Retrieving the ground state of spin glasses using thermal noise: Performance of quantum annealing at finite temperatures

Kohji Nishimura,[1] Hidetoshi Nishimori,[1] Andrew J. Ochoa,[2] and Helmut G. Katzgraber[2,3]

[1]*Department of Physics, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo 152-8551, Japan*
[2]*Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843-4242, USA*
[3]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA*

We study the problem to infer the ground state of a spin-glass Hamiltonian using data from another Hamiltonian with interactions disturbed by noise from the original Hamiltonian, motivated by the ground-state inference in quantum annealing on a noisy device. It is shown that the average Hamming distance between the inferred spin configuration and the true ground state is minimized when the temperature of the noisy system is kept at a finite value, and not at zero temperature. We present a spin-glass generalization of a well-established result that the ground state of a purely ferromagnetic Hamiltonian is best inferred at a finite temperature in the sense of smallest Hamming distance when the original ferromagnetic interactions are disturbed by noise. We use the numerical transfer-matrix method to establish the existence of an optimal finite temperature in one- and two-dimensional systems. Our numerical results are supported by mean-field calculations, which give an explicit expression of the optimal temperature to infer the spin-glass ground state as a function of variances of the distributions of the original interactions and the noise. The mean-field prediction is in qualitative agreement with numerical data. Implications on postprocessing of quantum annealing on a noisy device are discussed.

I. INTRODUCTION

Quantum annealing [1–8] is the quantum-mechanical counterpart of simulated annealing [9], an optimization method used to find the ground state of, e.g., an Ising model. Because combinatorial optimization problems can be formulated as the ground-state search of an Ising model [10], typically with spin-glass-like complex interactions, the development of efficient methods to solve these computationally hard problems is an important target of current research activities. Quantum annealing has attracted a great deal of attention in recent years partly due to the introduction of its hardware implementation, the D-Wave quantum annealer [11]. Evidence has been provided that the machine indeed runs quantum mechanically [11–21]; however, there remain open questions before the device becomes practically useful. One of the problems is the control error, i.e., imperfections in the setting of parameter values of the Ising Hamiltonian in the device [22,23]. Because it is difficult to set the parameter values (interactions and local fields of the Hamiltonian) with a high precision, the device might be attempting to find the ground state of the wrong Hamiltonian, thus compromising the reliability of the final output. This phenomenon arises in any analog device like the D-Wave quantum annealer and it is crucial to devise and implement ingenious methods to mitigate the influence of control errors. One approach is quantum error correction [24–28], however, at the cost of decreasing the number of available logical qubits.

A closely related problem has been analyzed in the context of classical error-correcting codes, in which a message is encoded and then transmitted through a noisy channel. The receiver has to correct errors in the received noisy signal and retrieve the original message. It is known that a certain type of error-correcting codes can be formulated in terms of the theory of spin glasses [29,30]. In this formulation, the task to retrieve the original message is translated to the inference of the ground state of an Ising model with uniform ferromagnetic interactions using only the information of a spin-glass Hamiltonian derived from the original ferromagnetic Hamiltonian by the application of noise to the interactions. It has been shown numerically [31], as well as analytically [32–34], that the best performance to retrieve the ferromagnetic ground state from the spin-glass Hamiltonian is achieved at a finite temperature rather than at zero temperature, since the Hamming distance of the retrieved spin configuration to the true ground state is a nonmonotonic function of the temperature. In fact, recent experiments on the D-Wave quantum annealer [35] study the trivial ferromagnetic case and illustrate that decoding is more efficient at finite temperature.

In the present paper we generalize the above formulation for error-correcting codes to the situation where the original (noiseless) Hamiltonian already has randomness in the interactions, i.e., a spin-glass Hamiltonian. The random interactions are then disturbed by (Gaussian) noise, and the task is to find a spin configuration closest to the ground state of the original Hamiltonian using only the Hamiltonian with disturbed interactions. One of the important motivations to study such a problem lies in the noise-mitigation task of analog devices as explained above: One is faced with the problem to infer the correct ground state of a Hamiltonian with random interactions out of data produced from the Hamiltonian with noise in addition to the original random interactions. Because it is very difficult to develop a general theory for this situation as was done for error-correcting codes [30,32], we use numerical methods for one- and two-dimensional models supplemented by a mean-field-type approach. Our result shows clearly that the original ground state is better inferred at finite temperature rather than at zero temperature. Therefore, tuning the D-Wave quantum annealer to the optimal decoding temperature might actually assist in mitigating the effects of analog noise in the device.

The paper is organized is as follows. In Sec. II, we formulate the problem. The results of numerical calculations are described in Sec. III. An analysis using mean-field theory is given in Sec. IV, followed by concluding remarks in Sec. V. Technical details are delegated to the Appendix.

II. FORMULATION OF THE PROBLEM

The goal is to infer the original ground state of a Hamiltonian whose interactions $J_{i_1,\ldots,i_p}$ follow a Gaussian distribution and are disturbed by noise, $J_{i_1,\ldots,i_p} \to \tilde{J}_{i_1,\ldots,i_p}$. We write the original spin-glass Hamiltonian with general many-body interactions as

$$H(\boldsymbol{S}) = -\sum_{i_1,\ldots,i_p} J_{i_1,\ldots,i_p} S_{i_1} \cdots S_{i_p}, \tag{1}$$

where $\boldsymbol{S}(= S_1, S_2, \ldots, S_N)$ is the set of Ising spins, i.e., $S_i \in \{\pm 1\}$. The original interactions $J_{i_1,\ldots,i_p}$ are generated from a Gaussian distribution $P(J_{i_1,\ldots,i_p})$ with mean unity and variance σ^2. The ground state of this Hamiltonian is given by the sign of the zero-temperature limit of the thermal expectation value, i.e.,

$$\begin{aligned} \mathcal{S}_i^{(0)} &= \lim_{\beta_0 \to \infty} \text{sgn}\langle S_i \rangle_{\beta_0} \\ &= \lim_{\beta_0 \to \infty} \text{sgn}\left(\frac{\text{Tr}_{\boldsymbol{S}}\, S_i \exp\left[-\beta_0 H(\boldsymbol{S})\right]}{\text{Tr}_{\boldsymbol{S}} \exp\left[-\beta_0 H(\boldsymbol{S})\right]} \right). \end{aligned} \tag{2}$$

