Guiding Principle for Minor-Embedding in Simulated-Annealing-Based Ising Machines

TATSUHIKO SHIRAI [iD][1]**, SHU TANAKA**[2,3]**, AND NOZOMU TOGAWA** [iD][1]**, (Member, IEEE)**
[1]Department of Computer Science and Communications Engineering, Waseda University, Tokyo 169-8555, Japan
[2]Department of Applied Physics and Physico-Informatics, Keio University, Kanagawa 223-8522, Japan
[3]Green Computing System Research Organization, Waseda University, Tokyo 162-0042, Japan

Corresponding author: Tatsuhiko Shirai (tatsuhiko.shirai@aoni.waseda.jp)

ABSTRACT We propose a novel type of minor-embedding (ME) in simulated-annealing-based Ising machines. The Ising machines can solve combinatorial optimization problems. Many combinatorial optimization problems are mapped to find the ground (lowest-energy) state of the logical Ising model. When connectivity is restricted on Ising machines, ME is required for mapping from the logical Ising model to a physical Ising model, which corresponds to a specific Ising machine. Herein we discuss the guiding principle of ME design to achieve a high performance in Ising machines. We derive the proposed ME based on a theoretical argument of statistical mechanics. The performance of the proposed ME is compared with two existing types of MEs for different benchmarking problems. Simulated annealing shows that the proposed ME outperforms existing MEs for all benchmarking problems, especially when the distribution of the degree in a logical Ising model has a large standard deviation. This study validates the guiding principle of using statistical mechanics for ME to realize fast and high-precision solvers for combinatorial optimization problems.

INDEX TERMS Annealing machine, graph minor-embedding, Ising model, optimization method, simulated annealing, statistical mechanics.

I. INTRODUCTION
A. MOTIVATION
Combinatorial optimization problems find the optimal combination of decision variables to minimize or maximize the objective function under given constraints. Solving a combinatorial optimization problem with a large number of decision variables is difficult because the number of solution candidates increases exponentially with the number of decision variables. Typical examples of combinatorial optimization problems found in textbooks include the satisfiability problem, the traveling salesman problem, and the knapsack problem. In our daily life, combinatorial optimization problems are ubiquitous. Common examples include the shift-planning optimization, the logistics optimization, and the traffic route optimization. Consequently, the development of efficient solvers for combinatorial optimization problems has attracted attention both in academia and in industry.

The associate editor coordinating the review of this manuscript and approving it for publication was Guijun Li [iD].

Ising machines have been developed as fast and high-precision solvers for combinatorial optimization problems [1]–[12]. They employ three phases to solve problems. In the first phase, a combinatorial optimization problem is mapped as an Ising problem. The Ising problem finds the ground (lowest-energy) state of the logical Ising model, which was originally introduced in statistical mechanics to describe the nature of phase transition materials [13], [14]. The Ising model consists of spins with values of $+1$ or -1. As described in Sec. III-A, the logical Ising model is defined on an undirected graph with unrestricted connectivity between vertices. The objective function and the constraints in a combinatorial optimization problem are encoded in the Ising model [15]–[17]. Different encoding methods have been proposed: machine learning [18], portfolio optimization [17], [19], traffic optimization [20], optimization in an integrated design circuit [21], [22], and material design [23]. In the second phase, the logical Ising model formulated in the first phase is mapped onto a physical Ising model. The model corresponds to the Ising machine considered. Here, the physical

Ising model is defined on an undirected graph where the connectivity between vertices may be restricted. For Ising machines with restricted connectivity such as D-Wave [1], [2] and CMOS annealing machines [3], [8], the mapping called minor-embedding (ME) [24] is necessary. In ME, a single spin in the logical Ising model is expressed by several spins in the physical Ising model. The set of spins is called a chain since chains are often formed in an actual ME. In the third phase, the Ising machine searches for the lowest-energy state according to its operation principle.

ME can be classified into two types according to the number of spins in each chain. In the first type, each chain has the same number of spins (i.e., a uniform chain length). This type of ME is often called clique ME or complete-graph ME because the logical Ising model with all-to-all coupling can be embedded. The algorithms for this type of ME have been developed for D-Wave [25]–[27] and CMOS annealing machines [28]. In the second type, each chain has a different number of spins. The total number of spins in the physical Ising model is usually smaller in the second type if the logical Ising model is not fully connected. Thus, the second type can embed a larger number of logical spins. Heuristic algorithms for finding this type of ME have been developed [29]–[36]. In existing MEs, the spins in a chain interact with ferromagnetic coupling. In both types of MEs, the chains have the same coupling strength. Here, we call the two types of MEs "uniform-length and uniform-coupling ME (ME *i*)" and "nonuniform-length and uniform-coupling ME (ME *ii*)", respectively.

B. SUMMARY OF CONTRIBUTIONS

Herein, we discuss the guiding principle of ME design to achieve a higher performance in simulated-annealing (SA)-based Ising machines. The main contributions are:

- A novel type of ME is proposed where the lengths are nonuniform and the coupling strength of each chain depends on the chain length. The formula between the coupling strength and the chain length is derived from a viewpoint of statistical mechanics. The coupling strength increases with the chain length. This type of ME, which is herein called "nonuniform-length and nonuniform-coupling ME (ME *iii*)", has not been discussed in previous studies.
- The performance of our proposed ME is compared to two existing types of ME through SA. The results demonstrate that the proposed ME has the best performance for all the problems. In particular, it outperforms the others when the degree of the logical Ising model is widely distributed. The results are general and independent of the distribution of the coupling strengths and biases in logical Ising models.

The rest of the paper is organized as follows. Section II briefly introduces the SA and thermal equilibrium states. The idea of thermal equilibrium states is necessary to derive the proposed ME (ME *iii*). Section III discusses ME to fix the

---

Algorithm 1 Simulated Annealing Implemented by Markov Chain Monte Carlo

1: **for** each run **do**
2:     initialize to a random initial state
3:     **for** each temperature T **do**
4:         **for** each Monte Carlo sweep at the temperature **do**
5:             choose a candidate site
6:             calculate the energy difference ΔE given by (1)
7:             generate a random number r such that $0 \leq r < 1$
8:             if r is less than the transition probability $W(\Delta E, T)$, update the state
9:         **end for**
10:        lower the temperature
11:    **end for**
12: **end for**

---

notation. Then a physical Ising model is presented to tune the chain lengths and intra-chain-coupling strengths in ME. With this model, we show the new type of ME as well as the two existing types of ME. Section IV explains the experimental setup. Section V demonstrates the numerical results, and Sec. VI discusses our results. Section VII concludes with a summary of the results and future research directions. The Appendices give supplemental information for the derivation of the proposed ME (Appendix A) and the experimental results (Appendix B and Appendix C).

II. SIMULATED ANNEALING AND THE THERMAL EQUILIBRIUM STATE

SA is a heuristic algorithm. It is useful in a wide range of application [37]–[39]. It has been employed to find the optimal solution of an objective function in combinatorial optimization problems. To explain SA as an operation principle of Ising machines in the language of statistical mechanics, we consider the objective function as an energy function, which is referred to as the Hamiltonian of the Ising model. As explained in Sec. I, the Ising model consists of spins with values of $+1$ and -1. Let $H(\{\sigma_i\})$ be the Hamiltonian of the Ising model, where $\{\sigma_i\}$ is a combination of decision variables called spins. In this case, the ground (lowest-energy) state corresponds to the spin combination (spin configuration) $\{\sigma_i\}$ that minimizes the value of $H(\{\sigma_i\})$.

Algorithm 1 shows the SA algorithm implemented by Markov Chain Monte Carlo (MCMC). The algorithm starts from a completely random initial state. That is, the spin configuration $\{\sigma_i\}$ is arbitrarily selected. Then the spin configuration is repeatedly updated. Let us consider a transition from the current state $\{\sigma_i'\}$ to a candidate state $\{\sigma_i\}$. The probability of making the transition is specified by a transition probability $W(\Delta E, T)$, which depends on temperature T and the energy difference between the two states defined by

$$\Delta E = H(\{\sigma_i\}) - H(\{\sigma_i'\}). \tag{1}$$

According to the principle of MCMC, the transition probability $W(\Delta E, T)$ must satisfy the balance condition, which is given as

$$\sum_{\{\sigma_i\}} W(\Delta E, T) \mathrm{P}_T^{\mathrm{eq}}(\{\sigma_i'\}) = \sum_{\{\sigma_i\}} W(-\Delta E, T) \mathrm{P}_T^{\mathrm{eq}}(\{\sigma_i\}), \quad (2)$$

where the summation means the summation of all the spin configurations. Two well-known choices of the transition probability satisfying the above equation are the heat-bath method and the Metropolis method. Here, $\mathrm{P}_T^{\mathrm{eq}}(\{\sigma_i\})$ is the probability distribution of the thermal equilibrium state at temperature T and is given by

$$\mathrm{P}_T^{\mathrm{eq}}(\{\sigma_i\}) = \frac{\exp\left[-\frac{H(\{\sigma_i\})}{T}\right]}{\sum_{\{\sigma_i'\}} \exp\left[-\frac{H(\{\sigma_i'\})}{T}\right]}. \quad (3)$$

Here, we set the Boltzmann constant, which is a physical constant, to unity. When the temperature T is fixed, Algorithm 1 is used to sample spin configurations in a thermal equilibrium state [16], [40]. The thermal equilibrium state at high temperature is a random state where the population is almost the same for all spin configurations. By contrast, the thermal equilibrium state at low temperature has a large population in the lower-energy states. In SA, by gradually lowering the temperature, the state should make transitions from a high-temperature state to a low-temperature state while annealing. After performing SA, a lower-energy state, ideally the ground state of $H(\{\sigma_i\})$, is identified.

