Typical approximation performance for maximum coverage problem

Satoshi Takabe,[1,2,*] Takanori Maehara,[2] and Koji Hukushima[3]

[1]*Department of Computer Science, Nagoya Institute of Technology, Gokiso-cho, Showa-ku, Nagoya, Aichi 466-8555, Japan*
[2]*RIKEN Center for Advanced Intelligence Project, 1-4-1 Nihonbashi, Chuo-ku, Tokyo 103-0027, Japan*
[3]*Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan*

This study investigated the typical performance of approximation algorithms known as belief propagation, the greedy algorithm, and linear-programming relaxation for maximum coverage problems in sparse biregular random graphs. After we used the cavity method for a corresponding hard-core lattice-gas model, results showed that two distinct thresholds of replica-symmetry and its breaking exist in the typical performance threshold of belief propagation. In the low-density region, the superiority of three algorithms in terms of a typical performance threshold is obtained by some theoretical analyses. Although the greedy algorithm and linear-programming relaxation have the same approximation ratio in worst-case performance, their typical performance thresholds are mutually different, indicating the importance of typical performance. Results of numerical simulations validate the theoretical analyses and imply further mutual relations of approximation algorithms.

I. INTRODUCTION

Approximation algorithms, which are important for hard optimization problems, have attracted researchers' interest because, for NP-hard optimization problems, one encounters difficulty when solving optimal solutions exactly in the polynomial time of the problem size. Since the P versus NP problem arose, development and performance analyses of approximation algorithms have persisted as a central issue of computer science and operations research.

Two performance evaluations of approximation algorithms exist: worst-case performance and typical (or average-case) performance. As described in this paper, we specifically examine the latter mainly using statistical-mechanical methods, although the former has been investigated mainly in the literature of theoretical computer science [1]. Worst-case performance is defined by the pair of an optimization problem and its approximation algorithm. An approximation ratio is then defined by the maximal ratio of an optimal value to an approximation value over all instances if the problem is a maximization problem (and vice versa, otherwise). It is important to provide a strict performance guarantee of the algorithm, although the worst-case instance is sometimes pathological. However, the typical performance is defined as the average performance of the approximation algorithm for a given optimization problem over randomized instances. It is sometimes useful for practical use. It sheds light on properties of optimization problems and approximation algorithms in a perspective that is different from worst-case analysis.

An interesting point of the typical property of optimization problems is found in a close relation to the spin-glass theory in statistical physics. Extensive studies of various computational problems have revealed that the concept of replica symmetry (RS) and its breaking (RSB) in spin-glass theory reflects average computational complexity [2] and structure of the solution space [3] of the problems.

The typical performance of approximation algorithms often exhibits a phase transition on typical goodness or accuracy of approximation, which is called the threshold phenomenon in the literature of theoretical computer science [4]. Some approximation algorithms for minimum vertex covers (min-VCs) are good examples of threshold phenomena. The problem is defined in a graph. The randomized problem is characterized using a random graph ensemble with c being the average degree. From a statistical-mechanical perspective, the problem in Erdös-Rényi random graphs has the RS-RSB threshold at $c = e = 2.71\ldots$ [5]. Moreover, three approximation algorithms have been investigated: belief propagation (BP) (or message passing) [6], the greedy leaf-removal algorithm [7], and linear-programming (LP) relaxation [8,9]. Approximation algorithms other than BP naively have no direct connection to the spin-glass theory. Nevertheless, for Erdös-Rényi random graphs, these algorithms have the same performance threshold as the RS-RSB threshold. However, subsequent studies of general random graphs [10] indicate that their thresholds are not equivalent for some random graphs because of their graph structure. Using statistical-mechanical techniques, typical performance has been studied for variants of leaf removal [11–15] and relaxation technique such as LP relaxation [16,17] and semidefinite-programming relaxation [18]. These studies have revealed not only typical approximate performance itself but also a suggestive connection to typical properties of optimization problems, random graph structure, and the spin-glass theory.

As described in this paper, we examine the unweighted maximum coverage (max-COV) problem defined in the next section. Although the max-COV belongs to the class of NP-hard, it has several practical applications such as pan and scan problems [19], multitopic blog watch [20], and text summa-

*s_takabe@nitech.ac.jp

rization [21]. Our main purpose is to examine the typical performance of approximation algorithms for the problem and to compare their performance thresholds. We specifically examine three approximation algorithms: belief propagation, the greedy algorithm, and LP relaxation. Some statistical-mechanical methods applied to both analytical and numerical analyses together with mathematical rigorous discussions clarify the typical performance of those algorithms and a suggestive mutual relation among approximation algorithms.

This paper is organized as follows. In the next section, we define details of the max-COV and its approximation algorithms. As a random graph ensemble, biregular random graphs are also defined. In Sec. III, a hard-core lattice-gas model for the problem is introduced. Its BP equations are obtained based on the Bethe-Peierls approximation. In Sec. IV, we present a study of the typical performance of BP using the RS cavity method. Calculation of the spin-glass susceptibility provides the threshold below which BP typically approximates max-COV with high accuracy. In Sec. V, the greedy algorithm is analyzed based on a mean-field rate equation of its deletion process. Using the obtained solution, we evaluate the typical performance threshold of the greedy algorithm. In Sec. VI, LP-relaxed approximate values are evaluated rigorously. The theorem is proved using the weak duality theorem. In Sec. VII, we describe some numerical results which support the validity of theoretical analyses presented in the previous sections. We also execute some additional simulations to consider the typical performance of a modified greedy algorithm and randomized rounding of LP relaxation. These results provide a suggestive relation between approximation algorithms. The last section is devoted to a summary and discussion of the results.

II. MAX-COVER PROBLEM AND APPROXIMATION ALGORITHMS

A. Maximum coverage problem

The max-COV is defined as follows: Let S be a set of M elements and $\mathcal{S} = \{S_1, \cdots, S_N\}$ be a collection of subsets of S. For a given positive integer $K(\leqslant N)$, the problem is to choose at most K subsets to maximize the total number of elements in the union of chosen subsets. An example of an instance with $M = 4$ and $N = 3$ is shown in Fig. 1(a). Given that $K = 2$, one can select subsets S_1 and S_3 such that all elements are included in the union of the subsets.

