Statistica Sinica Preprint No: SS-2021-0228	
Title	Statistical Analysis of Quantum Annealing
Manuscript ID	SS-2021-0228
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202021.0228
Complete List of Authors	Yazhen Wang, Shang Wu and Hongzhi Liu
Corresponding Author	Shang Wu
E-mail	shangwu@fudan.edu.cn

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Statistica Sinica

STATISTICAL ANALYSIS OF QUANTUM ANNEALING

Yazhen Wang[1], Shang Wu[2], and Hongzhi Liu[1]

University of Wisconsin-Madison[1] *and Fudan University*[2]

October 1, 2021

To appear in Statistica Sinica

Abstract: Quantum computation is based on quantum physics to build quantum devices for performing calculations and processing information. Although large-scale general-purpose quantum computers are still many years away, special-purpose quantum computers, such as quantum annealers, are being built with capabilities that exceed those of classical computers. The quantum annealers are created to realize quantum annealing. This study explores quantum annealing and investigates its statistical properties. As such, we establish a lower bound on the probability of quantum annealing solving optimization problems. Furthermore, we examine physical devices and Monte Carlo simulations to implement quantum annealing and expand our understanding of the quantum annealing process.

Key words and phrases: Annealing, Combinatorial optimization, Hamiltonian, Ising model, Markov chain Monte Carlo, Quantum computation.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

1. Introduction

As an optimization method, classical annealing is based on an analogy between the energy behavior of a complex physical system and the objective function of an optimization problem. Treating the objective function as the energy of the physical system, we convert the problem of minimizing the objective function into one of searching for minimum energy configurations (called ground states) of the physical system.

Simulated annealing (SA) is a well-known computer-based Monte Carlo simulation that mimics a system's behavior in order to find its minimum energy configurations. The SA scheme is as follows. After identifying the objection function of the minimization problem with the energy of the physical system, we assign the physical system a temperature as an artificially introduced control parameter. We select an initial temperature that is high relative to the system energy scale in order to induce thermal fluctuations and sample the energy configurations. Markov chain Monte Carlo (MCMC) simulations are used to perform energy sampling and probabilistically explore the extremely large search space. As we decrease the temperature gradually from the initial value to zero, the system is driven to a state with the lowest energy value—namely, the minimum of the objective function—and, thus, we obtain a solution to the optimization problem; see Bertsimas

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

and Tsitsiklis (1993), Kirkpatrick et al. (1983), and Winker (2001) for more details.

Quantum annealing is the quantum analog of classical annealing. It is based on the process of a quantum physical system, the lowest energy of which provides a solution to the minimization problem under study. The quantum states corresponding to the lowest energy are called the ground states of the system. Quantum annealing starts with a simple quantum system initialized in its ground state, which is then drive slowly toward the target complex system. According to the adiabatic quantum theorem (Farhi et al. (2000), Farhi et al. (2001), and Farhi et al. (2002)), when a quantum system is initialized in its ground state and then gradually evolves, it has a tendency to stay at a ground state. Hence, at the end of the quantum annealing process, measuring the state of the quantum system renders an answer to the original optimization problem with a certain probability. More details can be found in Boixo et al. (2014), Brooke et al. (1999), Santoro et al. (2002), and Wang et al. (2016).

Both classical annealing and quantum annealing are powerful techniques for solving difficult optimization problems, whether they are used in computer-based simulations or physical machines. The simulation approach applies "escape" rules in Monte Carlo simulations to prevent the

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

system from becoming trapped in local minima of an energy (or objective) function, and eventually drives the system toward its lowest energy state with a certain probability. The physical scheme uses a physical system or builds a device to engineer a physical system, the ground states of which represent the sought-after solution of an optimization problem.

The systems in both situations enable a probabilistic exploration of their extremely large configuration spaces, and ultimately "freeze" in the global minima with a certain probability. With enough repeated tries, each approach can find the global minimum and solve the optimization problem. The key difference between classical annealing and quantum annealing is the thermal hopping used in classical annealing and the quantum tunneling used in quantum annealing to escape from local minima and reach the global minimum. Quantum annealing has long been studied in quantum computation for building specialized quantum computers, such as D-Wave annealers (Britton et al. (2012), Feynman (1982), Hu and Wang (2021), Martoňák et al. (2002), McGeoch (2014), Nielsen and Chuang (2010), Wang (2012), Wang et al. (2016), Wang and Song (2020), Wang (2021), and Wang and Liu (2022)). Here, we investigate quantum annealing from a statistical viewpoint. We provide a lower bound on the probability of the quantum annealing system staying at a ground state at the end of the

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

quantum annealing process, where the probability often refers to the success probability of quantum annealing. We also discuss quantum annealing implementations and illustrate quantum tunneling through the lens of data augmentation.

The rest of the paper proceeds as follows. Section 2 briefly reviews the classical Ising model and SA. Section 3 explores quantum annealing. Here, we establish its statistical properties and discuss its implementations using D-Wave devices and MCMC-based methods in the context of the Ising model. Section 4 concludes the paper. Here, we also discuss topics for future research on quantum annealing. In particular, we point out that classical computer-based MCMC simulations of quantum annealing can be used to provide intuitive explanations of quantum tunneling in quantum annealing through a data augmentation connection. All proofs are relegated to the Supplementary Appendix.

2. The Ising model and simulated annealing

We describe the Ising model using a graph $\mathcal{G}$, with site and edge sets denoted by $\mathcal{V}$ and $\mathcal{E}$, respectively. Each site is associated with a random variable taking values in $\{+1, -1\}$, and each edge specifies the interaction (or coupling) between the random variables on the two sites connected by

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

the edge. Denote by d the total number of sites in $\mathcal{G}$. For example, we may consider $\mathcal{G}$ as a lattice with d sites. Denote a configuration or state by $s = (s_1, s_2, \ldots, s_d)$, that is, a d-dimensional vector with site variables $s_i = \pm 1$. The classical Ising model has the following Hamiltonian (or energy):

$$\mathbf{H}_I^c(s) = -\sum_{\langle i,j\rangle \in \mathcal{E}} J_{ij} s_i s_j - \sum_{i \in \mathcal{V}} h_i s_i, \tag{2.1}$$

where J_{ij} denotes the strength of the interaction between sites i and j associated with edge $\langle i, j\rangle$ in graph $\mathcal{G}$, and h_i gives the strength of the external local fields imposed on site i. We refer to a set of fixed values $\{J_{ij}, h_i\}$ as one instance of the Ising model. The Boltzmann (or Gibbs) distribution specifies the probability of a given configuration s, as follows:

$$P_\beta(s) = \frac{e^{-\beta \mathbf{H}_I^c(s)}}{Z_\beta} \quad \text{and} \quad Z_\beta = \sum_s e^{-\beta \mathbf{H}_I^c(s)}, \tag{2.2}$$

where $\beta = \frac{1}{k_B T}$, T denotes the absolute temperature of the system, and k_B is a generic physical constant called the Boltzmann constant. The normalization constant Z_β refers to the partition function of the Boltzmann distribution.

We can represent a combinatorial optimization problem using the Ising model, with its objective function corresponding to the Hamiltonian $\mathbf{H}_I^c(s)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Minimizing the objective function is equivalent to finding a configuration s^* with the minimum energy; that is, s^* minimizes the Hamiltonian $\mathbf{H}_I^c(s)$ over all s. We refer to the configuration s^* as the ground state of the Ising model.

Combinatorial optimization is computationally very hard, because there are 2^d configurations in the search space, which has an exponential increase in the system size d. SA is often applied to solve such optimization problems. It involves a temperature $T = T(t)$ as a decreasing function of the time t. A relatively high initial temperature $T(0)$ is set to induce thermal fluctuations and facilitate an exploration of the large search space. As the temperature gradually decreases, MCMC simulations are used to sample the configurations. Eventually, we drive the system to a ground state, which renders a solution to the minimization problem.

3. Quantum annealing

3.1 Theoretical analysis

We describe a quantum system by its quantum state and the dynamic evolution of the state, where the quantum state is characterized by a unit vector in a complex vector space, and the dynamic evolution of the state is governed by a Hermitian matrix called the quantum Hamiltonian. To spec-

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

ify a quantum Hamiltonian for the quantum system that drives quantum annealing, we need to introduce some notation. Define

$$I_j = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_j^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \text{and } \sigma_j^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad j = 1, \ldots, d, \tag{3.1}$$

where σ_j^x and σ_j^z are Pauli matrices in the x and z axes, respectively. The matrices and site index j serve as the quantum counterparts of the values $\{+1, -1\}$ for the binary random variables associated with the sites in the classical Ising model.

