Transforming generalized Ising models into Boltzmann machines

Nobuyuki Yoshioka,[*] Yutaka Akagi, and Hosho Katsura
Department of Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

We find an exact mapping from the generalized Ising models with many-spin interactions to equivalent Boltzmann machines, i.e., the models with only two-spin interactions between physical and auxiliary binary variables accompanied by local external fields. More precisely, the appropriate combination of the algebraic transformations, namely the star-triangle and decoration-iteration transformations, allows one to express the model in terms of fewer-spin interactions at the expense of the degrees of freedom. Furthermore, the benefit of the mapping in Monte Carlo simulations is discussed. In particular, we demonstrate that the application of the method in conjunction with the Swendsen-Wang algorithm drastically reduces the critical slowing down in a model with two- and three-spin interactions on the Kagomé lattice.

I. INTRODUCTION

Inspired by the tremendous success in machine learning fields such as the image or speech recognition, climate prediction, and market analysis, the increasing number of studies have shown the validity of artificial neural network in physical problems. For instance, physicists have recognized that the phase classification can be done in a parallel fashion to the ordinary machine learning tasks by using the data of the classical, quantum, or auxiliary degrees of freedom [1–12]. Another important direction focuses on the representability itself. The applications such as the variational wave functions describing ground states turn out to outperform the state-of-the-art methods [13–17], and are also valid in estimating the Boltzmann factor of a system to accelerate the classical Monte Carlo (MC) simulations [3,18]. Furthermore, numerous studies reveal the ability to precisely express nonlinear functions such as the wave functions of stabilizer states and chiral topological states [19–24]. This is not only insightful from the perspective of exploring soluble models, but also fruitful in constructing new numerical techniques including optimization of variational functions, MC simulations, and so on.

Since the invention of the MC method, physicists have long made efforts to develop versatile and efficient simulation methods to investigate statistical models. In classical lattice systems, the single-spin flip (SSF) algorithm is undoubtedly one of the most widely used techniques as it is model independent. The locality of the variables involved in a single update procedure, however, inevitably leads to a severe slowing down near critical points or at low temperature for nonordering systems. One of the solutions is to apply the global updates such as the cluster algorithms [25,26], worm algorithm [27], and loop algorithm [28], but they are mostly restricted to two-body interacting systems. Developing a generic technique applicable to a wide variety of systems involving many-body interactions is highly challenging.

In this work, we establish a mapping from the generalized Ising model to the Boltzmann machine (BM). The former, which includes many-spin interactions, is used to describe magnetic and thermodynamic properties of solids, the effective model of alloys, spin glass models, and so on [29–31]. The BM, on the other hand, is an expression of probability distribution by two-spin interactions between physical (visible) and auxiliary (hidden) binary degrees of freedom. The mapping procedure goes as follows: one first decomposes a p-spin Ising interaction into a sum of three- and $(p-1)$-spin interactions by adding an auxiliary spin. This is repeated for the latter many-spin interaction until the total interaction is expressed in terms of two- and three-spin interactions. Finally, each three-spin interaction is transformed into two-spin interactions and single-spin terms. The idea of such consecutive decomposition can also be seen in the context of quantum annealing [32–35].

After constructing the rigorous mapping, we take advantage of it by presenting a novel global update scheme; the application of the Swendsen-Wang (SW) algorithm to the exactly surrogate BM. By comparing the autocorrelation times of the visible spin configurations, we demonstrate in a model with two- and three-spin interactions on the Kagomé lattice that the sampling efficiency of the cluster algorithm on the BM at the critical temperature is drastically improved compared to that of SSF performed on the original Hamiltonian. While the mapping introduced in this paper is applicable to the generalized Ising model in any number of dimensions, our results are presented in the above-mentioned model for clarity and simplicity.

The remainder of this paper is organized as follows. In Sec. II we first introduce the most primary transformation techniques, namely the decoration-iteration and star-triangle transformations. Embedding the hidden spins by combining the two, an arbitrary many-spin interaction is shown to be mappable to the BM. Furthermore, multiple interaction terms can be taken into account by simply considering the embedding procedures independently. In Sec. III, we see that the transformation is numerically beneficial since the existing

---

*nysocloud@g.ecc.u-tokyo.ac.jp

cluster update can be applied to the BM. Finally, the summary for the current work and discussion concerning the future direction are given in Sec. IV. For the completeness of the paper, in Appendix A we derive the explicit expression of the BM mapped from the generalized Ising model. In Appendix B, we qualitatively see that the block Gibbs sampling for the RBM equivalent to the ferromagnetic Ising model is highly inefficient. The partition function of the pure three-spin interacting model on the Kagomé lattice, from which the absence of the phase transition follows, is calculated in Appendix C. Also, two schemes to perform the cluster update under magnetic field are given in Appendix D, and the observation of the physical quantity in the extended space is discussed in Appendix E.

II. ALGEBRAIC TRANSFORMATION OF BOLTZMANN FACTORS

In this section, we find the equivalence of the generalized Ising models and the BMs with specific architectures, i.e., the restricted Boltzmann machine (RBM) and the deep Boltzmann machine (DBM). [See Fig. 1 for the graphic representation.] First, we state the definition of the RBM and DBM, which give some probability distributions of visible spins with auxiliary binary degrees of freedom, or the hidden spins. The Boltzmann factor of an RBM is given as [36,37]

$$\pi(\boldsymbol{\sigma}) = \sum_{\boldsymbol{h}} \widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}), \tag{1}$$

$$\widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}) = \exp\left\{ \sum_{i,j} W_{ij}\sigma_i h_j + \sum_i a_i \sigma_i + \sum_j b_j h_j \right\}, \tag{2}$$

where $\pi(\boldsymbol{\sigma})$ and $\widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h})$ are the Boltzmann factor of N_v visible spins and the RBM with additionally N_h hidden spins, respectively. Here, σ_i and h_j are the ith visible and jth hidden spins that take either $+1$ or -1, coupled via the interaction W_{ij}. The local external field is given as a_i and b_j for visible and hidden spins, respectively. We denote the visible and hidden spin configurations by $\boldsymbol{\sigma} := (\sigma_1, \sigma_2, \ldots, \sigma_{N_v}) \in \mathcal{S}$ and
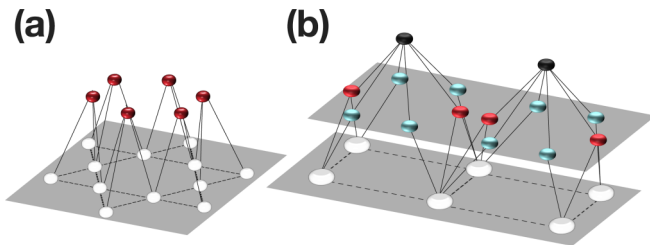

FIG. 1. Schematic picture of (a) restricted Boltzmann machine (RBM) that is equivalent to a model with three-spin Ising-type interaction, (b) deep Boltzmann machine (DBM) that is equivalent to a model with four-spin Ising-type interaction. The white and black objects denote visible and deep spins, respectively. Also the blue and red objects are the hidden spins introduced by the decoration-iteration transformation (DIT) and star-triangle transformation (STT), which are given by Eqs. (5) and (7). The presence of the layers is denoted by the gray planes. Note that the mapping to the RBM or DBM is applicable irrespective of the spatial dimension.

$\boldsymbol{h} := (h_1, h_2, \ldots, h_{N_h}) \in \mathcal{H}$, where $\mathcal{S} = \{-1, 1\}^{N_v}$ and $\mathcal{H} = \{-1, 1\}^{N_h}$ are the sets of all possible binary spin configurations for visible and hidden spins, respectively. Also, to discriminate between the spaces with and without the hidden spins, we call $\mathcal{S}$ as the "original space" and $\mathcal{S} \cup \mathcal{H}$ as the "extended space." Note that the absence of the intralayer couplings is reflected in the bipartite structure as is depicted in Fig. 1(a). The BM without any restriction on the connectivity is not considered in the following.

