New Physics: Sae Mulli,
Vol. 67, No. 8, August 2017, pp. 1023∼1027
http://dx.doi.org/10.3938/NPSM.67.1023

Finite-Size Scaling in q-Coloring Problems

Meesoon Ha*
Department of Physics Education, Chosun University, Gwangju 61452, Korea

(Received 21 March 2017 : revised 5 July 2017 : accepted 5 July 2017)

We numerically investigate q-coloring (q-COL) problems in random networks in terms of a stochastic-local-search (SLS) algorithm. Random q-COL problems involve finding solutions where all nodes consist of different colors from their neighbors' colors, among q colors. For a fixed number of colors, say $q = 3$ or larger, various phase transitions have been reported as the average degree of links (the number density of constraints), $c = \langle k \rangle$, increases. This is because the set of solutions undergoes several types of phase transitions similar to those observed in the mean-field theory of spin glasses at zero temperature. Eventually, a dynamic coloring threshold is found to exist, above which no more solutions exist. Using the finite-size scaling (FSS) technique for nonequilibrium absorbing phase transitions, we analyze critical behaviors in the dynamic phase transition of q-COL problems by using the SLS algorithm, where we test both Erdös-Rényi and regular random networks. Finally, we discuss the extended FSS in q-COL problems compared to random k-satisfiability ones.

PACS numbers: 05.40.-a, 02.70.-c, 64.60.Ht, 89.20.Ef
Keywords: Heuristic, Random constraint, q-COL, Threshold behavior, Noise, Finite-size scaling

I. INTRODUCTION

Graph coloring is a basic problem in combinatorics, which consists of a graph and q colors to be considered coloring the vertices (nodes) without the neighbors that have the same color [1–3]. Depending on the number of colors and neighbors, graph coloring becomes one of famous *nondeterministic polynomial-solving time* (NP)-complete class problems. Although there is the most celebrated four-colors theorem [4,5] for planar graphs, q-COL problems for general graphs can be extremely hard to solve, known to be NP-complete [6], which is believed that no proper algorithm can decide whether an arbitrary instance of graphs is colorable or not within a polynomial time with respect to the total number of nodes. Since the NP-complete problem can be taken as a benchmark to evaluate the performance of various algorithms in computer science, such as time tabling of practical applications, registering allocations of complies, assigning frequencies in mobiles radios, and so on.

In this paper, we numerically study the q-coloring (q-COL) problems of random graphs, and focus on the role of the degree fluctuations in scaling properties of dynamic phase transitions with respect to both Erdös-Rényi random (ER) and regular random (RR) networks. In particular, we employ a heuristic algorithm, namely stochastic-local-search (SLS), and the extended finite-size scaling (FSS) theory of nonequilibrium absorbing phase transitions (APTs), where it is quite conventional to analyze two major physical quantities, the average rescaled solving time for a given instance from the solved (SOL) samples, $\langle t_{\rm sol} \rangle$, and the density of unsatisfied nodes from the unsolved (UNSOL) samples, $\rho_{\rm all}$, against two control parameters, the average of degree per node, $c = \langle k \rangle$, and the noise (temperature-like) parameter p. These quantities are useful to identify critical behaviors in random constraint satisfaction problems (CSPs) [7].

The objective of this paper is to propose the universal scaling properties of dynamic phase transitions in random CSPs and to discuss the interplay of the internal disorder for the structure of neighbors and the external noise for the SLS algorithm in such universality issues.

*E-mail: msha@chosun.ac.kr

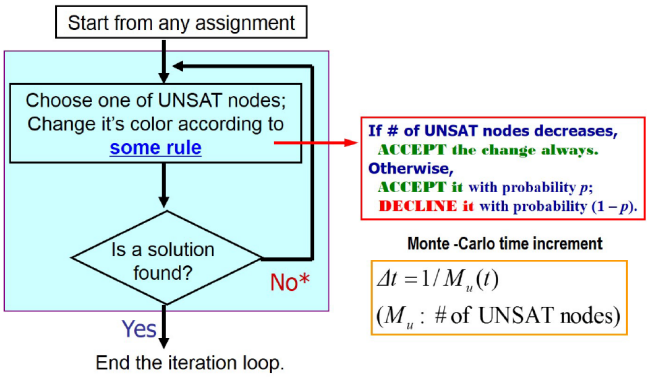

Fig. 1. (Color online) Flowchart of the SLS pseudo-algorithm of A-COL heuristics for random q-COL problems. The iteration loop terminates after the maximum time $T_{\max}$ even if a solution is not found. Note that this is almost the same as the A-SAT for random K-Satisfiability problems [7].

