Check for
updates

Proof of Avoidability of the Quantum First-Order Transition in Transverse Magnetization in Quantum Annealing of Finite-Dimensional Spin Glasses

Mizuki Yamaguchi[1] · Naoto Shiraishi[1] · Koji Hukushima[1]

Abstract

It is rigorously shown that an appropriate quantum annealing for any finite-dimensional spin system has no quantum first-order transition in transverse magnetization. This result can be applied to finite-dimensional spin-glass systems, where the ground state search problem is known to be hard to solve. Consequently, it is strongly suggested that the quantum first-order transition in transverse magnetization is not fatal to the difficulty of combinatorial optimization problems in quantum annealing.

Keywords Quantum annealing · Quantum first-order transition · Hard optimization problems · Spin glass · Self-averaging

1 Introduction

Solving combinatorial optimization problems efficiently is a major topic in theoretical computer science. From the viewpoint of physics, this problem can be described as energy minimization with a given classical spin system. Inspired by this, the simulated annealing [1] and the quantum annealing (QA) [2, 3] were invented as generic heuristic methods for solving these optimization problems. Given a classical Hamiltonian $\hat{U}^{N:J}$ of system size N and for the set of quenched coupling constants J, the QA solves the energy minimization problem as follows: we set the Hamiltonian of the quantum system as

$$\hat{H}^{N:J}(\gamma) = \hat{U}^{N:J} + \gamma \hat{D} , \tag{1}$$

where $\hat{D}$ is noncommutative with $\hat{U}^{N:J}$ and γ is a parameter controlled in QA. This Hamiltonian $\hat{D}$ is called the *driver Hamiltonian* and has a known ground state. Starting with the system with large γ and varying slowly to $\gamma = 0$, one would expect the ground state of

---

Communicated by Aernout van Enter.

✉ Mizuki Yamaguchi
yamaguchi-q@g.ecc.u-tokyo.ac.jp

1   Graduate School of Arts and Sciences, The University of Tokyo, Meguro-ku 3-8-1, Tokyo 1530041, Japan

$\hat{H}^{N:J}(\gamma)$ to change from the known ground state of $\hat{D}$ to the desired solution state, the ground state of $\hat{U}^{N:J}$. The quantum adiabatic theorem [4] guarantees that the desired ground state is attainable by taking a sufficiently long time to change γ. A natural question raised here is whether this process is time efficient, i.e., whether it takes only polynomial time of the system size N to reach the true ground state of $\hat{U}^{N:J}$. Roughly speaking, the time cost is proportional to the inverse square of the minimum gap between the ground-state energy and the first excited state energy of the Hamiltonian in the annealing process [3, 5].

The first proposal of QA employed the transverse magnetic term $-\hat{M}_x^N = -\sum_{i=1}^N \hat{\sigma}_i^x$ as the driver Hamiltonian $\hat{D}$, and argued based on numerical experiments that the QA with this driver Hamiltonian is superior to simulated annealing [2]. On the other hand, this type of QA was shown to fail in the p-spin model (p-body mean-field ferromagnetic model with $p > 2$) [6, 7], where the gap is exponentially small and the computation time is exponentially large. In the case of the p-spin model, the transverse magnetization undergoes a discontinuous jump in the annealing process, which we call the quantum first-order transition in transverse magnetization (QFOT for short) analogous to the first-order transition in thermodynamics. This is a challenging phenomenon for QA, since the sudden change in the transverse magnetization causes an exponential collapse of the gap, implying the failure of annealing.

The fundamental origin of the failure of the QA for hard optimization problems is under discussion in this field. One possible argument is that the failures mainly are due to the QFOT [8], based on the observed fact that the QFOT frequently appears when the QA fails [9]. In contrast, another argument is that the failure of the QA has origins other than the QFOT. For example, it was reported in Ref [10–14] that QA in some models shows exponentially small gaps at points other than the point of QFOT. However, the latter observation is based on specific models and does not provide a general argument on the relation between the QFOT and the failure of the QA.

To resolve this controversy, we adopt the QA with antiferromagnetic fluctuations (QA-AFF) first proposed in Ref [7] and set the entire Hamiltonian as $\hat{H}^{N:J}(\gamma; \alpha) = \hat{U}^{N:J} - \gamma \hat{M}_x^N + \frac{\alpha}{N}(\hat{M}_x^N)^2$. The additional fluctuations term makes the QFOT more avoidable. Previous studies based on specific models show that in the ferromagnetic p-spin model (without quenched randomness) and the Hopfield model, the QA-AFF succeeds in avoiding QFOT under some conditions, while the QFOT appears unavoidable under others [7, 9]. In spite of these previous investigations, the potential effectiveness of the QA-AFF in general systems has not yet been uncovered.

In this paper, we rigorously prove that the QA-AFF for finite-dimensional spin-glass systems avoids the QFOT. Our result applies to general finite-dimensional spin-glass systems as long as i.i.d. quenched random variables are used. Even for systems that exhibit the QFOT under the conventional QA with only the transverse magnetic field, the addition of the antiferromagnetic fluctuations term always removes this singularity and makes the transverse magnetization in this QA continuous as a function of γ. In other words, the QFOT in QA can be completely avoided by adding antiferromagnetic fluctuations.

We note that the search for the ground state of three-dimensional Ising spin-glass systems is considered to be a computationally hard task since it belongs to an NP-hard problem [15], which is a class of the most difficult combinatorial optimization problems. It is believed that even quantum computers cannot solve NP-hard problems efficiently, which suggests that the QA-AFF in fact fails at some point. Based on our findings, we assert that the QFOT is not fatal to the difficulty of combinatorial optimization problems in QA.

In our proof, the self-averaging plays a pivotal role to derive the absence of singularity. First, we show that the ground-state energy $E_g^{N:J}(\gamma)$ is self-averaging in the conventional

QA only with the transverse magnetic field. There remains the possibility that the function obtained by taking the average of the quenched randomness has a singularity (i.e., non-differentiable points). Then, it can be shown that the addition of the antiferromagnetic fluctuations term does indeed remove the singularity for any finite-dimensional system. However, for some long-range interaction systems, this claim may not hold, which is consistent with the previous studies showing that QFOT cannot be avoided in the mean-field models.

This paper is organized as follows. In Sect. 2, we explain our setup and main claim. In Sect. 3, we describe the outline of the proof, presenting several key ideas in this proof. In Sect. 4, which consists of four subsections, we prove the main theorem. We briefly review the Legendre transformation in Sect. 4.1. Sections 4.2 and 4.3 are devoted to the investigation of self-averaging for a fixed parameter and uniform self-averaging for a function, respectively. We finally introduce the antiferromagnetic fluctuations in Sect. 4.4, which completes the proof of the avoidance of the QFOT.

2 Setup and Main Claim

2.1 Setup

We deal with the energy minimization problem of a classical spin-1/2 system on a finite-dimensional lattice with N spins. Pairs of spins $\hat{\sigma}_i^z$ and $\hat{\sigma}_j^z$ interact with each other through the coupling constant J_{ij}, which is a quenched random variable. The Hamiltonian of this classical system is thus expressed as

$$\hat{U}^{N:J} = -\sum_{i=1}^{N}\sum_{j=1}^{N} J_{ij}\hat{\sigma}_i^z\hat{\sigma}_j^z \,, \tag{2}$$

where the set of coupling constants J_{ij} is denoted simply by J.

To solve the energy minimization problem by QA, we add a transverse magnetic field with strength $\gamma \in (-\infty, \infty)$, which leads to the following Hamiltonian for the quantum system:

$$\hat{H}^{N:J}(\gamma) = \hat{U}^{N:J} - \gamma \hat{M}_x^N = \hat{U}^{N:J} - \gamma \sum_{i=1}^{N} \hat{\sigma}_i^x \,. \tag{3}$$

As discussed in detail in Sect. 4.4, dealing with the QA-AFF, $\hat{U}^{N:J}$ in this formula is replaced by $\hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2$.

Throughout this paper, we consider models in finite dimension with $\hat{U}^{N:J}$ generated by a shift-invariant probability distribution for quenched random variables. We here clarify the meaning of *finite dimension* for later use. If the lattice is placed in $\mathbb{R}^n$ space and each site interacts only with its neighbors, the notion of dimension has no confusion. The problem may arise when distant sites also interact with forces that decay with distance. To cover these systems, we define the finite dimensionality for $D \geq 2$ with D being the spatial dimension as follows[1]:

Definition 2.1.1 (*Finite dimensionality*) We say that a Hamiltonian $\hat{U}^{N:J}$ (or a system) is in a finite dimension if the following two conditions are satisfied[2]:

---

[1] In one dimension, the same argument holds but the order evaluation of the result differs.

[2] Our results hold for finite-dimensional systems with general p-body interactions. For example, in three-body interacting system $\hat{U}^{N:J} = -\sum_{h=1}^{N}\sum_{i=1}^{N}\sum_{j=1}^{N} J_{hij}\hat{\sigma}_h^z\hat{\sigma}_i^z\hat{\sigma}_j^z$, finite-dimensionality is the condition

- For any site i, the sum of interactions with i is bounded from above as

$$\sum_{j=1}^{N} \sqrt{\left[J_{ij}^2\right]} \leq c_{\text{bulk}} , \tag{4}$$

where $[\cdot]$ is the random average with respect to the quenched randomness (see Sect. 4.2 for details), and c_{bulk} is a constant independent of the system size N.
- For any size A of the system, there exists a decomposition of sites into $B = N/A$ subsystems of size A denoted by $A[1], \ldots, A[B]$, such that the sum of interactions between inside and outside of any subsystem $A[k]$ is bounded from above as

$$\sum_{i \in s(A[k])} \sum_{j \notin s(A[k])} \sqrt{\left[J_{ij}^2\right]} \leq c_{\text{surface}} A^{1-1/D} , \tag{5}$$

which means that the surface energy is insignificant compared to the bulk energy for a sufficiently large system. Here $s(A[k])$ is the set of the sites in $A[k]$, and c_{surface} is a constant independent of the size of the system and the subsystems.