A crucial notion in introducing the DBM is the "layer," for which we provide a description based on the graphical understanding of Fig. 1. As aforementioned and also obvious from Fig. 1(a), no coupling is present between the visible or hidden spins in the RBM. Such a bipartite structure is understood as a BM with the "visible layer and one hidden layer." Namely, the visible layer contains all the visible spins, and the hidden layer consists of hidden spins that are interacting with one or more visible spins. We may analogously construct another hidden layer on top as depicted in Fig. 1(b), which is a BM with the "visible layer and two hidden layers" in turn. In principle, one may add as many hidden layers as desired. Such an architecture is generally referred to as the DBM. The maximum number of layers obtained by mapping from the generalized Ising model is two as we discuss in Sec. II B 3, and therefore we discriminate the first and second hidden layers as the "hidden layer" and "deep layer" in the following. The spins included in these layers are correspondingly referred to as the "hidden spins" and "deep spins." To be concrete, the Boltzmann factor of a DBM is given as

$$\pi(\boldsymbol{\sigma}) = \sum_{\boldsymbol{h},\boldsymbol{d}} \widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{d}), \tag{3}$$

$$\widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{d}) = \exp\left\{ \sum_{i,j} W_{ij}\sigma_i h_j + \sum_{j,k} W'_{jk} h_j d_k \right.$$
$$\left. + \sum_i a_i \sigma_i + \sum_j b_j h_j + \sum_k b'_k d_k \right\}, \tag{4}$$

where $\widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}, \boldsymbol{d})$ is the Boltzmann factor for each spin configuration, d_k is the kth deep spin with the local field b'_k, and W'_{jk} is the interaction between the jth hidden spin and the kth deep spin. As is the case for the visible and hidden spins, a configuration of the deep spins is denoted by $\boldsymbol{d} := (d_1, d_2, \ldots, d_{N_d}) \in \mathcal{D}$, where $\mathcal{D} = \{-1, 1\}^{N_d}$ is the set of all possible deep spin configurations. The union $\mathcal{S} \cup \mathcal{H} \cup \mathcal{D}$ is also referred to as the extended space in the following.

The probabilistic network given by quadratic terms as Eqs. (2) and (4) is known to be powerful to express nonlinear functions and in fact applied widely in the field of machine learning, condensed matter physics, quantum physics, and so on [13–15,19–23]. In the remainder of this section, we find the exact mapping from the generalized Ising model to the RBM or DBM.

A. Transformation techniques

We introduce two mapping techniques to embed hidden spins as is graphically described in Fig. 2: the decoration iteration transformation (DIT) and star-triangle transformation (STT)
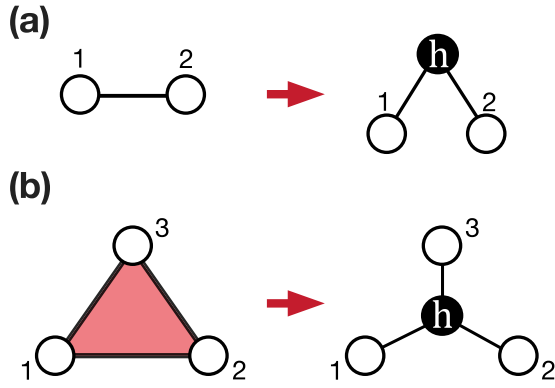
(a)

(b)

FIG. 2. Schematic description of the mapping techniques: (a) Decoration-iteration transformation and (b) star-triangle transformation. The white and black circles correspond the visible and hidden spins, respectively, and the numbers denote the labels of the visible spins. The solid lines and the filled region denote the two- and three-spin interactions, respectively.

(STT). Note that the newly embedded spins are auxiliary and the original interactions are realized by tracing out such degrees of freedom.

The DIT, depicted in Fig. 2(a), is a very simple transformation which embeds a hidden spin h between two interacting visible spins as follows [38,39]:

$$e^{J\sigma_1\sigma_2} = \Delta \sum_{h=\pm 1} \exp\{W[\sigma_1 + \text{sgn}(J)\sigma_2]h\}, \tag{5}$$

where Δ is the normalization factor. The new interaction W is given as

$$W = \text{arc cosh}(e^{2|J|})/2. \tag{6}$$

Since the DIT can be carried out for any J, an arbitrary Ising model with two-spin interactions including random spin-glass, frustrated system, and fully connected models can be mapped into an equivalent RBM. The application of such a transformation technique allows one to obtain the exact solution for a model on, for instance, two-spin interacting Ising model on a bond-decorated lattice that can be transformed into the soluble model on an undecorated lattice [38,40–44].

The other technique, known as the STT, embeds a hidden spin h into three visible spins with both two- and three-spin interactions as is illustrated in Fig. 2(b) [39,45–48]. Expressed in the form of the Boltzmann weight, this can be written as [49]

$$\exp[M\sigma_1\sigma_2\sigma_3 + J_1\sigma_2\sigma_3 + J_2\sigma_3\sigma_1 + J_3\sigma_1\sigma_2]$$
$$= \Delta \sum_{h=\pm 1} \exp\left[\sum_{i=1}^{3}(W_i h + a_i)\sigma_i + bh\right], \tag{7}$$

where M and J_i are the amplitudes of three- and two-spin interactions, respectively. The interaction between the visible and the hidden spins in the extended space is denoted by W_i, and the local fields are denoted by a_i and b, respectively. It

can be shown that Eq. (7) amounts to eight nonequivalent conditions which yields the solutions as

$$\exp(4\chi_i a_i) = \frac{\sinh[2(|J_i| + M)]}{\sinh[2(|J_i| - M)]}, \tag{8}$$

$$\cosh(2W_i)$$
$$= \frac{e^{2|J_i|}\cosh[2(|J_j|+|J_k|)] - e^{-2|J_i|}\cosh[2(|J_j|-|J_k|)]}{[2\cosh(4|J_i|) - 2\cosh(4M)]^{1/2}}, \tag{9}$$

$$\sinh(2b)$$
$$= \frac{-\sinh(2\chi_i W_i)\sinh(4M)}{\{[\cosh(4J_j) - \cosh(4M)][\cosh(4J_k) - \cosh(4M)]\}^{1/2}}, \tag{10}$$

where $\chi_i = \text{sgn}(J_i)$ is the sign of the two-spin interaction. The subscripts in Eqs. (9) and (11), i.e., i, j, and k, must be chosen such that none of them are identical to each other. Importantly, the STT is valid under the following conditions [50]:

$$|M| < |J_i|, \quad (i = 1, 2, 3), \tag{11}$$

$$\text{sgn}(J_1 J_2 J_3) = 1. \tag{12}$$

The notion of the DIT and STT can be generalized to include many-spin interactions. A system in the original space with both four- and two-spin interactions, for instance, can be mapped to a model with three-spin interactions that involves a single hidden spin. This mapping, used to obtain the solution of the zero-field eight-vertex model, is referred to as the star-square transformation [39,51–53]. While the transformation from the extended space into the original space, known as "the star-polygon transformation," is achieved by tracing out the hidden spins and is, in general, tractable [43,54], its inverse mapping exists in very limited cases. We note in passing that the DIT and STT have been extended to models with local quantum degrees of freedom such as Heisenberg spins and itinerant electrons [55–57].

B. Generalized Ising model as Boltzmann machine

Here we show that the generalized Ising model, which consists of many-spin interactions, can be mapped to an equivalent model with two-spin interactions and local fields. First, let us consider some bare many-spin interaction within the model. The Boltzmann factor is given as

$$\pi_{|\mathcal{C}|}(\boldsymbol{\sigma}_{\mathcal{C}}; M) := \exp\left(M \prod_{j\in\mathcal{C}} \sigma_j\right), \tag{13}$$

where M is the amplitude of the interaction. Here the set of sites of the visible spins involved, dubbed as the "cell" in the following, is denoted as $\mathcal{C}$. Correspondingly, a visible spin configuration of the cell is denoted by $\boldsymbol{\sigma}_{\mathcal{C}} \in \mathcal{S}_{\mathcal{C}}$, where $\mathcal{S}_{\mathcal{C}}$ is the set of all possible configurations of the visible spins included in $\mathcal{C}$. The subscript of the Boltzmann factor $|\mathcal{C}|$ is the number of the visible spins included in the cell. Note that arbitrary hidden and deep spins can be uniquely labeled due to the construction procedure we introduce in the following.