Converting an instance to a bipartite graph, one obtains an integer-programming (IP) representation of the max-COV. Let $G = (V_1, V_2, E)$ be a bipartite graph where each edge in

E connects to a vertex in V_1 and a vertex in V_2. Vertices in V_1 and V_2 are labeled respectively by $i \in \{1, \ldots, N\}$ and $a \in \{1, \ldots, M\}$. V_1 corresponds to a collection of subsets $\mathcal{S}$, whereas V_2 represents S. Each edge (i, a) is set if subset S_i includes element a. In Fig. 1, the example in panel (a) is converted to a bipartite graph in panel (b). We then introduce binary variables $\{x_i\}$ and $\{y_a\}$, respectively, to V_1 and V_2. Also, x_i is set to 1 if vertex i (or subset S_i) is selected and 0 otherwise. Similarly, y_a is set to 1 if vertex a is connected to a selected vertex (or element a belongs to the union of chosen subsets) and 0 otherwise. The problem is therefore represented by the following IP problem:

$$
\begin{aligned}
\text{Max.} \quad & \sum_{a=1}^{M} y_a, \\
\text{subject to} \quad & y_a \leqslant \sum_{i; (i,a) \in E} x_i \qquad \forall a \in V_2, \\
& \sum_{i=1}^{N} x_i \leqslant K, \\
& x_i \in \{0,1\} \, \forall i \in V_1, \quad y_a \in \{0,1\} \, \forall a \in V_2.
\end{aligned}
\tag{1}
$$

The inequality in the second constraint can be replaced with an equality sign because the problem is a maximization problem.

As described in this paper, the typical-case property of the max-COV is analyzed by randomizing its instance. Then, we introduce $\rho_x = K/N$ and assume that $\rho_x \in [0,1]$ is constant. For random bipartite graphs, we specifically examine (L, R)-biregular random graphs where degrees of each vertex in V_1 and V_2 respectively denote L and R. When we use this simple random graph ensemble, our statistical-mechanical analyses reveal interesting properties of the problem.

To take the large-graph limit, the number of vertices in V_2 is rescaled as $M = \alpha N$ with a constant factor α. We assume that a random graph is sparse, i.e., L and R are constant with respect to N. For the randomized max-COV, we define an average optimal cover ratio over random graphs with cardinality $N + M$ and a given ρ_x by

$$
\rho_y(\rho_x; N) = \frac{1}{M} \overline{\sum_{a=1}^{M} y_a^{\text{opt}}(G, \rho_x)},
\tag{2}
$$

where $\{y_a^{\text{opt}}(G, \rho_x)\}$ represents optimal solutions in V_2 and $\overline{(\cdots)}$ represents an average over random bipartite graphs with size $N + M$. Its limiting value to $N \to \infty$ is denoted by ρ_y. We simply call it the average optimal cover ratio.

B. Approximation algorithms

We introduce three approximation methods for max-COVs. The first algorithm is BP. The recursive equations called BP equations are derived from the Bethe-Peierls approximation for a spin system corresponding to a given optimization problem. For systems on trees, graphs with no cycles, the Bethe-Peierls approximation and BP are exact. In general, however, there exist cycles in a graph yielding correlations between variables or spins, which result in inexact estimation of solutions (or configurations). BP in graphs with some cycles is regarded as an approximation method. It is called loopy BP [22].
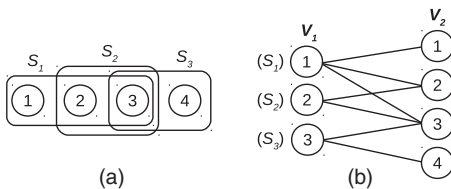

FIG. 1. Example of the maximum coverage problem. (a) An example with $M = 4$ and $N = 3$. S_1, S_2, and S_3 represent subsets of $S = \{1,2,3,4\}$. (b) A bipartite graph corresponding to panel (a). Each vertex in V_1 stands for a subset in panel (a).

The second one is a simple greedy algorithm. At each step, this algorithm has the following procedure: (i) choose one vertex named i with the maximum degree in V_1, (ii) delete vertices neighboring to vertex i from V_2, and (iii) update V_1 to $V_1 \backslash i$ and return to step (i) if $|V_1| > N - K$. This simple algorithm gives an approximation ratio of $1 - 1/e$ [23]. The problem cannot be approximated within this ratio unless P $=$ NP [24].

The last is LP relaxation. An integer programming problem including the max-COV is relaxed to LP problems by replacing integral constraints with real constraints. The LP-relaxed max-COV is given as

$$
\begin{aligned}
\text{Max.} \quad & \sum_{a=1}^{M} y_a, \\
\text{subject to} \quad & y_a \leqslant \sum_{i;(i,a)\in E} x_i \qquad \forall a \in V_2, \\
& \sum_{i=1}^{N} x_i \leqslant K, \\
& x_i \in [0,1]\, \forall i \in V_1, \quad y_a \in [0,1]\, \forall a \in V_2.
\end{aligned}
\tag{3}
$$

Actually, the LP approximation value gives the upper bound of the problem. Because LP problems are solvable exactly in polynomial time, LP relaxation is a widely used approximation technique. The approximate solution obtained by LP relaxation usually involves nonintegers. One must round those numbers appropriately to obtain an approximate integral solution for the IP problem. Here, we consider randomized rounding [25]. Using the obtained approximation solution $\{x_i^{\mathrm{LP}}\}$, one selects vertex $i \in V_1$ to set $I \subset V_1$ with probability x_i^{LP}/K up to K vertices. Then, the rounded solution $\{x_i^{\mathrm{LPr}}\}$ is set to $x_i^{\mathrm{LPr}} = 1$ for $i \in I$ and $x_i^{\mathrm{LPr}} = 0$ otherwise. The rounded approximation value is readily calculated from $\{x_i^{\mathrm{LPr}}\}$. In terms of the worst-case performance, LP relaxation and its randomized rounding in expectation has the same approximation ratio as the greedy algorithm.

As described in this paper, we study the typical performance of these approximation algorithms. It is evaluated by approximate values averaged over randomized max-COVs defined in the last subsection. Similar to the average optimal cover ratio, the average cover ratio is defined as the average ratio of the approximate value to the cardinality M of vertex set V_2. It is regarded as exhibiting good typical performance if the average cover ratio obtained by an approximation algorithm is equal to the average optimal cover ratio in the large-N limit. The main aim of this paper is to evaluate the typical performance threshold of ρ_x below which an approximate algorithm exhibits good typical performance with L and R fixed. Evaluation of the approximation algorithms is accomplished by comparing their typical performance thresholds.