The paper is organized as follows: In Sec. II, we present the q-COL dynamics and the SLS algorithm with the average coloring (A-COL) heuristics, where two physical quantities are proposed against two control parameters. Numerical simulation results are provided in Sec. III with both ER and RR networks, where we test the extended FSS theory of nonequilibrium APTs. Finally, the paper is concluded in Sec. IV with some remarks.

II. MODEL

Consider a large sparse random network (graph), which consists of N nodes (vertices) and L links (edges). With a given number of colors, say $q(\geq 3)$, each node is assigned as one of q colors. Since the coloring problem is to find the configuration that each node has a different color from its nearest neighboring nodes' colors, one can investigate the statistical physics for q-COL problems [8–11] as a Pott [12] spin model with antiferromagnetic interactions on random networks, where each node has a spin variable s_i with $i = 1, 2, ..., N$ and $s_i = 1, 2, ..., q$. From equilibrium concepts, it is natural to define the Hamiltonian as

$$\mathcal{H}(\{s\}) = \frac{1}{2}\sum_{i,j=1}^{N} a_{ij}\delta(s_i - s_j), \qquad (1)$$

where $a_{ij} = 1$ if two nodes are connected; otherwise, 0, and $\delta(s-s') = 1$ if $s = s'$; otherwise, 0. This implies that no energy contribution for the connected pair having the different colors. The ground-state energy is 0 if and only if the network is q-colorable.

The q-COL problems can be also studied through a SLS algorithm as presented in this paper. We set $q = 3$ and employ the SLS algorithm, the A-COL heuristics where the color of the randomly chosen unsatisfied node is changed with an effective temperature (noise) p. As described in Fig. 1, the A-COL heuristics keeps updating until the Monte-Carlo (MC) simulation time step $t_{\rm MC} = T_{\max}$. Instead of measuring directly the energy from Eq. (1), we here analyze two physical quantities,

$$\langle t_{\rm sol}\rangle \text{ (average solving time)}, \qquad (2)$$

$$\rho_{\rm all} = \frac{\langle M_u\rangle}{N} \text{ (density of unsatisfied nodes)}, \qquad (3)$$

for the fixed number of colors, $q = 3$, as the average degree per node $c = \langle k\rangle = L/2N$ and p increase.

The idea of the A-COL heuristics was first suggested by Zdeborová and Krzakala [13] as the extension of Walk-COL that is the slight different adaptation, closed to the work of Walk-SAT by Selman and coworkers [14]. The detailed analysis of Walk-SAT also discussed [7,15], in terms of the A-SAT heuristics [16]. The approach of SLS heuristics is unsophisticated, compared to complete solvers, message passing algorithms (belief propagation and survey propagation), and cavity formalisms at the replica symmetric level [13,17,18]. However, it is simple and easy, so it is popular to investigate dynamic phase transitions in random CSPs. In the A-COL heuristics, an arbitrary configuration moves onto a neighboring one according simple rules (see Fig. 1) against the noise level. How to identify and understand the typical hardness in q-COL problems is particularly very similar to that of epidemic models that exhibit nonequilibrium absorbing phase transitions (APTs) [19].

So far, it is known that the following analogies exist between MC simulations of dynamic processes and A-COL heuristics: (i) Local search heuristic corresponds to dynamic object in configuration space; (ii) An instance is a sample, and a solution is the global minimum (ground-state) energy; (iii) The difference between the numbers of unsatisfied nodes among neighboring configurations is the energy gradient; (iv) SLS algorithms correspond to the MC simulations of nonequilibrium processes without the detailed balance.

