] and the physical picture that the configuration space can be regarded as a collection of macroscopic states (each macroscopic state itself contains a set of microscopic configurations). The cavity method can also be understood from a more mathematical framework of partition function expansion.[70,71] This method is particularly convenient for investigating single problem instances. Here we will describe two message-passing algorithms inspired by the cavity method, namely belief propagation and survey propagation.

6.1. Spin glass model

Since each vertex i has two candidate covering states $s_i = 0$ (uncovered) and $s_i = 1$ (covered), the total number of microscopic configurations is 2^N. We introduce the following partition function $Z(\beta)$ for the vertex cover problem:[27,64]

$$Z(\beta) = \sum_{\underline{s}} \prod_{i=1}^{N} \mathrm{e}^{-\beta s_i} \prod_{(j,k)\in G} \left[1-(1-s_j)(1-s_k)\right], \qquad (9)$$

where the summation is over all the 2^N microscopic configurations $\underline{s} \equiv \{s_1, s_2, \ldots, s_N\}$. The edge product term of Eq. (9) guarantees that $s_j + s_k \geq 1$ for each edge (j,k) of the graph (i.e., at least one of the two vertices j and k is in the covered state), otherwise the microscopic configuration s has no contribution to $Z(\beta)$. Therefore, only VCs contribute to $Z(\beta)$. The positive reweighting parameter β emphasizes VCs of smaller cardinality. In the limit of $\beta \to \infty$, the partition function $Z(\beta)$ is contributed exclusively by the optimal VCs.

6.2. Belief propagation

Let us denote by $p_i^{(0)}$ the marginal probability that a vertex i is in the uncovering state $s_i = 0$. Assuming the covering states of vertex i's neighboring vertices are independent of each other in the absence of i (i.e., the Bethe–Peierls approximation), we obtain the following expression for $p_i^{(0)}$:

$$p_i^{(0)} = \frac{\prod_{j\in\partial i}(1-p_{j\to i}^{(0)})}{\mathrm{e}^{-\beta} + \prod_{j\in\partial i}(1-p_{j\to i}^{(0)})}, \qquad (10)$$

where ∂i denotes the set of neighboring vertices of vertex i, and $p_{j\to i}^{(0)}$ is the probability of vertex j being in the uncovering state $s_j = 0$ in the absence of the edge (i,j). To understand the above expression, we notice that, when the covering state of vertex i is $s_i = 0$, all its neighboring vertices j must be in the covering state $s_j = 1$. Such a requirement leads to the product term of the denominator (and also that of the numerator). On the other hand, if vertex i is in the covered state, the constraints on all the edges attached to i are simultaneously satisfied, but the VC cardinality increases by 1, leading to a Boltzmann factor $\mathrm{e}^{-\beta}$ in the denominator of the above expression.

Under the same Bethe–Peierls approximation, we can write down the equation for $p_{j\to i}^{(0)}$ as

$$p_{j\to i}^{(0)} = \frac{\prod_{k\in\partial j\backslash i}(1-p_{k\to j}^{(0)})}{\mathrm{e}^{-\beta} + \prod_{k\in\partial j\backslash i}(1-p_{k\to j}^{(0)})}, \qquad (11)$$

where $\partial k\backslash i$ denotes the set of neighboring vertices of vertex k (excluding vertex i). Equation (11) is referred to as the belief propagation (BP) equation for the vertex cover problem. This equation can also be derived from the framework of partition function expansion.[70,71] For a graph with M edges, there are $2M$ such equations. We can try to solve this set of equations by numerical iteration. If a fixed point can be reached by this iteration process, the energy density x (i.e., the mean fraction of vertices in the covered state) is then evaluated as

$$x = 1 - \frac{1}{N}\sum_{i=1}^{N} p_i^{(0)}. \qquad (12)$$

The entropy density of VCs at the energy density x can also be computed.[28]