Quantum annealing has the following quantum Hamiltonian:

$$\mathbf{H}_{QA}(t) = A(t)\mathbf{H}_I^q + B(t)\mathbf{H}_X, \tag{3.2}$$

$$\mathbf{H}_I^q = -\sum_{\langle i,j\rangle \in \mathcal{E}} J_{ij}\boldsymbol{\sigma}_i^z\boldsymbol{\sigma}_j^z - \sum_{i\in\mathcal{V}} h_i\boldsymbol{\sigma}_i^z, \qquad \mathbf{H}_X = -\sum_{i\in\mathcal{V}} \boldsymbol{\sigma}_i^x, \tag{3.3}$$

where $A(t)$ and $B(t)$ are time-dependent smooth functions controlling the annealing schedules, $\mathcal{G}$ is the graph specified in the definition of the classical Ising model with site set $\mathcal{V}$ and edge set $\mathcal{E}$, J_{ij} represents the interaction between sites i and j associated with the edge $\langle i, j\rangle \in \mathcal{E}$, and h_i is the strength of the external local fields imposed on site $i \in \mathcal{V}$. Here, we use the convention in the quantum literature that $\boldsymbol{\sigma}_i^z\boldsymbol{\sigma}_j^z$ denotes the tensor product

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

of σ_i^z and σ_j^z, along with identity matrices, in such a way that

$$\boldsymbol{\sigma}_i^z \boldsymbol{\sigma}_j^z \equiv I_1 \otimes \cdots \otimes I_{i-1} \otimes \sigma_i^z \otimes I_{i+1} \otimes \cdots \otimes I_{j-1} \otimes \sigma_j^z \otimes I_{j+1} \otimes \cdots \otimes I_d;$$

similarly, $\boldsymbol{\sigma}_i^x$ and $\boldsymbol{\sigma}_i^z$ denote the following tensor products of d matrices of size two:

$$\boldsymbol{\sigma}_i^x \equiv I_1 \otimes \cdots \otimes I_{i-1} \otimes \sigma_i^x \otimes I_{i+1} \otimes \cdots \otimes I_d,$$

$$\boldsymbol{\sigma}_i^z \equiv I_1 \otimes \cdots \otimes I_{i-1} \otimes \sigma_i^z \otimes I_{i+1} \otimes \cdots \otimes I_d.$$

The Pauli matrices σ_i^x and σ_i^z in the tensor products are defined in (3.1).

Suppose that quantum annealing starts at $t = 0$ and ends at $t = t_f$, where t_f refers to the annealing duration. The quantum annealing schedules $A(t)$ and $B(t)$ typically satisfy that $A(t_f) = B(0) = 0$, $A(t)$ is decreasing, and $B(t)$ is increasing. By controlling the annealing schedules $A(t)$ and $B(t)$, we allow quantum annealing to gradually move the Hamiltonian from $\mathbf{H}_{QA}(0) = A(0)\mathbf{H}_X$ to $\mathbf{H}_{QA}(t_f) = B(t_f)\mathbf{H}_I^q$. Because $A(0)$ and $B(t_f)$ are known scalars, $\mathbf{H}_{QA}(t)$ shares the same eigenvectors as $\mathbf{H}_X$ at the initial time $t = 0$, and the same as $\mathbf{H}_I^q$ at the final time t_f. In addition, their corresponding eigenvalues differ by factors of $A(0)$ and $B(t_f)$, respectively. Note that the lowest energy of a quantum system is equal to the smallest eigenvalue of its Hamiltonian, with its ground state(s) being the eigenvector(s) corresponding to the smallest eigenvalue. The Hamiltonian

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

$\mathbf{H}_X = -\sum_{i=1} \boldsymbol{\sigma}_i^x$ is a simple Hermitian matrix with explicit expressions for its smallest eigenvalue and the corresponding eigenvector, and the quantum system governed by the Hamiltonian $\mathbf{H}_X$ can be easily prepared in its ground state.

The quantum annealing procedure is as follows. We begin with an initial quantum system prepared in its ground state. The selected annealing schedules $A(t)$ and $B(t)$ allow us to engineer the quantum system to gradually move from $\mathbf{H}_{QA}(0) = A(0)\mathbf{H}_X$ toward $\mathbf{H}_{QA} = B(t_f)\mathbf{H}_I^q$. Therefore, the quantum annealing evolution driven by $\mathbf{H}_{QA}(t)$ essentially evolves the quantum system from the initial system $\mathbf{H}_X$ initialized at its ground state to the final system $\mathbf{H}_I^q$. The quantum adiabatic theorem indicates that if the quantum system initially starts in its ground state, it tends to remain in the ground states of the instantaneous Hamiltonian during the Hamiltonian evolution. Thus, at the end of the quantum annealing evolution, we measure the system to find the lowest energy of $\mathbf{H}_I^q$ if the quantum system is in its ground state.

For the quantum Hamiltonian $\mathbf{H}_I^q$ defined in (3.3), its lowest energy is equal to its smallest eigenvalue. Here $\mathbf{H}_I^q$ involves only commuting diagonal matrices $\boldsymbol{\sigma}_i^z$, and its eigenvalues are equal to its diagonal entries. These, in turn, are exactly all the 2^d values of the classical Hamiltonian $\mathbf{H}_I^c(\mathbf{s})$ in

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

(2.1) corresponding to the 2^d configurations ordered lexicographically. The following theorem describes the relationship between the classical Hamiltonian $\mathbf{H}_I^c(\mathbf{s})$ and the quantum Hamiltonian $\mathbf{H}_I^q$.

Theorem 1. *The eigenvalues of the quantum Hamiltonian* $\mathbf{H}_I^q$ *in (3.3) are given by the* 2^d *values of the classical Hamiltonian* $\mathbf{H}_I^c(\mathbf{s})$ *in (2.1) evaluated at the* 2^d *configurations* $\mathbf{s} \in \{+1, -1\}^d$. *In particular, the minimum of* $\mathbf{H}_I^c(\mathbf{s})$ *over* $\mathbf{s} \in \{+1, -1\}^d$ *is equal to the smallest eigenvalue of* $\mathbf{H}_I^q$.

Theorem 1 shows that finding the minimal energy of the classical Ising model described by $\mathbf{H}_I^c$ is mathematically identical to finding the minimal energy of the quantum Hamiltonian $\mathbf{H}_I^q$. Thus, at the end of the quantum annealing process, measuring the quantum system renders a solution to the combinatorial minimization problem with the objective function $\mathbf{H}_I^c(s)$. As in SA, each quantum annealing run produces a solution to the optimization problem with some probability, and running quantum annealing many times enables us to solve the optimization problem.

According to the quantum adiabatic theorem (Aharonov et al. (2007), Born and Fock (1928), McGeoch (2014), Morita and Nishimori (2008), and Wang et al. (2016)), for appropriately chosen $A(t)$ and $B(t)$, we have that, with some probability, the quantum annealing driven by (3.2) can find the global minimum of $\mathbf{H}_I^c(\mathbf{s})$ in (2.1) and solve the minimization

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

problem at the final annealing time t_f. The following theorem provides a probability bound on successfully solving the optimization problem at the final annealing time t_f using quantum annealing.

Theorem 2. *Suppose that the quantum system associated with quantum annealing is driven by* $\mathbf{H}_{QA}(t)$*, as defined in (3.2). Then, the probability that the lowest energy of* $\mathbf{H}_I^c$ *in (2.1) is obtained by measuring the system at the end of quantum annealing is bounded from below by*

$$\max\left\{\left[\left(1-\int_0^1\left\|\frac{d}{du}\left(|v_1(u)\rangle,\cdots,|v_r(u)\rangle\right)\right\|du\right)_+\right]^2,\{v_j\}_{1\le j\le r},1\le r\le\zeta\right\},\tag{3.4}$$

where ζ *denotes the number of ground states for the quantum Hamiltonian* $\mathbf{H}_I^q$ *in (3.3),* $(x)_+$ *denotes the positive part of* x*, that is,* $(x)_+$ *is equal to* x *if* $x\ge 0$*, and zero otherwise, and* $(|v_1(u)\rangle,\cdots,|v_r(u)\rangle)$ *is a matrix formed by the* r *column vectors* $|v_1(u)\rangle,\cdots,|v_r(u)\rangle$*, defined as follows. Denote by* $\xi_1(u)\le\xi_2(u)\le\cdots\le\xi_{2^d}(u)$ *the* 2^d *instantaneous eigenvalues of* $\mathbf{H}_{QA}(ut_f)$*, listed in increasing order, along with the corresponding* 2^d *normalized eigenvectors* $v_1(u),v_2(u),\cdots,v_{2^d}(u)$*. Here, for any eigenvalue with multiplicity greater than one, the eigenvalue is repeated in the list, with the number of repetitions equal to its multiplicity, and the multiple eigenvectors corresponding to the same eigenvalue are ordered as a group; that is, their positions in the list are interchangeable, and the maximum in (3.4) is taken*

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

over $1 \le r \le \zeta$ *and possible group orderings of* $v_j(u)$.