Herein the expectation value of a physical quantity in the thermal equilibrium state at temperature T is referred to as the thermal average and is denoted as

$$\langle f \rangle_T := \sum_{\{\sigma_i\}} f(\{\sigma_i\}) \mathrm{P}_T^{\mathrm{eq}}(\{\sigma_i\}; T), \quad (4)$$

where $f(\{\sigma_i\})$ is an arbitrary function of the spin configuration.

III. MINOR-EMBEDDING
In this section, we describe our proposed ME. First, we introduce the concept of ME and a physical Ising model to systematically tune the chain length and the intra-chain-coupling strength.

A. BRIEF INTRODUCTION OF MINOR-EMBEDDING
ME is the mapping from a logical Ising model to a physical Ising model. The symbols L and P denote a logical Ising model and a physical Ising model, respectively. The logical Ising model is defined on an undirected graph $G_\mathrm{L} = (V_\mathrm{L}, E_\mathrm{L})$, where V_L and E_L are sets of vertices and edges, respectively. Herein we refer to G_L as a logical graph. The number of vertices is denoted by N_L. As mentioned in Sec. II, the Hamiltonian of the Ising model is an objective function and is given by

$$H_\mathrm{L}(\{\sigma_i\}) = -\sum_{(i,j)\in E_\mathrm{L}} J_{ij}\sigma_i\sigma_j - \sum_{i\in V_\mathrm{L}} h_i\sigma_i, \quad (5)$$

where $\sigma_i \in \{-1, 1\}$ is the logical spin, J_{ij} is the interaction between spins σ_i and σ_j, and h_i is the bias on the spin σ_i. Both J_{ij} and h_i are real values. $J_{ij} > 0$ indicates ferromagnetic coupling, whereas $J_{ij} < 0$ denotes antiferromagnetic coupling. Many combinatorial optimization problems can be mapped as problems to find the ground state of $H_\mathrm{L}(\{\sigma_i\})$. The interaction strengths J_{ij} and the biases h_i are specified by the objective function and constraints of the combinatorial optimization problem.

In a similar manner, the physical Ising model is defined on an undirected graph $G_\mathrm{P} = (V_\mathrm{P}, E_\mathrm{P})$, where a physical spin with a binary variable is put on each vertex. [For the specific form of the Hamiltonian in this study, see eq. (6).] Hereafter, G_P is referred to as the physical graph and it corresponds to the graph determined by the Ising machine architecture. In general, the physical graph G_P has a degree constraint where each vertex can have at most a constant degree. For example, the degree is 6 for the Chimera graph [2], 15 for the Pegasus graph in the D-Wave machines [41], and 5 (1st generation prototype [3]) and 8 (2nd generation prototype [8]) in the CMOS annealing machines. Due to the connectivity restriction among vertices, the logical graph G_L is not typically a subgraph of G_P.

ME enables G_L to be expressed in G_P even when G_L is not a subgraph of G_P. Each vertex in the logical graph, $i \in V_\mathrm{L}$, is mapped to a set of several vertices in the physical graph, $\phi(i) \subset V_\mathrm{P}$. ME is defined by mapping $\phi : V_\mathrm{L} \to V_\mathrm{P}$, which satisfies the following conditions [29]:

1) For each vertex $i \in V_\mathrm{L}$, the vertices in $\phi(i) \subset V_\mathrm{P}$ are connected and the connection is called chain;
2) For all $i \neq j$ in V_L, $\phi(i)$ and $\phi(j)$ are disjointed;
3) For each pair $(i, j) \in E_\mathrm{L}$, the corresponding pair exists in the physical graph (i.e., a pair of vertices, $k \in \phi(i)$ and $\ell \in \phi(j)$, satisfying $(k, \ell) \in E_\mathrm{P}$).

The physical spins in a chain interact with a ferromagnetic coupling. When the ferromagnetic coupling is sufficiently large, the ground state of the logical Ising model and that of the physical Ising model have a one-to-one correspondence [24]. This implies that the ground state of $H_\mathrm{L}(\{\sigma_i\})$ is obtained by searching the ground state of the embedded physical Ising model.

B. PHYSICAL ISING MODEL TO TUNE CHAIN LENGTHS AND INTRA-CHAIN-COUPLING STRENGTHS IN MINOR-EMBEDDING
This subsection describes a physical Ising model to systematically tune the chain lengths and the intra-chain-coupling strengths. The upper panel of Fig. 1 represents the logical Ising model with $N_\mathrm{L} = 5$. The logical spins σ_i and σ_j are connected with coupling strength J_{ij} when there is an edge between the corresponding vertices, and the bias with the strength h_i is applied on each spin i. The lower panel shows the physical Ising model in which the logical Ising model is embedded. For simplicity, we assume that each chain is a ring of vertices in G_P. Each vertex in the logical graph, $i \in V_\mathrm{L}$,

Logical Ising model

Physical Ising model
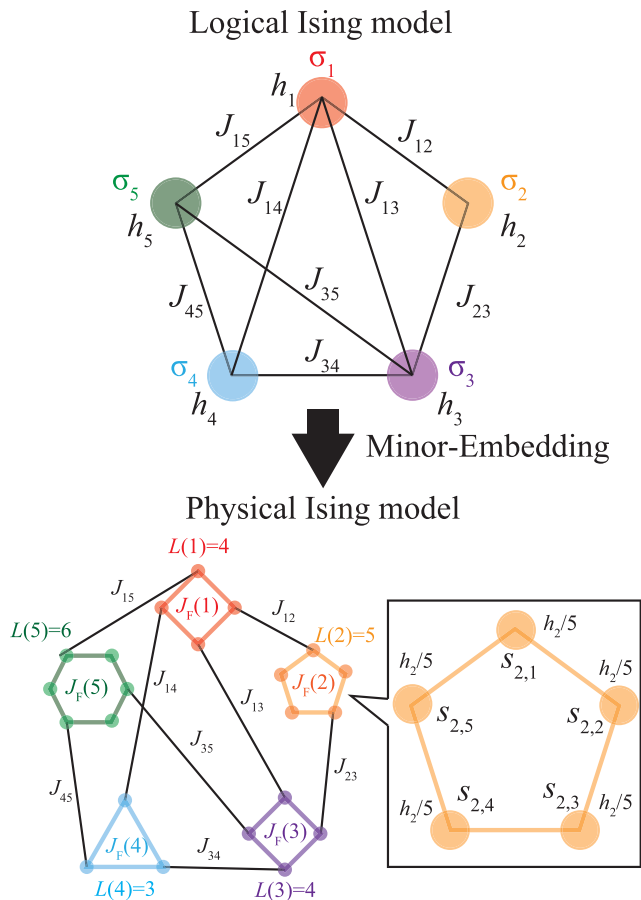
FIGURE 1. Example of mapping from a logical Ising model to a physical Ising model. (Upper) Logical Ising model with logical spins $\{\sigma_i\}_{i=1}^5$. $\{J_{ij}\}$ and $\{h_i\}$ denote the coupling strengths and the biases in the logical Ising model, respectively. (Lower) Physical Ising model, which is mapped by the ME of the logical Ising model. Each spin in the logical Ising model σ_i is mapped to a ring with length $L(i)$ and ferromagnetic coupling strength $J_F(i)$. Physical spins are labeled by $s_{i,k}$, where i is the ring label and k runs from 1 to $L(i)$. This model can tune the length of the ring $L(i)$ and the intra-ring-coupling strength $J_F(i)$.

is mapped to the ring with the length $L(i)$, and the physical spins in the ring are connected through a ferromagnetic coupling with the strength $J_F(i)$. The Hamiltonian of the physical Ising model is explicitly given by

$$H_P(\{s_{i,j}\}) = - \sum_{(i,j)\in E_L} J_{ij} s_{i,v_i(j)} s_{j,v_j(i)} - \sum_{i\in V_L} \frac{h_i}{L(i)} \sum_{k=1}^{L(i)} s_{i,k}$$
$$- \sum_{i\in V_L} J_F(i) \sum_{k=1}^{L(i)} s_{i,k} s_{i,k+1}, \quad (6)$$

where $s_{i,j} \in \{-1, 1\}$ is the j-th physical spin in the ring $\phi(i)$ and the periodic boundary condition is imposed (i.e., $s_{i,L(i)+1} = s_{i,1}$). There is an interaction with the strength J_{ij} between a physical spin in a ring $\phi(i)$ and a physical spin in a ring $\phi(j)$, and $v_i(j)$ denotes the physical spin in a ring $\phi(i)$. We assume that each physical spin interacts with one physical spin in other rings, at most. Therefore, $v_i(j) \neq v_i(k)$ if $j \neq k$. The bias on each physical spin in a ring

$\phi(i)$ is set as $h_i/L(i)$. In this way, the biases applied to the spins in a ring become uniform. Since intra-ring-couplings are ferromagnetic couplings, $J_F(i) > 0$ for all i. When the strength of $J_F(i)$ is sufficiently large, the ground states of the logical Ising model and the physical Ising model have a one-to-one correspondence.

C. TYPES OF MINOR-EMBEDDING
We consider three types of ME: ME *i*, ME *ii*, and ME *iii*. These depend on the choice of the ring length $L(i)$ and the intra-ring-coupling strength $J_F(i)$. ME *i* and ME *ii* have been studied previously [25]–[29], [31]–[36]. ME *iii* is a new type proposed in this study. Equation (6) can systematically express the three types of ME.