It is straightforward to see that the Boltzmann weight of the original system can be written as

$$\pi(\boldsymbol{\sigma}) = \prod_{\mathcal{C}} \pi_{|\mathcal{C}|}(\boldsymbol{\sigma}_{\mathcal{C}}; M_{\mathcal{C}}) = \prod_{\mathcal{C}} \left(\sum_{\boldsymbol{h}_{\mathcal{C}}, \boldsymbol{d}_{\mathcal{C}}} \widetilde{\pi}_{\mathcal{C}}(\boldsymbol{\sigma}_{\mathcal{C}}, \boldsymbol{h}_{\mathcal{C}}, \boldsymbol{d}_{\mathcal{C}}) \right), \quad (14)$$

where $\widetilde{\pi}_{\mathcal{C}}(\boldsymbol{\sigma}_{\mathcal{C}}, \boldsymbol{h}_{\mathcal{C}}, \boldsymbol{d}_{\mathcal{C}})$ is the weight of the RBM obtained by transformation. Note that although a visible spin may be included in multiple cells, the auxiliary spins are uniquely allocated to their corresponding cells by construction. A single realization of hidden and deep spins labeled by $\mathcal{C}$ is denoted as $\boldsymbol{h}_{\mathcal{C}} \in \mathcal{H}_{\mathcal{C}}$ and $\boldsymbol{d}_{\mathcal{C}} \in \mathcal{D}_{\mathcal{C}}$, where $\mathcal{H}_{\mathcal{C}}$ and $\mathcal{D}_{\mathcal{C}}$ are the sets of all possible configurations of the hidden and deep spins included in $\mathcal{C}$, respectively.

As we see from Eq. (14), the Boltzmann factor of any model can be decomposed into products over cells and the independence of hidden spins holds. For simplicity, we focus on bare p-spin interactions in the following.

1. Three-spin interaction as RBM

First, we discuss the three-spin interaction, which is the simplest possible case. The STT cannot be applied straightforwardly to a bare three-spin interaction due to the conditions given in Eqs. (11) and (12). To avoid this problem, we introduce the "virtual two-spin interactions" that cancel each other out as follows:

$$\pi_3(\boldsymbol{\sigma}; M) = \exp(M\sigma_1\sigma_2\sigma_3)$$
$$= \exp\left[M\sigma_1\sigma_2\sigma_3 + J_1\sigma_2\sigma_3 + J_2\sigma_3\sigma_1 + J_3\sigma_1\sigma_2\right]$$
$$\times \exp\left[-(J_1\sigma_2\sigma_3 + J_2\sigma_3\sigma_1 + J_3\sigma_1\sigma_2)\right]. \quad (15)$$

As is depicted in Fig. 3(a), this can be mapped into the RBM by applying Eq. (7) to the first and subsequently Eq. (5) to the second factors. Note that the amplitudes of the virtual two-spin interaction, J_i, can be taken arbitrarily as long as Eqs. (11) and (12) are satisfied.

Next, let us consider two sets of interacting spins as denoted in Fig. 3(b). Although the naive application of Eq. (15) yields eight hidden spins, two of them on the shared edges can be eliminated by modifying the signs of virtual two-spin interactions. For instance, by considering the virtual two-spin interactions with a homogeneous amplitude, we obtain

$$\pi_3(\boldsymbol{\sigma}_{\mathcal{C}_1}; M)\pi_3(\boldsymbol{\sigma}_{\mathcal{C}_2}; M)$$
$$= \exp(M\sigma_1\sigma_2\sigma_3)\exp(M\sigma_2\sigma_3\sigma_4)$$
$$= \exp\left[M\sigma_1\sigma_2\sigma_3 + J\sigma_2\sigma_3 + J\sigma_3\sigma_1 + J\sigma_1\sigma_2\right]$$
$$\times \exp\left[M\sigma_2\sigma_3\sigma_4 + J\sigma_3\sigma_4 - J\sigma_4\sigma_2 - J\sigma_2\sigma_3\right]$$
$$\times \exp\left[-(J\sigma_3\sigma_1 + J\sigma_1\sigma_2 + J\sigma_3\sigma_4 - J\sigma_4\sigma_2)\right], \quad (16)$$

from which six hidden spins emerge.

Furthermore, a system with pure three-spin interaction on the triangular lattice, known to be exactly soluble and dubbed as the Baxter-Wu model [58], can also be mapped into the RBM merely without hidden spins generated by the DIT. In other words, we may choose the signs of the virtual two-spin interactions on the triangles carefully so that the sum at each edge would cancel out, resulting in a reduced number of the hidden spins. Still, there are an exponential number of ways to represent this model by tuning the amplitudes and the signs
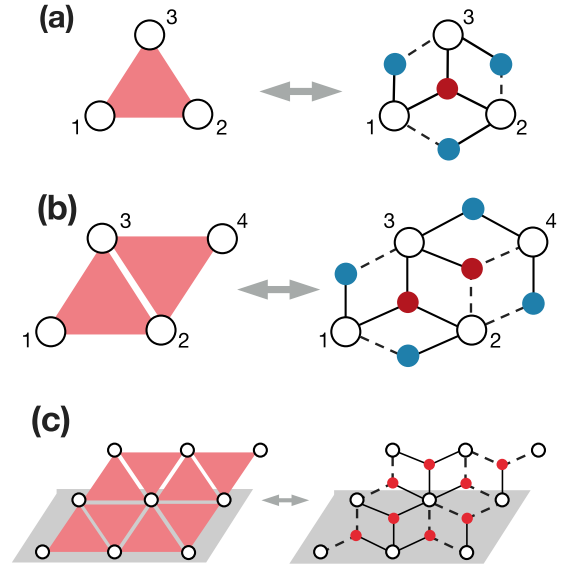

FIG. 3. (a) Transforming a pure three-spin interaction into an RBM. The red filled circles in the right-hand side are generated by the STT and the blue ones by the DIT. The black solid and dotted lines denote the positive and negative two-spin interactions, respectively. The amplitudes of the virtual two-spin interactions are taken as $J_i > 0$ in the figure. (b) Transforming a couple of three-spin interactions. The signs of the virtual two-spin interactions are modified so that the number of the hidden spins is reduced. (c) Transforming the Baxter-Wu model into an RBM.

of the virtual two-spin interactions. Shown in Fig. 3(c) is the mapping with 4-fold periodicity along the x axis.

2. Four-spin interaction as DBM

Next, we show that four-spin interaction can be expressed by introducing the second hidden layer, or the "deep" layer. The illustration of the two-step transformation is shown in Fig. 4(a). In the first step, we interpret the product of Ising
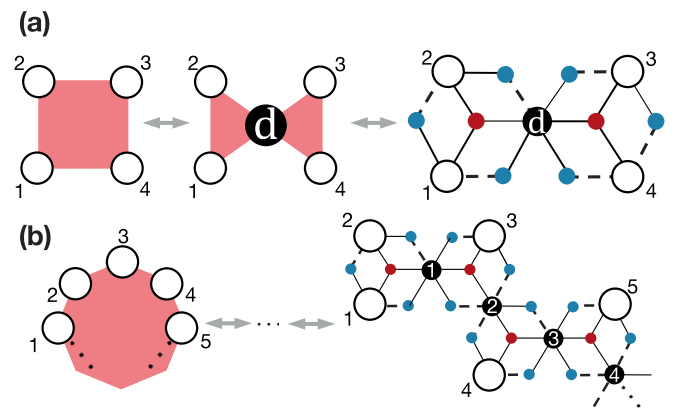

FIG. 4. (a) Transforming four-spin interaction into a DBM. The black filled circle is the spin in the deep layer and the other notation follows that of Fig. 3. (b) Transforming p-spin interaction into a DBM. The architecture of the BM is determined from the number of the spins involved in the interaction, but is irrelevant to the actual spatial dimension.

variable as a single new binary variable and apply the DIT as

$$\pi_4(\boldsymbol{\sigma}; M^{(0)}) = \exp[M^{(0)}\sigma_1\sigma_2\sigma_3\sigma_4]$$

$$= \Delta \sum_{d=\pm 1} \exp[M^{(1)}(\sigma_1\sigma_2 + \sigma_3\sigma_4)d]$$

$$= \Delta \sum_{d=\pm 1} \pi_3(\sigma_1, \sigma_2, d; M^{(1)})\pi_3(\sigma_3, \sigma_4, d; M^{(1)}),$$