Furthermore, assume that $\lambda_1(u) - \lambda_0(u)$ *is bounded below from zero uniformly over* $u \in [0,1]$, *where* $\lambda_0(u)$ *and* $\lambda_1(u)$ *denote the smallest and second smallest instantaneous eigenvalues of* $\mathbf{H}_{QA}(ut_f)$, *respectively. Then, the probability that the quantum annealing procedure can find the lowest energy of* $\mathbf{H}_I^c$ *in (2.1) is bounded from below by*

$$1 - 2^d \zeta \max_{u\in[0,1]} \left\{ \frac{1}{\lambda_1(u) - \lambda_0(u)} \left\| \frac{d\mathbf{H}_{QA}(u\, t_f)}{du} \right\| \right\}^2, \tag{3.5}$$

where $\|\cdot\|$ *denotes the matrix spectral norm.*

The ground-state success probability for quantum annealing is usually derived under the unique ground-state condition, in the asymptotic sense that we obtain some expressions or bounds for the leading terms of the ground-state success probability by taking t_f to infinity (Aharonov et al. (2007), Born and Fock (1928), McGeoch (2014), and Morita and Nishimori (2008)). The probability lower bounds in (3.4) and (3.5) are for finite t_f, without the unique ground-state restriction. The results established in Theorems 1 and 2, together with existing asymptotic results, provide the theoretical foundation that makes it possible for the quantum annealing process, driven by (3.2), to find the global minimum of $\mathbf{H}_I^c(\mathbf{s})$ and solve the minimization problem with a certain probability.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

3.2 Implementation using D-Wave machines

D-Wave machines are commercially available computing hardware devices that have been built using superconducting technology to implement quantum annealing. They are analog computers made specifically to handle combinatorial optimization linked to the classical Ising model. Their quantum processor chips are based on associated graphs that specify quantum annealing with possible adjustments to the standard 20 μs duration and annealing schedules, as described in Sections 2 and 3. Since 2011, five generations of D-Wave computing machines have been developed, with the number of graph sites equal to 128, 512, 1152, 2048, and 5640, respectively. The first four generations used the Chimera graph, and the fifth generation used the Pegasus graph. D-Wave machines have been applied to solve combinatorial optimization problems in research studies and real applications. Although there is no quantum speedup found in D-Wave machines, it has been demonstrated that quantum annealing can be much faster than classical annealing for solving certain optimization problems; see Denchev et al. (2015) and Farhi et al. (2002) for specific examples that illustrate the advantage of quantum annealing over classical annealing. For further information, see also Boixo et al. (2014), Boixo et al. (2015 a), Boixo et al. (2015 b), Dattani et al. (2019), Denchev et al. (2015), Farhi et

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

al. (2002), Hastings (2020), Hen et al. (2015), Johnson et al. (2011), Katzgraber et al. (2014), Lanting et al. (2014), O'Gorman et al. (2014), Perdomo-Ortiz et al. (2012), Perdomo-Ortiz et al. (2014), Rieffel et al. (2014), and Wang et al. (2016).

3.3 Implementation using path-integral and MCMC simulations

Various methods have been established that use asymptotic expansion to approximately implement quantum annealing using MCMC simulations on classical computers. These approaches use the path-integral formulation with the Trotter formula (see Kato (1978), Suzuki (1976), and Trotter (1959)) to show that the quantum system driven by the quantum annealing Hamiltonian $\mathbf{H}_{QA}(t)$ in (3.2) is asymptotically equivalent to a classical anisotropic Ising model. For simplicity, consider the case of $h_i = 0$. The classical anisotropic Ising model with temperature τT has the following Hamiltonian:

$$\mathbf{H}_{aI}^{c}(\mathbf{s}) = -\sum_{l=1}^{\tau}\left[B(t)\sum_{(i,j)\in\mathcal{E}(\mathcal{G})} J_{ij}s_{il}s_{jl} + J(t)\sum_{j\in\mathcal{V}(\mathcal{G})} s_{jl}s_{j,l+1}\right], \qquad (3.6)$$

where τ is an integer, s_{il} are random variables taking values in $\{+1, -1\}$, J_{ij} are the regular couplings along the direction of the original Ising model, and l is the index for an extra new direction, often referred to as the imaginary-

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

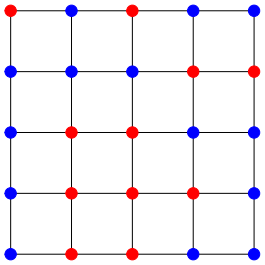
time direction, with

$$J(t) = -\frac{\tau T}{2} \ln \left[\tanh \left(\frac{A(t)}{\tau T} \right) \right]$$

as the coupling along the imaginary-time direction. Let $\mathbf{s}_l = \{s_{il}, i = 1, \ldots, b\}$, for $l = 1, \ldots, \tau$. We call $\mathbf{s}_l$ the lth Trotter slice. Similar to the SA case, standard MCMC techniques are used to perform simulations of the classical anisotropic Ising model with the Hamiltonian $\mathbf{H}^c_{aI}$ on classical computers. With the Trotter slices s_{il}, for $i = 1, \ldots, d$ and $l = 1, \ldots, \tau$, generated from the MCMC simulations, we use the majority rule to produce each site value $s_i =$ sign of the sum $s_{i1} + \cdots + s_{i\tau}$ and form a configuration $s = \{s_i, i = 1, \ldots, d\}$ as a solution to minimize $\mathbf{H}^c_I(\cdot)$ defined in (2.1). The classical computer-based simulation approach refers to simulated quantum annealing (SQA). Figure 1 illustrates a lattice structure for a classical Ising model and a Trotter slice structure for its corresponding classical anisotropic Ising model. Because the Trotter slice graph is much more complex than the lattice, MCMC simulations of the two Ising models with different structures evidently indicate that SQA is much slower than SA. Consequently, SQA is often used for benchmarking in quantum computation and to gain some insightful understanding of quantum annealing; see Hu and Wang (2021), Martoňák et al. (2002), Morita and Nishimori (2008), Wang et al. (2016), and the references therein for more details.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

(a) Lattice Structure



(b) Trotter Slices

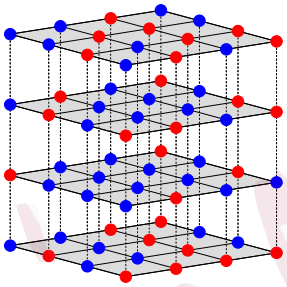


Figure 1: Plots of lattice structures for a classical Ising system and its corresponding classical anisotropic Ising system, where (a) and (b) illustrate a lattice as a simple graph for the Ising system and its corresponding graph with four Trotter slices for the anisotropic Ising system, respectively.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

3.4 Discussion of theoretical analysis and practical implementation

Quantum annealing was proposed as a potential way of solving optimization problems. Theoretical results have established the foundation for quantum annealing, and implementations by physical quantum annealers and path-integral-based MCMC simulations have demonstrated its practical usefulness. It is interesting, but challenging to accurately connect the theoretical results with the practical performance of implementations by physical or simulation means, because quantum annealers are noisy, and MCMC simulations are approximation methods based on asymptotics.

Specifically, theoretical results indicate certain probabilities for quantum annealing to find ground states, and quantum annealers and SQA practically confirm some success probabilities for obtaining ground states using quantum annealing. The asymptotic justification derives positive probabilities for quantum annealing to find ground states by letting the annealing time t_f go to infinity. The probability lower bound in (3.5) for finite t_f is given by $1 - 2^d \zeta \aleph$, where

$$\aleph = \max_{u \in [0,1]} \left\{ \frac{1}{\lambda_1(u) - \lambda_0(u)} \left\| \frac{d\mathbf{H}_{QA}(u\, t_f)}{du} \right\| \right\}^2 . \tag{3.7}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Note that

$$\frac{d\mathbf{H}_{QA}(u\,t_f)}{du} = \frac{dA(u\,t_f)}{du}\,\mathbf{H}_X + \frac{dB(u\,t_f)}{du}\,\mathbf{H}_I^q, \tag{3.8}$$

which depends on u only through the derivatives of the annealing schedules $A(t)$ and $B(t)$. By choosing appropriate $A(t)$ and $B(t)$, we can ensure that the probability lower bound in (3.5) is positive, and thus guarantee that quantum annealing can find the lowest energy of $\mathbf{H}_I^c$, with some probability. Indeed, from (3.7) and (3.8), we obtain

$$\begin{aligned}
\aleph &\le \max_{u\in[0,1]} \left\{ [\lambda_1(u) - \lambda_0(u)]^{-2} \left[\left| \frac{dA(u\,t_f)}{du} \right| \|\mathbf{H}_X\| + \left| \frac{dB(u\,t_f)}{du} \right| \|\mathbf{H}_I^q\| \right]^2 \right\} \\
&\le \left\{ \min_{u\in[0,1]} [\lambda_1(u) - \lambda_0(u)] \right\}^{-2} [\|\mathbf{H}_X\| + \|\mathbf{H}_I^q\|]^2 \\
&\quad \max_{u\in[0,1]} \left\{ \left| \frac{dA(u\,t_f)}{du} \right| \bigvee \left| \frac{dB(u\,t_f)}{du} \right| \right\}^2,
\end{aligned} \tag{3.9}$$

where $\|\mathbf{H}_X\|$ and $\|\mathbf{H}_I^q\|$ are the spectral norms of $\mathbf{H}_X$ and $\mathbf{H}_I^q$, respectively, and $\vee$ stands for the maximum. For a given quantum annealing setup, we have fixed d, ζ, $\|\mathbf{H}_X\|$, and $\|\mathbf{H}_I^q\|$, and specified the minimum of $\lambda_1(u) - \lambda_0(u)$ over $u \in [0,1]$. Hence, (3.9) indicates that it is theoretically possible to choose schedule functions $A(\cdot)$ and $B(\cdot)$ with small enough absolute derivatives in order to make $\aleph < 1/[\zeta 2^d]$, and thus produce a positive probability lower bound $1 - 2^d\zeta\aleph$. More precisely, if the schedule functions $A(\cdot)$ and $B(\cdot)$ satisfy that, for $u \in [0,1]$,

$$\left| \frac{dA(u\,t_f)}{du} \right| \bigvee \left| \frac{dB(u\,t_f)}{du} \right| < \frac{\min_{u\in[0,1]} [\lambda_1(u) - \lambda_0(u)]}{2^{d/2}\sqrt{\zeta}\,[\|\mathbf{H}_X\| + \|\mathbf{H}_I^q\|]},$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

then $1 - 2^d \zeta \aleph > 0$.