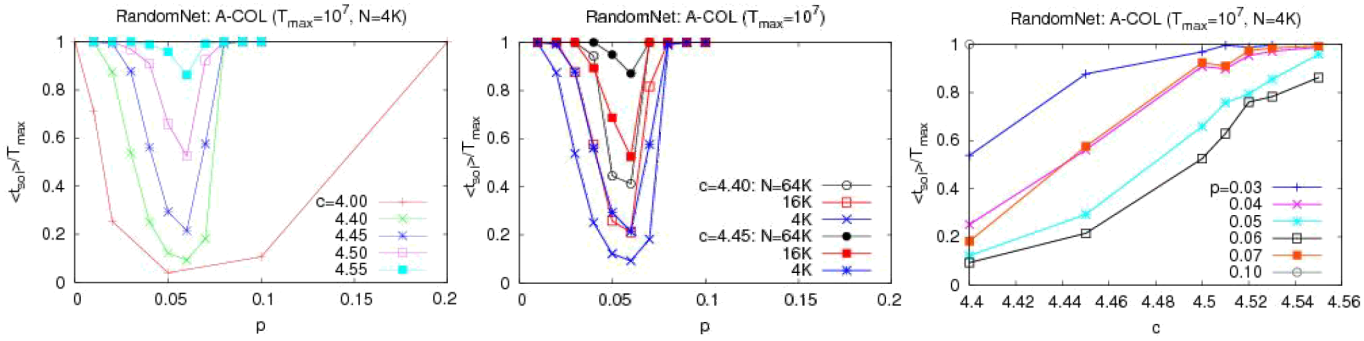

Fig. 2. (Color online) The average rescaled solving time on ER networks are plotted against p and c, respectively, from which one can indicate the optimal value of p, $p_{\rm opt} \simeq 0.06$, and the threshold value of c, $c_{\rm th,ER} \simeq 4.55$. At $p = p_{\rm opt}$, $\tau \equiv \langle t_{\rm sol}\rangle / T_{\rm max}$ is the minimum. Up to $c = c_{\rm th}$, τ remains as the minimum in thermodynamic limit.

However, it is hardly discussed about the universal scaling properties and FSS tests in q-COL problems. So we need to speculate them in details. In particular, we here focus on how the degree fluctuations of random networks affect the scaling properties of dynamic phase transitions in 3-COL problems. For random networks, two possible ensembles exist: (i) the Poissonian degree distribution $P(k) = e^{-c}c^k/k!$, namely ER networks [20], and (ii) the uniform degree distribution $P(k) = \delta(k-c)$, namely RR networks.

III. NUMERICAL RESULTS

We perform the MC simulations of 3-COL problems using the A-COL heuristics on both ER and RR networks of N nodes, where two control parameters are considered, the temperature-like noise parameter p and the average degree per node c. The update of the A-COL heuristics lasts up to $t_{\rm MC} = T_{\rm max}$.

First, we define the optimal noise value, $p_{\rm opt}$, at which the solving time is the fastest, and the threshold value of $c_{\rm th}$ is also the largest, at which the solving time $\langle t_{\rm sol}\rangle$ is still smaller than $T_{\rm max}$. Fig. 2 shows the average rescaled solving time on ER networks, which implies that $p_{\rm opt} \simeq 0.06$ and $c_{\rm th,ER} \simeq 4.55$ for various $N = 4\times10^3$, 16×10^3, 64×10^3 and $T_{\rm max} = 10^7$.

To interpret the solvability of the A-COL heuristics, the FSS theory of nonequilibrium APTs is employed as follows: From the solvable (SOL, absorbing) phase to the unsolvable (UNSOL, active) phase, the average rescaled solving time, $\tau \equiv \langle t_{\rm sol}\rangle / T_{\rm max}$, is obtained only from solved samples, while, from the UNSOL phase to the SOL phase, the density of unsatisfied nodes, $\rho_u(t,N) \equiv \langle M_u(t,N)\rangle_u / N$, is obtained from unsolved samples. Here the MC time increment $\Delta t = 1/M_u(t)$, where $M_u(t)$ is the number of unsatisfied nodes at time t. The essential part of APTs is how to identify that a sample is dynamically active with two major physical quantities near and at the criticality, the density of active nodes and the survival probability of samples, which correspond to the density of UNSAT nodes ($\rho_{\rm all}$) and the fraction of the UNSAT samples to the all samples ($P_{\rm surv}$) in q-COL problems, respectively.