We also clarify the meaning of *shift-invariant probability distribution* for quenched random variables. Let $P_{ij}(J_{ij})$ be the probability distribution for the coupling constants J_{ij}. This probability distribution is shift-invariant if $P_{ij} = P_{kl}$ holds for any i, j, k, l such that $\mathbf{r}_i - \mathbf{r}_j = \mathbf{r}_k - \mathbf{r}_l$, where $\mathbf{r}_i$ is a D dimensional vector representing the lattice position of site i. If the lattice consists of several sublattices, the above characterization further requires that i and k belong to the same sublattice. The system without quenched randomness is considered a special case of a shift-invariant system.

2.2 Main Result

We shall define the quantum first-order transition in transverse magnetization (QFOT) in systems with quenched random variables. We first provide the definition of the absence of the QFOT for a system without quenched randomness, i.e., $\hat{U}^{N:J}$ is deterministically constructed depending on N. In this case, we say that the QA does not exhibit a QFOT if the transverse magnetization density $\frac{\langle g(\gamma)|\hat{M}_x^N|g(\gamma)\rangle}{N}$ of a ground state $|g(\gamma)\rangle$ of $\hat{H}^{N:J}(\gamma)$ converges uniformly to a continuous function $m_*(\gamma)$ in thermodynamic limit ($N \to \infty$):

$$\lim_{N \to \infty} \sup_{\gamma \in [0,\infty)} \left| \frac{\langle g(\gamma)|\hat{M}_x^N|g(\gamma)\rangle}{N} - m_*(\gamma) \right| = 0 . \tag{6}$$

Here, the ground state of $\hat{H}^{N:J}(\gamma)$ is implicitly assumed to be unique. If the ground state degenerates, we regard that our condition should be satisfied for any state $|g\rangle$ in the state space of minimum energy: $G(\hat{H}^{N:J}(\gamma)) = \arg\min_{|\psi\rangle}\langle\psi|\hat{H}^{N:J}(\gamma)|\psi\rangle$. Or equivalently, the above definition can be replaced by

$$\lim_{N \to \infty} \sup_{\gamma \in [0,\infty)} \max_{|g\rangle \in G(\hat{H}^{N:J}(\gamma))} \left| \frac{\langle g|\hat{M}_x^N|g\rangle}{N} - m_*(\gamma) \right| = 0 . \tag{7}$$

Footnote 2 continued

that $\sum_{i=1}^{N}\sum_{j=1}^{N} \sqrt{[J_{hij}^2]} \leq c_{\text{bulk}}$ instead of (4) and that $\sum_{h \in s(A[k])} \sum_{i \notin s(A[k])} \sum_{j \in s(A[k])} \sqrt{[J_{hij}^2]} \leq c_{\text{surface}} A^{1-1/D}$ instead of (5).

In the case of spin glasses with quenched randomness, we shall define the QFOT as its typical behavior, replacing the convergence in the above definition with stochastic convergence (especially convergence in mean square).

Definition 2.2.1 We say that a given QA shows no quantum first-order phase transition in transverse magnetization if there exists a continuous function $m_*(\gamma)$ such that

$$\lim_{N \to \infty} \left[\sup_{\gamma \in [0,\infty)} \left(\frac{\langle g(\gamma) | \hat{M}_x^N | g(\gamma) \rangle}{N} - m_*(\gamma) \right)^2 \right] = 0 . \tag{8}$$

Here, $[\cdot]$ means the average for the quenched random variable.

We remark that the function $m_*(\gamma)$ is independent of the quenched random variables. Thus, the above definition states that for almost all $\hat{U}^{N:J}$ the transverse magnetization density in the ground state converges to the same function.

We prove that the QFOT in the above definition is always avoidable in finite-dimensional spin-glass systems.

Theorem 2.2.2 *For any classical Hamiltonian in finite dimension with quenched random variables sampled from a shift-invariant probability distribution, the QA-AFF for this Hamiltonian does not have a QFOT.*

This is the main result of this paper. Notice that this theorem only describes the absence of a jump in the transverse magnetization density during the QA process, and does not evaluate the size of the energy gap that determines the annealing time required for successful computation.

2.3 Remark

We here put two remarks related to our main result.

The first remark is on the connection to the computational hardness, in particular computational complexity. It is known that the energy minimization problem in three-dimensional spin glass is an NP-hard problem. Thus, from the perspective of computational complexity, the finite-dimensional spin glass with $D \geq 3$ is as complex as the Sherrington-Kirkpatrick model. Therefore, our subject is indeed the efficiency of QA for hard combinatorial optimization problems.

This paper rules out the possibility that the difficulty in QA for finite-dimensional spin glasses lies in the QFOT. Our result, however, does not state that the QA-AFF actually succeeds as a method for solving combinatorial optimization problems and that NP $\subseteq$ BQP in terms of computational complexity. A more plausible scenario would be that the QA-AFF suffers from causes other than the QFOT. We will discuss this point in detail in Sect. 5.

The second remark is on the form of $\hat{U}^{N:J}(; \alpha)$ in QA-AFF. In QA-AFF, the Hamiltonian at $\gamma = 0$ is $\hat{U}^{N:J}(; \alpha) = \hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2$, not the classical spin glass Hamiltonian $\hat{U}^{N:J}$ we wish to solve. However, this discrepancy does not matter for the success or the failure of the QA for the following reason: if the coupling constants do not take continuous real numbers but discrete numbers of decimal places, by setting α smaller than the smallest unit of energy and measuring the final state with the computational basis, we can observe one of the true ground states with a finite probability[3]. Indeed, under the restriction that the coupling

---

[3] Let $|\text{ans}\rangle$ be a classical ground state of $\hat{U}^{N:J}$ and d be the smallest unit of energy of $\hat{U}^{N:J}$. Since we have $\langle \text{ans}| \hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2 |\text{ans}\rangle = \langle \text{ans}| \hat{U}^{N:J} |\text{ans}\rangle + \alpha$, the ground-state energy of $\hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2$ is

constants are integers, ground-state search problem for three-dimensional spin glasses has been proven to be NP-hard [15].

3 Outline of the Proof

Our first simple but important step is inspired by thermodynamics, where thermodynamic functions (e.g., the Helmholtz free energy and the Gibbs free energy) are connected via the Legendre transformation with respect to an extensive variable (e.g., volume) and an intensive variable (e.g., pressure). We regard the ground-state energy $E_g^{N:J}(\gamma)$ as a *thermodynamic function* with an *intensive variable* γ. Then, its inverse Legendre transformation [16, 17] yields the constrained minimum energy $U_g^{N:J}(M_x)$ with an *extensive variable* M_x.

The functional forms of $E_g^{N:J}(\gamma)$ and $U_g^{N:J}(M_x)$ depend on $\hat{U}^{N:J}$. However, if the probability distribution generating $\hat{U}^{N:J}$ is shift-invariant, we can show that

$$\frac{E_g^{N:J}(\gamma)}{N} \sim e_g(\gamma) \,,$$

$$\frac{U_g^{N:J}(Nm_x)}{N} \sim u_g(m_x) \,, \tag{9}$$

where $\sim$ stands for stochastic convergence (especially convergence in mean square), which is nothing but self-averaging in terms of physics. In the conventional QA only with a transverse magnetic field, $-e_g(\gamma)$ and $u_g(m_x)$ are shown to be convex, while $u_g(m_x)$ might be not strictly convex, which leads to a singularity as an non-differentiable point in $e_g(\gamma)$ at the QFOT.

To remove this singularity, we add the antiferromagnetic fluctuations term to the Hamiltonian $\hat{H}^{N:J}(\gamma)$, which results in $\hat{H}^{N:J}(\gamma; \alpha) = \hat{U}^{N:J} - \gamma \hat{M}_x^N + \frac{\alpha}{N}(\hat{M}_x^N)^2$. In this case, the corresponding $u_g(m_x; \alpha)$ is strictly convex, since $u_g(m_x)$ is convex and the square function αm_x^2 provides a small convex curve. Consequently, it follows that the QFOT does not occur in a typical evaluation with the addition of the antiferromagnetic fluctuations term.

The effort for this proof is mainly devoted to proving self-averaging. In this paper, we use the finite dimensionality for the proof, while some studies employ other methods [18]. To show the self-averaging of $E_g^{N:J}(\gamma)$, we decompose the system into $B = N/A$ subsystems of size A and use a version of the law of large numbers recalling that these subsystems are i.i.d. Subsequently, we show the uniform self-averaging (uniform convergence) of $E_g^{N:J}(\gamma)$. The assertion of the main Theorem 2.2.2 is described as uniform self-averaging of the transverse magnetization.

4 Proof

4.1 Preparation

We introduce two functions $E_g^{N:J}(\gamma)$ and $U_g^{N:J}(M_x)$ analogous to *thermodynamic functions* and demonstrate that they are Legendre transformation and inverse Legendre transformation

---

Footnote 3 continued
lower than that of $\hat{U}^{N:J}$ plus α. Consequently, measuring the ground state of $\hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2$ with the computational basis yields the ground state(s) of $\hat{U}^{N:J}$ with probability at least $1 - \alpha/d$ if $\alpha < d$.