III. BP EQUATIONS FOR MAX-COV

As explained in this section, BP equations for max-COV are derived based on the statistical-mechanical model for the problem. We set particles on vertices in V_1 and V_2, which respectively occupy vertex i and a if $x_i = 1$ and $y_a = 1$. The hard-core lattice-gas model for max-COVs in bipartite graph

G is then naively given as the following partition function:

$$
\Xi_0(\mu; G) = \sum_{\boldsymbol{x}\in\{0,1\}^N} \sum_{\boldsymbol{y}\in\{0,1\}^M} \exp\left(\mu \sum_{a=1}^{M} y_a\right) H\left(K - \sum_{i=1}^{N} x_i\right)
$$
$$
\times \prod_{a=1}^{M} H\left(\sum_{i\in\partial a} x_i - y_a\right),
\tag{4}
$$

Therein, $H(x) = 1\,(x \geqslant 0), 0\,(x < 0)$ is the Heaviside step function and $\partial a = \{i \in V_1 \mid (i,a) \in E\}$ stands for a set of neighbors of vertex $a \in V_2$. In the partition function, μ represents a chemical potential for particles on V_2. One can construct BP equations for this model. However, it is inconvenient for practical use because of the constraint $\sum_{i=1}^{N} x_i \leqslant K$.

Therefore, we use the following alternative partition function:

$$
\Xi(\mu; G) = \sum_{\boldsymbol{x}\in\{0,1\}^N} \sum_{\boldsymbol{y}\in\{0,1\}^M} \exp\left(\mu' \sum_{i=1}^{N} x_i + \mu \sum_{a=1}^{M} y_a\right)
$$
$$
\times \prod_{a=1}^{M} H\left(\sum_{i\in\partial a} x_i - y_a\right),
\tag{5}
$$

where $\mu' = \mu'(\mu,\rho_x; G)$ is a chemical potential for particles on V_1. The ratio of μ' with respect to μ is defined as

$$
\kappa = -\frac{\mu'}{\mu}.
\tag{6}
$$

This parameter is regarded as a Lagrange multiplier for the constraint on the number of selected vertices. The appropriate value of κ must satisfy the condition given by

$$
\rho_x(\mu,\kappa) = \rho_x, \quad \rho_x(\mu,\kappa) \equiv \left\langle \frac{1}{N}\sum_{i=1}^{N} x_i \right\rangle_\mu,
\tag{7}
$$

where $\langle \cdots \rangle_\mu$ is a grand-canonical average with a given μ. To consider ground states, we take the large-μ limit with the parameter κ being fixed.

First, we construct BP equations for Eq. (5). Using the Bethe-Peierls approximation, we give the single-spin probability $P_i(x)$, in which x_i takes x, as

$$
P_i(x) \simeq Z_i^{-1} e^{-\mu\kappa x} \prod_{a\in\partial i} P_{a\to i}(x),
\tag{8}
$$

where $\partial i = \{a \in V_2 \mid (i,a) \in E\}$ is a set of neighbors of vertex $i \in V_1$. Z_* is a normalization factor hereinafter. $P_{a\to i}(x)$ represents the marginal probability of $\boldsymbol{x}_{\partial a \backslash i}$ and y_a under the condition $x_i = x$. Similarly, the single-spin probability $P_a(y)$, in which y_a takes y, reads

$$
P_a(y) \simeq Z_a^{-1} e^{\mu y} \sum_{\boldsymbol{x}_{\partial a}} H\left(\sum_{i\in\partial a} x_i - y\right) \prod_{i\in\partial a} P_{i\to a}(x),
\tag{9}
$$

where $P_{i\to a}(x)$ is the probability of x_i taking x in the cavity graph $G\backslash a$. These probabilities satisfy the following recursive

relations under the same approximation:

$$P_{i \to a}(x) \simeq Z_{i \to a}^{-1} e^{-\mu \kappa x} \prod_{b \in \partial i \backslash a} P_{a \to i}(x), \qquad (10)$$

$$P_{a \to i}(x) \simeq Z_{a \to i}^{-1} \sum_y e^{\mu y} \sum_{\boldsymbol{x}_{\partial a \backslash i}} H\left(x + \sum_{j \in \partial a \backslash i} x_j - y\right)$$

$$\times \prod_{j \in \partial a \backslash i} P_{j \to a}(x). \qquad (11)$$

To take the large-μ limit later, it is convenient to introduce the cavity fields $\{h_{ia}\}$ and $\{\hat{h}_{ai}\}$ defined respectively by $P_{i \to a}(x) \propto \exp(\mu h_{ia} x)$ and $P_{a \to i}(y) \propto \exp(\mu \hat{h}_{ai} y)$. Then, BP equations for cavity fields are given as

$$h_{ia} = -\kappa + \sum_{b \in \partial i \backslash a} \hat{h}_{bi}, \qquad (12)$$

$$\hat{h}_{ai} = -\frac{1}{\mu} \ln \left[1 - \frac{1}{1 + e^{-\mu}} \prod_{j \in \partial a \backslash i} \frac{1}{1 + e^{\mu h_{ja}}} \right]. \qquad (13)$$

By rescaling the single-spin probabilities as $P_i(x) \propto \exp(\mu \xi_i x)$ and $P_a(y) \propto \exp(\mu \hat{\xi}_a y)$ using local fields $\{\xi_i\}$ and $\{\hat{\xi}_a\}$, respectively, Eqs. (8) and (9) then read

$$\xi_i = -\kappa + \sum_{a \in \partial i} \hat{h}_{ai}, \qquad (14)$$

$$\hat{\xi}_a = -\frac{1}{\mu} \ln \left[1 - \frac{1}{1 + e^{-\mu}} \prod_{i \in \partial a} \frac{1}{1 + e^{\mu h_{ia}}} \right]. \qquad (15)$$

One can estimate the single-spin probability by solving the BP equations (12) and (13) as the loopy belief propagation.