4. Conclusion

We have investigated quantum annealing along with classical annealing for solving combinatorial optimization problems. We have established a probability lower bound for quantum annealing to find a solution to an optimization problem. We discuss implementations of quantum annealing via D-Wave physical devices and path-integral-based MCMC simulations. Our study provides a theoretical foundation for using quantum annealing to solve combinatorial optimization problems.

4.1 Future research topics in quantum annealing

Quantum annealing plays an important role in quantum computation, and is relatively new in statistics. This paper raises statistical issues related to quantum annealing and leaves open problems for future research. These include investigating the sharpness of the lower bounds in Theorem 2 and their relationship with the asymptotic results on the ground-state success probability, as well as their practical implications. It would also be worth studying the quantum processes related to the unitary dynamic evolution of quantum annealing and classical processes associated with MCMC im-

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

plementations of quantum annealing, as well as the statistical relationship between the two types of processes. Furthermore, we may explore the performance of D-Wave machines and, in particular, the impact of noise on solving optimization problems. As a case in point, below, we explore a key difference between thermal fluctuations in classical annealing and quantum tunneling in quantum annealing, and propose combining the data augmentation scheme and the Monte Carlo implementation approach to intuitively illustrate quantum tunneling from a statistical perspective.

4.2 Quantum tunneling and data augmentation

Unlike classical annealing, which uses thermal fluctuations to make the Ising system jump from state to state over intermediate energy barriers and search for the desired lowest-energy state, quantum annealing relies on quantum mechanical fluctuations instead of thermal jumps for state transitions. The two terms $\mathbf{H}_I^q$ and $\mathbf{H}_X$ of $\mathbf{H}_{QA}$ in (3.2) and (3.3) are non-commutable matrices, and represent the potential and kinetic energies, respectively, of the underlying quantum system. The move from $\mathbf{H}_X$ to $\mathbf{H}_I^q$ through $\mathbf{H}_{QA}$ during the quantum annealing process can be physically accomplished by engineering magnetic fields to induce quantum fluctuations via quantum tunneling. Quantum tunneling refers to the quantum

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

phenomenon that particles tunnel through a barrier in an impossible condition under classical physics. It permits the annealing process to search for distinct states by traveling directly through energy barriers, instead of hopping over them thermally, as in the classical annealing case. Quantum tunneling can be explained by the Heisenberg uncertainty principle and the wave–particle duality of matter in quantum physics, but cannot be adequately explained by classical physics (see Sakurai and Napolitano (2010) and Shankar (1994)).

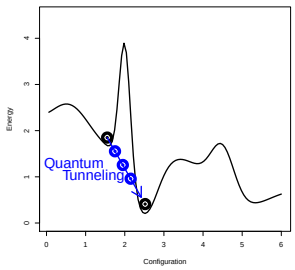
Sections 3.2 and 3.3 indicate that quantum annealing may be implemented by physical devices or MCMC simulations. D-Wave physical devices aim to realize the unitary dynamic evolution of quantum annealing by natural Schrödingier dynamics, and path-integral-based MCMC simulations approximate the dynamic evolution of quantum annealing using artificial time evolutions of Monte Carlo dynamics. The two approaches may be connected through some statistical sampling distributions for the quantum probability model associated with the unitary dynamic evolution in annealing and the statistical model linked to SQA. In particular, although it is difficult to explain quantum tunneling in a physical implementation of quantum annealing, we can use SQA to provide an intuitive statistical illustration of tunneling. Note that the key difference between the clas-

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

sical Ising Hamiltonians $\mathbf{H}_I^c(s)$ in (2.1) and $\mathbf{H}_{aI}^c(\mathbf{s})$ in (3.6) is the extra imaginary-time direction in $\mathbf{H}_{aI}^c(\mathbf{s})$. The well-known data augmentation scheme can be used to accommodate the extra direction, as follows. We introduce an augmented random variable to represent the imaginary-time direction, and describe SQA in an augmented search space. The tunneling, which is difficult to explain in the original search space, may have an intuitive explanation in the larger search space. For example, Figure 2 illustrates that tunneling is a possible way of traveling through a barrier in two dimensions; however, in the augmented three dimensions, there is a natural route to bypass the barrier. The representation of SQA using data augmentation may offer an intuitive explanation for the tunneling effect in quantum annealing.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

(a) Quantum Tunneling



(b) Front and Back Views of Tunneling in the Augmented Space

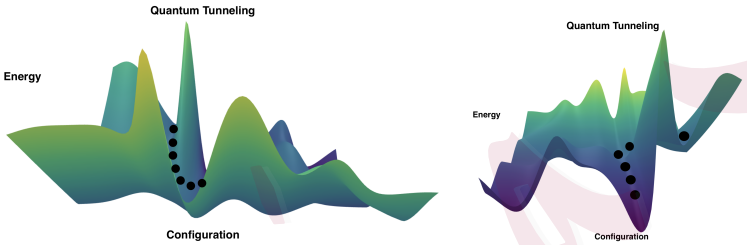


Figure 2: Plots of quantum tunneling, where (a) and (b) illustrate quantum tunneling in quantum annealing for getting over an energy barrier and a corresponding possible tunneling effect in the augmented space, respectively.

Acknowledgments

The research of Yazhen Wang was supported in part by NSF grants DMS-1707605 and DMS-1913149. The authors thank the co-editor, Xiaotong

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Shen, an associate editor, and two anonymous referees for their helpful comments and suggestions.

References

Aharonov, D., van Dam, W., Kempe, J., Landau, Z., Lloyd, S. and Regev, O. (2007). Adiabatic quantum computation is equivalent to standard quantum computation. SIAM J. Comp. 37, 166.

Albash, T., Rønnow, T. F., Troyer, M. and Lidar, D. A. (2014). Reexamining classical and quantum models for the D-Wave One processor: the role of excited states and ground state degeneracy. arXiv:1409.3827v1.

Arute, F., Arya, K., Babbush, R. et al. (2019). Quantum supremacy using a programmable superconducting processor. Nature 574, 505-510. https://doi.org/10.1038/s41586-019-1666-5.

Aspuru-Guzik, A., Dutoi, A. D., Love, P. J. and Head-Gordon, M. (2005). Simulated quantum computation of molecular energies. Science 309, 1704-1707.

Bertsimas, D. and Tsitsiklis, J. (1993). Simulated annealing. Statistical Science 8, 10-15.

Blanes, S., Casas, F., Oteo, J. A. and Ros, J. (2009). The Magnus expansion and some of its applications. Physics Reports 470, 151-238.

Boixo, S., Albash, T., Spedalieri, F. M., Chancellor, N. and Lidar, D. A. (2014).

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Experimental signature of programmable quantum annealing. Nature Communications 4, 2067.

Boixo, S., Rønnow, T. F., Isakov, S. V., Wang, Z., Wecker, D., Lidar, D. A., Martinis, J. M. and Troyer, M. (2014). Evidence for quantum annealing with more than one hundred qubits. Nature Physics 10, 218-224. DOI: 10.1038/NPHYS2900.

Boixo, S., Smelyanskiy, V. N., Shabani, A., Isakov, S. V., Dykman, M., Denchev, V. S., Amin, M., Smirnov, A., Mohseni, M. and Neven, H. (2015 a). Computational role of collective tunneling in a quantum annealer. arXiv:1411.4036v2.

Boixo, S., Smelyanskiy, V. N., Shabani, A., Isakov, S. V., Dykman, M., Denchev, V. S., Amin, M., Smirnov, A., Mohseni, M. and Neven, H. (2015 b). Computational role of multiqubit tunneling in a quantum annealer. arXiv:1502.05754v1.