[16] of each other. For completeness, below we will describe several basic results of the Legendre transformation in terms of $E_g^{N:J}(\gamma)$ and $U_g^{N:J}(M_x)$. Readers who are familiar with these techniques can skip this subsection.

Definition 4.1.1 The ground-state energy of $\hat{H}^{N:J}(\gamma)$ is defined as

$$E_g^{N:J}(\gamma) = \min_{|\psi\rangle} \langle\psi|\hat{H}^{N:J}(\gamma)|\psi\rangle . \tag{10}$$

Here, $|\psi\rangle$ runs all possible pure states.

The domain of minimization in the above definition can be extended from pure states to general mixed states.

Proposition 4.1.2 *The ground-state energy* $E_g^{N:J}(\gamma)$ *is also the minimum expectation energy of general mixed states:*

$$E_g^{N:J}(\gamma) = \min_{\hat{\rho}} Tr\hat{\rho}\hat{H}^{N:J}(\gamma) . \tag{11}$$

Proof Fix $\hat{\rho}' \in \arg\min_{\hat{\rho}} \mathrm{Tr}\hat{\rho}\hat{H}^{N:J}(\gamma)$. Decomposing this state as $\hat{\rho}' = \sum_t p_t |\psi_t\rangle\langle\psi_t|$, we obtain

$$\min_{\hat{\rho}} \mathrm{Tr}\hat{\rho}\hat{H}^{N:J}(\gamma) = \mathrm{Tr}\hat{\rho}'\hat{H}^{N:J}(\gamma)$$

$$= \sum_t p_t \langle\psi_t|\hat{H}^{N:J}(\gamma)|\psi_t\rangle$$

$$\geq E_g^{N:J}(\gamma) . \tag{12}$$

The inverse inequality is obvious. □

We next introduce the minimum energy conditioned by the transverse magnetization.

Definition 4.1.3 The minimum energy of $\hat{U}^{N:J}$ conditioned by x-magnetization at $M_x = \langle\hat{M}_x^N\rangle$ is defined as

$$U_g^{N:J}(M_x) = \min_{\hat{\rho}|\mathrm{Tr}\hat{\rho}\hat{M}_x^N=M_x} \mathrm{Tr}\hat{\rho}\hat{U}^{N:J} . \tag{13}$$

Notably, $E_g^{N:J}(\gamma)$ and $U_g^{N:J}(M_x)$ are connected through the Legendre transformation.

Proposition 4.1.4 $E_g^{N:J}(\gamma)$ *is the Legendre transformation of* $U_g^{N:J}(M_x)$.

Proof Combining the definitions of $E_g^{N:J}(\gamma)$ and $U_g^{N:J}(M_x)$, we find

$$E_g^{N:J}(\gamma) = \min_{M_x\in[-N,N]} \min_{\hat{\rho}|\mathrm{Tr}\hat{\rho}\hat{M}_x^N=M_x} \mathrm{Tr}\hat{\rho}(\hat{U}^{N:J} - \gamma\hat{M}_x^N)$$

$$= \min_{M_x\in[-N,N]} (U_g^{N:J}(M_x) - \gamma M_x) . \tag{14}$$

This means the Legendre transformation[4] of $U_g^{N:J}(M_x)$ in terms of M_x. □

Proposition 4.1.5 $E_g^{N:J}(\gamma)$ *is a concave function.*

---

[4] In the mainstream style of mathematics, $E_g^{N:J}(\gamma)$ is *minus* the Legendre transformation. Our notation is based on that widely used in thermodynamics [17].

Proof The Legendre transformation provides a concave function. □

Proposition 4.1.6 $U_g^{N:J}(M_x)$ *is a convex function.*

Proof For any $\lambda(0 \leq \lambda \leq 1)$, $M_-, M_+(M_- < M_+)$, we fix $\hat{\rho}_+ \in \arg\min_{\hat{\rho}|\mathrm{Tr}\hat{\rho}\hat{M}_x^N = M_+} \hat{U}^{N:J}$ and $\hat{\rho}_- \in \arg\min_{\hat{\rho}|\mathrm{Tr}\hat{\rho}\hat{M}_x^N = M_-} \hat{U}^{N:J}$, which are density matrices minimizing $\hat{U}^{N:J}$ under the constraint that the x-magnetization is $M_\pm$, respectively. Then, putting $M(\lambda) := (1 - \lambda)M_- + \lambda M_+$, we have

$$
(1 - \lambda)U_g^{N:J}(M_-) + \lambda U_g^{N:J}(M_+) = \mathrm{Tr}((1 - \lambda)\hat{\rho}_- + \lambda\hat{\rho}_+)\hat{U}^{N:J}
$$
$$
\geq \min_{\hat{\rho}|\mathrm{Tr}\hat{\rho}\hat{M}_x^N = M(\lambda)} \mathrm{Tr}\hat{\rho}\hat{U}^{N:J}
$$
$$
= U_g^{N:J}((1 - \lambda)M_- + \lambda M_+) , \tag{15}
$$

which means the convexity of $U_g^{N:J}(M_x)$. □

Proposition 4.1.7 $U_g^{N:J}(M_x)$ *is the inverse Legendre transformation of* $E_g^{N:J}(\gamma)$.

Proof It is known that if a function f is the Legendre transformation of a convex function g, then the inverse Legendre transformation of f is g [16, 19]. □

We remark that although the domain of M_x is a finite region $[0, N]$ and that of γ is a semi-infinite region $[0, \infty)$, all the aforementioned results are valid for these domains.

4.2 Self-Averaging

In this subsection, we shall show self-averaging of the ground-state energy

$$
\frac{E_g^{N:J}(\gamma)}{N} \sim e_g(\gamma) . \tag{16}
$$

with respect to quenched randomness. Namely, almost all Hamiltonians obtained by random quench have the same ground-state energy density. Self-averaging allows us to discuss quenched systems only by considering the averaged quantity $e_g(\gamma)$, not each $E_g^{N:J}(\gamma)$.

To this end, we divide the system with Hamiltonian $\hat{U}^{N:J}$ into B copies of subsystems with equal size A as $N = AB$. We denote the k-th subsystem by $A[k](k = 1, \ldots, B = N/A)$. We define the Hamiltonian of $A[k]$ denoted by $\hat{U}^{A[k]:J}$ as a restriction of $\hat{U}^{N:J}$ to subsystem $A[k]$ with removing all the bonds from $A[k]$ to outside $A[k]$. We introduce a *block-decomposed Hamiltonian* on the same system, which is a product of $\hat{U}^{A[k]:J}$ denoted by $\hat{U}^{A\#B:J} := \sum_{k=1}^B \hat{U}^{A[k]:J}$. The difference between $\hat{U}^{N:J}$ and $\hat{U}^{A\#B:J}$ is the interaction terms between different subsystems.

The block-decomposed Hamiltonian $\hat{U}^{A\#B:J}$ plays a central role in our proof. In particular, an argument similar to the law of large numbers is applicable to $\hat{U}^{A\#B:J}$, which follows from the fact that $\hat{U}^{A\#B:J}$ is a product of i.i.d. random Hamiltonians; $\hat{U}^{A[k]:J}(k = 1, \ldots, B)$. Since $\hat{U}^{A\#B:J}$ is close to $\hat{U}^{N:J}$, we can derive several self-averaging results in systems with $\hat{U}^{N:J}$. Note that this proof idea is a standard technique in the statistical mechanics of random systems (see Ref [20]).

We first introduce symbols describing an average over quenched randomness and its fluctuation:

Definition 4.2.1 Consider a system with quenched random variables J. Let $X^{N:J}$ be a stochastic variable depending on the quenched random variables J. We denote by $[X^{N:J}]$ the average of $X^{N:J}$ with respect to the quenched randomness J. We also denote its root mean square $\sqrt{[(X^{N:J})^2]}$ by $\overline{\overline{X^{N:J}}}$.

Note that $\overline{\overline{X^{N:J}}}$ is a norm (i.e., it satisfies the triangle inequality), which is a direct consequence of Schwarz inequality.

We first bound the root mean square of operator norms of the system Hamiltonian $\hat{U}^{N:J}$ and the difference between the Hamiltonian and its block-decomposed one; $\hat{U}^{N:J} - \hat{U}^{A\#B:J}$. The latter quantity can be regarded as surface energy.

Proposition 4.2.2 *Suppose that $\hat{U}^{A[k]:J}$ ($k = 1, \ldots, B$) are i.i.d. random Hamiltonians of D-dimensional systems. Then, the operator norm of the bulk energy $\hat{U}^{N:J}$ and the surface energy $\hat{U}^{N:J} - \hat{U}^{A\#B:J}$ are bounded respectively as*

$$\overline{\overline{\left\|\hat{U}^{N:J}\right\|_{\text{op}}}} \le c_{\text{bulk}} N \ ,$$

$$\overline{\overline{\left\|\hat{U}^{N:J} - \hat{U}^{A\#B:J}\right\|_{\text{op}}}} \le c_{\text{surface}} A^{1-1/D} B \ , \tag{17}$$

where c_{bulk} and c_{surface} are constants independent of N, A, and B.