IV. REPLICA-SYMMETRIC SOLUTION

In this section, the typical performance of BP is studied using the RS cavity method based on the simplest RS ansatz. The RS ansatz assumes that cavity fields $\{h\}$ and $\{\hat{h}\}$ are independent random variables respectively following probability distributions $P(h)$ and $\hat{P}(\hat{h})$. For biregular random graphs, it is apparent that these distributions have no variance because of the absence of fluctuation of degree in V_1 and V_2. We therefore introduce the cavity fields h and $\hat{h}$ in (L, R)-biregular random graphs. When we use the BP equations (12) and (13) they satisfy the following RS cavity equations in the large-N limit:

$$h = -\kappa + (L - 1)\hat{h}, \qquad (16)$$

$$\hat{h} = -\frac{1}{\mu} \ln \left[1 - \frac{1}{(1 + e^{-\mu})(1 + e^{\mu h})^{R-1}} \right]. \qquad (17)$$

Similarly, the local fields in (L, R)-biregular random graphs are set to ξ and $\hat{\xi}$, which satisfy

$$\xi = -\kappa + L\hat{h}, \qquad (18)$$

$$\hat{\xi} = -\frac{1}{\mu} \ln \left[1 - \frac{1}{(1 + e^{-\mu})(1 + e^{\mu h})^R} \right]. \qquad (19)$$

According to Eq. (7), for a given ρ_x, the local field ξ is determined as

$$\frac{e^{\mu \xi}}{1 + e^{\mu \xi}} = \rho_x. \qquad (20)$$

Then, using Eqs. (18) and (20), we represent the appropriate parameter κ as

$$\kappa = L\hat{h} - \frac{1}{\mu} \ln r, \qquad (21)$$

where $r = \rho_x/(1 - \rho_x)$.

By substituting Eqs. (16) and (21) in Eq. (17) and by introducing $x = e^{-\mu \hat{h}}$, one obtains the self-consistent equation as

$$(1 + rx)^{R-1}(1 - x) = \frac{1}{1 + e^{-\mu}}. \qquad (22)$$

The order of this solution changes depending on the value of ρ_x. The solution x vanishes as μ becomes large if $\rho_x < 1/R$ holds. Using Taylor expansion of the left-hand side of Eq. (22), we estimate it as

$$x = \frac{e^{-\mu}}{1 - (R - 1)r} + O(e^{-2\mu}). \qquad (23)$$

Using Eq. (19), we obtain an average density of particles on V_2 as follows:

$$\lim_{M \to \infty} \overline{\left\langle \frac{1}{M} \sum_{i=1}^M y_a \right\rangle_\mu} = \frac{1 - (1 + rx)^{-R}}{1 - (1 + rx)^{-R} + e^{-\mu}} \qquad (24)$$

$$= R\rho_x + O(e^{-\mu}). \qquad (25)$$

Taking the large-μ limit, we obtain the average cover ratio for $\rho_x < 1/R$ as

$$\rho_y^{\mathrm{RS}} = R\rho_x. \qquad (26)$$

This relation indicates that each vertex in V_2 is connected to at most one chosen vertex in V_1. However, if ρ_x is larger than $1/R$ ($R \geqslant 2$), then the solution x remains constant for sufficiently large μ. In this case, a simple solution $\rho_y^{\mathrm{RS}} = 1$ is obtained by using Eq. (24) because the average cover ratio in Eq. (26) touches to 1 at $\rho_x = 1/R$.

Before closing this section, we discuss the stability of the RS solutions using the spin-glass susceptibility defined as

$$\chi_{\mathrm{SG}}(\mu) = \frac{1}{N + M} \overline{\sum_{i \in V_1} \sum_{a \in V_2} (\langle x_i y_a \rangle_\mu - \langle x_i \rangle_\mu \langle y_a \rangle_\mu)^2}. \qquad (27)$$

Another representation of χ_{SG} using cavity fields is known based on the fluctuation-dissipation theorem [26]. As shown in Ref. [27], in the case of biregular random graphs, it reads

$$\chi_{\mathrm{SG}}(\mu) \simeq \sum_{d=0}^\infty \lambda^d, \quad \lambda = \mathbb{E}\left[\sum_{j \in \partial a \backslash i; b \in \partial j \backslash a} \left(\frac{\partial \hat{h}_{ai}}{\partial \hat{h}_{bj}} \right)^2 \right], \qquad (28)$$

where $\mathbb{E}[\cdots]$ represents an average over vertices in V_1 and random graphs. The cavity fields are mutually correlated in general. Actually, BP cannot converge any more [28] if the susceptibility diverges. In this sense, the divergence of χ_{SG} means not only the instability of the RS solution but also poor typical performance of BP as an approximation algorithm in the static sense. In the low-density region where $\rho_x < 1/R$
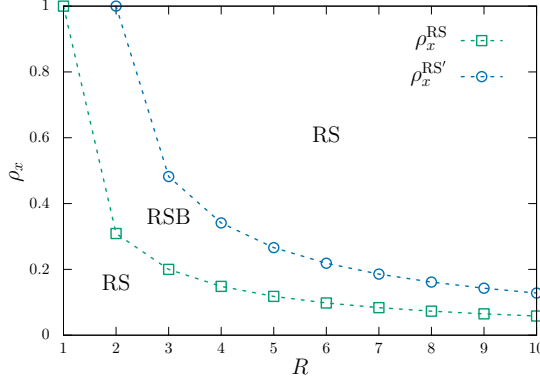
FIG. 2. RS-RSB thresholds ρ_x^{RS} (squares) and $\rho_x^{\mathrm{RS}'}$ (circles) in $(3R, R)$-regular random graphs.

holds, it is evaluated as

$$
\begin{aligned}
\lambda &= (L-1)(R-1)\left[\frac{r(1-x)}{1+rx}\right]^2 \\
&\rightarrow (L-1)(R-1)r^2 \quad (\mu \rightarrow \infty).
\end{aligned} \tag{29}
$$

The RS-RSB threshold in the large-μ limit therefore reads

$$
\rho_x^{\mathrm{RS}} = \frac{1}{1 + \sqrt{(L-1)(R-1)}}. \tag{30}
$$

Otherwise, the other threshold $\rho_x^{\mathrm{RS}'} (> 1/R)$ satisfies

$$
(L-1)(R-1)\left[\frac{\rho_x^{\mathrm{RS}'}(1-x^*)}{(1-\rho_x^{\mathrm{RS}'})(1+x^*)}\right]^2 = 1, \tag{31}
$$

where x^* is a solution of

$$
\left(1 + \frac{\rho_x^{\mathrm{RS}'}}{1-\rho_x^{\mathrm{RS}'}}x^*\right)^{R-1}(1-x^*) = 1. \tag{32}
$$

In Fig. 2, RS-RSB thresholds ρ_x^{RS} and $\rho_x^{\mathrm{RS}'}$ in $(3R, R)$-regular random graphs are shown as a function of R. Except for the case $R = 1$, the RS-RSB thresholds separate RS and RSB regions. The RSB region remains for a finite ρ_x while the RS-RSB thresholds converge to 0 in the large-R limit.