Based on the extended FSS theory of APTs [7,19] in random q-COL problems at the criticality, $\epsilon \equiv \frac{c-c_{\rm th}}{c_{\rm th}} = 0$, dynamic scaling can be defined as follows:

$$P_{\rm surv} = \frac{\mathcal{N}_u}{\mathcal{N}} = \phi(t/N^{\bar z}), \tag{4}$$

$$\rho_{\rm surv}(t,N) = t^{-\delta} f(t/N^{\bar z}), \tag{5}$$

where the dynamic exponent $\bar z = \nu_{||}/\bar\nu$ and the decay exponent $\delta = \beta/\nu_{||}$ from $\rho_{\rm surv}(\epsilon > 0) \sim \epsilon^{\beta}$, the correlated volume $\xi_v \sim \epsilon^{\bar\nu}$, and the correlated time $\xi_t \sim \epsilon^{-\nu_{||}}$ with $\nu_{||} = 1$ on networks. Moreover, $\mathcal{N}$ is the total number of samples, $\mathcal{N}_u$ is the number of unsolved samples, and $t_{\rm H}(\sim N^{\bar z})$ is the time when $P_{\rm surv} = 1/2$ (see the detailed relations of the critical exponents in [19]). In the subcritical regime ($\epsilon < 0$, SOL/absorbing) where all samples are solved, $t_{\rm H} = o(1) \ll 1$ and $\rho_{\rm surv}(t) \sim \exp(-At)$ as $N \to \infty$ while $\rho_{\rm surv}(N) \sim 1/N$ as $t \to \infty$. In the supercritical regime ($\epsilon > 0$, UNSOL/active) where all samples are survived (unsolved), $t_{\rm H}$ cannot exist since $\tau = 1$ and $\rho_{\rm surv}(t) = o(1)$ as $N \to \infty$.

Fig. 3 and Fig. 4 show how the degree fluctuations of the networks affect the critical threshold of the average

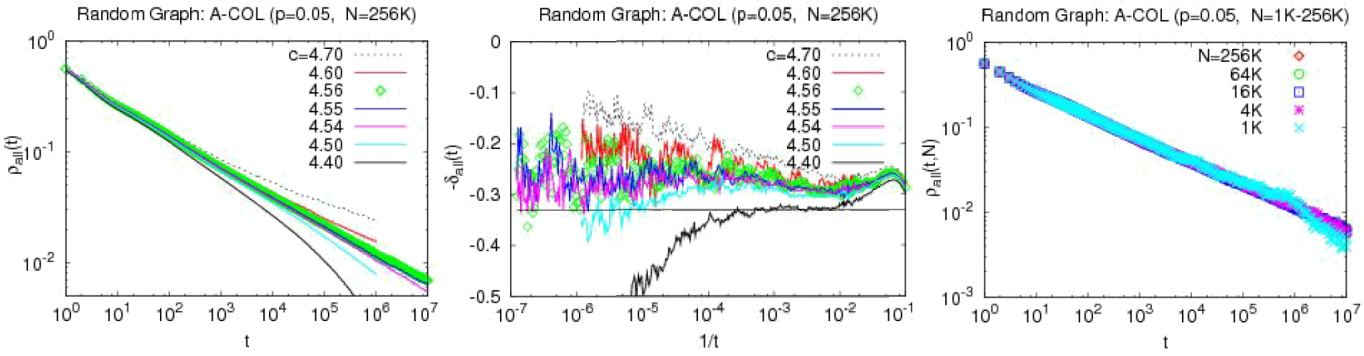

Fig. 3. (Color online) For ER networks at $p = 0.05 \simeq p_{\rm opt}$, the temporal behavior of the UNSAT density, $\rho_{\rm all}(t) \sim t^{-\delta}$, is plotted as t increases, where the case of $c = c_{\rm th,ER} = 4.56(1)$ exhibits $\delta_{\rm ER} \simeq 0.25$ for $N = 256 \times 10^3$. In the middle panel, the effective exponent $\delta_{\rm all}(t)$ is plotted against $1/t$, compared to various $c = 4.40, ..., 4.70$. At $c = 4.56 \simeq c_{\rm th,ER}$, finite effects are tested for $N = 256 \times 10^3, ..., 1 \times 10^3$ up to $T_{\rm max} = 10^7$.