Born, M. and Fock, V. (1928). Beweis des Adiabatensatzes. Zeitschrift für Physik A 51, 165-180.

Britton, J. W., Sawyer, B.C., Keith, A., Wang, C.-C.J., Freericks, J. K., Uys, H., Biercuk, M. J. and Bollinger, J. J. (2012). Engineered 2D Ising interactions on a trapped-ion quantum simulator with hundreds of spins. Nature 484, 489-492.

Brooke, J., Bitko, D., Rosenbaum, T. F. and Aeppli, G. (1999). Quantum annealing of a disordered magnet. Science 284, 779-781.

Clarke, J. and Wilhelm, F. (2008). Superconducting quantum bits. Nature 453, 1031-1042.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Dattani, N., Szalay, S. and Chancellor, N. (2019). Pegasus: The second connectivity graph for large-scale quantum annealing hardware. arXiv:1901.07636.

Denchev, V., Boixo, S., Isakov, S., Ding, N., Babbush, R., Smelyanskiy, V., Martinis, J. and Neven, H. (2015). What is the computational value of finite range tunneling? Physical Review X 6, 031015.

Farhi, E., Goldstone, J., Gutmann, S., Lapan, J., Lundgren, A. and D. Preda, D. (2001). A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. Science 292, 472-476.

Farhi, E., Goldstone, J., Gutmann, S. and Sipser, M. (2000). Quantum computation by adiabatic evolution. arXiv:0001106v1.

Farhi, E., Goldstone, J., Gutmann, S. and Sipser, M. (2002). Quantum adiabatic evolution algorithms versus simulated annealing. arXiv:0201031v1.

Farhi, E., Goldstone, J. and Gutmann, S. (2014). A quantum approximate optimization algorithm. arXiv:1411.4028.

Feynman, R. P. (1982). Simulating physics with computers. International Journal of Theoretical Physics 21, 467-488.

Hajek, B. (1988). Cooling schedules for optimal annealing. Mathematics of Operations Research 13, 311-329.

Hastings, M. B. (2020). The power of adiabatic quantum computation with no sign problem. arXiv:2005.03791.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Hen, I., Job, J., Albash, T., Rønnow, T. F., Troyer, M. and Lidar, D. A. (2015). Probing for quantum speedup in spin glass problems with planted solutions. arXiv:1502.01663v2

Holevo, A. S. (1982). Probabilistic and Statistical Aspects of Quantum Theory. Springer, North-Holland, Amsterdam.

Hu, J. and Wang, Y. (2021). Quantum annealing via path-integral Monte Carlo with data augmentation. Journal of Computational and Graphical Statistics 30, 284-296.

Johnson, M. W., M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich1, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson and G. Rose (2011). Quantum annealing with manufactured spins. Nature 473, 194-198.

Katzgraber, H. G., Hamze, F. and Andrist, R. S. (2014). Glassy Chimeras could be blind to quantum speedup: Designing better benchmarks for quantum annealing machines. Physical Review X 4, 021008.

Kato, T. (1978). Trotter's product formula for an arbitrary pair of self-adjoint contraction semi- groups. In Topics in Functional Analysis (Essays Dedicated to M. G. Krein on the Occasion of His 70th Birthday). Adv. in Math. Suppl. Stud. 3 185-195. Academic Press, Boston, MA.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Kirkpatrick, S., Gelatt, C. D., Vecchi, M. P. (1983). Optimization by simulated annealing. Science 220, 671-680.

Lanting, T., A. J. Przybysz, A. Yu. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, S. Uchaikin, A. B. Wilson and G. Rose (2014). Entanglement in a quantum annealing processor. Physical Review X 4, 021041.

Majewski, J., Li, H. and Ott, J. (2001). The Ising model in physics and statistical genetics. Am. J. Hum. Genet. 69 853-862.

Martin-Mayor, V. and Hen, I. (2015). Unraveling quantum annealers using classical hardness. arXiv:1502.02494v1.

Martoňák, R., Santoro, G. E. and Tosatti, E. (2002). Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model. Physical Review B 66, 094203.

McGeoch, C. C. (2014). Adiabatic Quantum Computation and Quantum Annealing. Synthesis Lectures on Quantum Computing. Morgan & Claypool.

Morita, S. and Nishimori, H. (2008). Mathematical foundation of quantum annealing. Journal of Mathematical Physics 49, 125210.

Nielsen, M. and Chuang, I. (2010). Quantum Computation and Quantum Information. 10th Anniversary edition. Cambridge: Cambridge University Press.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

O'Gorman, B., Babbush, R., Perdomo-Ortiz, A., Aspuru-Guzik, A., and Smelyanskiy, V. (2014). Bayesian network structure learning using quantum annealing. arXiv:1407.3897v2.

Preskill, J. (2018). Quantum computing in the NISQ era and beyond. Quantum 2 (2018): 79.

Perdomo-Ortiz, A., Dickson, N., Drew-Brook, M., Rose, G. and Aspuru-Guzik, A. (2012). Finding low-energy conformations of lattice protein models by quantum annealing. Scientific Reports 2, 571.

Perdomo-Ortiz, A., Fluegemann, J., Narasimhan, S., Biswas and R., Smelyanskiy, V. N. (2014). A quantum annealing approach for fault detection and diagnosis of graph-based systems. arXiv:1406.7601v2

Pudenz, K. L., Albash, T. and Lidar, D. A. (2014). Error corrected quantum annealing with hundreds of qubits. Nature Communications 5, 3243, arXiv:1307.8190

Rieffel, E., Venturelli, D., O'Gorman, B., Do, M., Prystay, E., and Smelyanskiy, V. (2014). A case study in programming a quantum annealer for hard operational planning problems. arXiv:1407.2887v1.

Santoro, G. E., Martoňák, R., Tosatti, E. and Car, R. (2002). Theory of quantum annealing of an Ising spin glass. Science 295, 2427.

Sakurai, J. J. and Napolitano, J. (2010). Modern Quantum Mechanics. Reading: Addison-Wesley. Second edition.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Shankar, R. (1994). Principles of Quantum Mechanics. Springer. Second edition.

Smolin, J. A. and Smith, G. (2014). Classical signature of quantum annealing. Frontiers in Physics 2, 52. arXiv: 1305.4904v1.

Suzuki, M. (1976). Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems. Comm. Math. Phys. 51: 183-190.

Trotter, H. F. (1959). On the product of semi-groups of operators. Proc. Amer. Math. Soc. 10, 545-551.

Venturelli, D., Mandrá, S., Knysh, S., O'Gorman, B., Biswas, R., Smelyanskiy, V. N. (2014). Quantum optimization of fully-connected spin glasses. arXiv:1406.7553v1.

Vinci, W., Albash, T., Mishra, A., Warburton, P. A., and Lidar, D. A. (2014). Distinguishing classical and quantum models for the D-Wave device. arXiv: 1403.4228v1.

Wang, Y. (2012). Quantum computation and quantum information. Statistical Science 27, 373-394..

Wang, Y. (2021). When quantum computation meets data science: Making data science quantum. Harvard Data Science Review (under review).

Wang, Y. and Liu, H. (2022). Quantum computing in a statistical context. Annual Review of Statistics and Its Application (to appear).

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

REFERENCES

Wang, Y. and Song, X. (2020). Quantum science and quantum technology. Statistical Science 35, 51-74.

Wang, Y., Wu, S., and Zou, J. (2016). Quantum annealing with Markov chain Monte Carlo simulations and D-Wave quantum computers. Statistical Science 31, 3, 362-398.

Winker, P. (2001). Optimization Heuristics in Econometrics: Applications of Threshold Accepting. Wiley.

Yazhen Wang and Hongzhi Liu, University of Wisconsin-Madison

E-mail: yzwang@stat.wisc.edu, hliu438@wisc.edu

Shang Wu, Fudan University

E-mail: shangwu@fudan.edu.cn

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

5. Supplementary Appendix: Proofs

5.1 Mathematical notations and quantum concepts

Denote by superscripts $*$, $'$ and $\dagger$ the conjugate of a complex number, the transpose of a vector or matrix, and the conjugate transpose operation, respectively. Denote by $\mathbb{R}$ the set of all real numbers. We use $\mathbb{C}^d$ to denote the d-dimensional complex space. For a vector ψ in $\mathbb{C}^d$, we comply with the convention in quantum mechanics and quantum computation that uses Dirac notations ket $|\cdot\rangle$ and bra $\langle\cdot|$ to indicate that $|\psi\rangle$ and $\langle\psi|$ are column and row vectors, respectively. We define a natural inner product in $\mathbb{C}^d$ to be $\langle u|v\rangle = \sum_{j=1}^d u_j^* v_j = (u_1^*, \cdots, u_d^*)(v_1, \cdots, v_d)'$, where $\langle u| = (u_1, \cdots, u_d)$ and $|v\rangle = (v_1, \cdots, v_d)'$, and the modulus (or norm) $\|u\| = \sqrt{\langle u|u\rangle}$. We call a matrix $\mathbf{A}$ Hermitian if $\mathbf{A} = \mathbf{A}^\dagger$, and a matrix $\mathbf{U}$ is said to be unitary if $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$, where $\mathbf{I}$ is an identity matrix. For any matrix $\mathbf{A}$, we denote its spectral norm by $\|\mathbf{A}\|$.