V. TYPICAL ANALYSIS OF THE GREEDY ALGORITHM

In this section, typical behavior of the greedy algorithm in (L, R)-biregular random graphs is investigated in the low-density region of $\rho_x < 1/R$. The corresponding RS solution indicates that one can choose vertices in V_1 without overlapping their neighbors. It is therefore sufficient to analyze the fraction of vertices in V_1 with the maximum degree during the deletion process. Here we analyze a rate equation that is used frequently for analyses of similar greedy algorithms [4,29]. Let $V(T)$ be the expected number of vertices with the maximum degree in V_1 at the Tth step of the algorithm. By the definition of the algorithm in Sec. II B, we find that

$$
NV(T+1) = NV(T) - 1 - L(R-1)\frac{NV(T)}{N-RT} + O(N^{-1}), \tag{33}
$$

where the assumption $L = O(1)$ is used. By introducing $v(t) = V(T)/N$ and $t = T/N$, we obtain the following differential equation in the large-graph limit:

$$
\frac{dv(t)}{dt} = -1 - \frac{L(R-1)}{1-Rt}v(t). \tag{34}
$$

Under the initial condition $v(0) = 1$, the solution reads

$$
v(t) = \frac{(L-1)(R-1)(1-Rt)^{\frac{L(R-1)}{R}} - (1-Rt)}{LR - L - R}. \tag{35}
$$

Let ρ_x^{g} be a threshold below which vertices with the maximum degree are left at the end of the algorithm. Considering that t represents a fraction of chosen vertices in V_1, we find $v(\rho_x^{\mathrm{g}}) = 0$, that is,

$$
\rho_x^{\mathrm{g}} = \frac{1}{R}\left\{1 - [(L-1)(R-1)]^{-\frac{R}{LR-L-R}}\right\}. \tag{36}
$$

If $\rho_x < \rho_x^{\mathrm{g}}$, then no chosen vertex in V_1 has overlapped neighbors in V_2, resulting in a good typical performance of the algorithm, i.e., $\rho_y^{\mathrm{g}} = R\rho_x$. However, the algorithm typically mistakes selections of vertices in V_1 and underestimates the average cover ratio. Results show that $\rho_x^{\mathrm{g}} \leqslant \rho_x^{\mathrm{RS}}$ where the equality holds for $R = 1$, indicating that BP is better than the greedy algorithm in terms of the typical performance threshold.

VI. ANALYSIS OF LP RELAXATION

An LP-relaxed value of the max-COV in any biregular graph is evaluated exactly using the LP duality [30] as follows.

Theorem 1. An LP-relaxed value of the max-COV in any (L, R)-biregular graph is LK if $K \leqslant 1/R$ holds.

Proof. $x_i = K/N$ $(1 \leqslant i \leqslant N)$ and $y_a = RK/N$ $(1 \leqslant a \leqslant M)$ is a feasible solution of LP-relaxed max-COV (3). The value of the cost function is then $RKM/N = LK$.

The Lagrangian function of Eq. (3) is written as

$$
\begin{aligned}
L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q}) &= \sum_{a=1}^{M} y_a + p\left(K - \sum_{i=1}^{N} x_i\right) \\
&\quad + \sum_{a=1}^{M} q_a\left(\sum_{i \in \partial a} x_i - y_a\right),
\end{aligned} \tag{37}
$$

where $p \in \mathbb{R}$ and $\boldsymbol{q} \in \mathbb{R}^M$. Using the weak duality theorem, one finds the following inequalities:

$$
\begin{aligned}
\max_{\boldsymbol{x}, \boldsymbol{y}} \min_{p, \boldsymbol{q}} L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q}) &\leqslant \min_{p, \boldsymbol{q}} \max_{\boldsymbol{x}, \boldsymbol{y}} L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q}) \\
&\leqslant \max_{\boldsymbol{x}, \boldsymbol{y}} L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q}).
\end{aligned} \tag{38}
$$

Consequently, $\max_{\boldsymbol{x}, \boldsymbol{y}} L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q})$ is an upper bound of the LP-relaxed value. Because Eq. (37) is represented by

$$
L(\boldsymbol{x}, \boldsymbol{y}, p, \boldsymbol{q}) = \sum_{i=1}^{N}\left(-p + \sum_{a \in \partial i} q_a\right)x_i + \sum_{a=1}^{M}(1 - q_a)y_a + pK, \tag{39}
$$

the solution $x_i = y_a = 1$ $(1 \leqslant i \leqslant N, 1 \leqslant a \leqslant M)$ realizes the maximum of the function with p and $\boldsymbol{q}$ satisfying $-p +$

$\sum_{a \in \partial i} q_a \geqslant 0$ and $1 - q_a \geqslant 0 (1 \leqslant a \leqslant M)$. Assuming that $q_a = q$ holds for $1 \leqslant a \leqslant M$, one finds that

$$\max_{x,y} L(x, y, p, q) = -pN + LNq + M - Mq + pK. \quad (40)$$

On the right-hand side, setting $(p, q) = (M/N, K/N)$ results in LK, which indicates that the upper bound of the LP-relaxed max-COV in an (L, R)-biregular graph is LK. Because the approximate solution denoted above achieves this bound, the LP-relaxed value is equivalent to the bound. ∎

This theorem claims that $\rho_y = R\rho_x (\rho_x \leqslant 1/R)$, suggesting that LP relaxation typically finds good approximate values in the RS regime. The LP-relaxed value is equivalent to ρ_y^{RS} in the high-density regime where $\rho_x > 1/R$.