- **ME *i*: uniform-length and uniform-coupling ME**
  In the first type of ME, all rings have the same length and coupling strength. We set the number of spins in a ring as $N_L - 1$, which is the maximum degree for each vertex. That is

$$L(i) = N_L - 1. \quad (7)$$

We set $v_i(j)$ as

$$v_i(j) = \begin{cases} j & \text{if } i > j, \\ j-1 & \text{if } i < j. \end{cases} \quad (8)$$

In ME *i*, the logical Ising model with all-to-all coupling can be embedded. We introduce a hyperparameter J_c for the intra-ring-coupling, which is expressed as

$$J_F(i) = J_c. \quad (9)$$

The number of vertices in the physical graph (i.e., the number of spins in the physical Ising model) is provided as

$$|V_P| = N_L(N_L - 1). \quad (10)$$

- **ME *ii*: nonuniform-length and uniform-coupling ME**
  In the second type of ME, the total number of physical spins is set as small as possible. For a given logical Ising model, it is sufficient to take the number of spins in a ring $\phi(i)$ as the degree of vertex denoted by k_i. Namely,

$$L(i) = k_i. \quad (11)$$

We set $v_i(j)$ as

$$v_i(j) = n_i(j) \text{ for } (i, j) \in E_L, \quad (12)$$

where $n_i(j) \in \mathbb{N}$ is an integer given for vertex $i \in V_L$. The integer is incremented by one when there is an edge between the logical spins i and j. That is, $(i, j) \in E_L$. For example, if a logical spin labeled by 2 interacts with spins labeled by 1, 5, and 6, then $n_2(1) = 1$, $n_2(5) = 2$,

and $n_2(6) = 3$. Similar to ME *i*, a uniform ferromagnetic coupling strength is assumed inside the ring and

$$J_{\text{F}}(i) = J_{\text{c}}. \tag{13}$$

In this ME, every physical spin is connected to a spin in another ring. Hence, the number of vertices in the physical Ising model is given by

$$|V_{\text{P}}| = \sum_{i \in V_{\text{L}}} \frac{k_i}{2}. \tag{14}$$

- *ME iii: nonuniform-length and nonuniform-coupling ME*

  We propose a new type of ME where the intra-ring-coupling strength depends on the ring length. Similar to ME *ii*, the length of ring $L(i)$ is equal to the degree of the vertices $i \in V_{\text{L}}$,

$$L(i) = k_i, \tag{15}$$

and $v_i(j)$ is set by eq. (12). The intra-ring-coupling strength $J_{\text{F}}(i)$ is given by

$$J_{\text{F}}(i) = -\frac{J_{\text{c}}}{2} \log \left[\tanh \left(\frac{1}{2L(i)} \right) \right]. \tag{16}$$

Here, $J_{\text{F}}(i)$ is a monotonically increasing function of $L(i)$, and it asymptotically behaves as

$$J_{\text{F}}(i) \sim \log L(i) \text{ for } L(i) \gg 1. \tag{17}$$

Below, we derive the formula in eq. (16). First, consider the local Hamiltonian of the *i*-th ring,

$$H_{\text{ring}}^{(i)}(\{s_{i,j}\}) = -J_{\text{F}}(i) \sum_{j=1}^{L(i)} s_{i,j} s_{i,j+1}. \tag{18}$$

Here, the effect due to inter-ring couplings between rings $\{J_{ij}\}$ and the biases on spins $\{h_i/L(i)\}$ is neglected. The correlation length $\xi_i(T)$ of this model at temperature T is given by [13] (see Appendix A for a detailed derivation)

$$\xi_i(T)^{-1} = -\log \left[\tanh \left(\frac{J_{\text{F}}(i)}{T} \right) \right], \tag{19}$$

where $\xi_i(T)$ is defined by

$$C_i(j) = \langle s_{i,k} s_{i,j+k} \rangle_T = \exp \left(-\frac{j}{\xi_i(T)} \right). \tag{20}$$

Here, $C_i(j)$ is called the correlation function. It describes the thermal average of the products of spins $s_{i,k}$ and $s_{i,j+k}$. The correlation function is independent of k due to the periodic boundary condition of the ring. As the distance between the spins increases, the value of $C_i(j)$ decays exponentially. The correlation length $\xi_i(T)$ determines the decay length scale. $\xi_i(T)$ is a monotonically decreasing function of T. At sufficiently low temperatures, the correlation length is much larger than the ring length $\xi_i(T) \gg L(i)$. Hence, all the spins in the ring tend to have the same values. On the other hand,

at sufficiently high temperatures, $\xi_i(T) \ll L(i)$. In this case, each spin in the ring randomly has values $+1$ or -1. The crossover occurs at $T_{\text{c}}(i)$, where

$$L(i) = \xi_i(T_{\text{c}}(i)). \tag{21}$$

Here, we assume that $T_{\text{c}}(i)$ of all the rings have the same value. That is,

$$T_{\text{c}}(i) = J_{\text{c}}. \tag{22}$$

Substituting eqs. (19) and (22) into (21) gives (16).

The guiding principle of ME design to achieve a high performance in Ising machines is that the intra-ring-coupling strength must be tuned according to eq. (16). In SA, the temperature T decreases from a high temperature to a low temperature. The physical spins in each ring randomly take values of $+1$ or -1 when $T > J_{\text{c}}$. By contrast, they are aligned in the same direction when $T < J_{\text{c}}$. The physical spins in each ring are aligned along the same direction simultaneously at $T = J_{\text{c}}$.

When the lengths of rings are uniform and $L(i) = N_{\text{L}} - 1$, ME *iii* is reduced to ME *i*. As such, the case with uniform-length and nonuniform-coupling ME is not considered in this study. Next, we compared the performance of the three MEs.

IV. EXPERIMENTAL SETUP
A. BENCHMARKING PROBLEMS

We considered four types of benchmarking problems (i.e., logical Ising models). Each benchmarking problem has its own distribution of the degree in G_{L} or of $\{J_{ij}\}$ and $\{h_i\}$.

- *Binomial-Bimodal problem*

  The logical graph G_{L} is created by connecting the vertices i and j by an edge with half probability. The degree distribution is given by the binomial distribution. The coupling strengths and the biases are chosen according to a bimodal distribution. That is, J_{ij} and h_i take values from $\{-1, +1\}$ with equal probability.

- *Binomial-Gaussian problem*

  The logical graph G_{L} is created by connecting the vertices i and j by an edge with half probability. The coupling strengths and the biases are chosen according to a Gaussian distribution with a mean of zero and a standard deviation of unity.

- *Power-Bimodal problem*

  The logical graph G_{L} with a scale-free network is created by the algorithm of the Barabasi-Albert (BA) model [42]. The degree distribution is given by a power-law distribution. The coupling strengths and the biases are chosen according to a bimodal distribution. That is, J_{ij} and h_i take values from $\{-1, +1\}$ with an equal probability.

- *Power-Gaussian problem*

  The logical graph G_{L} with a scale-free network is created by the algorithm of the BA model. The coupling strengths and the biases are chosen according to a Gaussian distribution with a mean of zero and a standard deviation of unity.

For each benchmarking problem, we prepared a hundred scenarios by creating ten connected logical graphs. For each connected graph, we generated ten sets of different coupling strengths $\{J_{ij}\}$ and biases $\{h_i\}$.

The BA model was originally introduced to explain the mechanism responsible for the emergence of power-law degree distributions of networks in various fields. In the algorithm of the BA model, the graph begins from a fully connected graph with m_0 vertices. At every step, a new vertex with m edges is added to m different vertices already present in the graph with a certain probability. The probability of connecting a new vertex and vertex i depends on the degree k_i, and is given as

$$P(i) = \frac{k_i}{\sum_j k_j}. \tag{23}$$

The algorithm ends when the number of vertices is N_{L}. Since the number of edges increases by m in every step, the number of edges in a graph with N_{L} vertices is approximately

$$\sum_{i \in V_{\mathrm{L}}} k_i \simeq m N_{\mathrm{L}}. \tag{24}$$

Numerical simulations and analytic results [42] have demonstrated that the graph evolves into a scale-free network. Namely, the histogram of the degree k denoted by $n(k)$ follows a power-law scaling. In the BA model, the exponent is 3 and is independent of m_0 and m.

Figure 2 shows histograms of the degree k in the Binomial-Bimodal problem and the Binomial-Gaussian problem [Fig. 2 (a)] and in the Power-Bimodal problem and the Power-Gaussian problem [Fig. 2 (b)] for the model with $N_{\mathrm{L}} = 1000$. The error bars denote the standard deviation of the 10 realizations of the logical graphs in each benchmarking problem. In the Binomial-Bimodal problem and the Binomial-Gaussian problem, there is a peak around $k = N_{\mathrm{L}}/2$ because the vertices are connected by an edge with half probability. The peak width is the order of $N_{\mathrm{L}}^{1/2}$. The histogram is well described by a scaled Gaussian distribution with a mean of $N_{\mathrm{L}}/2$ and a standard deviation of $\sqrt{N_{\mathrm{L}}}/2$. On the other hand, in the Power-Bimodal problem and the Power-Gaussian problem, the degree is more widely distributed, and the histogram follows a power law. The power-law scaling of $n(k) \propto k^{-3}$ is consistent with our data. Here, we set $m_0 = 3$ and $m = 3$.