Proof These bounds are direct consequences of the finite dimensionality of the system. The bulk energy is bounded as

$$\overline{\overline{\left\|\hat{U}^{N:J}\right\|_{\text{op}}}} \le \overline{\overline{\sum_{i=1}^{N}\sum_{j=1}^{N}\left|J_{ij}\right|}} \le \sum_{i=1}^{N}\sum_{j=1}^{N}\overline{\overline{J_{ij}}} \le c_{\text{bulk}} N \ , \tag{18}$$

where we used the triangle inequality in the second inequality. The surface energy is bounded as

$$\overline{\overline{\left\|\hat{U}^{N:J} - \hat{U}^{A\#B:J}\right\|_{\text{op}}}} \le \overline{\overline{\sum_{k=1}^{B}\sum_{i\in s(A[k])}\sum_{j\notin s(A[k])}\left|J_{ij}\right|}}$$

$$\le \sum_{k=1}^{B}\sum_{i\in s(A[k])}\sum_{j\notin s(A[k])}\overline{\overline{J_{ij}}}$$

$$\le c_{\text{surface}} A^{1-1/D} B \ , \tag{19}$$

where $s(A[k])$ is a set of sites in subsystem $A[k]$. $\qquad\square$

Now we shall bound the fluctuation of the ground-state energy in terms of quenched randomness. We first show a slightly weak inequality, and then tighten the inequality by applying the obtained inequality iteratively.

Proposition 4.2.3 *The standard deviation of $E_g^{N:J}(\gamma)$ satisfies the bound*

$$\overline{\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]}} \le c_{\text{bulk}} N \ . \tag{20}$$

Proof For any Hamiltonian $\hat{U}^{N:J}$, we have the following bound:

$$
\begin{aligned}
\left| E_g^{N:J}(\gamma) + \gamma N \right| &= \left| \min_{\hat{\rho}} \mathrm{Tr}[\hat{\rho}\hat{H}^{N:J}(\gamma)] - \min_{\hat{\rho}} \mathrm{Tr}[\hat{\rho}(-\gamma \hat{M}_x^N)] \right| \\
&\le \left\| \hat{H}^{N:J}(\gamma) + \gamma \hat{M}_x^N \right\|_{\mathrm{op}} \\
&= \left\| \hat{U}^{N:J} \right\|_{\mathrm{op}} .
\end{aligned}
\tag{21}
$$

Here, the second line follows from an elementary inequality

$$
\left| \min_{\hat{\rho}} \mathrm{Tr}\hat{\rho}\hat{X} - \min_{\hat{\rho}} \mathrm{Tr}\hat{\rho}\hat{Y} \right| \le \max_{\hat{\rho}} \left| \mathrm{Tr}\hat{\rho}\hat{X} - \mathrm{Tr}\hat{\rho}\hat{Y} \right| = \left\| \hat{X} - \hat{Y} \right\|_{\mathrm{op}} .
\tag{22}
$$

Thus we arrive at the desired inequality:

$$
\begin{aligned}
\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} &= \overline{E_g^{N:J}(\gamma) + \gamma N - [E_g^{N:J}(\gamma) + \gamma N]} \\
&\le \overline{E_g^{N:J}(\gamma) + \gamma N} \le \overline{\left\| \hat{U}^{N:J} \right\|_{\mathrm{op}}} \le c_{\mathrm{bulk}} N ,
\end{aligned}
\tag{23}
$$

where the first inequality follows from an elementary fact that $\overline{X - x_0}$ is minimized when $x_0 = [X]$, and the last inequality follows from Proposition 4.2.2. □

The above inequality can be tightened by applying the above result to the block-decomposed system $A\#B$ iteratively.

Proposition 4.2.4 *The fluctuation of $\frac{E_g^{N:J}(\gamma)}{N}$ vanishes in the thermodynamic limit. In particular, for any positive $\epsilon > 0$, we have*

$$
\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} = O(N^{1-1/D+\epsilon}) .
\tag{24}
$$

Proof We start with

$$
\begin{aligned}
\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} &\le \overline{E_g^{N:J}(\gamma) - [E_g^{A\#B:J}(\gamma)]} \\
&\le \overline{E_g^{N:J}(\gamma) - E_g^{A\#B:J}(\gamma)} + \overline{E_g^{A\#B:J}(\gamma) - [E_g^{A\#B:J}(\gamma)]} .
\end{aligned}
\tag{25}
$$

The first inequality follows from that $\overline{X - x_0}$ is minimized when $x_0 = [X]$, and the second inequality follows from the triangle inequality. We first evaluate the first term of the right-hand side of (25) as

$$
\begin{aligned}
\overline{E_g^{N:J}(\gamma) - E_g^{A\#B:J}(\gamma)} &\le \overline{\left\| \hat{H}^{N:J}(\gamma) - \hat{H}^{A\#B:J}(\gamma) \right\|_{\mathrm{op}}} \\
&= \overline{\left\| \hat{U}^{N:J} - \hat{U}^{A\#B:J} \right\|_{\mathrm{op}}} \\
&\le c_{\mathrm{surface}} A^{1-1/D} B .
\end{aligned}
\tag{26}
$$

Here, we used (22) in the first inequality, and used Proposition 4.2.2 in the last inequality. We next bound the second term of the right-hand side of (25). Since $E_g^{A[k]:J}(k = 1, \ldots, B)$ are independent random variables, Proposition 4.2.3 applies to each subsystem, which yields

$$
\overline{E_g^{A\#B:J}(\gamma) - [E_g^{A\#B:J}(\gamma)]}^2 = \sum_{k=1}^{B} \overline{E_g^{A[k]:J}(\gamma) - [E_g^{A[k]:J}(\gamma)]}^2
$$

$$\leq Bc_{\text{bulk}}^2 A^2 . \tag{27}$$

Combining these two inequalities, we obtain

$$\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} \leq c_{\text{surface}} A^{1-1/D} B + c_{\text{bulk}} A B^{1/2} . \tag{28}$$

Setting $A = N^a$ and $B = N^{1-a}$ with $a = D/(D+2)$, we obtain

$$\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} \leq (c_{\text{surface}} + c_{\text{bulk}}) N^{1-1/(D+2)}$$
$$= O(N^{1-1/(D+2)}) . \tag{29}$$

We notice that the above inequality (29) is stronger than (20). Therefore, by replacing (20) in the derivation of (27) by (29) (i.e., we use $\overline{E_g^{A[k]:J}(\gamma) - [E_g^{A[k]:J}(\gamma)]} = O(A^{1-1/(D+2)})$ instead of $\overline{E_g^{A[k]:J}(\gamma) - [E_g^{A[k]:J}(\gamma)]} = c_{\text{bulk}} A = O(A)$ in (27)), we can obtain a further stronger inequality on $\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]}$. By repeating this operation[5], we finally arrive at

$$\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} = O(N^{1-1/D+\epsilon}) . \tag{30}$$

$\square$

In addition, the existence of the ground-state energy density in the thermodynamic limit can be shown.

Proposition 4.2.5 *The averaged ground-state energy density converges in the thermodynamic limit*

$$e_g(\gamma) := \lim_{N \to \infty} \frac{[E_g^{N:J}(\gamma)]}{N} . \tag{31}$$

Moreover, the speed of convergence is evaluated as

$$\frac{[E_g^{N:J}(\gamma)]}{N} - e_g(\gamma) = O(N^{-1/D}) . \tag{32}$$

Proof Since $\hat{U}^{A[k]:J}$ $(k = 1, \ldots, B)$ are i.i.d. random Hamiltonians, we have $[E_g^{A\#B:J}(\gamma)] = B[E_g^{A[1]:J}(\gamma)]$, which implies

$$\left| \frac{[E_g^{N:J}(\gamma)]}{N} - \frac{[E_g^{A[1]:J}(\gamma)]}{A} \right| = \left| \frac{[E_g^{N:J}(\gamma)]}{N} - \frac{[E_g^{A\#B:J}(\gamma)]}{N} \right| \leq c_{\text{surface}} A^{-1/D} . \tag{33}$$

This shows that $a_N := \frac{[E_g^{N:J}(\gamma)]}{N}$ is a Cauchy sequence and hence converges. $\square$

We finally prove the self-averaging of the ground-state energy.

---

[5] Once $\overline{E_g^{A[k]:J}(\gamma) - [E_g^{A[k]:J}(\gamma)]} = O(A^{1-n_m})$ is shown, we can get $\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]} = O(A^{1-1/D}B) + O(A^{1-n_m}B^{1/2}) = O(N^{1-\frac{1}{D+2-2Dn_m}})$ for $a = \frac{D}{D+2-2Dn_m}$. The recurrence formula $n_{m+1} = \frac{1}{D+2-2Dn_m}$ with the initial term $n_0 = 0$ has a limit $\lim_{m \to \infty} n_m = 1/\max(D, 2)$.

Theorem 4.2.6 (Self-averaging of $E_g^{N:J}(\gamma)$) *For any γ, the ground-state energy density $\frac{E_g^{N:J}(\gamma)}{N}$ converges to $e_g(\gamma)$ in mean square:*

$$\overline{\overline{\frac{E_g^{N:J}(\gamma)}{N} - e_g(\gamma)}} = O(N^{-1/D+\epsilon}) . \tag{34}$$

Proof Combining Proposition 4.2.4 with Proposition 4.2.5, we easily have

$$\overline{\overline{E_g^{N:J}(\gamma) - Ne_g(\gamma))}} \leq \overline{\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]}} + \overline{\overline{[E_g^{N:J}(\gamma)] - Ne_g(\gamma)}}$$
$$\leq \overline{\overline{E_g^{N:J}(\gamma) - [E_g^{N:J}(\gamma)]}} + c_{\text{surface}} N^{1-1/D}$$
$$= O(N^{1-1/D+\epsilon}) , \tag{35}$$

which is equivalent to the desired result.   □

4.3 Uniform Self-Averaging

In this subsection, we will show uniform self-averaging (i.e., self-averaging as a function of γ), which is a stronger condition than the self-averaging discussed in the previous subsection. The key idea for the proof of uniform self-averaging is to put many *regularity checkpoints* on the γ-axis. To demonstrate self-averaging for any γ, we employ self-averaging at the nearest regularity checkpoint of γ and evaluate the speed of convergence at γ. Since it is not easy to show uniform self-averaging in the half-infinite region $[0, \infty)$ directly, we set the domain of γ in $E_g^{N:J}(\gamma)$ as $[0, \tilde{\gamma}]$ for $\tilde{\gamma}$ that diverges slowly as N increases.