VII. NUMERICAL RESULTS

This section presents a description of some numerical results for validation of the theoretical analyses. Here, we set $(L, R) = (9, 3)$ as an example. Then, the RS-RSB threshold and the typical performance threshold of the greedy algorithm are evaluated as $\rho_x^{RS} = 0.2$ and $\rho_x^g = (1 - 2^{-4/5})/3 = 0.1418 \ldots$. Biregular random graphs are generated based on implementation of the configuration model [31]. At least 400 random graphs are used to take a random graph average.

A. Average cover ratio

To examine the validity of theoretical analyses on the typical performance of approximation algorithms, the average cover ratio ρ_y is evaluated using several methods. We employ a BP-guided decimation (BPD) algorithm [22] as a variant of loopy BP to obtain a feasible solution satisfying all the constraints. The algorithm fixes a value of a variable based on a solution of the BP equation (11) until all variables are fixed to either 0 or 1. When a variable is fixed, BP equations are updated by applying the fixed value. If the number of variables in V_1 fixed to 1 reaches to K, the remainders of x are immediately fixed to 0 to satisfy the constraint. An approximate value of the algorithm is thus a function of the parameter κ, which depends on inputs, i.e., μ, K, and a graph. Here, the BP equations with $\mu = 20$ are solved iteratively up to 150 steps, which enables the algorithm to fix a variable practically even if the iterations cannot reach to the RS fixed point. The parameter κ is tuned to maximize an approximate value while the detail will be reported elsewhere. In addition, the RS estimation of the average cover ratio ρ_y^{RS} in the RS regime $\rho_x < \rho_x^{RS}$ is used for comparison to the BPD algorithm, though its typical performance is possibly evaluated directly as in Ref. [32].

The theoretical result presented in Sec. VI is used for LP relaxation, which is valid for arbitrary biregular graphs. Average approximate values of the greedy algorithm are estimated numerically with $N = 10^3$.

We also evaluate the average optimal value using the replica exchange Monte Carlo (EMC) method [33]. Results show that single-spin-flip updates of the model (4) take a long relaxation time for equilibration even using the EMC method, indicating the existence of deep valleys of the free energy. Consequently, it is necessary to accelerate equilibration for a system with sufficiently small chemical potential μ. To avoid such slow relaxation, we consider an alternative lattice-gas model in
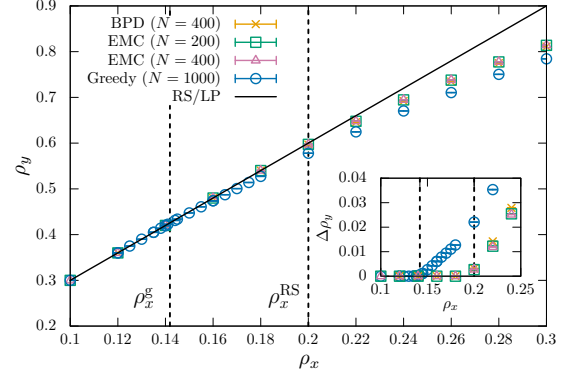

FIG. 3. Maximum cover ratio ρ_y obtained by some approximation algorithms as a function of ρ_x. Average optimal cover ratios are estimated for the cardinality $N = 200$ (squares) and 400 (triangles) by the replica EMC method, for $N = 1000$ by the greedy algorithm (circles) and for $N = 400$ by the BPD algorithm (cross marks). The solid line is the RS solution, equivalent to the LP-relaxed value. Two vertical lines respectively represent the analytical performance threshold ρ_x^g of the greedy algorithm and the RS-RSB threshold ρ_x^{RS}. Inset: The difference $\Delta\rho_y$ between the RS optimal cover ratio and numerical estimations by corresponding approximation algorithms as a function of ρ_x.

bipartite graph G represented as

$$\Xi_1(\mu; G) = \sum_{x \in \{0,1\}^N} \exp\left[\mu \sum_{a=1}^{M} \theta\left(\sum_{i \in \partial a} x_i\right)\right] \delta\left(K, \sum_{i=1}^{N} x_i\right), \quad (41)$$

where $\theta(x)$ takes 1 if $x > 0$ and 0 otherwise, and $\delta(x, y)$ is Kronecker's δ. In this model, variables y are eliminated because $y_a = \theta(\sum_{i \in \partial a} x_i)$ holds in the large-μ limit. The ground states of the alternative model therefore correspond to the optimal solutions of the max-COV in the same graph. Moreover, single-spin-flip updates in the model are substantially equivalent to multispin updates in the original model (4), which makes the relaxation time to equilibrium states markedly short. Because optimal values are invariant by replacing the inequality constraint $\sum_{i=1}^{N} x_i \leqslant K$ with an equality one, the equality constraint is adopted and Kawasaki dynamics for a density-conserved system [34] is applied.

The results are presented in Fig. 3. Numerical results are compatible to the RS estimation nearly up to the RS-RSB threshold $\rho_x^{RS} = 0.2$ for the BPD algorithm and the EMC method, and up to the typical performance threshold ρ_x^g predicted analytically for the greedy algorithm. As shown in the inset of Fig. 3, the average cover ratio by the greedy algorithm deviates from the average optimal cover ratio above ρ_x^g, while other estimates split above ρ_x^{RS}, indicating that the Bethe-Peierls approximation and LP relaxation are no longer appropriate. In fact, as shown in Fig. 4, the loopy BP for $N = 400$ fails its convergence above $\rho_x \simeq 0.195$, which may imply the existence of a dynamical one-step RSB phase. It is expected that the BPD algorithm finds approximate solutions with good accuracy if a corresponding loopy BP converges to a fixed point. Otherwise, decimation in the BPD algorithm is performed based on a wrong estimation by the Bethe-Peierls ap-
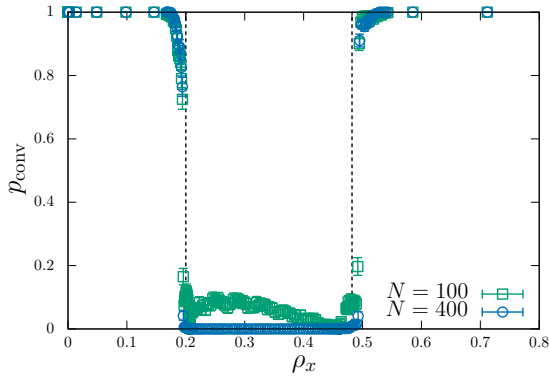
FIG. 4. Probability of successfully convergent samples in the loopy BP for $(9,3)$-biregular random graphs as a function of ρ_x. The dotted vertical lines represent RS-RSB thresholds shown in Fig. 2.

proximation. We emphasize, however, that the BPD algorithm finds relatively good approximate solutions even in the RSB phase as shown in Fig. 3. From these observations, we confirm that statistical-mechanical analyses successfully predict typical performance thresholds by approximation algorithms.