B. SIMULATION DETAILS

We applied SA to the physical Ising models by adopting the single-spin flip Monte Carlo method. In each update of the spin configuration, the spin is randomly selected and the energy difference ΔE is calculated between the current state and the candidate state in which the chosen spin is flipped [see eq. (1)]. Here, the heat-bath transition probability at temperature T is used and is given as

$$W(\Delta E, T) = \left[1 + \exp\left(\frac{\Delta E}{T} \right) \right]^{-1}. \tag{25}$$
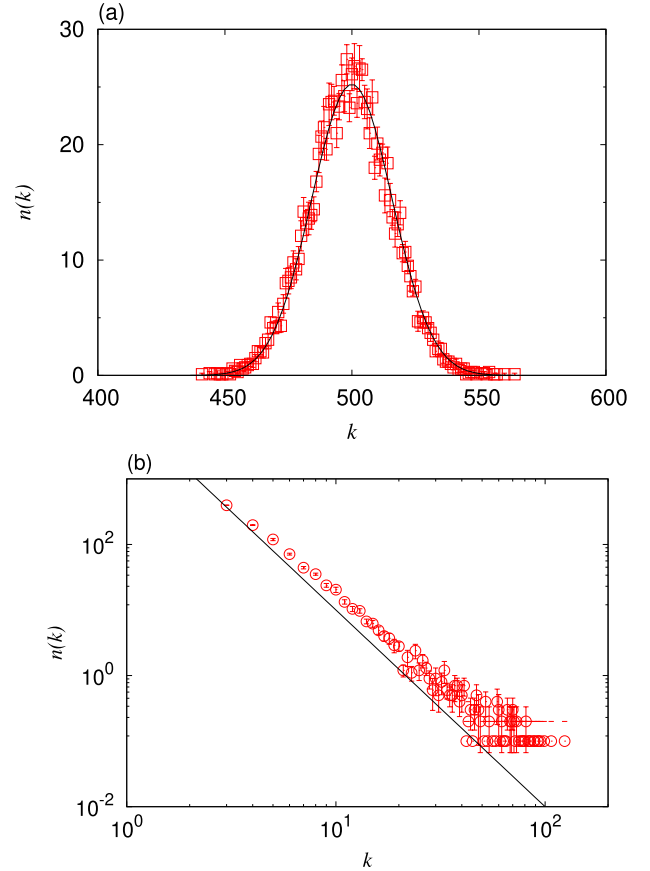
FIGURE 2. Histograms of the degree for the logical Ising model with 1000 spins in (a) the Binomial-Bimodal problem and the Binomial-Gaussian problem and (b) the Power-Bimodal problem and the Power-Gaussian problem. In Fig. (a), the line is drawn using a Gaussian distribution. In Fig. (b), the line is a guide to show the power-law scaling of $n(k) \propto k^{-3}$. Error bars are the standard deviation of 10 realizations of the logical graphs for each benchmarking problem.

Equation (25) satisfies the balance condition [see eq. (2)], and each Monte Carlo step (MCS) repeats the updates $|V_{\mathrm{P}}|$ times. The temperature is initially set to $T_{\mathrm{ini}} = 10$, which is larger than the typical energy scale, and decreases by 10^{-4} in every MCS. The temperature at the end of the annealing is zero. Appendix B shows the result using a different type of annealing schedule. Regardless of the annealing schedule, the same results are qualitatively produced.

After performing SA, the values of the logical spins $\{\sigma_i\}$ are determined from the spin configuration of the physical Ising model. If all the physical spins in the ring have the same value, the value is the same as that for the logical spin. If not, the value of the logical spin $+1$ or -1 is determined by the majority vote. Namely, when five physical spins take $+1$ and three physical spins take -1 in a ring $\phi(i)$, σ_i is determined as $+1$. If $(+1)$-spins and (-1)-spins are the same, the value of the corresponding logical spin is set to $+1$.

For each physical Ising model (i.e., the model mapped by an embedding), we performed SA one hundred times to estimate the average and standard deviation of the quantities described in Sec. V.

V. NUMERICAL RESULTS

We compared the performances of ME *i*, ME *ii*, and ME *iii* for each benchmarking problem. We measured two quantities. The first one is the step to solution (STS). The STS is the number of steps required for the algorithm to obtain the ground state at least once with a probability of 0.99, and it is defined by [9], [43]

$$R_{99}^{(k)} = \frac{\log(1 - 0.99)}{\log(1 - P_s^{(k)})}, \quad (26)$$

where $P_s^{(k)}$ is the success probability and k is a label of the logical Ising models. Herein k is the run, which ranges from 1 to 100. A small $R_{99}^{(k)}$ value indicates a good performance. Here, the success probability is estimated as $P_s^{(k)} = N_s^{(k)}/100$, where $N_s^{(k)}$ is the number of obtained ground states in one hundred simulations of SA. We measured STS for relatively small-size systems up to $N_L = 28$ because it is difficult to obtain the ground state of $H_L(\{\sigma_i\})$ for a larger system size.

For the N_L-dependence of the performance in a larger-sized system, we calculated the energy density (i.e., the value of $H_L(\{\sigma_i\})/N_L$), where σ_i is determined by the majority vote after SA (see subsection IV-B). $\epsilon^{(k)}$ denotes the average of the energy density for 100 simulations of SA, where k is the label of the logical Ising model. A smaller $\epsilon^{(k)}$ indicates a better performance. Note that the energy density can be evaluated on the order of N_L^2 steps. Thus, this quantity is useful to study larger-sized systems.

To investigate the performance of ME in the benchmarking problems, we used the median of the STSs and the median of energy densities. These are denoted as R_{99} and ϵ, respectively.

Figure 3 plots the J_c-dependences of R_{99} (a) and ϵ (b) in ME *i* for each benchmarking problem. Here, the optimal values of J_c that minimizes R_{99} or ϵ are found. We estimated the optimal values of J_c in each ME for different sized systems, where 0.1 is used as the precision threshold of J_c (see Appendix C for the N_L-dependences). In the following, we show the data of R_{99}, ϵ, and $\epsilon^{(k)}$ at the optimal value of J_c.

A. STEP TO SOLUTION (STS)

First, we used STS to compare the performance among ME *i*, ME *ii*, and ME *iii*. Figure 4 shows the N_L-dependences of STS for each benchmarking problem. In all cases, STS increases exponentially with N_L. For small system sizes, $N_L = 8$ or 12, there is not a clear difference in performance. However, a clear difference appears as the number of logical spins increases. For all benchmarking problems, ME *i* shows the poorest performance. For the Binomial-Bimodal problem [Fig. 4 (a)] and the Binomial-Gaussian problem [Fig. 4 (b)], the performances of ME *ii* and ME *iii* are the same within the margin of error. On the other hand, for the Power-Bimodal problem [Fig. 4 (c)] and the Power-Gaussian problem [Fig. 4 (d)], ME *iii* outperforms ME *ii*.
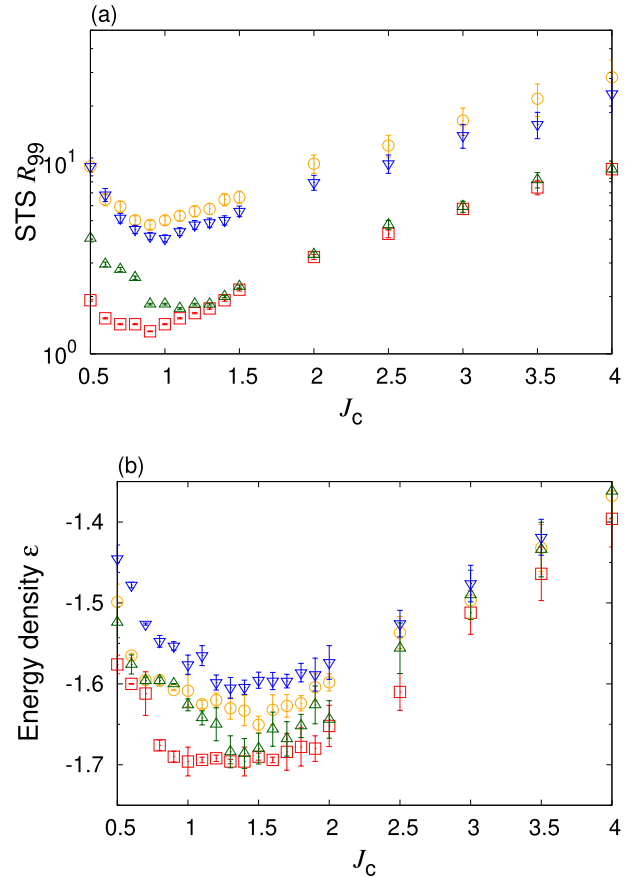

FIGURE 3. (a) J_c-dependences of the STS R_{99} and (b) the energy density ϵ in ME *i*. Squares (red), circles (orange), up-triangles (green), and down-triangles (blue) denote the Binomial-Bimodal problem, the Binomial-Gaussian problem, the Power-Bimodal problem, and the Power-Gaussian problem, respectively. Number of logical spins used is (a) $N_L = 8$ and (b) $N_L = 10$.