(17)

where $\pi_4(\boldsymbol{\sigma}; M^{(0)})$ is the Boltzmann factor for four-spin interaction with the amplitude $M^{(0)}$. The interaction amplitudes $M^{(0)}$ and $M^{(1)}$ are related by Eq. (6) as

$$M^{(1)} = \text{arc } \cosh(e^{2|M^{(0)}|})/2. \quad (18)$$

Next, we use Eq. (15), the result for the three-spin interaction, under homogeneous virtual two-spin interactions $J > |M^{(1)}|$ for simplicity. Here, we obtain the expression as

$$\pi_4(\boldsymbol{\sigma}; M^{(0)}) = \Delta \sum_d \sum_{h_a, h_b} \sum_{h_1, \dots h_6} \widetilde{\pi}_2^{\text{STT}}(\boldsymbol{\sigma}, \boldsymbol{h}, d)$$

$$\times \widetilde{\pi}_2^{\text{DIT}}(\boldsymbol{\sigma}, \boldsymbol{h}, d)\widetilde{\pi}_1(\boldsymbol{\sigma}, \boldsymbol{h}, d), \quad (19)$$

where

$$\widetilde{\pi}_2^{\text{STT}}(\boldsymbol{\sigma}, \boldsymbol{h}, d) = \exp[W(\sigma_1 + \sigma_2 + d)h_a$$

$$+ W(\sigma_3 + \sigma_4 + d)h_b],$$

$$\widetilde{\pi}_2^{\text{DIT}}(\boldsymbol{\sigma}, \boldsymbol{h}, d) = \exp\left[\sum_{i=1}^4 W'(\sigma_i - d)h_i + W'(\sigma_1 - \sigma_2)h_5\right.$$

$$\left. + W'(\sigma_3 - \sigma_4)h_6\right],$$

$$\widetilde{\pi}_1(\boldsymbol{\sigma}, \boldsymbol{h}, d) = \exp\left[a\sum_i \sigma_i + a_d d + b(h_a + h_b)\right]. \quad (20)$$

Here, $\widetilde{\pi}_2^{\text{STT(DIT)}}(\boldsymbol{\sigma}, \boldsymbol{h}, d)$ is the Boltzmann factor for the two-spin interaction in the extended space obtained by the STT (DIT), and $\widetilde{\pi}_1(\boldsymbol{\sigma}, \boldsymbol{h}, d)$ is that for the local external fields. The interaction W (W') is the interlayer coupling introduced by the STT (DIT). The magnetic field for the visible, deep, and hidden spins are denoted by a, a_d, and b, respectively. The equivalence of amplitudes of the three-spin interactions $M^{(1)}$ leads to $a_d = 2a$. Also, the application of DIT to the virtual two-spin interaction yields $W' = \text{arc } \cosh(e^{2J})/2$. Note that the other parameters are obtained by substituting $M = M^{(1)}$ and homogeneous $J_i = J$ into Eq. (7) as

$$\exp(4a) = \frac{\sinh[2(J + M^{(1)})]}{\sinh[2(J - M^{(1)})]}, \quad (21)$$

$$\cosh(2W) = \frac{e^{2J}\cosh(4J) - e^{-2J}}{[2\cosh(4J) - 2\cosh(4M^{(1)})]^{1/2}}, \quad (22)$$

$$\sinh(2b) = \frac{-\sinh(2W)\sinh(4M^{(1)})}{|[\cosh(4J) - \cosh(4M^{(1)})]|}. \quad (23)$$

Another way to transform four-spin interaction is to apply the star-square transformation. Although it also requires a single deep spin, the architecture of the hidden spins would be symmetric and hence different from the aforementioned

transformation. To keep the number of auxiliary spins minimum, we will not use the star-square transformation in the following.

Note that N_h/N_d, or the ratio of the number of the hidden spins to that of the visible spins, may be reduced in a larger system as well as in Sec. II B 1; we may cancel out the virtual two-spin interactions by modifying their signs.

3. p-spin interaction as DBM

The discussion for four-spin interaction can be extended to p-spin interaction. Assume that Ising interactions up to $(p-1)$-spin terms can be mapped into a BM. The DIT splits the original model into two $(p/2+1)$-spin terms for even p and $[(p+1)/2+1]$- and $[(p-1)/2+1]$-spin terms for odd p. The mapping exists for $p = 3, 4$ as we showed in the previous sections, hence for arbitrary $p > 4$ as well. For further discussion and explicit expression in the form of BM, see Appendix A.

The number of the hidden and deep spins can be computed as well. We denote the numbers of hidden spins in the first layer generated by the STT, DIT, and that of the deep spins in the second hidden layer as n_h^{STT}, n_h^{DIT}, and n_d, which are shown in blue, red, and black filled circles in Fig. 4, respectively. It is straightforward to show that the following relations are satisfied:

$$n_h^{\text{STT}} = 3(p-2), \quad n_h^{\text{DIT}} = p - 2, \quad n_d = p - 3. \quad (24)$$

It is noteworthy from the perspective of the numerical cost that the number of hidden and deep spins increases only linearly with respect to p. The computational order remains to be the same for any update scheme.

III. MONTE CARLO SAMPLING ON BOLTZMANN MACHINE

In the following, we utilize the BM obtained by the transformation to classical MC sampling. The block Gibbs sampling, which is frequently used in the machine learning community, turns out to lack efficiency in terms of autocorrelation although the numerical cost per a single MC step is low [18,59]. We alternatively apply the Swendsen-Wang algorithm to the extended space and demonstrate the speed up compared to the SSF on the original space. First, we take the square-lattice Ising model for a simplified description of our scheme, and then proceed to show the results in the model with ferromagnetic two-spin interactions and alternating three-spin interactions on the Kagomé lattice, which is one of the the most comprehensible models that includes the many-spin interaction and also suffers from the slowing down at the critical temperature.

A. Ising model on square lattice

We consider a Hamiltonian with two-spin interactions on a square lattice

$$H = -\sum_{\langle i,j \rangle} \sigma_i\sigma_j, \quad (25)$$

which shows the ferroparamagnetic transition at $T_c = 2/\ln(\sqrt{2}+1)$ as is widely known in statistical physics [60].

The study by the SSF algorithm, i.e., the Glauber dynamics in a broader sense, suffers from severe critical slowing down [61,62]. The application of global updates, e.g., cluster updates [25,26], is one of the solutions under some circumstances.

Note that Eq. (25) itself can be regarded as the RBM since the square lattice is bipartite. Since the block Gibbs sampling method, which fix the configuration in a sublattice to enable independent sampling from another, is not beneficial as is qualitatively discussed in Appendix B, we alternatively apply the well-known Swendsen-Wang algorithm to take advantage of the expression. Although there is substantially no further gain for two-spin interacting model, in the following subsection, we demonstrate the speed up in terms of the autocorrelation time. As is shown in Sec. II, the number of the hidden and deep spins increases linearly with the system size, and therefore the computational order of a single MC step remains to be $O(N)$ in a model with short-range interactions, which is also the case in the current work.

B. Generalized Ising model on Kagomé lattice

Here, we consider a model with ferromagnetic two-spin interactions and also three-spin interactions on a Kagomé lattice. Let E be the set of edges and $\triangle(\triangledown)$ be an upward (downward) triangle on the lattice. The Hamiltonian is defined as

$$-\beta H(\boldsymbol{\sigma}) = \beta \sum_{\langle i,i'\rangle \in E} \sigma_i \sigma_{i'} + \sum_{\triangle} M_{\triangle} \tau_{\triangle} + \sum_{\triangledown} M_{\triangledown} \tau_{\triangledown}, \quad (26)$$

where $\tau_{\triangle} = \prod_{i \in \triangle} \sigma_i$ and $\tau_{\triangledown} = \prod_{i \in \triangledown} \sigma_i$ is the product of the spin variables, $M_{\triangle}/\beta$ ($M_{\triangledown}/\beta$) denotes the amplitude of the three-spin interactions for upward (downward) triangles, and $\langle i, i'\rangle$ is the edge connecting sites i and i'. The symmetry that combines the spin inversion and the mirror inversion is present when $M_{\triangle} + M_{\triangledown} = 0$, and the model exhibits second-order transition at finite temperature for finite $M_{\triangle}$ [see Appendix C for discussion when only three-spin interactions are present]. This can be understood from the correspondence between the present model and the antiferromagnetic Ising model with a uniform external field on the honeycomb lattice [63].