We next introduce disturbed interactions $\tilde{J}_{i_1,\ldots,i_p}$ by adding a noise term $\xi_{i_1,\ldots,i_p}$ to the original interactions,

$$\tilde{J}_{i_1,\ldots,i_p} \equiv J_{i_1,\ldots,i_p} + \xi_{i_1,\ldots,i_p}. \tag{3}$$

The variables $\xi_{i_1,\ldots,i_p}$ are distributed according to a Gaussian distribution $P(\xi_{i_1,\ldots,i_p})$ with zero mean and variance γ^2. The problem is to find a spin configuration closest to the ground state of the *original* Hamiltonian Eq. (1) from the *noisy* Hamiltonian, Eq. (5) below, at finite temperature T or the inverse temperature $\beta = 1/T$,

$$\mathcal{S}_i^{(\beta)} = \text{sgn}\langle S_i \rangle_\beta = \text{sgn}\left(\frac{\text{Tr}_{\boldsymbol{S}}\, S_i \exp\left[-\beta \tilde{H}(\boldsymbol{S})\right]}{\text{Tr}_{\boldsymbol{S}} \exp\left[-\beta \tilde{H}(\boldsymbol{S})\right]} \right), \tag{4}$$

where

$$\tilde{H}(\boldsymbol{S}) = -\sum_{i_1,\ldots,i_p} \tilde{J}_{i_1,\ldots,i_p} S_{i_1} \cdots S_{i_p} \tag{5}$$

represents the Hamiltonian with the added noise. Notice that we are interested only in the *sign* of the spin average, not the magnitude, as indicated in Eq. (4). As a measure of similarity between the ground state of the original Hamiltonian and the spin configuration of the noisy system, we define the overlap $M(T)$ as

$$M(T) = \int \prod dJ\, P(J) \int \prod d\xi\, P(\xi)\, \mathcal{S}_i^{(0)} \mathcal{S}_i^{(\beta)}, \tag{6}$$

where the products run over the set of interactions. The overlap $M(T)$ is closely related to the average Hamming distance $D(T)$,

$$D(T) = \int \prod dJ\, P(J) \int \prod d\xi\, P(\xi) \sum_{i=1}^{N} \frac{\left(\mathcal{S}_i^{(0)} - \mathcal{S}_i^{(\beta)}\right)^2}{4} \tag{7}$$

by the relation

$$D(T) = \frac{N}{2}(1 - M(T)). \tag{8}$$

A large overlap $M(T)$ means a small Hamming distance. The Hamming distance is a standard measure of the quality of error-correcting codes [30], and we adopt it as the quantity to be minimized [or the overlap $M(T)$ to be maximized] in the present paper.

Optimal decoding occurs at the point where $M(T)$ has a maximum. Here we analyze how this maximum depends on the strength of the disorder γ.

When $\sigma = 0$, the Hamiltonian in Eq. (1) represents the ferromagnetic Ising model. Then, $\mathcal{S}_i^{(0)} = 1$ for all i (or maybe $\mathcal{S}_i^{(0)} = -1\ \forall i$), and the overlap $M(T)$ is identical to that of error-correcting codes with all original bits being unity [30]. It is known in this case that the overlap in Eq. (6) takes a maximum value at $T_N = \gamma^2$, the so-called Nishimori temperature [30–34]. It is difficult to apply the same theory to the case where $\sigma \neq 0$ due to the lack of proper symmetry. We therefore use numerical methods in the following section to study the behavior of the disordered system.

III. NUMERICAL ANALYSIS

Following Ref. [31], we apply the numerical transfer-matrix method to a triangular ladder with two- and three-body interactions, as depicted in Fig. 1, using free boundary conditions,

$$H = -\sum_{i=1}^{N-2} \left(J_i^{(1)} S_i S_{i+1} S_{i+2} + J_i^{(2)} S_i S_{i+2} \right). \tag{9}$$

Furthermore, we study a quasi-two-dimensional system with triangular ladders stacked on top of each other as depicted in Fig. 2 using the transfer-matrix method.

The overlap for finite-size systems

$$M(T) = \frac{1}{N} \sum_{i=1}^{N} \mathcal{S}_i^{(0)} \mathcal{S}_i^{(\beta)} \tag{10}$$

is calculated for these lattices with the ground-state configuration $\mathcal{S}_i^{(0)}$ determined by the Viterbi algorithm [36], which is a zero-temperature transfer-matrix method. We take the configurational average over the distributions of the original interactions and noise by sampling 400 disorder realizations. The number of spins is $N = 10^5$ for the ladder and $N = LM$ with $L = 10^3$ and $M = 10$ for stacked ladders. The standard deviations of the original interactions are chosen to be $\sigma = 1.0$, 1.4, 2.0, and 3.0. For each value of σ, we apply noise with a
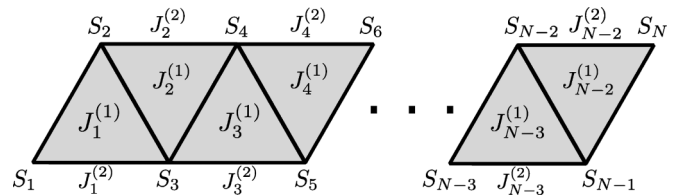

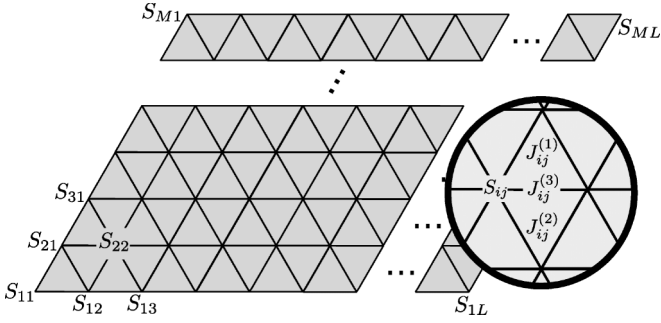FIG. 1. Ising spin glass on the triangular ladder.