We start by showing that $E_g^{N:J}(\gamma)$ is Lipschitz continuous.

Proposition 4.3.1 *For any γ_1, γ_2 and any instance, the difference between $E_g^{N:J}(\gamma)$ (and $e_g(\gamma)$) with γ_1 and γ_2 is bounded as*

$$\left| E_g^{N:J}(\gamma_1) - E_g^{N:J}(\gamma_2) \right| \leq N |\gamma_1 - \gamma_2| ,$$
$$\left| e_g(\gamma_1) - e_g(\gamma_2) \right| \leq |\gamma_1 - \gamma_2| . \tag{36}$$

Proof The first inequality of (36) follows from (22) as $\left| E_g^{N:J}(\gamma_1) - E_g^{N:J}(\gamma_2) \right| \leq \left\| \hat{H}^{N:J}(\gamma_1) - \hat{H}^{N:J}(\gamma_2) \right\|_{\text{op}} = N |\gamma_1 - \gamma_2|$. The same argument holds for their mean and under the thermodynamic limit.   □

Theorem 4.3.2 (*Uniform self-averaging of $E_g^{N:J}(\gamma)$*) *The ground-state energy density $\frac{E_g^{N:J}(\gamma)}{N}$ converges uniformly on $[0, \tilde{\gamma}]$ in mean square:*

$$\overline{\overline{\max_{\gamma \in [0,\tilde{\gamma}]} \left| \frac{E_g^{N:J}(\gamma)}{N} - e_g(\gamma) \right|}} = O(N^{-2/(5D)+\epsilon}) , \tag{37}$$

where we set $\tilde{\gamma} = N^{1/(5D)}$.

Proof For convenience, we suppose that $N^{1/D+1/(5D)}$ is an integer. Corresponding to integers $w = 1, \ldots, N^{1/D+1/(5D)}$, we define the regularity checkpoints and their covering intervals as

$$\gamma_w = (w - 1/2)N^{-1/D},$$

$$I_w = \left[(w-1)N^{-1/D}, \, wN^{1/D} \right] . \tag{38}$$

With noting that $|\gamma - \gamma_w| \leq \frac{N^{-1/D}}{2}$ for any $\gamma \in I_w$, we have

$$
\begin{aligned}
& \left| E_g^{N:J}(\gamma) - Ne_g(\gamma) \right| \\
& \leq \left| E_g^{N:J}(\gamma) - E_g^{N:J}(\gamma_w) \right| + \left| E_g^{N:J}(\gamma_w) - Ne_g(\gamma_w) \right| + \left| Ne_g(\gamma_w) - Ne_g(\gamma) \right| \\
& \leq \left| E_g^{N:J}(\gamma_w) - Ne_g(\gamma_w) \right| + N^{1-1/D} ,
\end{aligned}
\tag{39}
$$

where we used Proposition 4.3.1 in the second inequality. Hence the maximum deviation of ground-state energy is bounded as

$$\overline{\overline{\max_{\gamma \in I_w} \left| E_g^{N:J}(\gamma) - Ne_g(\gamma) \right|}} \leq \overline{\overline{E_g^{N:J}(\gamma_w) - Ne_g(\gamma_w)}} + N^{1-1/D} . \tag{40}$$

Using this relation, we arrive at the desired result:

$$
\begin{aligned}
\overline{\overline{\max_{\gamma \in [0, \tilde{\gamma}]} \left| E_g^{N:J}(\gamma) - Ne_g(\gamma) \right|}}^2 &= \overline{\overline{\max_{w=1}^{N^{1/D+1/(5D)}} \max_{\gamma \in I_w} \left| E_g^{N:J}(\gamma) - Ne_g(\gamma) \right|}}^2 \\
& \leq \sum_{w=1}^{N^{1/D+1/(5D)}} \overline{\overline{\max_{\gamma \in I_w} \left| E_g^{N:J}(\gamma) - Ne_g(\gamma) \right|}}^2 \\
& \leq \sum_{w=1}^{N^{1/D+1/(5D)}} \left(\overline{\overline{E_g^{N:J}(\gamma_w) - Ne_g(\gamma_w)}} + N^{1-1/D} \right)^2 \\
& = O(N^{2-1/D+1/(5D)+\epsilon}) .
\end{aligned}
\tag{41}
$$

In the first inequality we used the following simple relation for nonnegative L_w;

$$\overline{\overline{\max_w L_w}}^2 = \left[\max_w \overline{\overline{L_w^2}} \right] \leq \left[\sum_w \overline{\overline{L_w^2}} \right] = \sum_w \overline{\overline{L_w}}^2 , \tag{42}$$

in the second inequality we used (40), and in the last inequality we used Proposition 4.2.6. □

We proceed to the uniform self-averaging of $U_g^{N:J}(Nm_x)$. To prove this, we introduce the inverse Legendre transformation of $e_g(\gamma)$, to which the ground-state energy density $\frac{U_g^{N:J}(Nm_x)}{N}$ converges.

Definition 4.3.3 The inverse Legendre transformation of $e_g(\gamma)$ is defined as

$$u_g(m_x) = \sup_{\gamma \in (-\infty, \infty)} (e_g(\gamma) + \gamma m_x) . \tag{43}$$

Theorem 4.3.4 (*Uniform self-averaging of* $U_g^{N:J}(Nm_x)$) $\frac{U_g^{N:J}(Nm_x)}{N}$ *converges uniformly on* $\left[0, \tilde{m}_x^{N:J} \right]$ *in mean square:*

$$\overline{\overline{\max_{m_x \in [0, \tilde{m}_x^{N:J}]} \left| U_g^{N:J}(Nm_x) - Nu_g(m_x) \right|}} = O(N^{1-2/(5D)+\epsilon}) , \tag{44}$$

where $\tilde{m}_x^{N:J} = 1 - \frac{2}{N^{1+1/(5D)}} \left\| \hat{U}^{N:J} \right\|_{\mathrm{op}}.$

Proof Consider a pair (m_x, γ) satisfying

$$U_g^{N:J}(Nm_x) = E_g^{N:J}(\gamma) + \gamma N m_x . \tag{45}$$

Since $E_g^{N:J}(\gamma)$ is the Legendre transformation of $U_g^{N:J}(Nm_x)$ and its minimum is achieved at $M_x = Nm_x$, we have

$$U_g^{N:J}(Nm_x) - \gamma N m_x \leq U_g^{N:J}(N) - \gamma N , \tag{46}$$

and thus

$$\gamma(1 - m_x) \leq \frac{U_g^{N:J}(N) - U_g^{N:J}(Nm_x)}{N} \leq \frac{2 \left\| \hat{U}^{N:J} \right\|_{op}}{N} \tag{47}$$

holds. It follows that $m_x \leq \tilde{m}_x^{N:J}$, then the corresponding γ with (45) satisfies $\gamma \leq \tilde{\gamma} = N^{1/5D}$. Hence, the domain of γ in the maximization in the Legendre transformation of $E_g^{N:J}(\gamma)$ and $e_g(\gamma)$ can be narrowed to $[0, \tilde{\gamma}]$:

$$U_g^{N:J}(Nm_x) = \max_{\gamma \in [0, \tilde{\gamma}]} (E_g^{N:J}(\gamma) + \gamma N m_x) ,$$
$$u_g(m_x) = \max_{\gamma \in [0, \tilde{\gamma}]} (e_g(\gamma) + \gamma m_x) . \tag{48}$$

Then, the difference between energy of a single instance and its average after taking the thermodynamic limit is evaluated as

$$\left| U_g^{N:J}(Nm_x) - N u_g(m_x) \right|$$
$$= \left| \max_{\gamma \in [0, \tilde{\gamma}]} (E_g^{N:J}(\gamma) + \gamma N m_x) - N u_g(m_x) \right|$$
$$\leq \left| \max_{\gamma \in [0, \tilde{\gamma}]} (N e_g(\gamma) + \gamma N m_x) - N u_g(m_x) \right| + \left| \max_{\gamma \in [0, \tilde{\gamma}]} (E_g^{N:J}(\gamma) - N e_g(\gamma)) \right|$$
$$\leq \max_{\gamma \in [0, \tilde{\gamma}]} \left| E_g^{N:J}(\gamma) - N e_g(\gamma) \right| . \tag{49}$$

Plugging Proposition 4.3.2 into the above inequality, we have the desired result:

$$\overline{\max_{m_x \in [0, \tilde{m}_x^{N:J}]} \left| U_g^{N:J}(Nm_x) - N u_g(m_x) \right|} \leq \overline{\max_{\gamma \in [0, \tilde{\gamma}]} \left| E_g^{N:J}(\gamma) - N e_g(\gamma) \right|}$$
$$= O(N^{1-2/(5D)+\epsilon}) . \tag{50}$$

$\square$

4.4 Antiferromagnetic Fluctuations

Suppose that the x-magnetization M_x shows a first-order phase transition (i.e., discontinuous jump) at some γ. At this point $e_g(\gamma)$ is no longer differentiable, and $u_g(m_x)$ is convex but not strictly convex. Our idea to avoid the first-order phase transition in M_x, based on the above observation, is adding a strictly convex function to $u_g(m_x)$. By construction, the modified $u_g(m_x)$ is strictly convex, and $e_g(\gamma)$ has no singularity.