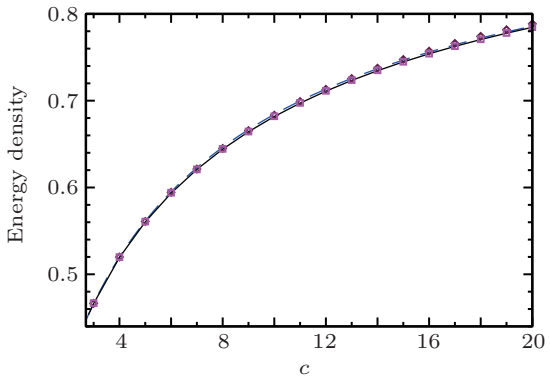


Fig. 3. The solid line and the dashed line are, respectively, the curve of energy density minimum $x_0(c)$ predicted by the 1RSB mean field theory and the long range frustration theory for random graphs of mean vertex degree c and $N = \infty$. Dimond symbols and square symbols are, respectively, the VC energy densities reached by the BPD algorithm ($\beta = 10$) and the SPD algorithm ($y = 3.05$) on a single random graph instance of $N = 10^5$ vertices.

Based on Eqs. (10) and (11), we have implemented a simple belief propagation-guided decimation (BPD) algorithm as follows. At a given value of β, we iterate the BP equation (11) a number of steps on a given graph G, and then compute the marginal probabilities $p_i^{(0)}$ for all the vertices. Then a small fraction of vertices i with the smallest values of $p_i^{(0)}$ are added to a VC and deleted from the graph G. We then simplify the

graph G and, if the simplified G still contains some edges, we repeat the above mentioned iteration-fixation process. Figure 3 demonstrates that the performance of this BPD algorithm is very good on random graphs. At each mean vertex degree c, the energy density of the constructed VC by BPD is very close to the energy density minimum $x_0(c)$ predicted by the 1RSB mean field theory.

6.3. Survey propagation

Although the BPD algorithm seems to work excellently on single random graph instances, the iteration of the BP equation (11) actually cannot converge to a fixed point when the reweighting parameter β is sufficiently large.[29] The reason behind this non-convergence is the breaking of ergodicity. When β is sufficiently large, we need to extend the Bethe–Peierls approximation to include the possibility of the existence of many macroscopic states.

As we are interested in the optimal VCs, we focus on the limiting case of $\beta = \infty$. Then each macroscopic state α of the configuration space is characterized by a minimal energy density $x^{(\alpha)}$. We can define a partition function at the level of macroscopic states as

$$\Xi(y) = \sum_{\alpha} \mathrm{e}^{-yNx^{(\alpha)}}, \tag{13}$$

where the summation is over all the macroscopic states, and y is a reweighting parameter at the macroscopic states level. A larger value of y favors macroscopic states of smaller minimal energies.

At $\beta = \infty$, the relevant microscopic configurations must all be energy local or global minimal points of the energy landscape. A macroscopic state is then composed of a set of minimal energy configurations of the same energy. In such a macroscopic state α, we can describe the covering state of each vertex i in a coarse-grained way: (i) if the covering state $s_i = 1$ in all the minimal energy configurations of α, then i is said to be frozen to the covered state, with coarse-grained state $S_i = 1$; (ii) if $s_i = 0$ in all the minimal energy configurations of α, then i is said to be frozen to the uncovered state, with coarse-grained state $S_i = 0$; (iii) in the remaining cases, $s_i = 0$ in some (but not all) minimal energy configurations and $s_i = 1$ in the remaining configurations of α, and then we regard i as being unfrozen, with coarse-grained state $S_i = *$.

Let us denote by $\pi_i^{(0)}$ the probability that vertex i is in the coarse-grained state $S_i = 0$. If we assume that the coarse-grained states of the neighboring vertices of i are all independent in the absence of i (this is the Bethe–Peierls approximation at the level of coarse-grained states), the following expression for $\pi_i^{(0)}$ can be written down:[26,27]

$$\pi_i^{(0)} = \frac{\prod_{j\in\partial i}(1-\pi_{j\to i}^{(0)})}{\mathrm{e}^{-y} + (1-\mathrm{e}^{-y})\prod_{j\in\partial i}(1-\pi_{j\to i}^{(0)})}, \tag{14}$$

where $\pi_{j\to i}^{(0)}$ is the probability that the neighboring vertex j is in the coarse-grained covering state $S_j = 0$ in the absence of vertex i. The above expression has a clear intuitive interpretation. If all the neighbors j of the central vertex i are not in the coarse-grained covering state $S_j = 0$ before vertex i is added to the graph, then there should exist at least one macroscopic state in which all these neighboring vertices are in the covered state. When i is added to the graph, then its covering state should be set to $s_i = 0$ to decrease energy. In all the other cases, the addition of vertex i will cause an increase of the minimal energy by 1, which explains the Boltzmann factor e^{-y} in the above equation.