FIG. 2. Ising spin glass on a stacked triangular ladder.

typical strength of $\gamma = 0.1, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.25, 1.5$, and 2.0, respectively.

Figure 3 shows typical results for triangular and stacked ladders for several pairs of σ and γ. The ordinate is the overlap $M(T)$ and the abscissa is the temperature T. The mean of the overlap is indicated by the red solid line, and the shaded area denotes the standard deviation. We observe a clear peak at finite temperature in each panel. Similar results were found for all pairs of σ and γ. These results clearly show that the ground state of the original Hamiltonian can be inferred with higher probability from data generated by the noisy Hamiltonian at finite temperature than at zero temperature.
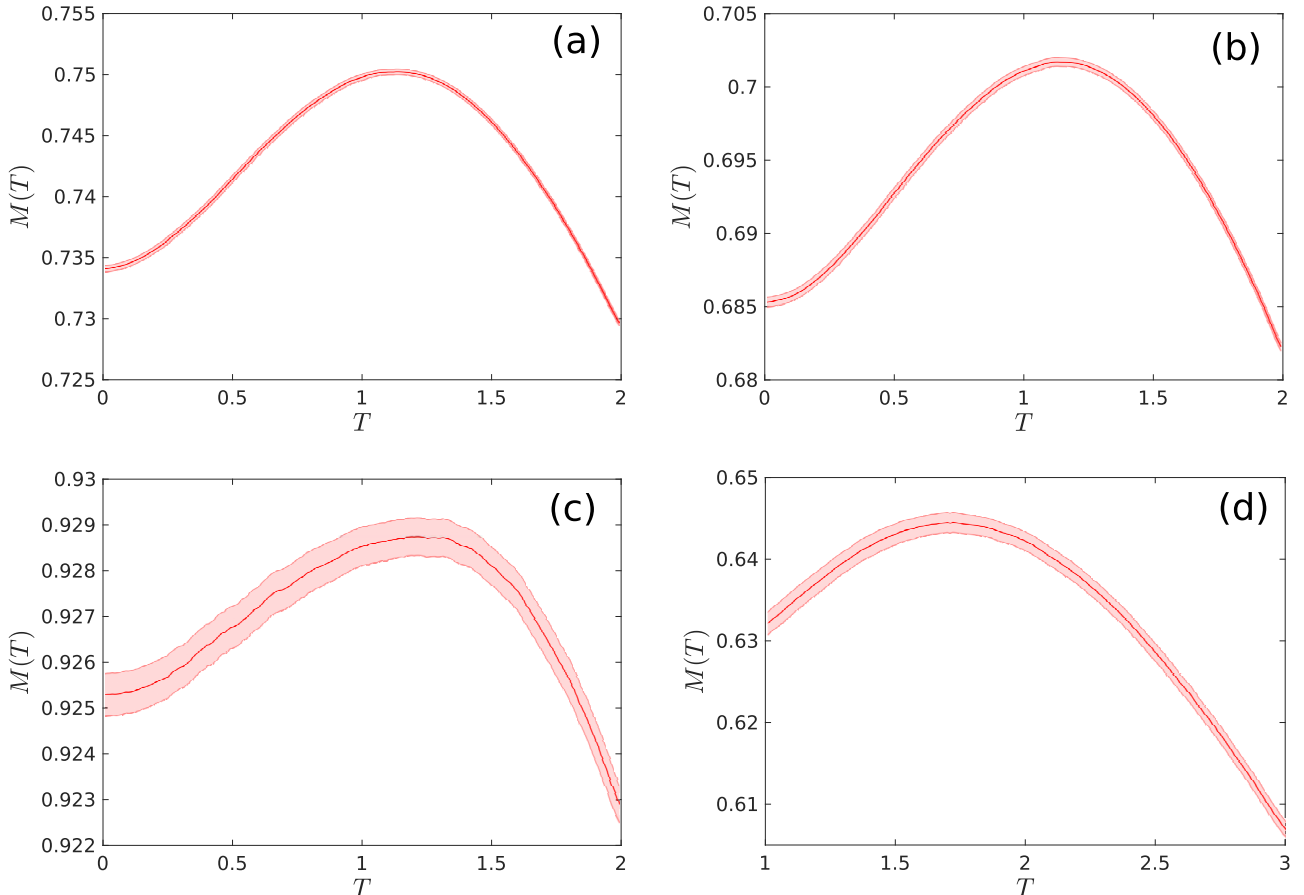
Figure 4 shows the optimal temperature T_{opt} for decoding, i.e., the peak position of $M(T)$, as a function of γ for each fixed value of σ. The green curve is the optimal temperature for $\sigma = 0$ ($T_{\text{N}} = \gamma^2$). The data indicate that the naive relation $T_{\text{opt}} = \gamma^2$ for $\sigma = 0$ does not hold when $\sigma > 0$. The functional form of $T_{\text{opt}}(\gamma)$ thus depends on the lattice structure, unlike the case of $\sigma = 0$.