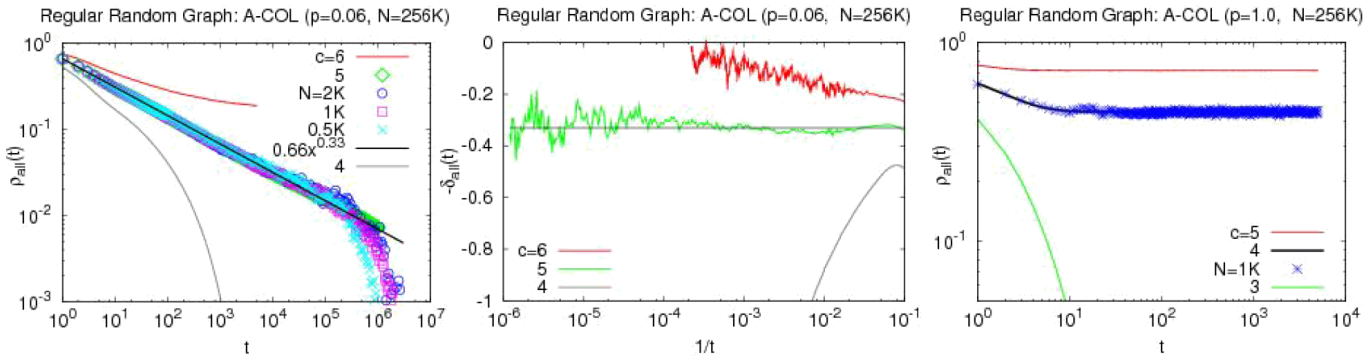

Fig. 4. (Color online) For RR networks at $p = 0.06$, the temporal behavior of the UNSAT density $\rho_{\rm all}(t) \sim t^{-1/3}$ is presented as t increases at $c_{\rm th,RR} = 5$ with various $N = 256 \times 10^3, ..., 0.5 \times 10^3$ and its effective exponent $\delta_{\rm all}(t)$ is plotted against $1/t$, compared to $c = 4, 5, 6$. At $p = 1$, $\rho_{\rm all}(t)$ are plotted for $N = 256 \times 10^3$ ($c = 3, 4, 5$) and 1×10^3 ($c = 4$).

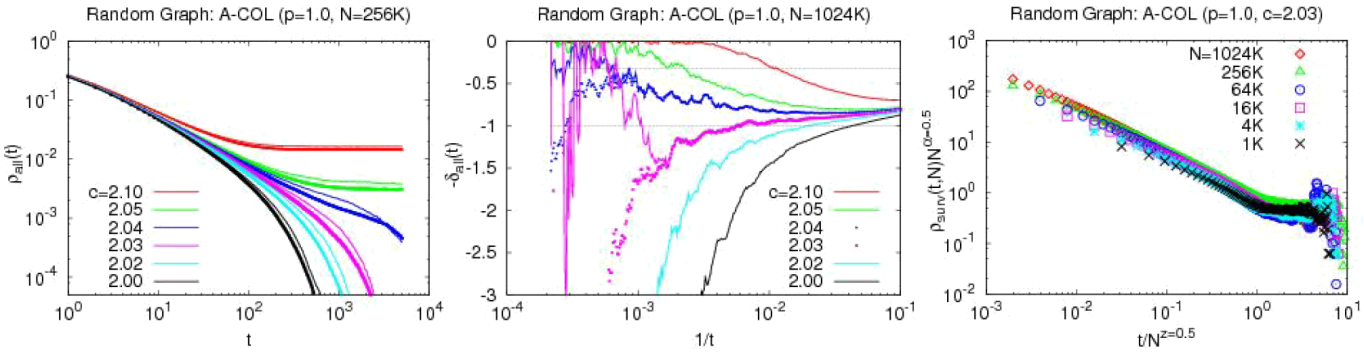

Fig. 5. (Color online) For ER networks at $p = 1.0$, the temporal behavior of the UNSAT density $\rho_{\rm all}(t) \sim t^{-1}$ is presented as t increases at $c_{\rm th,ER} = 2.03$ for $N = 256 \times 10^3$ and its effective exponent $\delta_{\rm all}(t)$ is plotted against $1/t$, compared to $c = 2.00, ..., 2.10$ for $N = 1024 \times 10^3$. Using the FSS analysis of $\rho_{\rm all}$, numerical data collapse well at $c_{\rm th,ER,p=1} = 2.03$ with $\delta_{\rm ER,p=1} = 1$, $\bar{z}_{\rm ER,p=1} = 1/2$, and $\alpha = 1/2$ from $\rho_{\rm surv}(t \to \infty) \sim N^{\alpha}$ for $N = 1024 \times 10^3, ..., 1 \times 10^3$.

degree per node and dynamic scaling at $p = p_{\rm opt}$ for both ER and RR networks, respectively. Our numerical data at $p = p_{\rm opt}$ are obtained from $P_{\rm surv} = 1$, so that $\rho_{\rm surv} = \rho_{\rm all}$. At $p = p_{\rm opt}$, we find that the decay exponent of the density of UNSAT nodes ρ_u on RR networks, $\delta_{\rm RR} \simeq 1/3$ is different from that on ER networks, $\delta_{\rm ER} \simeq 1/4$ (see

Fig. 3 and Fig. 4). While $p_{\rm opt,RR}$ is pretty similar to $p_{\rm opt,ER}$, $c_{\rm th,RR}(=5) > c_{\rm th,ER}(\simeq 4.56)$. It is noted that the value of c on RR networks allows the positive integer.