In the following, we assume $M_{\triangle} = -M_{\triangledown} = M > 0$. The Boltzmann weight is transformed as

$$\pi(\boldsymbol{\sigma}) = \exp\left[\beta \sum_{\langle i,i'\rangle \in E} \sigma_i \sigma_{i'} + \sum_{\triangle} M_{\triangle} \tau_{\triangle} + \sum_{\triangledown} M_{\triangledown} \tau_{\triangledown} \right]$$

$$(27)$$

$$= \Delta \sum_{\boldsymbol{h}} \exp\left[\sum_{\triangle} W_{\triangle} h_{\triangle} \sum_{i \in \triangle} \sigma_i + \sum_{\triangledown} W_{\triangledown} h_{\triangledown} \sum_{i \in \triangledown} \sigma_i \right.$$

$$\left. + b \sum_{\triangle,\triangledown} (h_{\triangle} + h_{\triangledown}) \right]$$

$$= \sum_{h} \widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h}), \quad (28)$$

where the parameters in the extended model are obtained by substituting $M_{\triangle(\triangledown)}$ and β into the STT, or Eq. (7). Note that
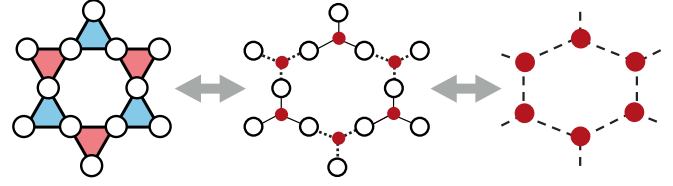

FIG. 5. Graphical understanding of the transformation for the spin model defined by Eq. (26) into RBM and hidden-spin-only model. The original space is defined on the Kagomé lattice, in which the triangles are colored to denote the signs of three-spin interactions. The white open and red filled circles correspond to the visible and hidden spins, respectively. Also the signs of the two-spin interactions are represented as solid and dotted lines for positive and negative, respectively.

the external fields on the visible spins is absent due to the cancellation caused by the alternating signs of the interactions. As is graphically described in Fig. 5, one hidden spin is embedded per triangle and denoted as $h_{\triangle}$ or $h_{\triangledown}$. The explicit expressions for the parameters can be read off from

$$\cosh(2W_{\triangle(\triangledown)}) = \frac{e^{2\beta} \cosh(4\beta) - e^{-2\beta}}{[2\cosh(4\beta) - 2\cosh(4M)]^{1/2}}, \quad (29)$$

$$\sinh(2b) = \frac{-\sinh(2W)\sinh(4M)}{|\cosh(4\beta) - \cosh(4M)|}, \quad (30)$$

where $W_{\triangle} = -W_{\triangledown} = W > 0$ and the signs of the parameters satisfy $\mathrm{sgn}(W_{\triangle(\triangledown)}) = \mathrm{sgn}(M_{\triangle(\triangledown)})$. Owing to the cancellation of the magnetic field on visible spins, we obtain a simple expression by tracing out the visible spins as

$$\widetilde{\widetilde{\pi}}(\boldsymbol{h}) = \sum_{\boldsymbol{\sigma}} \widetilde{\pi}(\boldsymbol{\sigma}, \boldsymbol{h})$$

$$= \Delta' \exp\left(-W_h \sum_{\langle j,j'\rangle \in \overline{E}} h_j h_{j'} + b \sum_{j \in \overline{V}} h_j \right), \quad (31)$$

where Δ' is another renormalization factor, $\overline{E}$ and $\overline{V}$ are the sets of edges and vertices in the honeycomb lattice as is shown in the rightmost panel of Fig. 5. Note that the interaction is antiferromagnetic, reflecting the alternating signs of the three-spin interactions in the original space. The amplitude of the two-spin interaction is obtained by the DIT as

$$W_h = \mathrm{arc}\cosh(e^{2W})/2. \quad (32)$$

Although the model defined by Eq. (31) is not soluble at $b \neq 0$, an approximate solution of the transition point can be obtained by imposing some assumption after mapping the original model to an eight-vertex model [63]. This turns out to be in a fairly good precision, but not exact, and hence we determine the transition point from the finite size scaling of the Binder ratio [64,65].

As was introduced by Binder, the renormalization group theory leads us to assume the scaling law for the Binder ratio

$$g := \frac{1}{2}\left(3 - \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2} \right), \quad (33)$$
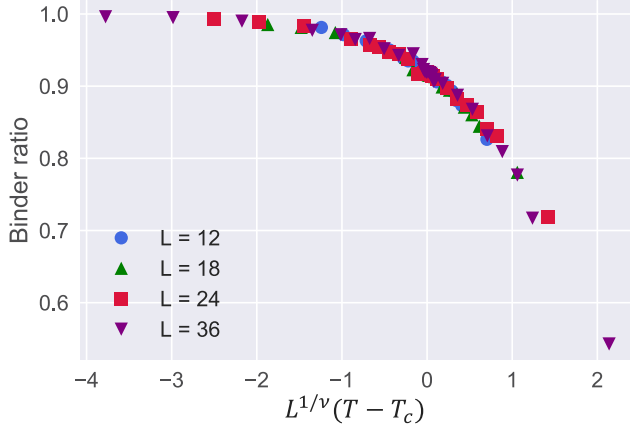
FIG. 6. Data collapse of the Binder ratio for the Ising model with two-spin and alternating-sign three-spin interaction on the Kagomé lattice. The critical temperature and the critical exponent for the correlation length are obtained as $T_c \sim 2.141$ and $\nu \sim 0.99$, respectively. The magnitude of two-spin interactions is set to unity while that of three-spin interactions is $M/\beta_c = 0.1$. The blue circles, green upward triangles, red squares, and purple downward triangles denote the data for the linear system sizes $L = 12, 18, 24$, and 36, respectively.

where m is the magnetization per site and $\langle \cdots \rangle$ denotes the thermal average. The scaling of this quantity in the vicinity of the critical temperature T_c is given as follows:

$$g \sim F(L^{1/\nu}(T - T_c)), \quad (34)$$

where L is the linear system size, ν is the critical exponent for the correlation length, and F is an appropriate polynomial function, which has been taken to be cubic in this work. Shown in Fig. 6 are the results for ferromagnetic two-spin and alternating-sign three-spin interacting model on the Kagomé lattice with $M/\beta = 0.1$. From the data collapse in Fig. 6, we see that the scaling analysis is valid. The critical exponent ν is confirmed for numerous M/β to be in good agreement with $\nu = 1$, suggesting that the transition falls into the two-dimensional Ising universality class [66]. The simultaneously estimated quantity, i.e., the critical temperature T_c, is summarized in Fig. 7(a) together with the approximate solution given by Ref. [63].

At the critical temperature, we use the cluster algorithm in the extended space. The detailed description is given in Appendix D. Our main result is summarized in Fig. 7(b), in which we compare the autocorrelation time τ of the magnetization measured in units of Monte Carlo steps per site for the whole system. [See Appendix E for the calculation of physical quantities in the extended space.] The magnetization at the tth Monte Carlo step is calculated from the spin configurations $\boldsymbol{\sigma}(t) = [\sigma_1(t), \sigma_2(t), \ldots, \sigma_{N_v}(t)]$ with N_v being the number of visible spins as

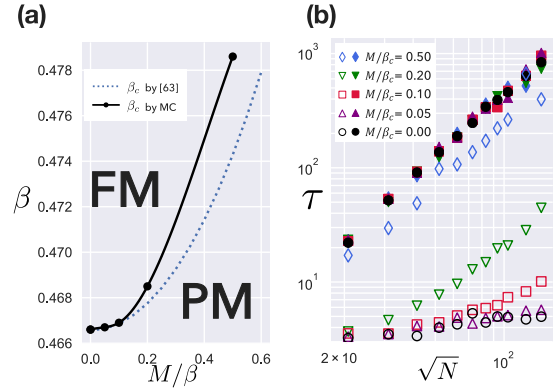$$m(t) = \sum_{i=1}^{N_v} \frac{\sigma_i(t)}{N_v}. \quad (35)$$

FIG. 7. (a) The autocorrelation time of the magnetization measured in units of Monte Carlo step. The black circle, purple upward triangle, red rectangle, green downward triangle, and blue diamond markers denote the magnitude of three-spin interactions to be $0, 0.05, 0.1, 0.2$, and 0.5, respectively. The filled (unfilled) markers represent the results by the single-spin flip in the original space (cluster update in the extended space). (b) The phase diagram of the model defined in Eq. (26). The boundary between the ferromagnetic (FM) and the paramagnetic (PM) phases are given here. The blue dashed line is calculated from the approximate solution in Ref. [63], and the black dots are given by the finite-size scaling of the Binder ratio. The numerically estimated inverse critical temperature at $M = 0$ approaches $\beta_c = \ln(3 + 2\sqrt{3})/4 \sim 0.4666$, which can be obtained from the exact solution [40].