IV. MEAN-FIELD ANALYSIS

Because it is difficult to develop a generic theory to derive $T_{\text{opt}}(\gamma)$ for the noisy case when $\sigma > 0$, we use mean-field theory to understand its behavior qualitatively. The original Hamiltonian is chosen to be the Sherrington-Kirkpatrick (SK) fully connected Ising spin glass [37],

$$H_{\text{SK}}(\boldsymbol{S}) = -\sum_{i<j} J_{ij} S_i S_j, \qquad (11)$$

where the sum runs over all distinct pairs of spins. The interactions are drawn from a Gaussian distribution

$$P(J_{ij}) = \frac{1}{\sigma}\sqrt{\frac{N}{2\pi}} \exp\left\{-\frac{N}{2\sigma^2}\left(J_{ij} - \frac{J_0}{N}\right)^2\right\}, \qquad (12)$$
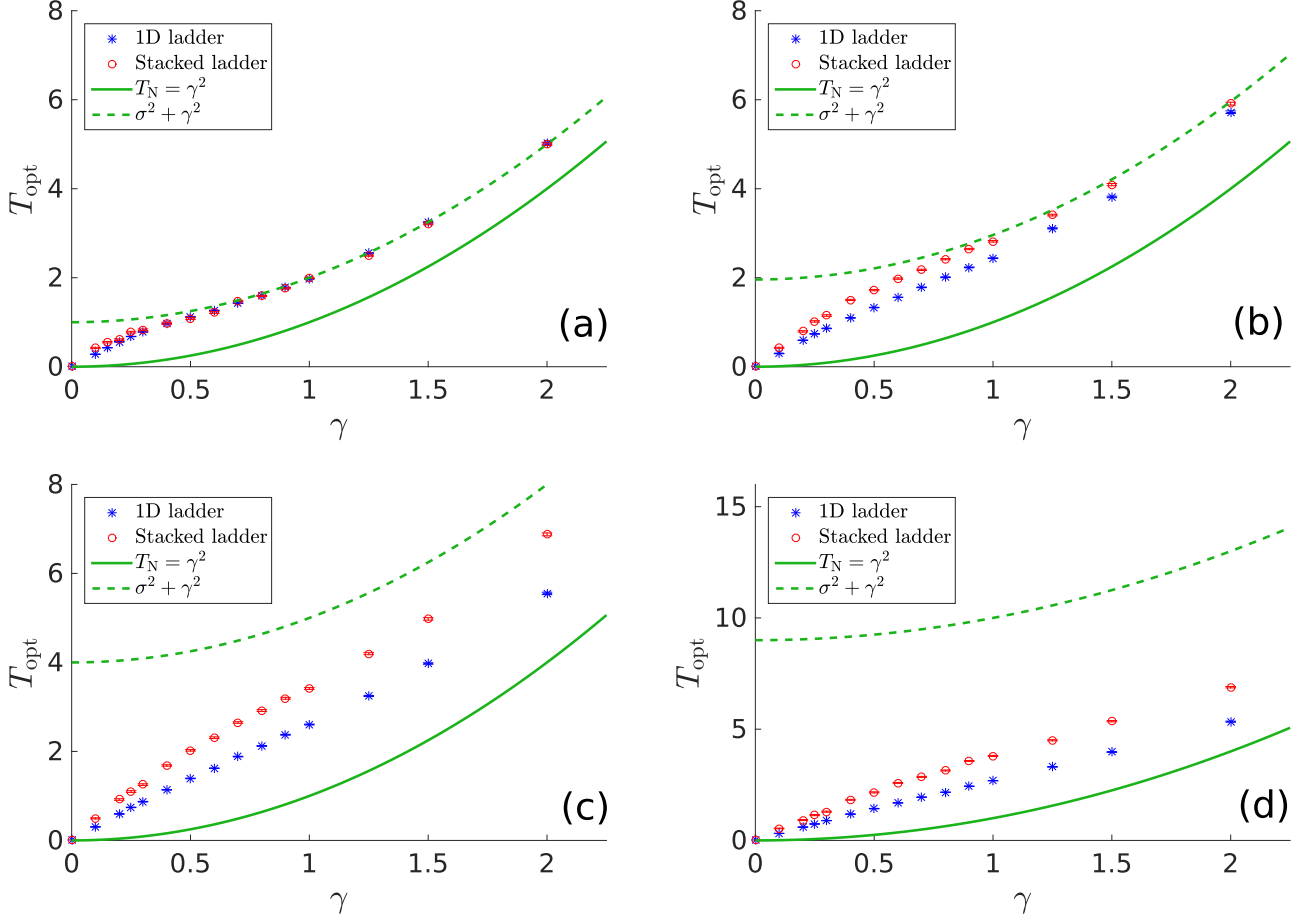

FIG. 3. Mean of the overlap $M(T)$ and its standard deviation as functions of the temperature T: data for the ladder (a) for $\sigma = 1$ and $\gamma = 0.5$ and (b) for $\sigma = 2$ and $\gamma = 0.4$; data for the stacked ladder (c) for $\sigma = 1$ and $\gamma = 0.5$ and (d) for $\sigma = 2$ and $\gamma = 0.4$.

FIG. 4. Optimal decoding temperature T_{opt} as a function of the standard deviation of the noise γ. The green solid curve in each panel represents the optimal temperature for $\sigma = 0$ ($T_{\text{N}} = \gamma^2$) and the green dashed line shows the result of the mean-field analysis ($\sigma^2 + \gamma^2$). (a) $\sigma = 1$, (b) $\sigma = 1.4$, (c) $\sigma = 2$, and (d) $\sigma = 3$.

where J_0 represents the mean of the distribution and σ^2 its variance. The Hamiltonian with noise is given by

$$\tilde{H}_{\text{SK}}(\boldsymbol{\tau}) = -\sum_{i<j} \tilde{J}_{ij} \tau_i \tau_j, \quad (\tau_i = \pm 1), \qquad (13)$$

where the interactions are affected by noise via

$$\tilde{J}_{ij} = J_{ij} + \xi_{ij}. \qquad (14)$$

As before, the noise follows a Gaussian distribution with variance γ^2,

$$P(\xi_{ij}) = \frac{1}{\gamma} \sqrt{\frac{N}{2\pi}} \exp\left(-\frac{N}{2\gamma^2} \xi_{ij}^2\right). \qquad (15)$$