In particular, we add a quantum antiferromagnetic fluctuations term $\frac{\alpha}{N}(\hat{M}_x^N)^2$ to the Hamiltonian, which we denote by $\hat{H}^{N:J}(\gamma; \alpha)$. Correspondingly, we denote the minimum energy conditioned by M_x by $U_g^{N:J}(M_x; \alpha)$. Then, Theorem 4.3.4 suggests

$$\frac{U_g^{N:J}(Nm_x;\alpha)}{N} \sim u_g(m_x;\alpha) := u_g(m_x) + \alpha m_x^2 . \tag{51}$$

Since $u_g(m_x;\alpha)$ is a strictly convex function, quantum first-order phase transition in M_x does not occur.

A nontrivial step in the aforementioned proof outline is connecting $\langle(\hat{M}_x^N)^2\rangle$ and $\langle\hat{M}_x^N\rangle^2$, since these two are in general not equal; $\langle(\hat{M}_x^N)^2\rangle \neq \langle\hat{M}_x^N\rangle^2$. In fact, these two are inequivalent in some long-range interacting systems (e.g., p-spin model with large p, discussed in [7]). On the other hand, we can prove $\langle(\hat{M}_x^N)^2\rangle \simeq \langle\hat{M}_x^N\rangle^2$ in short-range interacting systems. This is the main task in this subsection. We note that our argument does not hold for $\alpha < 0$.

Definition 4.4.1 For $\alpha \in (0,\infty)$, we introduce Hamiltonians and related quantities corresponding to QA-AFF denoted by

$$\hat{H}^{N:J}(\gamma;\alpha) = \hat{H}^{N:J}(\gamma) + \frac{\alpha}{N}(\hat{M}_x^N)^2 ,$$

$$\hat{U}^{N:J}(;\alpha) = \hat{U}^{N:J} + \frac{\alpha}{N}(\hat{M}_x^N)^2 ,$$

$$U_g^{N:J}(M_x;\alpha) = \min_{\hat{\rho}|\text{Tr}\hat{\rho}\hat{M}_x^N=M_x} \text{Tr}\hat{\rho}\hat{U}^{N:J}(;\alpha) ,$$

$$u_g(m_x;\alpha) = u_g(m_x) + \alpha m_x^2 . \tag{52}$$

We first prove the uniform self-averaging of $U_g^{N:J}(Nm_x;\alpha)$.

Proposition 4.4.2 *The ground-state energy density* $\frac{U_g^{N:J}(Nm_x;\alpha)}{N}$ *converges uniformly to* $u_g(m_x;\alpha)$ *on* $[0, \tilde{m}_x^{N:J}]$ *in mean square:*

$$\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]} \left| \frac{U_g^{N:J}(Nm_x;\alpha)}{N} - u_g(m_x;\alpha) \right|} = O(N^{-2/(5D)+\epsilon}) . \tag{53}$$

Proof We first derive the following bound:

$$\left| U_g^{A\#B:J}(Nm_x;\alpha) - U_g^{A\#B:J}(Nm_x) - \alpha Nm_x^2 \right| \le \alpha A . \tag{54}$$

The elementary inequality $\text{Tr}\hat{\rho}(\hat{M}_x^N)^2 - (\text{Tr}\hat{\rho}\hat{M}_x^N)^2 \ge 0$ implies

$$U_g^{A\#B:J}(Nm_x;\alpha) \ge U_g^{A\#B:J}(Nm_x) + \alpha Nm_x^2 . \tag{55}$$

Fix $\hat{\rho}' \in \arg\min_{\hat{\rho}|\text{Tr}\hat{\rho}\hat{M}_x^N=Nm_x} \text{Tr}\hat{\rho}\hat{U}^{A\#B:J}$, which has x-magnetization Nm_x and minimizes the energy $\hat{U}^{A\#B:J}$. We construct a state from $\hat{\rho}'$ by removing all the correlation between subsystems:

$$\hat{\rho}_k = \text{Tr}_{i\notin s(A[k])}\hat{\rho}' ,$$

$$\hat{\rho}_\otimes = \bigotimes_{k=1}^{B} \hat{\rho}_k . \tag{56}$$

By construction, $\hat{\rho}_\otimes$ is separable into subsystems, and $\hat{\rho}_\otimes$ also minimizes $\hat{U}^{A\#B:J}$ with x-magnetization equal to Nm_x: $\hat{\rho}_\otimes \in \arg\min_{\hat{\rho}|\text{Tr}\hat{\rho}\hat{M}_x^N=Nm_x} \text{Tr}\hat{\rho}\hat{U}^{A\#B:J}$. The AFF term in $\hat{\rho}_\otimes$ can be directly estimated as follows:

$$\text{Tr}\hat{\rho}_{\otimes}(\hat{M}_x^N)^2$$

$$= \sum_{k=1}^{B}\sum_{l=1}^{B}\text{Tr}\hat{\rho}_{\otimes}\hat{M}_x^{A[k]}\hat{M}_x^{A[l]}$$

$$= \sum_{k=1}^{B}\sum_{l=1}^{B}\text{Tr}\hat{\rho}_{\otimes}\hat{M}_x^{A[k]}\text{Tr}\hat{\rho}_{\otimes}\hat{M}_x^{A[l]} + \sum_{k=1}^{B}\left(\text{Tr}\left[\hat{\rho}_{\otimes}(\hat{M}_x^{A[k]})^2\right] - \left(\text{Tr}\hat{\rho}_{\otimes}(\hat{M}_x^{A[k]})\right)^2\right)$$

$$\leq N^2 m_x^2 + BA^2 , \tag{57}$$

which implies a relation evaluating the difference between $U_g^{A\#B:J}$ with and without the AFF term:

$$U_g^{A\#B:J}(Nm_x) = \text{Tr}\hat{\rho}_{\otimes}\hat{U}^{A\#B:J}$$

$$= \text{Tr}\hat{\rho}_{\otimes}\hat{U}^{A\#B:J}(Nm_x;\alpha) - \frac{\alpha}{N}\text{Tr}\left[\hat{\rho}_{\otimes}(\hat{M}_x^N)^2\right]$$

$$\geq U_g^{A\#B:J}(Nm_x;\alpha) - \alpha Nm_x^2 - \alpha A . \tag{58}$$

Using the two inequalities (55) and (58), we arrive at (54).

Now we fix $A = N^{a'}$ with $a' = 1/(D+1)$ in (54). Combining (54), Propositions 4.2.2 and 4.3.4 with recalling $Nu_g(m_x;\alpha) = Nu_g(m_x) + \alpha Nm_x^2$, we obtain the desired result:

$$\overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{N:J}(Nm_x;\alpha) - Nu_g(m_x;\alpha)\right|}}$$

$$\leq \overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{N:J}(Nm_x;\alpha) - U_g^{A\#B:J}(Nm_x;\alpha)\right|}}$$

$$+ \overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{A\#B:J}(Nm_x;\alpha) - U_g^{A\#B:J}(Nm_x) - \alpha Nm_x^2\right|}}$$

$$+ \overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{A\#B:J}(Nm_x) - U_g^{N:J}(Nm_x)\right|}}$$

$$+ \overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{N:J}(Nm_x) - Nu_g(m_x)\right|}}$$

$$\leq \overline{\overline{\max_{m_x\in[0,\tilde{m}_x^{N:J}]}\left|U_g^{N:J}(Nm_x) - Nu_g(m_x)\right|}} + \alpha A + 2c_{\text{surface}}A^{1-1/D}B$$

$$= O(N^{1-2/(5D)+\epsilon}) . \tag{59}$$

Here, the first inequality follows from the triangle inequality. $\square$

Definition 4.4.3 We denote by $m_*(\gamma;\alpha)$ the unique argument that minimizes $u_g(m_x;\alpha) - \gamma m_x$. We also define $M_*^{N:J}(\gamma;\alpha)$ as the expectation value of $\hat{M}_x^N$ in a ground state of $\hat{H}^{N:J}(\gamma;\alpha)$ (i.e., $\langle g|\hat{M}_x^N|g\rangle$, where $|g\rangle$ is a ground state of $\hat{H}^{N:J}(\gamma;\alpha)$). For brevity, we sometimes drop the arguments γ and α in $m_*(\gamma;\alpha)$ and $M_*^{N:J}(\gamma;\alpha)$, and simply write m_* and $M_*^{N:J}$.

Since $u_g(m_x; \alpha)$ is a strictly convex function, $m_*(\gamma; \alpha)$ is a continuous function[6] with respect to γ for any $\alpha > 0$. We shall show that $\frac{M_*^{N:J}}{N}$ converges to a continuous function $m_*(\gamma; \alpha)$, which completes the proof of our main result.