The estimation of τ is done by evaluating the decay of the equilibrium autocovariance [28]

$$A(t) = \frac{\langle |m(t_0 + t)m(t_0)| \rangle - \langle |m(t_0)| \rangle^2}{\langle |m(t_0)|^2 \rangle - \langle |m(t_0)| \rangle^2} = A_0 e^{-t/\tau}, \quad (36)$$

where $\langle \cdots \rangle$ denotes the average over t_0, namely the MC steps. The critical slowing down is constantly observed in the SSF performed on the original space, while the application of the cluster update to the BM significantly improves the situation. We observe that the dynamical exponent z, which is the slope of data in Fig. 7(b), is also reduced, while the possibility that it gradually grows in the larger system sizes cannot be ruled out.

The increase in the autocorrelation time along the three-spin interaction M/β_c is understood in the following way. The virtual magnetic field induced in the hidden spins by the STT is amplified as M/β_c is increased, and thus the virtual magnetization per cluster increases, modifying the flipping probability of each cluster to be unbalanced. While the cluster is flipped randomly when the magnetic field is absent, finite-valued Zeeman energy results in unbalanced flipping probability due to the detailed balance condition. [See Appendix D for further discussion.] Such a situation prevents the system from exploring the spin configurations efficiently, and thus shows a weaker speed up.

IV. CONCLUSION AND DISCUSSION

In the current work, we find an algebraic transformation of the Ising model with many-spin interactions into the Boltzmann machine in which only two-spin interactions and

virtual local fields are present. The decoration-iteration and star-triangle transformations are applied to embed hidden and deep spins, namely the auxiliary degrees of freedom to be traced out. At the expense of the dimension of the spin space, significant suppression of the critical slowing down is achieved by applying the cluster algorithm to the Boltzmann machine.

Our scheme is also capable of handling continuous classical spin systems with many-spin interactions. As in the case with two-spin interactions [26], we may consider projecting each variable on some axis. Namely, we rewrite a continuous variable S_i on site i by a new Ising variable σ_i as $S_i = \sigma_i |S_i \cdot n_i| n_i + S_i^\perp$, where n_i is the randomly chosen projection axis and $S_i^\perp$ is orthogonal to n_i. We can now apply our method by regarding the model for $\{\sigma_i\}$ obtained by projection as the generalized Ising model. The randomness of the projection axis at each Monte Carlo step would assure the ergodicity of the scheme.

Beyond our scope in the current work is the optimal transformation for simulation. The transformation is nonunique when the virtual interaction is required, and we may even consider infinitely strong coupling to express extreme situations such as decoupled or completely aligned pairs of spins. Nonuniqueness arises also when four-spin interaction is present. Although we excluded the application of the star-square transformation for clarity, comparison of the numerical efficiency between different transformations may be worth investigating. Switching into different Boltzmann machines for each step may allow us to explore the free energy landscape more efficiently.

In closing, we would like to note the applicability of the algebraic transformation to wider fields of research. One interesting direction is undoubtedly the pursuit of equilibrium statistical physics, which includes extending and exploring exactly soluble models and replacing the Swendsen-Wang algorithm with other global updates to tackle frustrated systems. Another problem lies in the field of computer science; the application to combinatorial optimization problems; our decompositions applicable also in the finite-temperature case may open a new way to introduce ancilla spins required in the experimental implementation of annealing process.

APPENDIX A: TRANSFORMATION OF *p*-SPIN INTERACTION

In the following, we discuss the transformation of a bare p-spin interaction into the BM. The procedure consists of two steps; embedding the deep spins, and then the hidden spins. A system with a bare p-spin interaction is mapped into an equivalent system with three-spin interactions in the first step and subsequently broken into systems with two-spin interactions and local fields in the second step. Since the DIT is applied repeatedly, we define

$$M^{(k+1)} = \operatorname{arc\,cosh}(e^{2|M^{(k)}|})/2, \tag{A1}$$

$$M^{(0)} = M, \tag{A2}$$

where M determines the amplitude of the interaction as $\pi_p(\sigma; M)$. The number of DITs applied, or k, is referred to the DIT transformation order.

Let us take four-spin interaction as the starting point. As we showed in the main text, the Boltzmann factor can be expressed by embedding the deep spin as

$$\pi_4(\sigma; M^{(0)}) = \exp[M^{(0)} \sigma_1 \sigma_2 \sigma_3 \sigma_4]$$
$$= \Delta \sum_{d=\pm 1} \pi_3(\sigma_1, \sigma_2, d; M^{(1)}) \pi_3(\sigma_3, \sigma_4, d; M^{(1)}). \tag{A3}$$

By replacing one of the Ising variables with a product of two, we obtain the expression for five-spin interaction as

$$\pi_5(\sigma; M^{(0)})$$
$$= \exp[M^{(0)} \sigma_1 \sigma_2 \sigma_3 \sigma_4 \sigma_5]$$
$$= \Delta \sum_{d_1 = \pm 1} \pi_3(\sigma_1, \sigma_2, d_1; M^{(1)}) \pi_4(\sigma_3, \sigma_4, \sigma_5, d_1; M^{(1)})$$
$$= \Delta \sum_{d_1, d_2 = \pm 1} \pi_3(\sigma_1, \sigma_2, d_1; M^{(1)}) \pi_3(\sigma_3, d_1, d_2; M^{(2)})$$
$$\times \pi_3(\sigma_4, \sigma_5, d_2; M^{(2)}), \tag{A4}$$

which is visually described in Fig. 8(a). Repeating the DIT such that the DIT transformation order is as homogeneous as possible, we can show *a posteriori* that the general expression is given as

$$\pi_p(\sigma; M) = \Delta \sum_{\mathbf{d}} \overbrace{\pi_3(\sigma_1, \sigma_2, d_1; M^{(n)}) \pi_3(\sigma_3, d_1, d_2; M^{(n)}) \cdots}^{2^{n+1}-(p-2)} \underbrace{\pi_3(\sigma_{p-1}, \sigma_p, d_{p-3}; M^{(n+1)}) \pi_3(\sigma_{p-2}, d_{p-3}, d_{p-4}; M^{(n+1)}) \cdots}_{2(p-2-2^n)}, \tag{A5}$$

where n is an integer satisfying $2^n \leqslant p - 2 < 2^{n+1}$. Note that there are two factors consisting of two visible spins, whereas the others contain only one. The number of factors with $M^{(n+1)}$ is zero if $p - 2 = 2^n$.
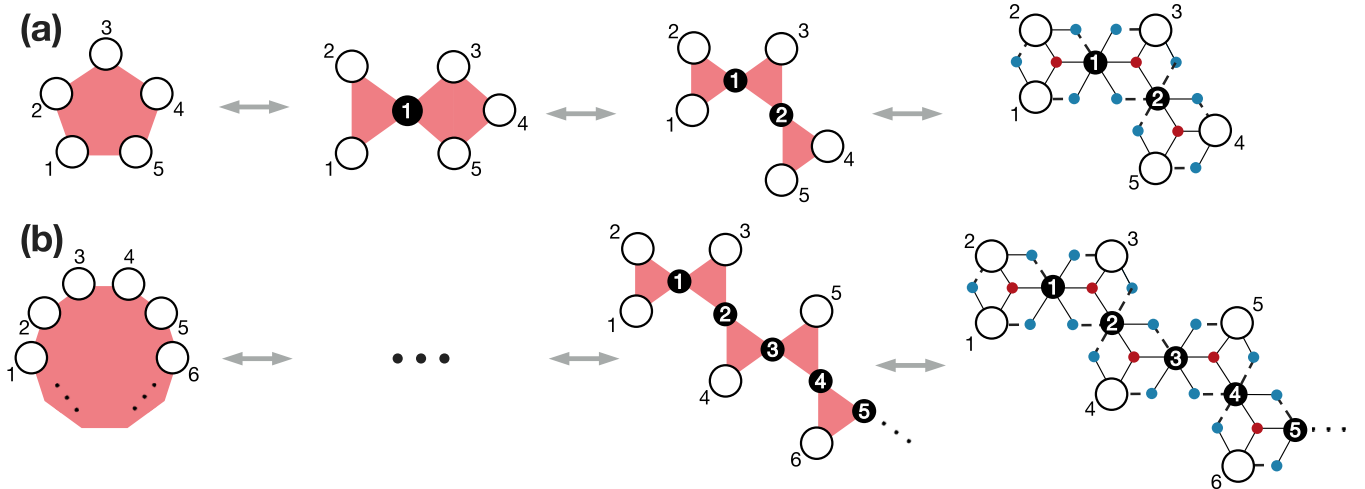
FIG. 8. The transformation of (a) five-spin interaction and (b) p-spin interaction. First mapped into a system with only three-spin interactions by applying DIT, the transformation technique for the three-spin interaction is used to break up into two-spin interactions with local external fields.