The overlap is

$$M(T) = \lim_{\beta_0 \to \infty} [\text{sgn} \langle S_i \rangle_{\beta_0} \, \text{sgn} \langle \tau_i \rangle_{\beta}], \qquad (16)$$

where the square brackets denote the configurational average over the distributions of interactions and noise,

$$\int \prod_{i<j} dJ_{ij} d\xi_{ij} \, P(J_{ij}) P(\xi_{ij})(\cdots) \equiv [\cdots]. \qquad (17)$$

The angular brackets represent the thermal average with respect to each Hamiltonian:

$$\begin{aligned}
\langle S_i \rangle_{\beta_0} &= \frac{\text{Tr}_S \, S_i \exp[-\beta_0 H(\boldsymbol{S})]}{\text{Tr}_S \exp[-\beta_0 H(\boldsymbol{S})]}, \\
\langle \tau_i \rangle_{\beta} &= \frac{\text{Tr}_{\boldsymbol{\tau}} \, \tau_i \exp[-\beta \tilde{H}(\boldsymbol{\tau})]}{\text{Tr}_{\boldsymbol{\tau}} \exp[-\beta \tilde{H}(\boldsymbol{\tau})]}.
\end{aligned} \qquad (18)$$

We assume β_0 to be finite in the course of the calculations and take the limit $\beta_0 \to \infty$ at the end. Our goal is to identify the temperature that maximizes the overlap. As detailed in the Appendix, the free energy per spin is calculated under the ansatz of replica symmetry as

$$
\begin{aligned}
-[f] = {}& \frac{\sigma^2 \beta_0^2}{4} q_0^2 + \frac{(\sigma^2 + \gamma^2)\beta^2}{4} q^2 - \frac{J_0 \beta_0}{2} m_0^2 - \frac{J_0 \beta}{2} m^2 \\
& + \frac{\sigma^2 \beta_0^2}{4} + \frac{(\sigma^2 + \gamma^2)\beta^2}{4} - \frac{\sigma^2 \beta_0^2}{2} q_0 - \frac{(\sigma^2 + \gamma^2)\beta^2 q}{2} \\
& + \int Dz_0 \ln(2 \cosh H_1(z_0)) \\
& + \int Dz \ln(2 \cosh H_2(z)), \qquad (19)
\end{aligned}
$$

where

$$H_1(z_0) = \sqrt{\sigma^2 \beta_0^2 q_0}\, z_0 + J_0 \beta_0 m_0, \qquad (20)$$

$$H_2(z_0) = \sqrt{(\sigma^2 + \gamma^2)\beta^2 q}\, z_0 + J_0 \beta m. \qquad (21)$$

The self-consistent equations for the order parameters read

$$m_0 = \int Dz_0 \tanh H_1(z_0) = [\langle S_i \rangle_{\beta_0}],$$

$$m = \int Dz \tanh H_2(z) = [\langle \tau_i \rangle_{\beta}],$$

$$q_0 = \int Dz_0 \tanh^2 H_1(z_0) = [\langle S_i \rangle_{\beta_0}^2], \qquad (22)$$

$$q = \int Dz \tanh^2 H_2(z) = [\langle \tau_i \rangle_{\beta}^2].$$

As shown in the Appendix, the two systems defined in Eqs. (11) and (13) decouple in the replica symmetric solution, and we obtain

$$M(T) = \lim_{\beta_0 \to \infty} [\mathrm{sgn}\,\langle S_i \rangle_{\beta_0}\, \mathrm{sgn}\,\langle \tau_i \rangle_{\beta}]$$

$$= \lim_{\beta_0 \to \infty} [\mathrm{sgn}\langle S_i \rangle_{\beta_0}]_{\sigma^2}[\mathrm{sgn}\langle \tau_i \rangle_{\beta}]_{\sigma^2 + \gamma^2}, \qquad (23)$$

where the subscript in the outer square brackets in the second line denotes the variance of the distribution of randomness. Since the first factor,

$$\lim_{\beta_0 \to \infty} [\mathrm{sgn}\,\langle S_i \rangle_{\beta_0}]_{\sigma^2}, \qquad (24)$$

is independent of β, only the second factor,

$$[\mathrm{sgn}\langle \tau_i \rangle_{\beta}]_{\sigma^2 + \gamma^2}, \qquad (25)$$

determines the temperature dependence of $M(T)$. The latter is nothing but the well-established case of vanishing original variance [30–34] because the expression involves a single system with noise. Therefore, we reach the conclusion that the overlap takes a maximum value at the temperature

$$T_{\mathrm{opt}} = \sigma^2 + \gamma^2. \qquad (26)$$

This result is naturally consistent with the case of vanishing original variance $\sigma = 0$, i.e., $T_{\mathrm{N}} = \gamma^2$.

The green dashed line depicted in Fig. 4 shows the mean-field prediction $T_{\mathrm{opt}} = \sigma^2 + \gamma^2$. The numerical data agree relatively satisfactorily with the mean-field prediction for small σ and large γ. The finite value of T_{opt} at $\gamma = 0$ should be an artifact of the replica symmetric solution because the limit $\gamma = 0$ represents the noiseless case, where the original and the noisy Hamiltonians coincide, and thus $T = 0$ should give the best result.

V. CONCLUSION

We have examined the effects of noise added to the interactions of an Ising spin-glass Hamiltonian. The goal is to infer the ground state of the original Ising Hamiltonian using only the output from the system with noisy interactions. The spin configuration of the latter system at finite temperature has been compared with the ground state of the original system.

It has been shown from numerical transfer-matrix calculations that the overlap of two spin configurations has a maximum at a *finite* temperature, thus giving the smallest Hamming distance. This means that the ground state of the original Hamiltonian is better inferred at finite temperature than at the ground state of the noisy system. An intuitive explanation of this result would be that the unperturbed ground state is an excited state of the noisy system. Therefore, a finite temperature that corresponds to the excitation energy might increase the chances to find the ground state of the original problem Hamiltonian. Nevertheless, the original ground state is just one of very many states to be realized at a specific finite temperature for the noisy system and our result is quite nontrivial.