Under similar considerations, we have the following equation for the probability $\pi_{j\to i}^{(0)}$:

$$\pi_{j\to i}^{(0)} = \frac{\prod_{k\in\partial j\setminus i}(1-\pi_{k\to j}^{(0)})}{\mathrm{e}^{-y} + (1-\mathrm{e}^{-y})\prod_{k\in\partial j\setminus i}(1-\pi_{k\to j}^{(0)})}. \tag{15}$$

This equation is referred to as the survey propagation (SP) equation for the vertex cover problem.[11,15,26,27]

The mean energy density and the entropy density Σ at the level of macroscopic states can both be expressed as functions of the probabilities $\{\pi_{i\to j}^{(0)}, \pi_{j\to i}^{(0)} : (i,j) \in G\}$. Such a theoretical procedure is referred to as the first-step replica-symmetry-breaking (1RSB) mean field procedure. It can be justified again through the framework of partition function expansion.[70,71] In the numerical calculations, the reweighting parameter y is set to be the largest value such that the entropy density Σ is non-negative. The energy density at this specific y is then regarded as the global minimum energy density. We show in Figs. 2 and 3 the ensemble-averaged minimum energy density obtained in such a way as a function of mean vertex degree c. These figures show that the 1RSB predictions are in close agreement with the predictions of the long range frustration theory and with the empirical results obtained by the message-passing algorithms.

For single graph instances, similar to the BPD algorithm, we can use the information obtained by Eqs. (14) and (15) to construct near-optimal VCs. A survey propagation-guided decimation (SPD) algorithm runs similarly to the BPD algorithm: We iterate the SP equation (15) for a number of steps and then determine the probability $\pi_i^{(0)}$ for each vertex i; then a small fraction of vertices i with the smallest values of $\pi_i^{(0)}$ are declared as being covered; we then simplify the graph G and repeat the iteration-fixation process as long as there are still edges in G. The results of this SPD algorithm are shown in Fig. 3 for random graphs. We find that SPD very slightly outperforms BPD.

7. Discussion

For purely random graphs, the 1RSB mean field theory and the long range frustration theory can both give very good

predictions about the global minimal cardinality of vertex covers. On the heuristic algorithms side, both the BPD algorithm (inspired by the replica-symmetric mean field theory) and the SPD algorithm (inspired by the 1RSB mean field theory) are able to construct vertex covers whose cardinalities reach the theoretically predicted values. These successes indicate that the mean field theories of statistical physics can give a good description about the statistical properties of the random vertex cover problem. The good performance of the BPD and SPD algorithms also means that near-optimal vertex covers can be efficiently constructed for random graph instances.

The BPD and the SPD algorithms are comparable in terms of implementation costs and computation time and memory space. As demonstrated in Fig. 3, the SPD algorithm slightly outperforms the BPD algorithm in terms of the cardinality of constructed vertex covers. Real-world instances of the vertex cover problem usually are not random graphs but have certain structural properties. We expect these two message-passing algorithms will also have very good performances for such instances.

The minimal vertex cover problem is complementary to the maximal independent set problem, which asks for the construction of the smallest set of vertices such that any two vertices of this set are not connected by an edge. This latter problem has significant applications in game theory and microeconomics,[72,73] and glass transition.[74] Since the independent set problem is equivalent to the vertex cover problem, the methods and algorithms described in this paper can be applied to this problem without much modification. For example, Ref. [30] has offered a detailed analysis of the entropy density of independent sets.

The hitting set problem[31] is a natural extension of the vertex cover problem. It is a vertex cover problem defined on a hypergraph, in which each edge may connect simultaneously with more than two vertices. This problem is also a NP-complete problem, and one of its important applications is in group testing.[32] Another interesting extension is the k-path vertex cover problem,[75] which asks for the construction of a set of vertices intersecting with every path of length $k \geq 2$ in a given graph (the case $k = 2$ is just the vertex cover problem). The k-path vertex cover problem originates from the field of communication protocol.[76] It can be regarded as a special type of hitting set problem.