Consider a quantum system in a state $|\psi\rangle$, where the quantum state $|\psi\rangle$ is a unit vector in $\mathbb{C}^d$. Measurements on the quantum system are described through observables, where an *observable* is defined as a Hermitian matrix on $\mathbb{C}^d$. Given an observable $\mathbf{M}$, we assume that it has the following eigen-decomposition,

$$\mathbf{M} = \sum_{a=1}^{d} x_a \, \mathbf{Q}_a, \tag{5.1}$$

where x_a are the real eigenvalues of $\mathbf{M}$, and $\mathbf{Q}_a$ are the corresponding projections onto the eigen-spaces of $\mathbf{M}$. When performing a measurement on $\mathbf{M}$ for the quantum system in state $|\psi\rangle$, we adopt a measure space $(\Omega, \mathcal{F})$ to accommodate its possible measurement outcomes, where the measurement result is treated as a

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

random variable X with a probability distribution P_ψ that is defined on $(\Omega, \mathcal{F})$. The random variable X takes values in $\{x_1, x_2, \cdots, x_d\}$, and the probability of obtaining measurement outcome x_a is given by

$$P_\psi(X = x_a) = \langle\psi|\mathbf{Q}_a|\psi\rangle, \quad a = 1, 2, \cdots, d.$$

For the quantum annealing case, the quantum system is described by the quantum Hamiltonian $\mathbf{H}_{QA}(t)$, and its state at time t is $|\psi(t)\rangle$. At the end of the annealing process, $t = t_f$, and $\mathbf{H}_{QA}(t_f) = B(t_f)\mathbf{H}_I^q$. As $B(t_f)$ is a known scalar, we may ignore it and treat $\mathbf{H}_{QA}(t_f)$ the same as $\mathbf{H}_I^q$. Measuring the quantum system to find its energy at t_f corresponds to performing a measurement on $\mathbf{H}_I^q$ for the quantum system in the state $|\psi(t_f)\rangle$. According to the described measuring scheme, we may find the lowest energy of the quantum system with probability $\langle\psi(t_t)|\mathbf{Q}|\psi(t_f)\rangle$, where $\mathbf{Q}$ denotes the projection onto the eigen-space of $\mathbf{H}_I^q$ corresponding to the smallest eigenvalue.

Denote by C's generic constants whose values are free of model parameters and may change from appearance to appearance.

5.2 Proof of Theorem 1

Let $\mathbf{e}_{j,+1} = (1, 0)^\dagger$ and $\mathbf{e}_{j,-1} = (0, 1)^\dagger$, $j = 1, \cdots, d$. Then $\mathbf{e}_{j,\pm1}$ are eigenvectors of Pauli matrix $\boldsymbol{\sigma}_j^z$ corresponding to eigenvalues ±1. For the classical Ising model, given a configuration $\mathbf{s} = (s_1, \cdots, s_b)$ with energy $\mathbf{H}_I^c(\mathbf{s})$ in (2.1), define a unit vector in $\mathbb{C}^{2^d}$, $\mathbf{e}_\mathbf{s} = \mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{b,s_d}$. We show that $\mathbf{e}_\mathbf{s}$ is an eigenvector of

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

$\mathbf{H}_I^q$ with corresponding eigenvalue $\mathbf{H}_I^c(\mathbf{s})$. Indeed,

$$\begin{aligned}
&\boldsymbol{\sigma}_j^z \mathbf{e}_{j,s_j} = s_j \mathbf{e}_{j,s_j},\\
&\left(\mathbf{I}_2 \otimes \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_j^z \otimes \mathbf{I}_2 \cdots \otimes \mathbf{I}_2\right)\left(\mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d}\right)\\
&= s_j \mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d} = s_j \mathbf{e}_{\mathbf{s}},\\
&\left(\mathbf{I}_2 \otimes \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_i^z \otimes \mathbf{I}_2 \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_j^z \otimes \mathbf{I}_2 \cdots \otimes \mathbf{I}_2\right)\left(\mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d}\right)\\
&= s_i s_j \mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d} = s_i s_j \mathbf{e}_{\mathbf{s}},
\end{aligned}$$

and

$$\begin{aligned}
\mathbf{H}_I^q \mathbf{e}_{\mathbf{s}} = &- \sum_{(i,j)\in\mathcal{E}(\mathcal{G})} J_{ij} \left(\mathbf{I}_2 \otimes \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_i^z \otimes \mathbf{I}_2 \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_j^z \otimes \mathbf{I}_2 \cdots \otimes \mathbf{I}_2\right)\\
&\qquad \left(\mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d}\right)\\
&- \sum_{j\in\mathcal{V}(\mathcal{G})} h_j \left(\mathbf{I}_2 \otimes \cdots \mathbf{I}_2 \otimes \boldsymbol{\sigma}_j^z \otimes \mathbf{I}_2 \cdots \otimes \mathbf{I}_2\right)\left(\mathbf{e}_{1,s_1} \otimes \mathbf{e}_{2,s_2} \cdots \otimes \mathbf{e}_{d,s_d}\right)\\
= &- \sum_{(i,j)\in\mathcal{E}(\mathcal{G})} J_{ij} s_i s_j \mathbf{e}_{\mathbf{s}} - \sum_{j\in\mathcal{V}(\mathcal{G})} h_j s_j \mathbf{e}_{\mathbf{s}}\\
= &\left[- \sum_{(i,j)\in\mathcal{E}(\mathcal{G})} J_{ij} s_i s_j - \sum_{j\in\mathcal{V}(\mathcal{G})} h_j s_j\right] \mathbf{e}_{\mathbf{s}} = \mathbf{H}_I^c(\mathbf{s}) \mathbf{e}_{\mathbf{s}}.
\end{aligned}$$

Thus, the 2^d eigenvalues of $\mathbf{H}_I^q$ are $\mathbf{H}_I^c(\mathbf{s})$, $\mathbf{s} \in \{+1, -1\}^d$, which are actually the diagonal entries of $\mathbf{H}_I^q$. If $\mathbf{s}_0$ achieves the global minimum of $\mathbf{H}_I^c(\mathbf{s})$, then $\mathbf{H}_I^c(\mathbf{s}_0)$ is the smallest eigenvalue of $\mathbf{H}_I^q$.

Remark 1. The quantum system governed by quantum Hamiltonian $\mathbf{H}_I^q$ in (3.3) can recover the Boltzmann distribution for observing configurations in the classical Ising model governed by $\mathbf{H}_I^c$ in (2.1). Indeed, the partition function and

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Gibbs state of the quantum Ising model are, respectively, given by

$$Z = tr[e^{-\beta \mathbf{H}_I^q}] \text{ and } \boldsymbol{\rho} = \frac{e^{-\beta \mathbf{H}_I^q}}{Z}.$$

Since $\mathbf{H}_I^q$ has eigenvalues $\mathbf{H}_I^c(\mathbf{s})$ with eigenvectors $\mathbf{e_s}$, it is easy to compute the partition function as follows,

$$Z = \sum_{\mathbf{s}} \langle \mathbf{e_s} | e^{-\beta \mathbf{H}_I^q} | \mathbf{e_s} \rangle = \sum_{\mathbf{s}} e^{-\beta \mathbf{H}_I^c(\mathbf{s})}.$$

The probability of observing configuration $\mathbf{s}$ is given by

$$\langle \mathbf{e_s} | \boldsymbol{\rho} | \mathbf{e_s} \rangle = \frac{1}{Z} \langle \mathbf{e_s} | e^{-\beta \mathbf{H}_I^q} | \mathbf{e_s} \rangle = \frac{1}{Z} e^{-\beta \mathbf{H}_I^c(\mathbf{s})},$$

which is equal to the classical Boltzmann distribution.

5.3 Proof of Theorem 2

Let $u = t/t_f$ be the dimensionless time, and set

$$\mathbf{H}(u) \equiv \mathbf{H}_{QA}(u\, t_f) = \mathbf{H}_{QA}(t) \text{ and } |\varphi(u)\rangle \equiv |\psi(u\, t_f)\rangle = |\psi(t)\rangle. \tag{5.2}$$

Since state $|\psi(t)\rangle$ of the quantum system in nature time t follows the Schrödingier equation with Hamiltonian $\mathbf{H}_{QA}(t)$, $|\varphi(u)\rangle$ satisfies

$$i \frac{d|\varphi(u)\rangle}{du} = t_f \mathbf{H}(u) |\varphi(u)\rangle, \tag{5.3}$$

where we reserve i for the unit imaginary number $\sqrt{-1}$ in the proof of Theorem 2.