A similar phenomenon has been known to exist for years for the case of uniform ferromagnetic interactions in the original Hamiltonian under the context of error-correcting codes [30–35]. The present work generalizes this old result to the case with randomness already in the original model. Although the former case of uniform original interactions has been able to be treated analytically with full generality in the sense that there is no restriction on the type of lattice or the range of interactions [30,32–34], it is difficult to develop a comparable analytical theory for the present case because of the difference in symmetries. We therefore used the numerical transfer-matrix method and a mean-field analysis. It is an important direction of further research to establish analytical results under general conditions for the case with random original interactions. In particular, it would be very useful to derive an explicit analytical expression for the optimal temperature that is applicable beyond mean-field theory.

Finiteness of the optimal temperature implies an important lesson for data analysis (postprocessing) of a real quantum annealer. Because any device operates at finite temperature, it is always a serious problem to keep temperature effects under control. However, our result suggests that thermal noise may be positively used to infer the ground state of the original Hamiltonian when the interactions are disturbed by noise. For example, if the original Hamiltonian has uniformly ferromagnetic interactions, the optimal temperature is given as

$$T_{\mathrm{N}} = J_0 \left(\frac{\gamma}{J_0} \right)^2, \qquad (27)$$

where J_0 is the original uniform interaction and γ^2 is the variance of noise. Notice that we set $J_0 = 1$ in Sec. II. If we insert the values roughly corresponding to those for the D-Wave machine, $J_0 = 5$ GHz and $\gamma/J_0 = 0.1$, the latter representing the standard deviation of control errors, the optimal temperature turns out to be approximately 2 mK. The operating temperature of the D-Wave machine is approximately 20 mK and, therefore, much higher than this theoretical optimal temperature. It should nevertheless be remarked that the numbers would easily change if the original interactions are random, if the real control errors are larger than $\gamma/J_0 = 0.1$, or if the noise is distributed according to a another distribution (e.g., $1/f$ or pink noise). It has indeed been shown in our numerical calculations that the optimal temperature for random original interactions is much higher than T_{N} as seen in

Fig. 4. It is, however, noteworthy that our theoretical estimate is not much different than the real operating temperature of the device. Therefore, it would be interesting to verify these theoretical predictions on the actual D-Wave 2X device for nontrivial spin-glass problems complementing previous results on the trivial ferromagnetic case [35].

We have chosen the overlap $M(T)$ as the measure of better inference of the ground state. This means to minimize the Hamming distance to the true ground state, which is the standard in error-correcting codes [30]. This measure is also useful in other circumstances including the spin-glass problem, where the spin configuration of the ground state is of as much interest as the ground-state energy. In the context of combinatorial optimization problems, however, it is often the case that the energy is a prime measure of performance. It is useful to remember here that a smaller Hamming distance to the ground state does not necessarily mean a smaller energy [27]. It may then happen that our present result does not apply as is, if it is desirable to minimize the energy, not to minimize the Hamming distance. This turns out to be a highly nontrivial problem and will be discussed in a forthcoming paper. It is also interesting to see how our conclusion would change if quantum effects are taken into account explicitly. We are studying this problem and results will be announced before too long.

APPENDIX: REPLICA-SYMMETRIC SOLUTION

In this Appendix, we explain the derivation of the replica-symmetric solution of the mean-field model discussed in Sec. IV. The calculation follows the standard *modus operandi*, but with two independent systems sharing part of the random interactions.

The configurational average of the nth power of the partition function is

$$[Z^n] = \int \prod_{i<j} dJ_{ij} d\xi_{ij} P(J_{ij}) P(\xi_{ij})$$
$$\times \mathrm{Tr} \exp \left\{ \sum_{i<j} \sum_{\alpha=1}^{n} \left(\beta_0 J_{ij} S_i^\alpha S_j^\alpha + \beta (J_{ij} + \xi_{ij}) \tau_i^\alpha \tau_j^\alpha \right) \right\}$$

$$= \mathrm{Tr} \int \left(\prod_{i<j} dJ_{ij} d\xi_{ij} \right) \exp \left\{ \sum_{i<j} \left(\beta_0 J_{ij} \sum_{\alpha=1}^{n} S_i^\alpha S_j^\alpha \right.\right.$$
$$+ \beta (J_{ij} + \xi_{ij}) \sum_{\alpha=1}^{n} \tau_i^\alpha \tau_j^\alpha$$
$$\left.\left. - \frac{N}{2\sigma^2} \left(J_{ij} - \frac{J_0}{N} \right)^2 - \frac{N}{2\gamma^2} \xi_{ij}^2 + \ln \left(\frac{1}{\sigma\gamma} \frac{N}{2\pi} \right) \right) \right\}.$$

$$(A1)$$

The standard mean-field calculation [30] leads to

$$-[f] = \lim_{n\to0} \left(-\frac{\sigma^2 \beta_0^2}{4n} \sum_{\alpha\neq\beta} q_{\alpha\beta}^{0}{}^2 - \frac{(\sigma^2+\gamma^2)\beta^2}{4n} \right.$$
$$\times \sum_{\alpha\neq\beta} q_{\alpha\beta}{}^2 - \frac{\sigma^2 \beta_0 \beta}{2n} \sum_{\alpha\beta} u_{\alpha\beta}{}^2$$
$$- \frac{J_0 \beta_0}{2n} \sum_{\alpha} m_\alpha^{0}{}^2 - \frac{J_0 \beta}{2n} \sum_{\alpha} m_\alpha{}^2$$
$$\left. + \frac{\sigma^2 \beta_0^2}{4} + \frac{(\sigma^2+\gamma^2)\beta^2}{4} + \frac{1}{n} \ln \mathrm{Tr} e^L \right), \quad (A2)$$

where

$$L \equiv \sigma^2 \beta_0^2 \sum_{\alpha<\beta} q_{\alpha\beta}^0 S^\alpha S^\beta + (\sigma^2+\gamma^2)\beta^2 \sum_{\alpha<\beta} q_{\alpha\beta} \tau^\alpha \tau^\beta$$
$$+ \sigma^2 \beta_0 \beta \sum_{\alpha\beta} u_{\alpha\beta} S^\alpha \tau^\beta$$
$$+ J_0 \beta_0 \sum_{\alpha} m_\alpha^0 S^\alpha + J_0 \beta \sum_{\alpha} m_\alpha \tau^\alpha. \quad (A3)$$