B. Greedy algorithm and its variant

Here we examine the greedy algorithm more closely. To validate the correctness of our analysis, we specifically examine a fraction of selected vertices without the maximum degree. The value r_g is evaluated by

$$r_g = 1 - \frac{\rho_x^g}{\rho_x}, \tag{42}$$

where ρ_x^g is given as Eq. (36). The algorithm finds an optimal solution of the problem if $r_g = 0$. As depicted in Fig. 5, the analytical estimations of r_g averaged over biregular random graphs agree well with the numerical results. The fraction arises at ρ_x^g corresponding to the typical performance threshold.

Additionally, we modify the greedy algorithm to reduce the gap of its typical performance to that of BP. As suggested above, the reason lies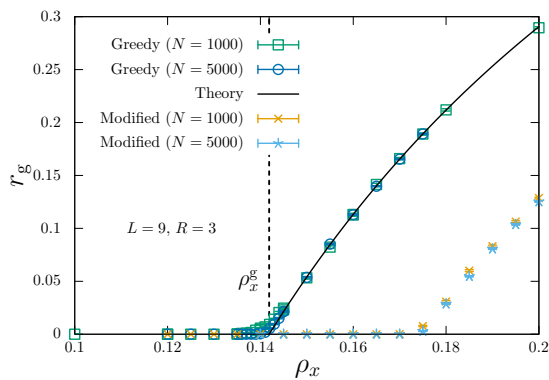 in the point that the greedy algorithm often chooses wrong vertices leading to a shortage of vertices with a maximum degree. To avoid the situation, one must select a vertex to preserve as many vertices to a maximum degree as possible, meaning that the optimal selection requires an exhaustive search. We therefore propose a modified algorithm in consideration of the influence of the selection on other vertices in V_1, the simplest improved algorithm toward the optimal selection. Let $\partial^2 i = \{j \in V_1 \backslash i \mid j \in \partial a \ (\forall a \in \partial i)\}$ be a set of the second neighbors of vertex $i \in V_1$. We also define the subset of vertices with the maximum degree in V_1 by W_1. The modified greedy algorithm is given as follows: at each step, (i) choose the vertex named $i \in W_1$ so that $|\partial^2 i \cap W_1|$ is minimized, (ii) delete vertices neighboring to vertex i from V_2, and (iii) update W_1 and V_1 and return to step (i) if $|V_1| > N - K$. In Fig. 5, the fraction r_g of the modified algorithm is also shown with that of the original greedy algorithm. As shown in Fig. 5, the typical performance threshold of the proposed algorithm is improved, although it is still below the RS-RSB threshold. This fact suggests that the typical performance of greedy algorithms depends on vertex choices in V_1 and BP selects vertices better than those greedy selections.

C. Randomized rounding of LP relaxation

In Sec. VII A, we examine the typical performance of LP relaxation in terms of its approximate value. In this subsection, randomized rounding, a practical means of constructing a feasible integer solution from LP relaxation, is applied to LP-relaxed solutions. It is noteworthy that the typical performance of randomized rounding probably depends on the selection of an LP solver and its setting. Here, we use IBM ILOG CPLEX with a default setting. The approximate value of the rounded solution is compared to that of the RS solution in the RS phase, which is LK. One considers that the rounding finds an optimal integer solution of the problem if two estimates mutually coincide. We define a success ratio p_r by the fraction of random graphs in which the rounding finds an optimal solution. Figure 6 presents the average approximation ratio ρ_y^r and the

FIG. 5. Fraction r_g of selected vertices without the maximum degree until the greedy algorithm stops as a function of ρ_x. Symbols represent numerical results of the greedy algorithm with cardinality $N = 1000$ (squares) and 5000 (circles) and the modified one with $N = 1000$ (cross marks) and 5000 (stars). The solid line exhibits the analytical estimation of the fraction r_g.
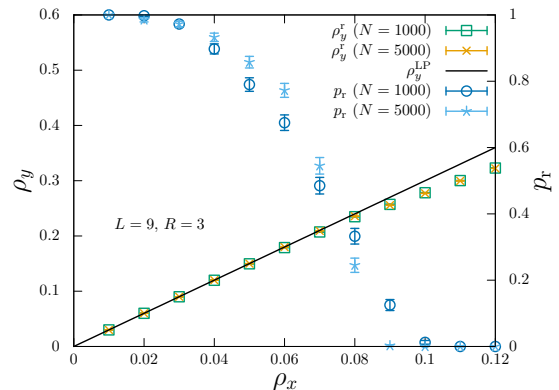

FIG. 6. Typical performance of randomized rounding of LP relaxation as a function of ρ_x (left vertical axis). Symbols represent approximation values ρ_y^r of rounded LP-relaxed solutions with cardinality $N = 1000$ (squares) and 5000 (cross marks). The solid line represents the average cover ratio equivalent to that of LP relaxation in the region shown. Squares and stars respectively represent the success ratio p_r (right vertical axis) of randomized rounding with $N = 1000$ and 5000.

success ratio p_r as a function of ρ_x, which strongly suggest that the randomized rounding exhibits a phase transition in terms of its typical performance. The threshold ρ_x^r is less than 0.1, which is much lower than ρ_x^g, whereas the LP-relaxed value is regarded as nearly optimal in the RS region $\rho_x < \rho_x^{RS} = 0.2$. We therefore conclude that the typical performance of the randomized rounding of LP-relaxed solution is inferior to that of the greedy algorithm.