Theorem 4.4.4 *(Absence of the QFOT)* $\frac{M_*^{N:J}}{N}$ *(as a function of γ) converges uniformly on* $[0, \infty)$ *in mean square:*

$$\overline{\overline{\sup_{\gamma \in [0,\infty)} \max_{|g\rangle \in G(\hat{H}^{N:J}(\gamma;\alpha))} \left| \frac{\langle g|\hat{M}_x^N|g\rangle}{N} - m_*(\gamma;\alpha) \right|}} = O(N^{-1/(5D)+\epsilon}) . \tag{60}$$

Proof Note that

$$\overline{\overline{1 - \tilde{m}_x^{N:J}}} = \frac{2 \overline{\overline{\left\| \hat{U}^{N:J} \right\|_{\text{op}}}}}{N^{1+1/(5D)}} \leq \frac{2c_{\text{bulk}}}{N^{1/(5D)}} . \tag{61}$$

We decompose the domain of M_x, $[0, N]$, into two regions, $I_1 := [0, N\tilde{m}_x^{N:J}]$ and $I_2 := [N\tilde{m}_x^{N:J}, N]$. Promising $\sup \emptyset = 0$ for convenience, we evaluate the square of the left-hand side of (60) (multiplied by N) as

$$\overline{\overline{\sup_{\gamma \in [0,\infty)} \max_{|g\rangle \in G(\hat{H}^{N:J}(\gamma;\alpha))} \left| \langle g|\hat{M}_x^N|g\rangle - Nm_*(\gamma;\alpha) \right|}}^2$$

$$\leq \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_1, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - Nm_* \right|}}^2 + \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_1, M_*^{N:J} \in I_2}} \left| M_*^{N:J} - Nm_* \right|}}^2$$

$$+ \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - Nm_* \right|}}^2 + \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_2}} \left| M_*^{N:J} - Nm_* \right|}}^2 . \tag{62}$$

Here, we used (42). We shall evaluate these four terms.

Before going to the evaluation, we introduce a useful relation that if functions P, Q, R, S, T satisfy $P \leq Q$ and $S + T^2 \leq R$, then

$$\overline{\overline{\max |T|}}^2 \leq \overline{\overline{\max |P - R|}} + \overline{\overline{\max |Q - S|}} \tag{63}$$

is satisfied. This relation is easily confirmed as

$$\overline{\overline{\max |T|}}^2 \leq [\max |R - S|]$$
$$\leq [\max(|P - R| + |Q - S|)]$$
$$\leq [\max |P - R|] + [\max |Q - S|]$$
$$\leq \overline{\overline{\max |P - R|}} + \overline{\overline{\max |Q - S|}} . \tag{64}$$

Now we evaluate the four terms in (62). To evaluate the first term of (62), we consider

$$P = U_g^{N:J}(M_*^{N:J}; \alpha) - \gamma M_*^{N:J} ,$$
$$Q = U_g^{N:J}(Nm_*; \alpha) - \gamma Nm_* ,$$

---

[6] In fact, $m_*(\gamma; \alpha)$ is Lipschitz continuous with constant $1/(2\alpha)$, which means that there is no quantum second-order transition in transverse magnetization either.

$$R = N u_g \left(\frac{M_*^{N:J}}{N}; \alpha \right) - \gamma M_*^{N:J},$$

$$S = N u_g (m_*; \alpha) - \gamma N m_*,$$

$$T = \left(\frac{\alpha}{N} \right)^{1/2} \left| M_*^{N:J} - N m_* \right|. \tag{65}$$

Here, $P \leq Q$ follows from the fact that $M_*^{N:J}$ is the minimizer of $U_g^{N:J}(M_x; \alpha) - \gamma M_x$. Also, $S + T^2 \leq R$ follows from the fact that m_* is the minimizer of $u_g(m_x; \alpha) - \gamma m_x$ and that the second derivative of $u_g(m_x; \alpha)$ is at least 2α. Thus, it follows from (63) that

$$\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ N m_* \in I_1, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - N m_* \right|}^2$$

$$\leq \frac{N}{\alpha} \overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ N m_* \in I_1, M_*^{N:J} \in I_1}} \left| U_g^{N:J}(M_*^{N:J}; \alpha) - N u_g \left(\frac{M_*^{N:J}}{N}; \alpha \right) \right|}$$

$$+ \frac{N}{\alpha} \overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ N m_* \in I_1, M_*^{N:J} \in I_1}} \left| U_g^{N:J}(N m_*; \alpha) - N u_g(m_*; \alpha) \right|}$$

$$\leq \frac{N}{\alpha} \overline{\sup_{\gamma, |g\rangle | M_*^{N:J} \in I_1} \left| U_g^{N:J}(M_*^{N:J}; \alpha) - N u_g \left(\frac{M_*^{N:J}}{N}; \alpha \right) \right|}$$

$$+ \frac{N}{\alpha} \overline{\sup_{\gamma, |g\rangle | N m_* \in I_1} \left| U_g^{N:J}(N m_*; \alpha) - N u_g(m_*; \alpha) \right|}$$

$$= O(N^{2-2/(5D)+\epsilon}). \tag{66}$$

In the last line, we used Proposition 4.3.4.

To evaluate the second term of (62), we use (63) with

$$P = U_g^{N:J}(N \tilde{m}_x^{N:J}; \alpha) - \gamma N \tilde{m}_x^{N:J},$$

$$Q = U_g^{N:J}(N m_*; \alpha) - \gamma N m_*,$$

$$R = N u_g(\tilde{m}_x^{N:J}; \alpha) - \gamma N \tilde{m}_x^{N:J},$$

$$S = N u_g(m_*; \alpha) - \gamma N m_*,$$

$$T = \left(\frac{\alpha}{N} \right)^{1/2} \left| N \tilde{m}_x^{N:J} - N m_* \right|. \tag{67}$$

Here, $P \leq Q$ follows from (i) the fact that $M_*^{N:J}$ is the minimizer of $U_g^{N:J}(M_x; \alpha) - \gamma M_x$, (ii) the convexity of $U_g^{N:J}(M_x; \alpha) - \gamma M_x$, and (iii) an assumption that $N m_* \leq N \tilde{m}_x^{N:J} \leq M_*$. Also, $S + T^2 \leq R$ holds by the same argument as in the first term. Thus, it follows from (63) that

$$\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ N m_* \in I_1, M_*^{N:J} \in I_2}} \left| N \tilde{m}_x^{N:J} - N m_* \right|}^2$$

$$\leq \frac{N}{\alpha} \overline{\left| U_g^{N:J}(N \tilde{m}_x^{N:J}; \alpha) - N u_g(\tilde{m}_x^{N:J}; \alpha) \right|}$$

$$+ \frac{N}{\alpha} \overline{\overline{\sup_{\gamma, |g\rangle |Nm_* \in I_1} \left| U_g^{N:J}(Nm_*; \alpha) - Nu_g(m_*; \alpha) \right|}}$$

$$= O(N^{2-2/(5D)+\epsilon}) . \tag{68}$$

In the last line, we used Proposition 4.3.2 and the fact that $N\tilde{m}_x^{N:J} \in I_1$. Using this inequality and (61), we have

$$\overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_1, M_*^{N:J} \in I_2}} \left| M_*^{N:J} - Nm_* \right|}}$$

$$\leq \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_1, M_*^{N:J} \in I_2}} \left| N\tilde{m}_x^{N:J} - Nm_* \right|}} + \overline{N(1 - \tilde{m}_x^{N:J})}$$

$$= O(N^{1-1/(5D)+\epsilon}) . \tag{69}$$

To evaluate the third term of (62), we use (63) with

$$P = U_g^{N:J}(M_*^{N:J}; \alpha) - \gamma M_*^{N:J} ,$$
$$Q = U_g^{N:J}(N\tilde{m}_x^{N:J}; \alpha) - \gamma N\tilde{m}_x^{N:J} ,$$
$$R = Nu_g(M_*^{N:J}/N; \alpha) - \gamma M_*^{N:J} ,$$
$$S = Nu_g(\tilde{m}_x^{N:J}; \alpha) - \gamma N\tilde{m}_x^{N:J} ,$$
$$T = \left(\frac{\alpha}{N}\right)^{1/2} \left| M_*^{N:J} - N\tilde{m}_x^{N:J} \right| . \tag{70}$$

Here, $P \leq Q$ holds by the same argument as in the first term. Also, $S + T^2 \leq R$ follows from (i) the fact that m_* is the minimizer of $u_g(m_x; \alpha) - \gamma m_x$, (ii) the fact that the second derivative of $u_g(m_x; \alpha)$ is at least 2α and (iii) an assumption that $M_* \leq N\tilde{m}_x^{N:J} \leq Nm_*$. Thus, it follows from (63) that

$$\overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - N\tilde{m}_x^{N:J} \right|^2}}$$

$$\leq \frac{N}{\alpha} \overline{\overline{\sup_{\gamma, |g\rangle |M_*^{N:J} \in I_1} \left| U_g^{N:J}(M_*^{N:J}; \alpha) - Nu_g\left(\frac{M_*^{N:J}}{N}; \alpha\right) \right|}}$$

$$+ \frac{N}{\alpha} \overline{\overline{\left| U_g^{N:J}(N\tilde{m}_x^{N:J}; \alpha) - Nu_g(\tilde{m}_x^{N:J}; \alpha) \right|}}$$

$$= O(N^{2-2/(5D)+\epsilon}) . \tag{71}$$

Consequently, we have

$$\overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - Nm_* \right|}}$$

$$\leq \overline{\overline{\sup_{\substack{\gamma, |g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_1}} \left| M_*^{N:J} - N\tilde{m}_x^{N:J} \right|}} + \overline{N(1 - \tilde{m}_x^{N:J})}$$

$$= O(N^{1-1/(5D)+\epsilon}) . \tag{72}$$

The last term of (62) is simply bounded as

$$\overline{\overline{\sup_{\substack{\gamma,|g\rangle \text{ s.t.} \\ Nm_* \in I_2, M_*^{N:J} \in I_2}} \left| M_*^{N:J} - Nm_* \right|}} \leq \overline{N(1 - \tilde{m}_x^{N:J})} = O(N^{1-1/(5D)}) \,. \tag{73}$$

Combining these four inequalities, we complete the proof of Theorem 2.2.2. □

5 Discussion

We have proved that the quantum annealing (QA) for finite-dimensional spin-glass systems does not show a quantum first-order transition in transverse magnetization (QFOT) by adding the antiferromagnetic fluctuations (AFF) term. This result holds for any spin-glass system as long as the system is in finite dimensions and its quenched randomness is sampled from a shift-invariant probability distribution. For simplicity of explanation, we assume that the interaction in $\hat{U}^{N:J}$ is two-body and the boundary is an open boundary condition, but our result applies to more general systems. In fact, our proof relies only on Proposition 4.2.2 (finite dimensionality) and the fact that subsystems are i.i.d. Thus, if these two conditions are satisfied, our result also applies to systems with the periodic and closed boundary conditions as well as those with local fields and short-range p-body interactions.