The order parameters in $[f]$ are determined by the saddle-point conditions

$$q_{\alpha\beta}^0 = \langle S^\alpha S^\beta \rangle_L = \left[\langle S_i^\alpha \rangle_{\beta_0} \langle S_i^\beta \rangle_{\beta_0} \right]$$
$$q_{\alpha\beta} = \langle \tau^\alpha \tau^\beta \rangle_L = \left[\langle \tau_i^\alpha \rangle_\beta \langle \tau_i^\beta \rangle_\beta \right]$$
$$m_\alpha^0 = \langle S^\alpha \rangle_L = \left[\langle S_i^\alpha \rangle_{\beta_0} \right] \quad (A4)$$
$$m_\alpha = \langle \tau^\alpha \rangle_L = \left[\langle \tau_i^\alpha \rangle_\beta \right]$$
$$u_{\alpha\beta} = \langle S^\alpha \tau^\beta \rangle_L = \left[\langle S_i^\alpha \rangle_{\beta_0} \langle \tau_i^\beta \rangle_\beta \right],$$

where

$$\langle \cdots \rangle_L \equiv \frac{\mathrm{Tr}(\cdots) e^L}{\mathrm{Tr} e^L}. \quad (A5)$$

If we assume the replica-symmetric ansatz [30], this free energy can be rewritten as follows:

$$-[f] = \lim_{n\to0} \left(-\frac{\sigma^2 \beta_0^2}{4n}(n^2-n)q_0^2 - \frac{(\sigma^2+\gamma^2)\beta^2}{4n}(n^2-n)q^2 \right.$$
$$- \frac{\sigma^2 \beta_0 \beta}{2n} n^2 u^2 - \frac{J_0 \beta_0}{2n} n m_0^2 - \frac{J_0 \beta}{2n} n m^2$$
$$\left. + \frac{\sigma^2 \beta_0^2}{4} + \frac{(\sigma^2+\gamma^2)\beta^2}{4} + \frac{1}{n} \ln \mathrm{Tr} e^{L_0} \right), \quad (A6)$$

where L_0 represents L under the replica-symmetric ansatz. The explicit form of $\ln \mathrm{Tr} e^{L_0}$ is [38]

$$
\begin{aligned}
\ln \mathrm{Tr} e^{L_0} = \ln \mathrm{Tr} \exp \Bigg\{ & \sigma^2 \beta_0^2 q_0 \sum_{\alpha<\beta} S^\alpha S^\beta + (\sigma^2 + \gamma^2)\beta^2 q \sum_{\alpha<\beta} \tau^\alpha \tau^\beta \\
& + \sigma^2 \beta_0 \beta u \left(\sum_\alpha S^\alpha \right) \left(\sum_\beta \tau^\beta \right) + J_0 \beta_0 m_0 \sum_\alpha S^\alpha + J_0 \beta m \sum_\alpha \tau^\alpha \Bigg\} \\
= & \, n \left(\int Dz_0 Dz Dw \ln \left[2 \cosh H_1(z_0,w) 2 \cosh H_2(z,w) \right] - \tilde{f} \right) + \mathcal{O}(n^2),
\end{aligned}
\tag{A7}
$$

where

$$
H_1(z_0,w) = \sqrt{a} z_0 + \sqrt{c} w + d, \quad H_2(z,w) = \sqrt{b} z + \sqrt{c} w + e,
\tag{A8}
$$

and

$$
\begin{aligned}
a &= \sigma^2 \beta_0^2 q_0 - \sigma^2 \beta_0 \beta u \\
b &= (\sigma^2 + \gamma^2)\beta^2 q - \sigma^2 \beta_0 \beta u \\
c &= \sigma^2 \beta_0 \beta u \\
d &= J_0 \beta_0 m_0 \\
e &= J_0 \beta m \\
\tilde{f} &= \frac{\sigma^2 \beta_0^2 q_0}{2} + \frac{(\sigma^2 + \gamma^2)\beta^2 q}{2}.
\end{aligned}
\tag{A9}
$$

Therefore, the free energy after taking the limit $n \to 0$ is

$$
\begin{aligned}
-[f] = & \, \frac{\sigma^2 \beta_0^2}{4} q_0^2 + \frac{(\sigma^2 + \gamma^2)\beta^2}{4} q^2 - \frac{J_0 \beta_0}{2} m_0^2 - \frac{J_0 \beta}{2} m^2 + \frac{\sigma^2 \beta_0^2}{4} + \frac{(\sigma^2 + \gamma^2)\beta^2}{4} \\
& + \int Dz_0 Dw \ln(2 \cosh H_1(z_0,w)) + \int Dz Dw \ln(2 \cosh H_2(z,w)) - \tilde{f}.
\end{aligned}
\tag{A10}
$$

This result reduces to Eq. (19) because the w dependence of H_1 and H_2 disappears as explained below. It turns out the averaged free energy does not depend on u, as can be verified by direct computations:

$$
-\frac{\partial [f]}{\partial u} = 0.
\tag{A11}
$$

We can therefore choose $u = 0$ without loss of generality. This implies that the original and noisy systems decouple completely because u is the parameter that connects these two systems as seen in Eq. (A3). When $u = 0$, $c = 0$ according to Eq. (A9) and hence the w dependence of $H_1(z_0,w)$ and $H_2(z,w)$ disappears as can be seen in Eq. (A8).