Finally, Fig. 5 shows the special limit of the A-COL heuristics, namely Random Walk-COL with $p = 1$, where at $c = c_{\rm th,ER} \simeq 2.03$, $\rho_{\rm all} \sim t^{-\delta}$ with $\delta = \delta_{\rm ER,p=1} \simeq 1$ and $t_{\rm H} \sim N^{\bar{z}}$ with $\bar{z} = \bar{z}_{\rm ER,p=1} \simeq 1/2$. This is quite distinct from the case of RR networks with $c_{\rm th,RR} = 4$, $\delta_{\rm RR,p=1} \simeq 1$, $\bar{z}_{\rm RR,p=1} = 0$ (see the right panel of Fig. 4).

IV. SUMMARY WITH REMARKS

We explored dynamic phase transitions in random 3-COL problems, where we employed the stochastic-local-search algorithm on random networks and the extended finite-size scaling analysis for nonequilibrium absorbing phase transitions. For the noise parameter (temperature-like, p) and the density of constraints (the average degree per node of random networks, c), we investigated the optimal value of $p_{\rm opt}$ and the critical threshold $c_{\rm th}$ on Erdös-Rényi (ER) and regular random networks. Moreover, we found that Random Walk-COL ($p = 1$) on ER random networks shows the same universality class as that of Random Walk-SAT for random 3-SAT problems [7,15], where the FSS exponent should be checked with more cautions.

Finally, our preliminary results shed the light on the universality issues for scaling properties of dynamic phase transitions in random q coloring problems. The further discussion on the origin of phase transitions will be provided as the future study in elsewhere for various q values, compared with the results obtained from other analytic approaches.

ACKNOWLEDGEMENTS

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF-2017R1D1A3A03000578).

REFERENCES

[1] S. Cook, *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing* (ACM, New York, 1971), pp. 151-158.

[2] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (Freeman, New York, 1979).

[3] C. H. Papadimitrious, *Computational Complexity* (Addison-Wesley, Reading, MA, 1994).

[4] K. Appel and W. Haken, Illinois. J. Math. **21**, 429 (1977).

[5] K. Appel, W. Haken and J. Koch, Illinois. J. Math. **21**, 491 (1977).

[6] R. M. Karp, *Complexity of Computer Computations*, edited by R. E. Miiler and J. W. Thacher (Plenum, New York, 1971), p. 85.

[7] M. Ha, New Phys.: Sae Mulli **63**, 995 (2013).

[8] D. J. Gross, I. Kanter and H. Sompolinsky, Phys. Rev. Lett. **55**, 304 (1985).

[9] J. van Mourik and D. Saad, Phys. Rev. E **66**, 056120 (2002).

[10] R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Phys. Rev. Lett. **89**, 268701 (2002).

[11] A. Braunstein, R. Mulet, A. Pagnani, M. Weigt and R. Zecchina, Phys. Rev. E **68**, 036702 (2003).

[12] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[13] L.Zdeboravá and F. Krzakala, Phys. Rev. E **76**, 031131 (2007).

[14] B. Selman, H. A. Kantz, and B. Cohen, *Proceedings of AAAI94, Seattle* (American Associate for Artificial Intelligence, Menlo Park, CA, 1994).

[15] S. H. Lee, M. Ha, C. Jeon and H. Jeong, Phys. Rev. E **82**, 061109 (2010).

[16] J. Ardelius and E. Aurell, Phys. Rev. E **74**, 037702 (2006).

[17] M. Mézard and G. Parisi, Eur. Phys. J. B **20**, 217 (2001).

[18] M. Mézard and G. Parisi, J. Stat. Phys. **111**, 1 (2003).

[19] J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattice Models* (Cambridge University Press, Cambridge, 1999), Chap. 6.

[20] P. Erdös and A. Rényi, Publ. Math. (Debrecen) **6**, 290 (1959).

[21] H. Hinrichsen, Adv. Phys. **49**, 815 (2000).