We also considered the time to solution (TTS) [9], [43]–[45]. TTS is the *total time* required for the algorithm to obtain the ground state at least once with a probability of 0.99. TTS is estimated as

$$\text{TTS} \simeq n_{\text{MCS}} \times \tau_{\text{MCS}} \times R_{99}, \quad (27)$$

where n_{MCS} is the number of the MCSs in the SA and τ_{MCS} is the time required to calculate one MCS. For all three MEs, n_{MCS} is the same as STS and is given as

$$n_{\text{MCS}} = 10^5. \quad (28)$$

In each MCS, we calculated the energy difference of a single-spin flip $|V_P|$ times. The time required for the calculation of the energy difference is $O(|V_P|^0)$ since the connectivity of the physical Ising model is sparse. Thus

$$\tau_{\text{MCS}} \sim |V_P|. \quad (29)$$

In the Binomial-Bimodal problem and the Binomial-Gaussian problem, the number of physical spins is on the order of N_L^2 for all three MEs. Thus, TTS qualitatively shows the same result as STS. In the Power-Bimodal problem and
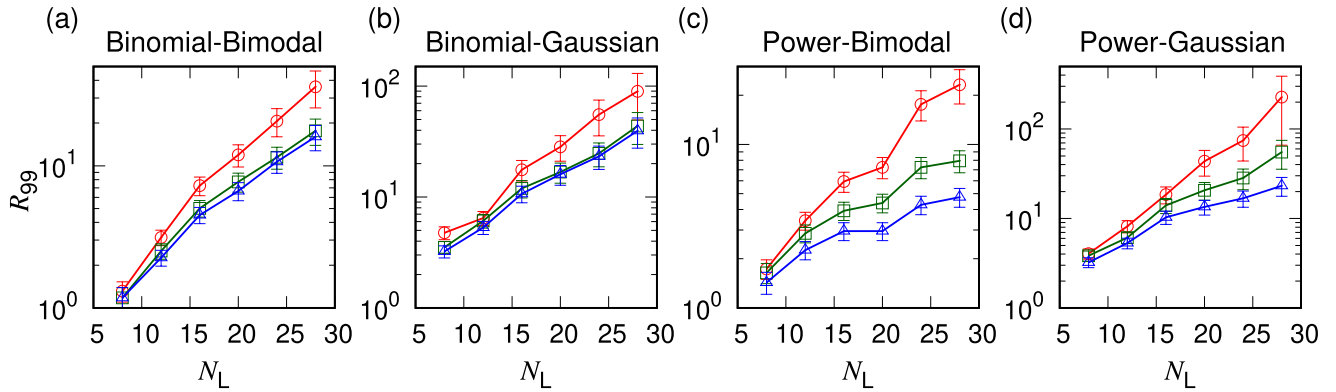
FIGURE 4. N_L-dependences of the STS for (a) the Binomial-Bimodal problem, (b) the Binomial-Gaussian problem, (c) the Power-Bimodal problem, and (d) the Power-Gaussian problem. ME *i*, ME *ii*, and ME *iii* are denoted by red circles, green squares, and blue triangles, respectively.
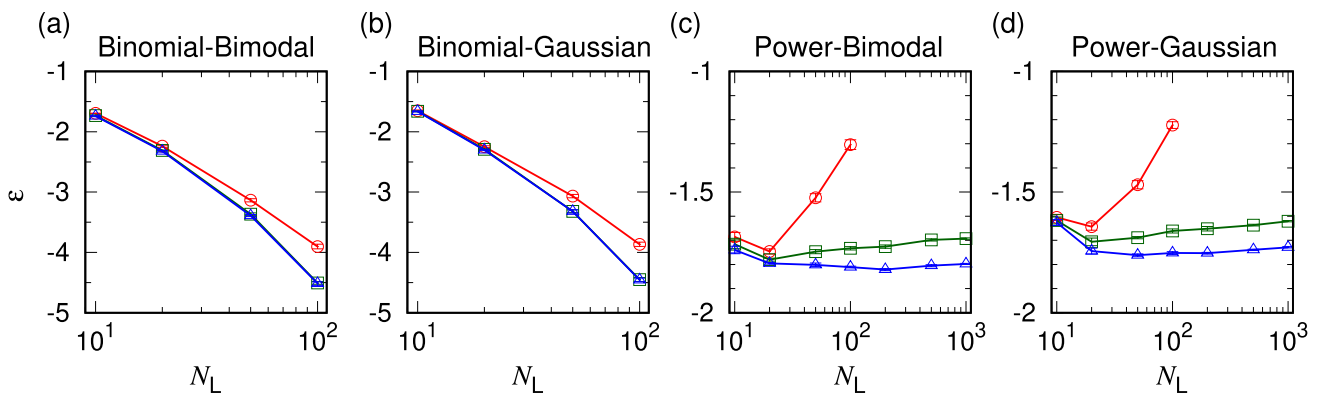
FIGURE 5. N_L-dependences of the energy density for (a) the Binomial-Bimodal problem, (b) the Binomial-Gaussian problem, (c) the Power-Bimodal problem, and (d) the Power-Gaussian problem. ME *i*, ME *ii*, and ME *iii* are denoted by red circles, green squares, and blue triangles, respectively. In Figs. (a) and (b), the data of ME *ii* and ME *iii* overlap with each other.

the Power-Gaussian problem, the number of physical spins is on the order of N_L^2 in ME *i*, while it is on the order of N_L in ME *ii* and ME *iii*. The scaling of $O(N_L)$ is obtained from eqs. (14) and (24). TTS shows larger performance differences between ME *i* and the other two compared to the STS.

B. ENERGY DENSITY

Next, we compared the performance among ME *i*, ME *ii*, and ME *iii* in terms of the energy density. The results are qualitatively the same as those of STS. Figure 5 shows the N_L-dependence of ϵ for each benchmarking problem. As N_L increases, the difference in performance among the MEs becomes clear. ME *i* has the poorest performance. ME *ii* and ME *iii* have the same performance for the Binomial-Bimodal problem [Fig. 5 (a)] and the Binomial-Gaussian problem [Fig. 5 (b)]. On the other hand, ME *iii* outperforms ME *ii* for the Power-Bimodal problem [Fig. 5 (c)] and the Power-Gaussian problem [Fig. 5 (d)]. Figures 5 (c) and (d) show that the differences in energy densities are almost the same for $N_L \geq 10^2$. This implies that ME *iii* will provide the best performance, even for larger-sized systems.

Figure 6 shows the scatterplot to compare the energy densities for each benchmarking problem (i.e., $\{\epsilon_k\}_{k=1}^{100}$, among

the three MEs) using the data at $N_L = 100$. The upper panel compares the energy densities between ME *i* and ME *ii*. All the points are plotted below the diagonal, indicating that ME *ii* outperforms ME *i* for all the logical Ising models in each benchmarking problem. The lower panel compares the energy densities between ME *ii* and ME *iii*. For the Binomial-Bimodal problem [Fig. 6 (a2)] and the Binomial-Gaussian problem [Fig. 6 (b2)], the points are plotted around the diagonal, implying that ME *ii* and ME *iii* have similar performance. On the other hand, for the Power-Bimodal problem [Fig. 6 (c2)] and the Power-Gaussian problem [Fig. 6 (d2)], all the points are plotted below the diagonal, indicating that ME *iii* is better suited for these benchmarking problems.

VI. DISCUSSION

The numerical studies demonstrate that ME *i* has the poorest performance. The poor performance of ME *i* is attributed to the large dimension of the solution space. In ME *i*, the logical Ising model with N_L-spins is mapped to the physical Ising model with $N_L(N_L-1)$-spins. On the other hand, in ME *ii*, the number of the physical spins is about $N_L^2/2$ for the Binomial-Bimodal problem and the Binomial-Gaussian problem, but is on the order of N_L for the Power-Bimodal problem and
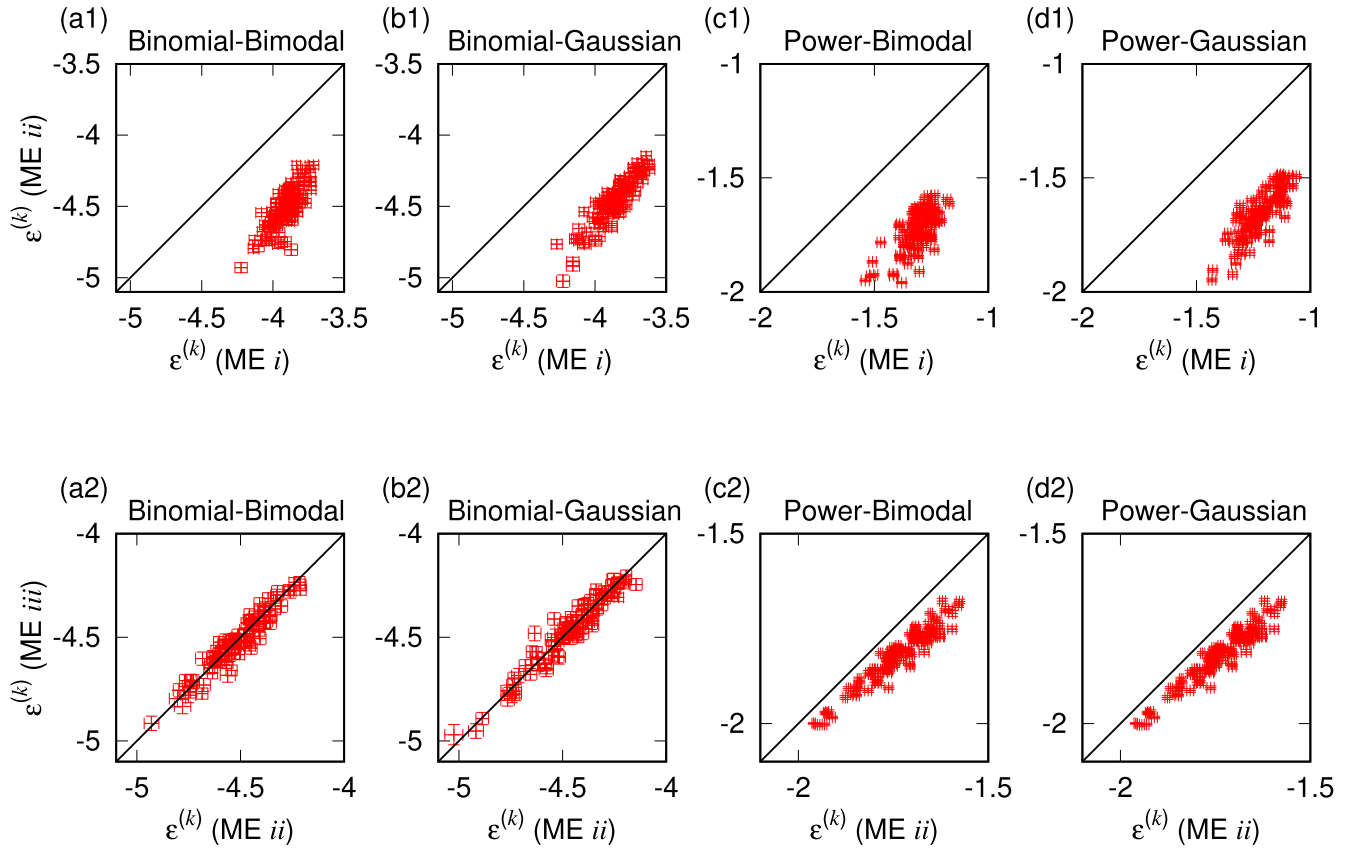
FIGURE 6. Scatterplots of the energy densities $\{\epsilon_k\}_{k=1}^{100}$ at $N_L = 100$ among ME *i*, ME *ii*, and ME *iii* for (a) the Binomial-Bimodal problem, (b) the Binomial-Gaussian problem, (c) the Power-Bimodal problem, and (d) the Power-Gaussian problem. Upper panel compares ME *i* and ME *ii*, and the lower panel compares ME *ii* and ME *iii*.