Key ideas in our proof
We first elucidate the power of uniform self-averaging. Uniform self-averaging is an important concept for discussing the absence of phase transitions. Applying an argument analogous to Chebyshev's inequality, we show that the function $\frac{M_*^{N:J}(\gamma)}{N}$ is in the sup-norm neighborhood of the function $m_*(\gamma)$ for almost all J. Conventional self-averaging alone, which corresponds to pointwise convergence, cannot eliminate the possibility that there is a discontinuous jump in each instance with different transition points depending on instances. On the other hand, uniform self-averaging indeed prohibits this unwanted possibility.

Next, we discuss the role of the AFF term. Thanks to the description with $U_g^{N:J}(M_x)$, our approach makes the meaning of the AFF term $((\hat{M}_x^N)^2$ term) much clearer than in the original paper of QA-AFF [7]. Namely, the AFF term strengthens the convexity in $u_g(m_x)$ and ensures that it is strictly convex, not merely convex, and this fact shows that the transverse magnetization is continuous with respect to γ. Similar arguments can be seen in some papers in statistical mechanics [21, 22], where the difficulties associated with first-order phase transitions are solved by devising the shape of the ensemble. However, we should notice the discrepancy $\langle(\hat{M}_x^N)^2\rangle \neq \langle\hat{M}_x^N\rangle^2$, which prevents a direct analogous argument. In particular, the procedure to obtain a narrowly convex function $u_g(m_x; \alpha)$ as presented in Sect. 4.4 does not always work well for long-range interacting systems, e.g., p-spin model with large p [7].

Implications to the hardness of QA
It is numerically well known that the QA for hard combinatorial optimization problems fails at some point in the QA process. As explained in the Introduction, the role of the QFOT in the failure of the QA is controversial. Our result says that the QFOT in QA for finite-dimensional spin glasses can be removed by adding the AFF term. We expect that the QFOT in *any* extensive sum of local observables $\hat{A}$ is avoidable by a slight modification of QA. If the observable $\hat{A}$ does not contain z-magnetization $\hat{\sigma}^z$, a slight extension of our argument leads to the desired consequence by simply adding the fluctuations term $\frac{\hat{A}^2}{N}$. On the other

hand, if $\hat{A}$ contains z-magnetization, our estimation $\langle \hat{A}^2 \rangle = O(1)$ for the optimal solution no longer holds, and some additional ideas are necessary, which is left for future research.

We emphasize that our result does not claim that the QA in finite-dimensional spin glasses succeeds and that the corresponding ground-state search problem can be efficiently solved. A more plausible scenario suggested by our result is that the QA in finite-dimensional spin glasses fails for different reasons from the QFOT. One candidate is the *glassy bottlenecks*, which are undetectable by the usual macroscopic observables. It is shown in Ref [13] that some models with a transverse magnetic field have exponentially small gaps in the glass phase rather than at the phase transition point. The arrangement of the ground state from one glassy state to another glassy state can make QA less efficient. Our result supports this picture.

However, it should be clarified that two statements can be reconciled: (i) the ground-state search problem for finite-dimensional spin glasses is efficiently solvable by QA-AFF, and (ii) no quantum computer can solve NP-hard problems efficiently. This apparent contradiction is resolved for the following reason: Statement (i) concerns the average-case hardness, which means that almost all instances of spin glasses can be solved efficiently. In contrast, statement (ii) concerns the worst-case hardness, claiming that for any quantum computer, there exists at least one instance that cannot be solved efficiently. Hence, it is possible that the ground state search problem for finite-dimensional spin glasses, which is an NP-hard problem, is typically easy and rarely has hard instances.

Future works

Note that it is difficult to simulate a QA-AFF classically because the AFF term is non-stoquastic [23] and gives rise to a negative sign problem. It is an open question whether we can avoid quantum first-order phase transition only by adding stoquastic terms to a QA.

Also, the exact minimization problem for finite-dimensional spin glass for $D \geq 3$ is NP-hard, though it is easy to solve the minimization problem that allows any errors proportional to the system size N. On the other hand, the PCP theorem [24–26] implies that there are problems for which even approximate minimization is NP-hard. It is an open question whether there exist QA-AFF or other extensions of QA for such problems where no quantum first-order phase transitions occur.

Data availability No datasets were generated or analyzed during the current study.

Declarations

Conflict of interest The authors declare no competing interest.

References

1. Kirkpatrick, S., Gelatt, C.D., Jr., Vecchi, M.P.: Optimization by simulated annealing. Science **220**(4598), 671–680 (1983)
2. Kadowaki, T., Nishimori, H.: Quantum annealing in the transverse Ising model. Phys. Rev. E **58**(5), 5355 (1998)
3. Farhi, E., Goldstone, J., Gutmann, S., Sipser, M.: Quantum computation by adiabatic evolution (2000). arXiv:quant-ph/0001106
4. Kato, T.: On the adiabatic theorem of quantum mechanics. J. Phys. Soc. Jpn. **5**(6), 435–439 (1950)
5. Albash, T., Lidar, D.A.: Adiabatic quantum computation. Rev. Mod. Phys. **90**(1), 015002 (2018)
6. Jörg, T., Krzakala, F., Kurchan, J., Maggs, A.C., Pujos, J.: Energy gaps in quantum first-order mean-field-like transitions: the problems that quantum annealing cannot solve. Europhys. Lett. **89**(4), 40004 (2010)
7. Seki, Y., Nishimori, H.: Quantum annealing with antiferromagnetic fluctuations. Phys. Rev. E **85**(5), 051112 (2012)
8. Jörg, T., Krzakala, F., Semerjian, G., Zamponi, F.: First-order transitions and the performance of quantum algorithms in random optimization problems. Phys. Rev. Lett. **104**(20), 207206 (2010)
9. Seki, Y., Nishimori, H.: Quantum annealing with antiferromagnetic transverse interactions for the Hopfield model. J. Phys. A **48**(33), 335301 (2015)
10. Young, A.P., Knysh, S., Smelyanskiy, V.N.: Size dependence of the minimum excitation gap in the quantum adiabatic algorithm. Phys. Rev. Lett. **101**(17), 170503 (2008)
11. Young, A., Knysh, S., Smelyanskiy, V.: First-order phase transition in the quantum adiabatic algorithm. Phys. Rev. Lett. **104**(2), 020502 (2010)
12. Farhi, E., Gosset, D., Hen, I., Sandvik, A., Shor, P., Young, A., Zamponi, F.: Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs. Phys. Rev. A **86**(5), 052334 (2012)
13. Knysh, S.: Zero-temperature quantum annealing bottlenecks in the spin-glass phase. Nat. Commun. **7**(1), 12370 (2016)
14. Takahashi, J., Hukushima, K.: Phase transitions in quantum annealing of an np-hard problem detected by fidelity susceptibility. J. Stat. Mech. **2019**(4), 043102 (2019)
15. Barahona, F.: On the computational complexity of Ising spin glass models. J. Phys. A **15**(10), 3241 (1982)
16. Shimizu, A.: Principles of Thermodynamics, vol. I, 2nd edn. University of Tokyo Press, Tokyo (2021). (**in Japanese**)
17. Callen, H.B.: Thermodynamics and an Introduction to Thermostatistics. Wiley, New Jersey (1991)
18. Chatterjee, S.: Absence of replica symmetry breaking in the random field Ising model. Commun. Math. Phys. **337**(1), 93–102 (2015)
19. Rockafellar, R.T.: Convex Analysis, vol. 18. Princeton University Press, New Jersey (1970)
20. Castellani, T., Cavagna, A.: Spin-glass theory for pedestrians. J. Stat. Mech. **2005**(05), 05012 (2005)
21. Hetherington, J.: Solid 3 he magnetism in the classical approximation. J. Low Temp. Phys. **66**, 145–154 (1987)
22. Yoneta, Y., Shimizu, A.: Squeezed ensemble for systems with first-order phase transitions. Phys. Rev. B **99**(14), 144105 (2019)
23. Albash, T.: Role of nonstoquastic catalysts in quantum adiabatic optimization. Phys. Rev. A **99**(4), 042334 (2019)
24. Arora, S., Lund, C., Motwani, R., Sudan, M., Szegedy, M.: Proof verification and the hardness of approximation problems. J. ACM **45**(3), 501–555 (1998)
25. Dinur, I.: The PCP theorem by gap amplification. J. ACM **54**(3), 12 (2007)
26. Håstad, J.: Some optimal inapproximability results. J. ACM **48**(4), 798–859 (2001)