Acknowledgment

Zhao Jin-Hua thanks Prof. Ou-Yang Zhong-Can for support.

References

[1] Cook S A 1971 *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing* edited by Lewis P M, Fischer M J, Hopcroft J E, Rosenberg A L, Thatcher J W and Young P R (New York) p. 151
[2] Karp R M 1972 *Complexity of Computer Computations* (New York: Plenum Press) p. 85
[3] Garey M and Johnson D S 1979 *Computers and Intractability: A Guide to the Theory of NP-Completeness* (San Francisco: Freeman)
[4] Papadimitriou C and Steiglitz K 1998 *Combinatorial Optimization: Algorithms and Complexity* (New York: Dover)
[5] Cheeseman P, Kanefsky B and Taylor W 1991 *Proceedings 12th Int. Joint Conf. on Artificial Intelligence* (Sydney) p. 163
[6] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
[7] Nishimori H 2001 *Statistical Physics of Spin Glasses and Information Processing* (Oxford: Oxford University Press)
[8] Mézard M and Montanari A 2009 *Information, Physics, and Computation* (New York: Oxford Unversity Press)
[9] Talagrand M 2003 *Spin Glasses* (Berlin: Springer)
[10] Hartmann A K and Weigt M 2005 *Phase Transitions in Combinatorial Optimization Problems* (Weinhei: Wiley-VCH)
[11] Mézard M and Parisi G 2001 *Eur. Phys. J. B* **20** 217
[12] Mézard M and Parisi G 2003 *J. Stat. Phys.* **111** 1
[13] Mézard M and Parisi G 1986 *J. Physique* **47** 1285
[14] Monasson R and Zecchina R 1997 *Phys. Rev. E* **56** 1357
[15] Mézard M, Parisi G and Zecchina R 2002 *Science* **297** 812
[16] Mézard M and Zecchina R 2002 *Phys. Rev. E* **66** 056126
[17] Montanari A, Ricci-Tersenghi F and Semerjian G 2008 *J. Stat. Mech.: Theor. Exp.* P04004
[18] Krzakala F, Montanari A, Ricci-Tersenghi F, Semerjian G and Zdeborová L 2007 *Proc. Natl. Acad. Sci. USA* **104** 10318
[19] Franz S, Leone M, Ricci-Tersenghi F and Zecchina R 2001 *Phys. Rev. Lett.* **87** 127209
[20] Cocco S, Dubois O, Mandler J and Monasson R 2003 *Phys. Rev. Lett.* **90** 047205
[21] Mézard M, Ricci-Tersenghi F and Zecchina R 2003 *J. Stat. Phys.* **111** 505
[22] Weigt M and Hartmann A K 2000 *Phys. Rev. Lett.* **84** 6118
[23] Weigt M and Hartmann A K 2001 *Phys. Rev. Lett.* **86** 1658
[24] Weigt M and Hartmann A K 2001 *Phys. Rev. E* **63** 056127
[25] Hartmann A K and Weigt M 2003 *J. Phys. A: Math. Gen.* **36** 11069
[26] Zhou H J 2003 *Eur. Phys. J. B* **32** 265
[27] Weigt M and Zhou H J 2006 *Phys. Rev. E* **74** 046110
[28] Zhou J and Zhou H J 2009 *Phys. Rev. E* **79** 020103
[29] Zhang P, Zeng Y and Zhou H J 2009 *Phys. Rev. E* **80** 021122
[30] Dall'Asta L, Pin P and Ramezanpour A 2009 *Phys. Rev. E* **80** 061136
[31] Mézard M and Tarzia M 2007 *Phys. Rev. E* **76** 041124
[32] Mézard M, Tarzia M and Toninelli C 2008 *J. Phys. C: Conf. Ser.* **95** 012019
[33] van Mourik J and Saad D 2002 *Phys. Rev. E* **66** 056120
[34] Mulet R, Pagnani A, Weigt M and Zecchina R 2002 *Phys. Rev. Lett.* **89** 268701
[35] Zdeborová L and Krzakala F 2007 *Phys. Rev. E* **76** 031131
[36] Karp R M and Sipser M 1981 *Proceedings of the 22nd Annual IEEE Symposium on Foundations of Computer Science* (NJ: Wiley-IEEE Computer Society Press) p. 364
[37] Zhou H J and Ou-Yang Z C 2003 arXiv:0309348v1 [cond-mat]
[38] Zdeborová L and Mézard M 2006 *J. Stat. Mech.: Theor. Exp.* P05003
[39] Zhou H J 2013 *Eur. Phys. J. B* **86** 455
[40] Breitbart Y, Chan C, Garofalakis M, Rastogi R and Silberschatz A 2001 *Proceedings of IEEE INFOCOM'* 2001 (Anchorage) p. 933
[41] Park K and Lee H 2001 *Proceedings ACM SIGCOM* 2001 (San Diego)
[42] Gomez-Gardenes J, Echenique P and Moreno Y 2006 *Eur. Phys. J. B* **49** 259
[43] Huang H, Raymond J and Wong K Y M 2014 *J. Stat. Phys.* (in press) (2012 arXiv:1209.4134 [cond-mat.dis-nn])
[44] Harant J 1998 *Discr. Math.* **188** 239
[45] Caro C 1979 *Technical Reports* (Tel-Aviv University)
[46] Wei V K 1981 *Techinical Reports* 81-11217-9 (Bell Laboratories NJ)
[47] Harant J 2011 *Disc. Appl. Math.* **159** 966
[48] Harant J 2006 *Discussiones Mathematicae Graph Theory* **26** 431
[49] Angel E, Campigotto R and Laforest C 2013 *Disc. Appl. Math.* **161** 847
[50] Bollobás B 2001 *Random Graphs* (2nd edn.) (Cambridge: Cambridge University Press)
[51] Erdös P and Rényi A 1960 *Publ. Math. Inst. Hung. Acad. Sci.* **5** 17
[52] Gazmuri P G 1984 *Network* **14** 367