VIII. SUMMARY AND DISCUSSION

As described in this paper, we investigate the typical performance of approximation algorithms called belief propagation, the greedy algorithm, and linear-programming relaxation for the maximum coverage problem in sparse biregular random graphs. The typical performance of BP is studied by application of the RS cavity method to a correspondent hard-core lattice-gas model. Results show that, in the large-μ limit, there exist two distinct RS-RSB thresholds regarded as typical performance thresholds of BP. In addition, the greedy algorithm performance and LP relaxation were studied especially in the low-density region. Results show that the typical performance threshold of the greedy algorithm is lower than that of BP and that LP-relaxed values are always equivalent to the RS solutions, leading to the threshold equivalent to that of BP. Those analytical results were validated by executing some numerical simulations. Results of additional numerical studies suggest that BP typically works better than the modified greedy algorithm and that randomized rounding of LP-relaxed solutions has a threshold lower than that of the greedy algorithm.

To assess the typical performance of BP as an approximation algorithm, we concentrated on statistical-mechanical analysis of max-COV up to the RS level. Further analyses based on the one-step RSB will be necessary to reveal statistical-mechanical properties of the problem and typical performance analysis of another algorithm called survey propagation. Another possible avenue of future work is the extension of our analyses to other random bipartite graphs. As with min-VCs, it is an attractive question whether the magnitude relation of typical performance thresholds changes depending on the random graph ensembles.

Our results provide not only respective typical performance of approximation algorithms but also their suggestive mutual relations. Specifically examining LP relaxation, it is worth emphasizing that the LP relaxation finds good approximate values compared to optimal values but the typical performance of its randomized rounding has the smallest threshold among the approximation algorithms studied here. To fill the gap of thresholds, it is important to examine modifications of LP relaxation such as the cutting-plane approach [35]. As for the greedy algorithm and its modification, numerical results suggest that evaluation of the influence of their deletion process affects the marked improvement of the typical performance threshold. The fact that BP is better than the greedy algorithm and its modification indicates that BP incorporates the influence more efficiently. These suggestions are expected to be of great help in understanding properties and relations of approximation algorithms in terms of typical performance. In addition, our analyses of the greedy algorithm and randomized rounding of LP relaxation illustrate that typical-case and worst-case evaluations capture different notions of approximate performance in optimization problems. This fact indicates the importance of the typical-case analysis of approximation algorithms. We hope that the arguments and results presented herein stimulate further studies and that the typical performance analyses of approximation algorithms will attract the interest of researchers in many diverse fields.

[1] V. V. Vazirani, *Approximation Algorithms* (Springer, Berlin, 2003).

[2] R. Monasson, R. Zecchina, S. Kirkpatrick, B. Selman, and L. Troyanskyk, Nature (London) **400**, 133 (1999).

[3] M. Mézard, F. Ricci-Tersenghi, and R. Zecchina, J. Stat. Phys. **111**, 505 (2003).

[4] R. M. Karp and M. Sipser, in *22nd Annual Symposium on Foundations of Computer Science (sfcs 1981)* (IEEE, New York, 1981), pp. 364–375.

[5] M. Weigt and A. K. Hartmann, Phys. Rev. Lett. **84**, 6118 (2000).

[6] M. Weigt and H. Zhou, Phys. Rev. E **74**, 046110 (2006).

[7] M. Bauer and O. Golinelli, Eur. Phys. J. B **352**, 339 (2001).

[8] T. Dewenter and A. K. Hartmann, Phys. Rev. E **86**, 041128 (2012).

[9] S. Takabe and K. Hukushima, J. Phys. Soc. Jpn. **83**, 043801 (2014).

[10] S. Takabe and K. Hukushima, J. Stat. Mech. (2016) 113401.

[11] Y.-Y. Liu, E. Csóka, H. Zhou, and M. Pósfai, Phys. Rev. Lett. **109**, 205703 (2012).

[12] S. Takabe and K. Hukushima, Phys. Rev. E **89**, 062139 (2014).

[13] C. Lucibello and F. Ricci-Tersenghi, Int. J. Stat. Mech. **2014**, 136829 (2014).

[14] Y. Habibulla, J.-H. Zhao, and H.-J. Zhou, in *Fronters in Algorithmics*, edited by J. Wang and C. K. Yap (Springer, New York, 2015), pp. 78–88.

[15] J. Takahashi, S. Takabe, and K. Hukushima, J. Phys. Soc. Jpn. **86**, 073001 (2017).

[16] H. Schawe and A. K. Hartmann, Eur. Phys. Lett. **113**, 30004 (2016).

[17] S. Takabe and K. Hukushima, Phys. Rev. E **93**, 053308 (2016).

[18] A. Javanmard, A. Montanari, and F. Ricci-Tersenghi, Proc. Natl. Acad. Sci. U.S.A. **113**, E2218 (2016).

[19] M. P. Johnson and A. Bar-Noy, in *2011 Proceedings IEEE INFOCOM* (IEEE, Shanghai, 2011), p. 1071.

[20] B. Saha and L. Getoor, in *Ninth SIAM International Conference on Data Mining* (Society for Industrial and Applied Mathematics, Philadelphia, 2009), pp. 697–708.

[21] H. Takamura and M. Okumura, in *Proceeding of the Conference of the European Chapter of the Association for Computational Linguistics* (Association for Computational Linguistics, Stroudsburg, PA, USA, 2009), pp. 781–789.

[22] M. Mézard and A. Montanari, *Information, Physics, and Computation* (Oxford University, Oxford, 2009).

[23] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, Math. Program. **14**, 265 (1978).

[24] U. Feige, J. Assoc. Comput. Mach. **45**, 634 (1998).

[25] P. Raghavan and C. D. Tompson, Combinatorica **7**, 365 (1987).

[26] O. Rivoire, G. Biroli, O. C. Martin, and M. Mézard, Eur. Phys. J. B **37**, 55 (2003).

[27] M. Mézard and M. Tarzia, Phys. Rev. E **76**, 041124 (2007).

[28] L. Zdeborová, Acta Phys. Slovaca Rev. **59**, 169 (2009).

[29] M. Weigt, Eur. Phys. J. B **28**, 369 (2002).

[30] R. T. Rockafellar, *Convex Analysis* (Princeton University, Princeton, NJ, 1972).

[31] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[32] F. Ricci-Tersenghi and G. Semerjian, J. Stat. Mech. (2009) P09001.

[33] K. Hukushima and K. Nemoto, J. Phys. Soc. Jpn. **65**, 1604 (1996).

[34] K. Kawasaki, Phys. Rev. **145**, 224 (1966).

[35] A. Schrijver, *Theory of Linear and Integer Programming* (Wiley & Sons, Chichester, 1998).