To consider the hidden spins, we substitute $M^{(n)}$ in the transformation introduced in Sec. II B 1 as, for instance,

$$
\begin{aligned}
&\pi_3(\sigma_1, \sigma_2, d_1; M^{(n)}) \\
&= \Delta \sum_{\boldsymbol{h}} \exp[(Wh + a)(\sigma_1 + \sigma_2 + d_1) + bh] \\
&\quad \times \exp\{W'[(\sigma_1 - \sigma_2)h_1 + (d_1 - \sigma_1)h_2 + (\sigma_2 - d_1)h_3]\},
\end{aligned}
\tag{A6}
$$

which is described in the rightmost panels of Fig. 8.

APPENDIX B: BLOCK GIBBS SAMPLING FROM RBM EQUIVALENT TO SQUARE ISING MODEL

In the following, we briefly introduce the block Gibbs sampling and discuss its application to the RBM equivalent to the ferromagnetic Ising model on the square lattice. Let A and B be the sublattices of the square lattice. The spin configuration on a sublattice is denoted as $\sigma_{A(B)} = \{\{\sigma_i\} | i \in V_{A(B)}\}$ where $V_{A(B)}$ is the set of $A(B)$-sublattice sites. The posterior distribution for spin configuration on the A sublattice is written as

$$
p(\sigma_A | \sigma_B) := \pi(\sigma_A, \sigma_B) / \sum_{\sigma_A} \pi(\sigma_A, \sigma_B),
\tag{B1}
$$

$$
\pi(\sigma_A, \sigma_B) = \exp\left(\beta \sum_{\langle i, j \rangle} \sigma_i \sigma_j\right),
\tag{B2}
$$

where $\pi(\sigma_A, \sigma_B)$ denotes the Boltzmann factor for the total system and $\langle i, j \rangle$ denotes the edge connecting the sites i and j. Since spins in the A sublattice do not couple to each other, Eq. (B1) can be factorized as

$$
p(\sigma_A | \sigma_B) = \prod_{i \in V_A} p(\sigma_i | \sigma_B),
\tag{B3}
$$

$$
p(\sigma_i | \sigma_B) = \frac{\exp\left[\beta \sigma_i \sum_{j \in \partial i} \sigma_j\right]}{2 \cosh\left[\beta \sum_{j \in \partial i} \sigma_j\right]},
\tag{B4}
$$

where ∂i denotes the set of sites adjacent to i. This allows us to sample each spin on A sublattice independently without rejection. The updated method that alternately switches the sublattice is called the block Gibbs sampling.

However, such an algorithm is not beneficial for the following reason. Consider a domain consisting of upward spins. In the bulk region of the domain, it is highly probable according to the Eq. (B3) that the newly sampled spins remain upwards as well. In other words, the limited range of the connection results in the mutual locking structure in the bulk region, allowing only the peripheral region to flip. Such a problematic situation is exacerbated as the domain size grows, and turns out that the slowing down is much worse than the single-spin flip.

APPENDIX C: PARTITION FUNCTION OF THREE-SPIN INTERACTING MODEL ON KAGOMÉ LATTICE

In this Appendix, we show that a model with only three-spin interactions on the Kagomé lattice is soluble. This model is a specific case of a broader class of models with crossing symmetry studied in Ref. [67], which do not exhibit a phase transition at finite temperature. The partition function is written as

$$
Z = \sum_{\sigma} \exp\left[\sum_{\triangle} M_{\triangle} \tau_{\triangle} + \sum_{\triangledown} M_{\triangledown} \tau_{\triangledown}\right],
\tag{C1}
$$

where $M_{\triangle}$ denotes the three-spin interaction and $\tau_{\triangle} := \prod_{j \in \triangle} \sigma_j$ the product of the Ising spins in an upward triangle in the lattice. Also $M_{\triangledown}$ and $\tau_{\triangledown}$ are defined similarly for a downward triangle. In order to compute Eq. (C1), we introduce the identity for a binary variable $x = \pm 1$ as follows:

$$
e^{Kx} = \cosh(K) \sum_{n=0,1} [x \tanh(K)]^n.
\tag{C2}
$$

Applying this identity to each triangle yields

$$Z = \sum_{\sigma} \sum_{n_\triangle, n_\triangledown} \prod_{\triangle} \prod_{\triangledown} \cosh(M_\triangle) \cosh(M_\triangledown)$$
$$\times [\tau_\triangle \tanh(M_\triangle)]^{n_\triangle} [\tau_\triangledown \tanh(M_\triangledown)]^{n_\triangledown}. \quad \text{(C3)}$$

Next, let us consider taking the sum over σ_j at some site j in Eq. (C1). Denoting the triangles touching the site j as $\triangle(j)$ and $\triangledown(j)$, the contribution from the spin at j can be given as $\sum_{\sigma_j} \sigma_j^{n_{\triangle(j)} + n_{\triangledown(j)}}$, which is nonzero only when $n_{\triangle(j)} = n_{\triangledown(j)}$. This argument holds for arbitrary j, and therefore the requirement $n_\triangle = n_\triangledown = 0, 1$ for all $\triangle$ and $\triangledown$ imposed for nonzero contribution. Accordingly, we obtain the concise expression of the partition function as

$$Z = C \left(1 + \prod_{\triangle, \triangledown} \tanh(M_\triangle) \tanh(M_\triangledown) \right), \quad \text{(C4)}$$

where $C = 2^{N_v} \prod_{\triangle} \prod_{\triangledown} \cosh(M_\triangle) \cosh(M_\triangledown)$ with N_v being the number of visible spins, i.e., the number of sites. The above expression clearly shows that the free energy in the thermodynamic limit is analytic, and hence the model does not show a phase transition at finite temperature. For $M_\triangle = M_\triangledown$, Eq. (C4) reproduces the partition function of the uniform model studied in Ref. [68]. We note that the present method is not limited to two-dimensional models. In fact, a similar model with four-spin interactions on a three-dimensional pyrochlore lattice is also soluble using the same technique [69].

APPENDIX D: CLUSTER UPDATES UNDER MAGNETIC FIELD

In this Appendix, we introduce two flavors of cluster updates accompanied with magnetic fields. We assume a model defined on a graph $G = (V, E)$ as follows:

$$H = -\sum_{\langle i,j \rangle \in E} J_{i,j} \sigma_i \sigma_j - \sum_{i \in V} h_i \sigma_i, \quad \text{(D1)}$$

where $J_{i,j}$ is the two-spin interaction between two binary degrees of freedom at sites i and j, or σ_i and σ_j, and h_i is the external field on site i. One way to take the external field into account is to modify the probability of flipping the cluster that is formed using the information on the interactions, and the other is to extend the space to express the field terms by interactions with the auxiliary space.

To introduce the first approach, let us remind that the detailed balance condition is given as

$$\frac{\pi(B)}{\pi(A)} = \frac{P(A \to B)\alpha(A \to B)}{P(B \to A)\alpha(B \to A)}, \quad \text{(D2)}$$

where $\pi(A)$ is the Boltzmann weight corresponding to a state A, $P(A \to B)$ is the trial proposal probability from the state A to B, and $\alpha(A \to B)$ is the corresponding acceptance

probability. The Swendsen-Wang algorithm applied in the main text adopts the Metropolis-Hastings rule

$$\alpha(A \to B) = \min\left(1, \frac{\pi(B)}{\pi(A)} \frac{P(B \to A)}{P(A \to B)}\right), \quad \text{(D3)}$$

which satisfies $\alpha = 1$ under $h_i = 0$. Since the external field modifies the Boltzmann weight as $\pi(A) \to e^{-\beta \sum_i h_i \sigma_i^A} \pi(A)$ at the inverse temperature β, the trial proposal must absorb such change to realize a rejection-free scheme under arbitrary external field. Concretely, the kth cluster C_k formed by the ordinary bonding process is flipped with probability $p_k = e^{-\beta m_k}/(e^{-\beta m_k} + e^{+\beta m_k})$ where $m_k = -\sum_{i \in C_k} h_i \sigma_i$ is the Zeeman energy by the external field. The additional computational effort per single MC step is ignorable.