the Power-Gaussian problem. The dimension of the solution space increases exponentially with respect to the number of physical spins. Thus, the dimension of the solution space in ME *i* rapidly increases with N_L compared to those in ME *ii* and ME *ii*.

ME *ii* and ME *iii* have the same performance for the Binomial-Bimodal problem and the Binomial-Gaussian problem. This can be understood as follows for large $L(i)$. For the Binomial problem, the degree distribution shows a peak around $k = N_L/2$ with the width on the order of $N_L^{1/2}$ [see Fig. 2(a)]. In ME *iii*, the length of the ring is equal to the degree. Thus, $L(i)$ is distributed with a mean of $L_{mean} = N_L/2$ and a standard deviation ΔL on the order of $N_L^{1/2}$. For large $L(i)$, the intra-ring-coupling strength $J_F(i)$ behaves as the $\log L(i)$ [see eq. (17)]. Then, the standard deviation of $J_F(i)$ scales as

$$\Delta J_F \sim \Delta(\log L) \sim \frac{\Delta L}{L_{mean}} \sim N_L^{-\frac{1}{2}}. \qquad (30)$$

The standard deviation ΔJ_F decays with N_L, implying that ME *iii* approaches ME *ii* as the system size increases.

On the other hand, ME *iii* outperforms ME *ii* for the Power-Bimodal problem and the Power-Gaussian problem for large N_L. In these problems, the degree of the logical Ising model is widely distributed, reflecting the distribution of $L(i)$ in the

physical Ising model. Namely, the lengths of some rings are on the order of 1, while others are on the order of N_L. In these problems, it is necessary to set the intra-ring-coupling strength according to eq. (16) to achieve a high performance in SA-based Ising machines.

VII. CONCLUSION AND OUTLOOK

Here, we discussed the guiding principle of ME design to achieve a high performance in SA-based Ising machines from a viewpoint of statistical mechanics. We proposed a new type of ME shown in eq. (16). In the proposed ME, the coupling strength inside a chain depends on the chain length. This is a unique approach that has not been discussed previously. We compared the performance of our proposed ME with the two existing MEs using four benchmarking problems. SA showed that the proposed ME has the best performance for all the benchmarking problems. In particular, it outperformed the others when the logical Ising model has a wide degree distribution. The results are independent of the distribution of coupling strengths and biases in the logical Ising model.

We demonstrated the importance of tuning the intra-chain coupling strengths in SA, which is regarded as an ideal Ising machine. In the future, we plan to apply eq. (16) to real Ising machines such as a CMOS annealing machine.

It is also important to compare our results with the case of quantum annealing (QA) [46], where the transverse-field strength Γ plays the role of the temperature T in SA. A recent paper [47] shows that ME *ii* outperforms ME *i*. This is consistent with the results in this study. Furthermore, our results imply that the performance of QA could be enhanced for some problems if the intra-chain coupling strength J_F is tuned according to the chain length. The D-Wave's report [48] evaluated the chain-length dependence of the tunneling energy between the all-up-spin state and the all-down-spin state of chains, and showed that Γ/J_F should be larger for a longer chain to achieve a high performance of QA. In this study, Γ was tuned instead of J_F. Interestingly, their result implied the opposite as ours using SA. This is a future problem to uncover the origin of the difference between QA and SA.

APPENDIX A
CORRELATION LENGTH IN A ONE-DIMENSIONAL ISING MODEL

The appendix provides a detailed derivation of the correlation length in a one-dimensional Ising model [see eq. (19) in the main text]. The correlation length is determined by the correlation function, which is given by

$$
\begin{aligned}
C_i(j) &= \langle s_{i,1} s_{i,j+1} \rangle_T, \\
&= \sum_{\{s_{i,l}\}} s_{i,1} s_{i,j+1} P_{\mathrm{eq}}(\{s_{i,l}, T\}),
\end{aligned}
$$

$$
= \sum_{\{s_{i,l}\}} s_{i,1} s_{i,j+1} \exp\left(K(i) \sum_{k=1}^{L(i)} s_{i,k} s_{i,k+1} \right) / Z. \quad (31)
$$

We use the Hamiltonian shown in eq. (18) and set $K(i) = J_F(i)/T$. Here, Z is called the partition function in statistical mechanics, which is given by

$$
Z = \sum_{\{s_{i,l}\}} \exp\left(K(i) \sum_{k=1}^{L(i)} s_{i,k} s_{i,k+1} \right). \quad (32)
$$

The transfer matrix method is a powerful tool in statistical mechanics. We applied it to calculate the partition function Z and the numerator on the right-hand side of eq. (31). First, we introduce

$$
T(s_{i,j}, s_{i,j+1}) = \exp\left(K(i) s_{i,j} s_{i,j+1} \right). \quad (33)
$$

Then the partition function is given by

$$
Z = \sum_{l=1}^{L(i)} \sum_{s_{i,l} \in \{+1,-1\}} T(s_{i,1}, s_{i,2}) T(s_{i,2}, s_{i,3}) \cdots
$$
$$
\cdots T(s_{i,L(i)-1}, s_{i,L(i)}) T(s_{i,L(i)}, s_{i,1}). \quad (34)
$$

Here, it is convenient to regard $T(s_{i,j}, s_{i,j+1})$ as a matrix element of T such as

$$
T = \begin{pmatrix} T(1,1) & T(1,-1) \\ T(-1,1) & T(-1,-1) \end{pmatrix} = \begin{pmatrix} e^{K(i)} & e^{-K(i)} \\ e^{-K(i)} & e^{K(i)} \end{pmatrix}. \quad (35)
$$

The matrix T is called the transfer matrix. Then the partition function is written as

$$
Z = \mathrm{Tr}\left(T^{L(i)} \right) = \lambda_+^{L(i)} + \lambda_-^{L(i)}, \quad (36)
$$

where $\lambda_\pm$ are the two eigenvalues of T [i.e., $\lambda_+ = 2\cosh K(i)$ and $\lambda_- = 2\sinh K(i)$]. Similarly, the numerator on the right-hand side of eq. (31) is evaluated as

$$
\begin{aligned}
\sum_{\{s_{i,l}\}} & s_{i,1} s_{i,j+1} \exp\left(K(i) \sum_{k=1}^{L(i)} s_{i,k} s_{i,k+1} \right) \\
&= \sum_{l=1}^{L(i)} \sum_{s_{i,l} \in \{+1,-1\}} s_{i,1} s_{i,j+1} \\
&\quad \times T(s_{i,1}, s_{i,2}) T(s_{i,2}, s_{i,3}) \cdots T(s_{i,L(i)}, s_{i,1}), \\
&= \mathrm{Tr}\left(\sigma T^j \sigma T^{L(i)-j} \right), \\
&= \lambda_+^j \lambda_-^{L(i)-j} + \lambda_+^{L(i)-j} \lambda_-^j, \quad (37)
\end{aligned}
$$

where

$$
\sigma = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (38)
$$

The correlation function is then given by

$$
\begin{aligned}
C_i(j) &= \frac{\lambda_+^j \lambda_-^{L(i)-j} + \lambda_+^{L(i)-j} \lambda_-^j}{\lambda_+^{L(i)} + \lambda_-^{L(i)}}, \\
&= \tanh^j K(i) \left(\frac{1 + \tanh^{L(i)-2j} K(i)}{1 + \tanh^{L(i)} K(i)} \right). \quad (39)
\end{aligned}
$$

For large $L(i)$, the parenthesis in eq. (39) can be approximated by one. Then

$$
C_i(j) = \tanh^j K(i). \quad (40)
$$

By combining this expression with

$$
C_i(j) = \exp\left(-\frac{j}{\xi_i(T)} \right), \quad (41)
$$

we obtain eq. (19) in the main text.