[1] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, Phys. Rev. E **58**, 5355 (1998).

[2] T. Kadowaki, Ph.D. thesis, Tokyo Institute of Technology, 1998.

[3] A. Finnila, M. Gomez, C. Sebenik, C. Stenson, and J. Doll, Quantum annealing: A new method for minimizing multidimensional functions, Chem. Phys. Lett. **219**, 343 (1994).

[4] A. Das and B. K. Chakrabarti, *Colloquium*: Quantum annealing and analog quantum computation, Rev. Mod. Phys. **80**, 1061 (2008).

[5] G. E. Santoro and E. Tosatti, Optimization using quantum mechanics: Quantum annealing through adiabatic evolution, J. Phys. A **39**, R393 (2006).

[6] S. Morita and H. Nishimori, Mathematical foundation of quantum annealing, J. Math. Phys. **49**, 125210 (2008).

[7] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective, Phys. Rep. **523**, 127 (2013).

[8] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, Science **292**, 472 (2001).

[9] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by simulated snnealing, Science **220**, 671 (1983).

[10] A. Lucas, Ising formulations of many NP problems, Front. Phys. **2**, 5 (2014).

[11] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk *et al.*, Quantum annealing with manufactured spins, Nature (London) **473**, 194 (2011).

[12] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Experimental signature of programmable quantum annealing, Nat. Commun. **4**, 2067 (2013).

[13] T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk *et al.*, Entanglement in a Quantum Annealing Processor, Phys. Rev. X **4**, 021041 (2014).

[14] T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, Consistency tests of classical and quantum models for a quantum annealer, Phys. Rev. A **91**, 042314 (2015).

[15] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, Computational multiqubit tunnelling in programmable quantum annealers, Nat. Commun. **7**, 10327 (2016).

[16] N. G. Dickson, M. W. Johnson, M. H. Amin, R. Harris, F. Altomare, A. J. Berkley, P. Bunyk, J. Cai, E. M. Chapple, P. Chavez *et al.*, Experimental signature of programmable quantum annealing, Nat. Commun. **4**, 1903 (2013).

[17] S. Boixo, T. F. Rønnow, S. V. Isakov, W. Zhihui, D. Wecker, D. A. Licar, J. M. Martinis, and M. Troyer, Evidence for quantum annealing with more than one hundred qubits, Nat. Phys. **10**, 218 (2014).

[18] T. Albash, T. Rønnow, M. Troyer, and D. Lidar, Reexamining classical and quantum models for the D-Wave One processor, Eur. Phys. J. Special Topics **224**, 111 (2015).

[19] P. J. D. Crowley, T. Ðurić, W. Vinci, P. A. Warburton, and A. G. Green, Quantum and classical dynamics in adiabatic computation, Phys. Rev. A **90**, 042317 (2014).

[20] H. G. Katzgraber, F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, Seeking Quantum Speedup Through Spin Glasses: The Good, the Bad, and the Ugly, Phys. Rev. X **5**, 031026 (2015).

[21] S. Mandrà, Z. Zhu, W. Wang, A. Perdomo-Ortiz, and H. G. Katzgraber, Strengths and weaknesses of weak-strong cluster problems: A detailed overview of state-of-the-art classical heuristics vs quantum approaches, arXiv:1604.01746.

[22] A. Perdomo-Ortiz, B. O'Gorman, J. Fluegemann, R. Biswas, and V. N. Smelyanskiy, Determination and correction of persistent biases in quantum annealers, Sci. Rep. **6**, 18628 (2015).

[23] Z. Zhu, A. J. Ochoa, S. Schnabel, F. Hamze, and H. G. Katzgraber, Best-case performance of quantum annealers on native spin-glass benchmarks: How chaos can affect success probabilities, Phys. Rev. A **93**, 012317 (2016).

[24] D. A. Lidar, Towards Fault Tolerant Adiabatic Quantum Computation, Phys. Rev. Lett. **100**, 160506 (2008).

[25] *Quantum Error Correction*, edited by D. A. Lidar and T. A. Brun (Cambridge University Press, Cambridge, U.K., 2013).

[26] K. L. Pudenz, T. Albash, and D. A. Lidar, Error-corrected quantum annealing with hundreds of qubits, Nat. Commun. **5**, 3243 (2014).

[27] K. L. Pudenz, T. Albash, and D. A. Lidar, Quantum annealing correction for random Ising problems, Phys. Rev. A **91**, 042302 (2015).

[28] W. Vinci, T. Albash, G. Paz-Silva, I. Hen, and D. A. Lidar, Quantum annealing correction with minor embedding, Phys. Rev. A **92**, 042310 (2015).

[29] N. Sourlas, Spin-glass models as error-correcting codes, Nature (London) **339**, 693 (1989).

[30] H. Nishimori, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, New York, 2001).

[31] P. Ruján, Finite Temperature Error-Correcting Codes, Phys. Rev. Lett. **70**, 2968 (1993).

[32] H. Nishimori, Optimum decoding temperature for error-correcting codes, J. Phys. Soc. Jpn. **62**, 2973 (1993).

[33] N. Sourlas, Spin glasses, error-correcting codes and finite-temperature decoding, Europhys. Lett. **25**, 159 (1994).

[34] Y. Iba, The Nishimori line and Bayesian statistics, J. Phys. A **32**, 3875 (1999).

[35] N. Chancellor, S. Szoke, W. Vinci, G. Aeppli, and P. A. Warburton, Maximum-entropy inference with a programmable annealer, Sci. Rep. **6**, 22318 (2016).

[36] A. Viterbi, Error bounds for convolutional codes and an asymptotically optimum decoding algorithm, IEEE Trans. Inf. Theory **13**, 260 (1967).

[37] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin-Glass, Phys. Rev. Lett. **35**, 1792 (1975).

[38] Note that $Dm \equiv \exp(-m^2/2)/\sqrt{2\pi}\, dm$ represents the Gaussian measure.