Because our goal is to study how close the quantum state is to the ground state, we naturally express the quantum state by the instantaneous eigenstates

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

of $\mathbf{H}(u)$. We denote by $\lambda_0(u) < \lambda_1(u) < \cdots < \lambda_k(u) < \cdots$ the eigenvalues of $\mathbf{H}(u)$ listed in a strictly increasing order, and $|k^{(\varpi)}(u)\rangle$ denotes the normalized instantaneous eigenstates of $\mathbf{H}(u)$ corresponding to the k-th eigenvalue $\lambda_k(u)$, where index ϖ is for eigenvalue multiplicity. Denote by $|0^{(\varpi)}(u)\rangle$, $\varpi = 1, \cdots, \zeta$, those normalized instantaneous eigenstates so that $\{|0^{(\varpi)}(1)\rangle\}_{1\leq\varpi\leq\zeta}$ are the ζ ground states of $\mathbf{H}_I^q$. That is, $u = 1$ corresponds to the end of the quantum annealing process, and there are ζ ground states $|0^{(\varpi)}(1)\rangle$, $\varpi = 1, \cdots, \zeta$, corresponding to the smallest eigenvalue $\lambda_0(1)$. Since our main analysis is targeted for the ground states, we focus on $|0^{(\varpi)}(u)\rangle$, their amplitudes and relationships with other eigenstates.

Note that we adopt two equivalent notations for the eigen-analysis of $\mathbf{H}(u)$: $\{\xi_j(u), |v_j(u)\rangle\}$ stated in Theorem 2 and $\{\lambda_k(u), |k^{(\varpi)}(u)\rangle\}$ defined above, where the former allows repetitions of eigenvalues $\xi_j(u)$ with multiplicity greater than 1, and the latter has strictly increasing eigenvalues $\lambda_k(u)$ and the multiplicity index ϖ.

The main idea of the long proof arguments is described as follows: (1) characterize the eigenstates $|k^{(\varpi)}(u)\rangle$; (2) obtain an expression for the quantum state $|\psi(u)\rangle$ in terms of these eigenstates; (3) derive the amplitudes in the expression corresponding to $|0^{(\varpi)}(u)\rangle$; and (4) find the norm of these amplitudes.

First we need to adjust the eigenstates to meet certain conditions in order to facilitate our analysis. Taking derivatives on both sides of $\mathbf{H}(u)|j^{(l)}(u)\rangle = \lambda_j(u)|j^{(l)}(u)\rangle$, we obtain

$$\frac{d\mathbf{H}(u)}{du}|j^{(l)}(u)\rangle + \mathbf{H}(u)\frac{d|j^{(l)}(u)\rangle}{du} = \frac{d\lambda_j(u)}{du}|j^{(l)}(u)\rangle + \lambda_j(u)\frac{d|j^{(l)}(u)\rangle}{du},$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

and thus for $k \neq j$,

$$\begin{aligned}
&\langle k^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|j^{(l)}(u)\rangle + \langle k^{(\varpi)}(u)|\mathbf{H}(u)\frac{d|j^{(l)}(u)\rangle}{du} \\
&= \langle k^{(\varpi)}(u)|\frac{d\lambda_j(u)}{du}|j^{(l)}(u)\rangle + \langle k^{(\varpi)}(u)|\lambda_j(u)\frac{d|j^{(l)}(u)\rangle}{du}. \qquad (5.4)
\end{aligned}$$

For orthonormal $|j^{(l)}(u)\rangle$ and $|k^{(\varpi)}(u)\rangle$, we have $\langle j^{(l)}(u)\rangle|k^{(\varpi)}(u)\rangle = 0$ and $\langle k^{(\varpi)}(u)|\mathbf{H}(u) = \lambda_k(u)\langle k^{(\varpi)}(u)|$. Substituting these into (5.4) yields

$$\begin{aligned}
&\langle k^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|j^{(l)}(u)\rangle + \lambda_k(u)\langle k^{(\varpi)}(u)|\frac{d|j^{(l)}(u)\rangle}{du} \\
&= \frac{d\lambda_j(u)}{du}\langle k^{(\varpi)}(u)|j^{(l)}(u)\rangle + \lambda_j(u)\langle k^{(\varpi)}(u)|\frac{d|j^{(l)}(u)\rangle}{du} \\
&= \lambda_j(u)\langle k^{(\varpi)}(u)|\frac{d|j^{(l)}(u)\rangle}{du},
\end{aligned}$$

which immediately leads to

$$\begin{aligned}
&\langle k^{(\varpi)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle \equiv \langle k^{(\varpi)}(u)|\frac{d|j^{(l)}(u)\rangle}{du} \\
&= \frac{1}{\lambda_j(u) - \lambda_k(u)}\langle k^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|j^{(l)}(u)\rangle, \quad j \neq k. \qquad (5.5)
\end{aligned}$$

Let $|\check{j}^{(l)}(u)\rangle = \exp\{i\eta_j^{(l)}(u)\}|j^{(l)}(u)\rangle$ be a time-dependent phase shift of $|j^{(l)}(u)\rangle$, that is, we add an accent mark $\check{}$ to mean a time-dependent shift for the eigenstates, where $\eta_j^{(l)}(u)$ satisfies

$$\langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + i\frac{d\eta_j^{(l)}(u)}{du} = 0,$$

which is possible since

$$\langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + \left(\langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle\right)^{\dagger} = \frac{d}{du}\langle j^{(l)}(u)|j^{(l)}(u)\rangle = 0,$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

and hence $\langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle$ is a pure imaginary number. Thus, we have

$$\begin{aligned}
&\langle \check{j}^{(l)}(u)|\frac{d}{du}|\check{j}^{(l)}(u)\rangle = e^{i\eta_j^{(l)}(u)}\langle \check{j}^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + i\frac{d\eta_j^{(l)}(u)}{du}\langle \check{j}^{(l)}(u)|\check{j}^{(l)}(u)\rangle \\
&= e^{i\eta_j^{(l)}(u)}e^{-i\eta_j^{(l)}(u)}\langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + i\frac{d\eta_j^{(l)}(u)}{du} \\
&= \langle j^{(l)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + i\frac{d\eta_j^{(l)}(u)}{du} = 0. \qquad (5.6)
\end{aligned}$$

Of course, $\{|\check{j}^{(l)}(u)\rangle, j = 0, 1, \cdots,\}$ remains to be orthonormal, the pair $(\lambda_j(u), |\check{j}^{(l)}(u)\rangle)$ still satisfies

$$\mathbf{H}(u)|\check{j}^{(l)}(u)\rangle = e^{i\eta_j^{(l)}(u)}\mathbf{H}(u)|j^{(l)}(u)\rangle = e^{i\eta_j^{(l)}(u)}\lambda_j(u)|j^{(l)}(u)\rangle = \lambda_j(u)|\check{j}^{(l)}(u)\rangle,$$

and for $k \neq j$,

$$\begin{aligned}
&\langle \check{k}^{(\varpi)}(u)|\frac{d}{du}|\check{j}^{(l)}(u)\rangle = e^{i\eta_j^{(l)}(u)}\langle \check{k}^{(\varpi)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle + i\frac{d\eta_j^{(l)}(u)}{du}\langle \check{k}^{(\varpi)}(u)|j^{(l)}(u)\rangle \\
&= e^{i[\eta_j^{(l)}(u)-\eta_k^{(\varpi)}(u)]}\langle k^{(\varpi)}(u)|\frac{d}{du}|j^{(l)}(u)\rangle \\
&= \frac{e^{i[\eta_j^{(l)}(u)-\eta_k^{(\varpi)}(u)]}}{\lambda_j(u)-\lambda_k(u)}\langle k^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|j^{(l)}(u)\rangle \\
&= \frac{1}{\lambda_j(u)-\lambda_k(u)}\langle \check{k}^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|\check{j}^{(l)}(u)\rangle, \qquad (5.7)
\end{aligned}$$

where the third equality is due to (5.5).