In the second approach, known as the "ghost spin method," one introduces an auxiliary spin that interacts with any spin exposed to the external (or virtual) field [70,71]. Defining $\widetilde{G} = (\widetilde{V}, \widetilde{E})$ with the ghost spin on the 0th site as

$$V \to \widetilde{V} = \{0\} \cap V, \quad \text{(D4)}$$

$$E \to \widetilde{E} = \{\langle 0, i \rangle | i \in V\}, \quad \text{(D5)}$$

we alternatively consider a Hamiltonian as follows:

$$\widetilde{H} = -\sum_{\langle i,j \rangle \in \widetilde{E}} \widetilde{J}_{i,j} \sigma_i \sigma_j, \quad \text{(D6)}$$

where

$$\widetilde{J}_{i,j} = \begin{cases} J_{i,j} & \text{if } \langle i, j \rangle \in E, \\ h_i & \text{if } j = 0, \\ h_j & \text{if } i = 0. \end{cases} \quad \text{(D7)}$$

Now that the new Hamiltonian consists solely of two-spin interactions, the ordinary cluster update can be applied.

APPENDIX E: OBSERVATION OF PHYSICAL QUANTITY IN EXTENDED MODEL

The transformation considered in the main text preserves the partition function, and moreover the Boltzmann factor for visible spin configurations. Therefore, to compute the expectation value of a physical observable $O(\sigma)$ in the extended space, one may simply consider the identical mapping $\widetilde{O}(\sigma, h) = O(\sigma)$ to obtain

$$\langle O \rangle = \frac{\sum_{\sigma, h} \widetilde{O}(\sigma, h) \widetilde{\pi}(\sigma, h)}{\sum_{\sigma, h} \widetilde{\pi}(\sigma, h)} = \frac{\sum_{\sigma} O(\sigma) \left[\sum_h \widetilde{\pi}(\sigma, h) \right]}{Z}$$
$$= \frac{\sum_{\sigma} O(\sigma) \pi(\sigma)}{Z}. \quad \text{(E1)}$$

In other words, one may simply ignore all the hidden spins and compute the quantities using the operator in the original space.

[1] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).

[2] L. Wang, Phys. Rev. B **94**, 195105 (2016).

[3] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).

[4] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).

[5] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **86**, 044708 (2017).

[6] N. Yoshioka, Y. Akagi, and H. Katsura, Phys. Rev. B **97**, 205110 (2018).

[7] E. P. L. Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).

[8] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).

[9] S. J. Wetzel, Phys. Rev. E **96**, 022140 (2017).

[10] P. Huembeli, A. Dauphin, and P. Wittek, Phys. Rev. B **97**, 134109 (2018).

[11] H. Araki, T. Mizoguchi, and Y. Hatsugai, Phys. Rev. B **99**, 085406 (2019).

[12] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, arXiv:1809.05519.

[13] G. Carleo and M. Troyer, Science **355**, 602 (2017).

[14] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Phys. Rev. B **96**, 205152 (2017).

[15] I. Glasser, N. Pancotti, M. August, I. D. Rodriguez, and J. I. Cirac, Phys. Rev. X **8**, 011006 (2018).

[16] G. Carleo, Y. Nomura, and M. Imada, Nat. Commun. **9**, 5322 (2018).

[17] N. Freitas, G. Morigi, and V. Dunjko, Int. J. Quantum Inf. **16**, 1840008 (2018).

[18] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).

[19] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. B **96**, 195145 (2017).

[20] D.-L. Deng, X. Li, and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

[21] X. Gao and L.-M. Duan, Nat. Commun. **8**, 662 (2017).

[22] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).

[23] Z.-A. Jia, Y.-H. Zhang, Y.-C. Wu, L. Kong, G.-C. Guo, and G.-P. Guo, Phys. Rev. A **99**, 012307 (2019).

[24] L. Wang, Phys. Rev. E **96**, 051301(R) (2017).

[25] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[26] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).

[27] N. Prokof'ev and B. Svistunov, Phys. Rev. Lett. **87**, 160601 (2001).

[28] H. Evertz, Adv. Phys. **52**, 1 (2003).

[29] G. Biroli and M. Mézard, Phys. Rev. Lett. **88**, 025501 (2001).

[30] V. Blum, G. L. W. Hart, M. J. Walorski, and A. Zunger, Phys. Rev. B **72**, 165113 (2005).

[31] L. Berthier and G. Biroli, Rev. Mod. Phys. **83**, 587 (2011).

[32] J. D. Biamonte, Phys. Rev. A **77**, 052331 (2008).

[33] W. Lechner, P. Hauke, and P. Zoller, Sci. Adv. **1**, e1500838 (2015).

[34] M. Leib, P. Zoller, and W. Lechner, Quantum Sci. Technol. **1**, 015008 (2016).

[35] V. Choi, Quant. Inf. Proc. **10**, 343 (2011).

[36] P. Smolensky, in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, MA, 1986), Vol. 1, pp. 194–281.

[37] G. Hinton, Neural Comput. **14**, 1771 (2002).

[38] M. E. Fisher, Phys. Rev. **113**, 969 (1959).

[39] F. J. Wegner, J. Math. Phys. **12**, 2259 (1971).

[40] I. Syozi, Prog. Theor. Phys. **6**, 306 (1951).

[41] I. Syozi, *Phase Transition and Critical Phenomena*, Vol. 1, edited by C. Domb and M. S. Green, (Academic, London, 1972).

[42] D. Antonosyan, S. Bellucci, and V. Ohanyan, Phys. Rev. B **79**, 014432 (2009).

[43] O. Rojas, J. Valverde, and S. de Souza, Physica A **388**, 1419 (2009).

[44] B. Lisnyi and J. Strečka, Phys. Status Solidi B **251**, 1083 (2014).

[45] F. Y. Wu, J. Phys. A **23**, 375 (1990).

[46] K. Lin and F. Wu, Mod. Phys. Lett. B **4**, 311 (1990).

[47] W. T. Lu and F. Y. Wu, Phys. Rev. E **71**, 046120 (2005).

[48] F. Caravelli and C. Nisoli, arXiv:1810.09190.

[49] K. Lin and F. Wu, Int. J. Mod. Phys. B **5**, 2125 (1991).

[50] It might seem that the mapping is inapplicable to frustrated systems due to the condition given in Eq. (12). However, we emphasize that any frustrated system can be mapped to RBM or DBM, for instance, by introducing virtual two-spin interactions.

[51] R. J. Baxter, Phys. Rev. Lett. **26**, 832 (1971).

[52] R. J. Baxter, Ann. Phys. (NY) **70**, 193 (1972).

[53] J. Strečka, Entropy **20**, 91 (2018).

[54] J. Strečka, Phys. Lett. A **374**, 3718 (2010).

[55] L. Čanová, J. Strečka, J. Dely, and M. Jaščur, Acta Phys. Pol. A **113**, 449 (2008).

[56] J. Strečka, L. Čanová, M. Jaščur, and M. Hagiwara, Phys. Rev. B **78**, 024427 (2008).

[57] J. Strečka, A. Tanaka, L. Čanová, and T. Verkholyak, Phys. Rev. B **80**, 174410 (2009).

[58] R. J. Baxter and F. Y. Wu, Phys. Rev. Lett. **31**, 1294 (1973).

[59] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, Cambridge, MA, 2016).

[60] L. Onsager, Phys. Rev. **65**, 117 (1944).

[61] R. J. Glauber, J. Math. Phys. **4**, 294 (1963).

[62] N. Ito, Physica A **196**, 591 (1993).

[63] F. Y. Wu, X. N. Wu, and H. W. J. Blöte, Phys. Rev. Lett. **62**, 2773 (1989).

[64] K. Binder, Phys. Rev. Lett. **47**, 693 (1981).

[65] S. Tanaka, R. Tamura, and B. Chakrabarti, *Quantum Spin Glasses, Annealing and Computation* (Cambridge University Press, Cambridge, MA, 2017).

[66] M. E. Fisher, Phys. Rev. **180**, 594 (1969).

[67] S. H. Simon and P. Fendley, J. Phys. A **46**, 105002 (2013).

[68] K. A. Muttalib, M. Khatun, and J. H. Barry, Phys. Rev. B **96**, 184411 (2017).

[69] J. H. Barry and F. Y. Wu, Int. J. Mod. Phys. B **3**, 1247 (1989).

[70] A. Coniglio, F. de Liberto, G. Monroy, and F. Peruggi, J. Phys. A **22**, L837 (1989).

[71] J. Kent-Dobias and J. P. Sethna, Phys. Rev. E **98**, 063306 (2018).