APPENDIX B
## PERFORMANCE COMPARISON OF MINOR-EMBEDDING FOR A DIFFERENT ANNEALING SCHEDULE

This appendix validates the proposed ME (ME *iii*) when a different annealing schedule in SA is used. Here, the temperature in the SA algorithm [see Algorithm 1] is taken as

$$
T_{k+1} = (1 - r) T_k, \quad (42)
$$

where T_k is the temperature at the k-th MCS. We set the total number of the MCSs as $n_{\mathrm{MCS}} = 10^4$, the initial temperature $T_{\mathrm{ini}} = 10$, and the final temperature $T_{\mathrm{fin}} = 0.01$. The initial temperature and the final temperature are the default values in the CMOS annealing machine. The decay rate of the temperature is determined by $T_1 = T_{\mathrm{ini}}$ and $T_{n_{\mathrm{MCS}}} = T_{\mathrm{fin}}$ as $r \simeq 6.9 \times 10^{-4}$. In this annealing schedule, the temperature is
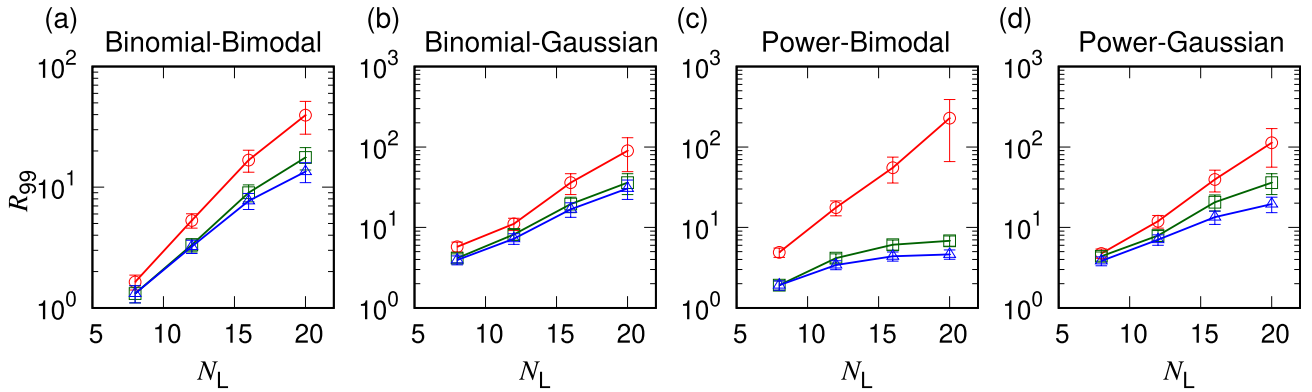
FIGURE 7. N_L-dependences of the STS for (a) the Binomial-Bimodal problem, (b) the Binomial-Gaussian problem, (c) the Power-Bimodal problem, and (d) the Power-Gaussian problem. ME *i*, ME *ii*, and ME *iii* are denoted by red circles, green squares, and blue triangles, respectively.
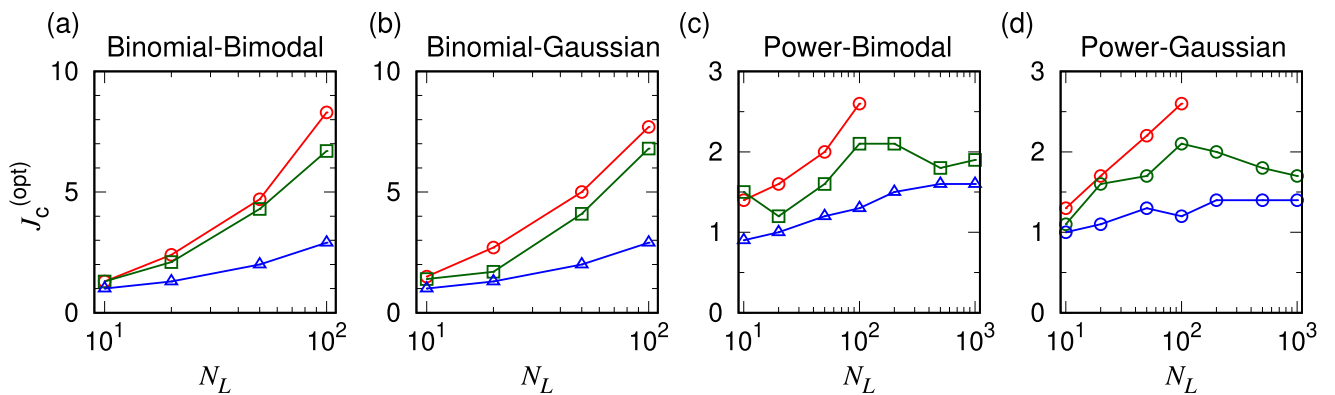
FIGURE 8. N_L-dependences of the optimal values of J_c for (a) the Binomial-Bimodal problem, (b) the Binomial-Gaussian problem, (c) the Power-Bimodal problem, and (d) the Power-Gaussian problem. ME *i*, ME *ii* and ME *iii* are denoted by red circles, green squares, and blue triangles, respectively.

lowered as an exponential function of MCSs. This is different from the annealing schedule in the main text, where the temperature linearly decays to zero.

We use the STS given by eq. (26) to compare the performances among ME *i*, ME *ii*, and ME *iii*. Figure 7 shows the N_L-dependences of the STS for each benchmarking problem. Similar to the main text, ME *iii* outperforms ME *i* and ME *ii* for all the benchmarking problems, implying that the results are independent of the annealing schedules.

APPENDIX C
N_L-DEPENDENCES OF THE OPTIMAL VALUES OF J_c

This appendix discusses the N_L-dependences of the optimal value J_c, which minimizes the energy density of the logical Ising model ϵ. Here, we denote the optimal value as $J_c^{(\text{opt})}$.

Figure 8 shows the N_L-dependences of $J_c^{(\text{opt})}$ for each benchmarking problem. In the Binomial-Bimodal problem and the Binomial-Gaussian problem [Figs. 8 (a) and (b)], $J_c^{(\text{opt})}$ increases with N_L. The increase of $J_c^{(\text{opt})}$ in ME *iii* is slow compared to those in ME *i* and ME *ii*. On the other hand, in the Power-Bimodal problem and the Power-Gaussian

problem [Figs. 8 (c) and (d)], the N_L-dependences of the optimal values are small compared to the Binomial-Bimodal problem and the Binomial-Gaussian problem. In ME *i*, $J_c^{(\text{opt})}$ increases with N_L. In ME *ii*, $J_c^{(\text{opt})}$ fluctuates between $J_c^{(\text{opt})} = 1$ and $J_c^{(\text{opt})} = 2$. In ME *iii*, $J_c^{(\text{opt})}$ gradually increases with N_L, but the increase is slower than that of the ME *i*.

REFERENCES
[1] M. W. Johnson *et al.*, "Quantum annealing with manufactured spins," *Nature*, vol. 473, no. 7346, pp. 194–198, May 2011.

[2] P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz, and J. Whittaker, "Architectural considerations in the design of a superconducting quantum annealing processor," *IEEE Trans. Appl. Supercond.*, vol. 24, no. 4, pp. 1–10, Aug. 2014.

[3] M. Yamaoka, C. Yoshimura, M. Hayashi, T. Okuyama, H. Aoki, and H. Mizuno, "A 20k-spin ising chip to solve combinatorial optimization problems with CMOS annealing," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 303–309, Jan. 2016.

[4] T. Inagaki, Y. Haribara, K. Igarashi, T. Sonobe, S. Tamate, T. Honjo, A. Marandi, P. L. Mcmahon, T. Umeki, K. Enbutsu, O. Tadanaga, H. Takenouchi, K. Aihara, K.-I. Kawarabayashi, K. Inoue, S. Utsunomiya, and H. Takesue, "A coherent ising machine for 2000-node optimization problems," *Science*, vol. 354, no. 6312, pp. 603–606, Nov. 2016.

[5] P. L. Mcmahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, R. L. Byer, M. M. Fejer, H. Mabuchi, and Y. Yamamoto, "A fully programmable 100-spin coherent ising machine with all-to-all connections," *Science*, vol. 354, no. 6312, pp. 614–617, Nov. 2016.

[6] M. Maezawa, K. Imafuku, M. Hidaka, H. Koike, and S. Kawabata, "Design of quantum annealing machine for prime factoring," in *Proc. 16th Int. Superconductive Electron. Conf. (ISEC)*, Jun. 2017, pp. 1–3.

[7] T. Okuyama, M. Hayashi, and M. Yamaoka, "An ising computer based on simulated quantum annealing by path integral Monte Carlo method," in *Proc. IEEE Int. Conf. Rebooting Comput. (ICRC)*, Nov. 2017, pp. 1–6.

[8] C. Yoshimura, M. Hayashi, T. Okuyama, and M. Yamaoka, "Implementation and evaluation of FPGA-based annealing processor for ising model by use of resource sharing," *Int. J. Netw. Comput.*, vol. 7, no. 2, pp. 154–172, 2017.

[9] M. Aramon, G. Rosenberg, E. Valiante, T. Miyazawa, H. Tamura, and H. G. Katzgraber, "Physics-inspired optimization for quadratic unconstrained problems using a digital annealer," *Frontiers Phys.*, vol. 7, p. 48, Apr. 2019.

[10] H. Goto, K. Tatsumura, and A. R. Dixon, "Combinatorial optimization by simulating adiabatic bifurcations in nonlinear Hamiltonian systems," *Sci. Adv.*, vol. 5, no. 4, Apr. 2019, Art. no. eaav2372.