[53] Dani V and Moore C 2011 *Lect. Notes Comput. Sci.* **6845** 472
[54] Frieze A M 1990 *Discr. Math.* **81** 171
[55] Bauer M and Golinelli O 2001 *Eur. Phys. J. B* **24** 339
[56] Newman M E J and Barkema G T 1999 *Monte Carlo Methods in Statistical Physics* (USA: Oxford University Press)
[57] Landau D and Binder K 2013 *A Guide to Monte Carlo Simulations in Statistical Physics* (3rd edn.) (Cambridge: Cambrige University Press)
[58] Bathel W and Hartmann A K 2004 *Phys. Rev. E* **70** 066120
[59] Marinari K and Paris G 1992 *Europhys. Lett.* **19** 45
[60] Hukushima K and Nemoto K 1996 *J. Phys. Soc. Jpn.* **65** 1604
[61] Zhou H J 2005 *Phys. Rev. Lett.* **94** 217203
[62] Zhou H J 2012 *Phys. Rev. Lett.* **109** 199901
[63] Liu Y Y, Csóka E, Zhou H J and Pósfai M 2012 *Phys. Rev. Lett.* **109** 205703
[64] Zhou H J 2014 *Spin Glasses and Message Passing* (in preparation)
[65] Wei W, Zhang R, Guo B and Zhang Z 2012 *Phys. Rev. E* **86** 016112
[66] Binder K and Young A P 1986 *Rev. Mod. Phys.* **58** 801
[67] Bethe H A 1935 *Proc. R. Soc. London A* **150** 552
[68] Peierls R 1936 *Proc. R. Soc. London A* **154** 207
[69] Peierls R 1936 *Proc. Camb. Phil. Soc.* **32** 477
[70] Xiao J Q and Zhou H J 2011 *J. Phys. A: Math. Theor.* **44** 425001
[71] Zhou H J and Wang C 2012 *J. Stat. Phys.* **148** 513
[72] Bramoullé Y and Kranton R 2007 *J. Econ. Theory* **135** 478
[73] Galeotti A, Goyal S, Jackson M, Vega-Redondo F and Yariv L 2010 *Rev. Econ. Stud.* **77** 218
[74] Ritort F and Sollich P 2003 *Add. Phys.* **52** 219
[75] Brešar B, Kardoš F, Katrenič J and Semanišin G 2011 *Disc. Appl. Math.* **159** 1189
[76] Novotný M 2010 *Lect. Notes Comput. Sci.* **6033** 106