[11] M. Maezawa, G. Fujii, M. Hidaka, K. Imafuku, K. Kikuchi, H. Koike, K. Makise, S. Nagasawa, H. Nakagawa, M. Ukibe, and S. Kawabata, "Toward practical-scale quantum annealing machine for prime factoring," *J. Phys. Soc. Jpn.*, vol. 88, no. 6, 2019, Art. no. 061012.

[12] T. Okuyama, T. Sonobe, K.-I. Kawarabayashi, and M. Yamaoka, "Binary optimization by momentum annealing," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 100, no. 1, Jul. 2019, Art. no. 012111, doi: 10.1103/PhysRevE.100.012111.

[13] H. Nishimori and G. Ortiz, *Elements of Phase Transitions and Critical Phenomena*. Oxford, U.K.: Oxford Univ. Press, 2010.

[14] M. Nakahara, *Lectures on Quantum Computing, Thermodynamics and Statistical Physics*, vol. 8. Singapore: World Scientific, 2013

[15] A. Lucas, "Ising formulations of many NP problems," *Frontiers Phys.*, vol. 2, p. 5, 2014.

[16] S. Tanaka, R. Tamura, and B. K. Chakrabarti, *Quantum Spin Glasses, Annealing and Computation*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[17] K. Tanahashi, S. Takayanagi, T. Motohashi, and S. Tanaka, "Application of ising machines and a software development for ising machines," *J. Phys. Soc. Jpn.*, vol. 88, no. 6, Jun. 2019, Art. no. 061010.

[18] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready, "Training a binary classifier with the quantum adiabatic algorithm," 2008, *arXiv:0811.0416*. [Online]. Available: http://arxiv.org/abs/0811.0416

[19] G. Rosenberg, P. Haghnegahdar, P. Goddard, P. Carr, K. Wu, and M. L. de Prado, "Solving the optimal trading trajectory problem using a quantum annealer," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 6, pp. 1053–1060, Sep. 2016.

[20] F. Neukart, G. Compostella, C. Seidel, D. von Dollen, S. Yarkoni, and B. Parney, "Traffic flow optimization using a quantum annealer," *Frontiers ICT*, vol. 4, p. 29, Dec. 2017.

[21] S. Kanamaru, D. Oku, M. Tawada, S. Tanaka, M. Hayashi, M. Yamaoka, M. Yanagisawa, and N. Togawa, "Efficient ising model mapping to solving slot placement problem," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*, Jan. 2019, pp. 1–6.

[22] K. Terada, D. Oku, S. Kanamaru, S. Tanaka, M. Hayashi, M. Yamaoka, M. Yanagisawa, and N. Togawa, "An ising model mapping to solve rectangle packing problem," in *Proc. Int. Symp. VLSI Design, Autom. Test (VLSI-DAT)*, Apr. 2018, pp. 1–4.

[23] K. Kitai, J. Guo, S. Ju, S. Tanaka, K. Tsuda, J. Shiomi, and R. Tamura, "Designing metamaterials with quantum annealing and factorization machines," *Phys. Rev. Res.*, vol. 2, no. 1, Mar. 2020, Art. no. 013319.

[24] V. Choi, "Minor-embedding in adiabatic quantum computation: I. the parameter setting problem," *Quantum Inf. Process.*, vol. 7, no. 5, pp. 193–209, Oct. 2008.

[25] V. Choi, "Minor-embedding in adiabatic quantum computation: II. minor-universal graph design," *Quantum Inf. Process.*, vol. 10, no. 3, pp. 343–353, Jun. 2011.

[26] C. Klymko, B. D. Sullivan, and T. S. Humble, "Adiabatic quantum programming: Minor embedding with hard faults," *Quantum Inf. Process.*, vol. 13, no. 3, pp. 709–729, Mar. 2014, doi: 10.1007/s11128-013-0683-9.

[27] T. Boothby, A. D. King, and A. Roy, "Fast clique minor generation in chimera qubit connectivity graphs," *Quantum Inf. Process.*, vol. 15, no. 1, pp. 495–508, Jan. 2016.

[28] D. Oku, K. Terada, M. Hayashi, M. Yamaoka, S. Tanaka, and N. Togawa, "A fully-connected ising model embedding method and its evaluation for CMOS annealing machines," *IEICE Trans. Inf. Syst.*, vol. 102, no. 9, pp. 1696–1706, 2019.

[29] J. Cai, W. G. Macready, and A. Roy, "A practical heuristic for finding graph minors," 2014, *arXiv:1406.2741*. [Online]. Available: http://arxiv.org/abs/1406.2741

[30] A. Perdomo-Ortiz, J. Fluegemann, R. Biswas, and V. N. Smelyanskiy, "A performance estimator for quantum annealers: Gauge selection and parameter setting," 2015, *arXiv:1503.01083*. [Online]. Available: http://arxiv.org/abs/1503.01083

[31] K. E. Hamilton and T. S. Humble, "Identifying the minor set cover of dense connected bipartite graphs via random matching edge sets," *Quantum Inf. Process.*, vol. 16, no. 4, p. 94, Apr. 2017.

[32] A. Zaribafiyan, D. J. J. Marchand, and S. S. Changiz Rezaei, "Systematic and deterministic graph minor embedding for Cartesian products of graphs," *Quantum Inf. Process.*, vol. 16, no. 5, p. 136, May 2017.

[33] R. Harris et al., "Phase transitions in a programmable quantum spin glass simulator," *Science*, vol. 361, no. 6398, pp. 162–165, 2018.

[34] A. D. King et al., "Observation of topological phenomena in a programmable lattice of 1,800 qubits," *Nature*, vol. 560, no. 7719, pp. 456–460, Aug. 2018.

[35] Y. Sugie, Y. Yoshida, N. Mertig, T. Takemoto, H. Teramoto, A. Nakamura, I. Takigawa, S.-I. Minato, M. Yamaoka, and T. Komatsuzaki, "Graph minors from simulated annealing for annealing machines with sparse connectivity," in *Theory and Practice of Natural Computing*, D. Fagan, C. Martín-Vide, M. O'Neill, and M. A. Vega-Rodríguez, Eds. Cham, Switzerland: Springer, 2018, pp. 111–123.

[36] S. Okada, M. Ohzeki, M. Terabe, and S. Taguchi, "Improving solutions by embedding larger subproblems in a D-Wave quantum annealer," *Sci. Rep.*, vol. 9, no. 1, p. 2098, Dec. 2019.

[37] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, 1983.

[38] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon, "Optimization by simulated annealing: An experimental evaluation; Part I, graph partitioning," *Oper. Res.*, vol. 37, no. 6, pp. 865–892, Dec. 1989.

[39] D. S. Johnson, C. R. Aragon, L. A. McGeoch, and C. Schevon, "Optimization by simulated annealing: An experimental evaluation; Part II, graph coloring and number partitioning," *Oper. Res.*, vol. 39, no. 3, pp. 378–406, Jun. 1991.

[40] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[41] N. Dattani, S. Szalay, and N. Chancellor, "Pegasus: The second connectivity graph for large-scale quantum annealing hardware," 2019, *arXiv:1901.07636*. [Online]. Available: http://arxiv.org/abs/1901.07636

[42] R. Albert and A. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.

[43] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, "Evidence for quantum annealing with more than one hundred qubits," *Nature Phys.*, vol. 10, no. 3, pp. 218–224, Mar. 2014, doi: 10.1038/nphys2900.

[44] T. F. Ronnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, "Defining and detecting quantum speedup," *Science*, vol. 345, no. 6195, pp. 420–424, Jul. 2014.

[45] M. Ohzeki, A. Miki, M. J. Miyama, and M. Terabe, "Control of automated guided vehicles without collision by quantum annealer and digital devices," *Frontiers Comput. Sci.*, vol. 1, p. 9, Nov. 2019.

[46] T. Kadowaki and H. Nishimori, "Quantum annealing in the transverse ising model," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 58, no. 5, pp. 5355–5363, Nov. 1998.

[47] R. Hamerly *et al.*, "Experimental investigation of performance differences between coherent ising machines and a quantum annealer," *Sci. Adv.*, vol. 5, no. 5, May 2019, Art. no. eaau0823.

[48] E. Andriyash, Z. Bian, F. A. Chudak, A. D. King, and W. G. Macready, "Boosting integer factoring performance via quantum annealing offsets," D-Wave, Burnaby, AB, Canada, Tech. Rep. 14-1002A-B, 2016.

SHU TANAKA received the B.Sci. degree from the Tokyo Institute of Technology, in 2003, and the M.Sci. and Dr.Sci. degrees from The University of Tokyo, in 2005 and 2008, respectively. He is currently an Associate Professor with the Department of Applied Physics and Physico-Informatics, Keio University. His research interests include quantum annealing, Ising machine, statistical mechanics, and materials science. He is a member of JPS.

TATSUHIKO SHIRAI received the B.Sci., M.Sci., and Dr.Sci. degrees from The University of Tokyo, in 2011, 2013, and 2016, respectively. He is currently an Assistant Professor with the Department of Computer Science and Communications Engineering, Waseda University. His research interests include quantum dynamics and statistical mechanics. He is a member of JPS.

NOZOMU TOGAWA (Member, IEEE) received the B.Eng., M.Eng., and Dr.Eng. degrees in electrical engineering from Waseda University, in 1992, 1994, and 1997, respectively. He is currently a Professor with the Department of Computer Science and Communications Engineering, Waseda University. His research interests include integrated system design, graph theory, information security, and quantum computing. He is a member of IEICE and IPSJ.

● ● ●