Now since (5.6) and (5.7) are satisfied by the instantaneous eigenstates $\check{j}^{(l)}(u)$ of $\mathbf{H}(u)$, we use them to express the quantum state $|\varphi(u)\rangle$ as follows:

$$|\varphi(u)\rangle = \sum_{j,l\geq 0}\alpha_j^{(l)}(u)|\check{j}^{(l)}(u)\rangle. \qquad (5.8)$$

Plugging expression (5.8) into the Schrödinger equation (5.3) we obtain

$$\sum_{j,l\geq 0} i\frac{d}{du}\left[\alpha_j^{(l)}(u)|\check{j}^{(l)}(u)\rangle\right] = \sum_{j,l\geq 0} t_f\mathbf{H}(u)\left[\alpha_j^{(l)}(u)|\check{j}^{(l)}(u)\rangle\right],$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

and simple manipulations lead to

$$\sum_{j,l\geq 0} i\left[\frac{d\alpha_j^{(l)}(u)}{du}|\check{j}^{(l)}(u)\rangle + \alpha_j^{(l)}(u)\frac{d}{du}|\check{j}^{(l)}(u)\rangle\right] = \sum_{j,l\geq 0} t_f\alpha_j^{(l)}(u)\mathbf{H}(u)|\check{j}^{(l)}(u)\rangle$$
$$= \sum_{j,l\geq 0} t_f\alpha_j^{(l)}(u)\lambda_j(u)|\check{j}^{(l)}(u)\rangle.$$

Taking products with state $\langle\check{k}^{(\varpi)}(u)|$ on both sides of above equation and noting the scalar nature of t_f, $\alpha_j^{(l)}(u)$ and $\lambda_j(u)$, we arrive at

$$\sum_{j,l\geq 0} i\left[\frac{d\alpha_j^{(l)}(u)}{du}\langle\check{k}^{(\varpi)}(u)|\check{j}^{(l)}(u)\rangle + \alpha_j^{(l)}(u)\langle\check{k}^{(\varpi)}(u)|\frac{d}{du}|\check{j}^{(l)}(u)\rangle\right]$$
$$= \sum_{j,l\geq 0} t_f\lambda_j(u)\alpha_j^{(l)}(u)\langle\check{k}^{(\varpi)}(u)|\check{j}^{(l)}(u)\rangle,$$

which can be simplified by using (5.6) and the orthonormality of $|\check{j}^{(l)}(u)\rangle$ as

$$\frac{d\alpha_k^{(\varpi)}(u)}{du} + \sum_{l\neq\varpi}\alpha_k^{(l)}\langle\check{k}^{(\varpi)}(u)|\frac{d}{du}|\check{k}^{(l)}(u)\rangle + \sum_{j\neq k}\sum_{l}\alpha_j^{(l)}(u)\langle\check{k}^{(\varpi)}(u)|\frac{d}{du}|\check{j}^{(l)}(u)\rangle$$
$$= -it_f\lambda_k(u)\alpha_k^{(\varpi)}(u). \tag{5.9}$$

5.3.1 Establishing the probability lower bound (3.4)

Recall that we list 2^d eigenvalues $\xi_1(u) \leq \xi_2(u) \leq \cdots \leq \xi_{2^d}(u)$ of $\mathbf{H}(u)$ in an increasing order along with the corresponding normalized eigenvectors $v_1(u), v_2(u), \cdots, v_{2^d}(u)$, where for any eigenvalue with a multiplicity greater than 1, the eigenvalue is repeated in the list with the number of repetitions equal to its multiplicity, and the multiple eigenvectors corresponding to the same eigenvalue are ordered as a group.

We use the ordered 2^d normalized eigenvectors $v_1(u), v_2(u), \cdots, v_{2^d}(u)$ to represent the quantum state $|\psi(u)\rangle$ defined in (5.2). Vector-matrix forms are

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

adopted to facilitate our analysis below. Denote by $\mathbf{\Lambda}(u)$ the diagonal matrix whose diagonal entries are the ordered 2^d eigenvalues $\xi_1(u), \xi_2(u), \ldots, \xi_{2^d}(u)$. Denote by $\mathbf{V}(u)$ the matrix whose columns are formed by the ordered 2^d normalized eigenvectors $v_1(u), v_2(u), \ldots, v_{2^d}(u)$, and denote by $\boldsymbol{\alpha}(u)$ the coefficient vector in the representation of $|\psi(u)\rangle$ by the ordered 2^d normalized eigenvectors. Then from the equations (5.8) and (5.9), we have

$$|\varphi(u)\rangle = \mathbf{V}(u)\boldsymbol{\alpha}(u), \tag{5.10}$$

and

$$\frac{d}{du}\boldsymbol{\alpha}(u) = -it_f\mathbf{\Lambda}(u)\boldsymbol{\alpha}(u) - \left(\mathbf{V}^\dagger(u)\frac{d}{du}\mathbf{V}(u)\right)\boldsymbol{\alpha}(u). \tag{5.11}$$

Since the probability lower bound (3.4) depends on the first section of $\boldsymbol{\alpha}(u)$, we need to break matrix $\mathbf{V}(u)$ and vector $\boldsymbol{\alpha}(u)$ into two parts. Take an integer r between 1 and 2^d, denote by $\boldsymbol{\alpha}_r(u)$ and $\boldsymbol{\alpha}_{r+}(u)$ the vectors formed by the first r components of $\boldsymbol{\alpha}(u)$ and the remaining $2^d - r$ components, respectively. Similarly, denote by $\mathbf{V}_r(u)$ and $\mathbf{V}_{r+}(u)$ the matrices formed by the first r columns of $\mathbf{V}(u)$ and the $2^d - r$ columns, respectively. Denote by $\mathbf{\Lambda}_r(u)$ the diagonal matrix formed by retaining the first r rows and columns of $\mathbf{\Lambda}(u)$. Then equation (5.11) immediately leads to

$$\begin{aligned}
\frac{d}{du}\boldsymbol{\alpha}_r(u) &= -it_f\mathbf{\Lambda}_r(u)\boldsymbol{\alpha}_r(u) - \left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}(u)\right)\boldsymbol{\alpha}(u) \\
&= -it_f\mathbf{\Lambda}_r(u)\boldsymbol{\alpha}_r(u) - \left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_r(u)\right)\boldsymbol{\alpha}_r(u) \\
&\quad - \left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u) \\
&= \mathbf{D}_r(u)\boldsymbol{\alpha}_r(u) - \left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u),
\end{aligned} \tag{5.12}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

where $\mathbf{D}_r(u) = -it_f\mathbf{\Lambda}_r(u) - \mathbf{V}_r^{\dagger}(u)\frac{d}{du}\mathbf{V}_r(u)$. The solution of the linear differential equation system (5.12) has an expression

$$\boldsymbol{\alpha}_r(u) = \mathbf{U}_r(u)\boldsymbol{\alpha}_r(0) - \mathbf{U}_r(u)\int_0^u [\mathbf{U}_r(x)]^{-1}\left(\mathbf{V}_r^{\dagger}(x)\frac{d}{dx}\mathbf{V}_{r+}(x)\right)\boldsymbol{\alpha}_{r+}(x)dx, \tag{5.13}$$

where $\mathbf{U}_r(u)$ is a fundamental matrix for the homogeneous linear differential equation system (5.12) with initial condition $\mathbf{U}_r(0) = \mathbf{I}$—namely, the columns of $\mathbf{U}_r(u)$ form a complete linearly independent set of solutions for the homogeneous linear differential equation system,

$$\frac{d}{du}\boldsymbol{\alpha}_r(u) = \mathbf{D}_r(u)\boldsymbol{\alpha}_r(u).$$

The fundamental matrix $\mathbf{U}_r(u)$ has an expression through the so-called Magnus expansion (Blanes et al. (2009)),

$$\mathbf{U}_r(u) = \exp\{\mathbf{\Xi}^{(r)}(u)\}, \quad \mathbf{\Xi}^{(r)}(u) = \sum_{k=1}^{\infty}\mathbf{\Xi}_k^{(r)}(u),$$

where $\mathbf{\Xi}_k^{(r)}(u)$ in the Magnus expansion can be computed by a recursive procedure through the matrices $\mathbf{\Upsilon}_k^{(j)}(u)$ as follows,

$$\mathbf{\Xi}_1^{(r)}(u) = \int_0^u \mathbf{D}_r(v)dv, \quad \mathbf{\Xi}_k^{(r)}(u) = \sum_{j=1}^{k-1}\frac{\nu_j}{j!}\int_0^u \mathbf{\Upsilon}_k^{(j)}(v)dv, \quad k \geq 2, \tag{5.14}$$

ν_j are Bernoulli numbers,

$$\mathbf{\Upsilon}_k^{(1)}(u) = [\mathbf{\Xi}_{k-1}^{(r)}(u), \mathbf{D}_r(u)], \quad \mathbf{\Upsilon}_k^{(j)}(u) = \sum_{l=1}^{k-j}[\mathbf{\Xi}_l^{(r)}(u), \mathbf{\Upsilon}_{k-l}^{(j-1)}(u)], \quad j = 2, \cdots, k-1, \tag{5.15}$$

and $[\mathbf{A}_1, \mathbf{A}_2] = \mathbf{A}_1\mathbf{A}_2 - \mathbf{A}_2\mathbf{A}_1$ is the matrix commutator of $\mathbf{A}_1$ and $\mathbf{A}_2$.

We therefore have the following lemma for $\mathbf{U}_r(u)$.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Lemma 1. *The fundamental matrix* $\mathbf{U}_r(u)$ *is unitary for integer* r *and* $u \in [0,1]$.

Proof. A square matrix $\mathbf{A}$ is said to be skew-Hermitian if and only if $\mathbf{A}^\dagger = -\mathbf{A}$. Denote by $\mathbf{SH}(r)$ the set formed by all skew-Hermitian matrice of size r. Then $\mathbf{SH}(r)$ is closed under any linear transformation with real coefficients, integration, and matrix commutator. Indeed, it is evident that

1. $\mathbf{B}_1, \mathbf{B}_2 \in \mathbf{SH}(r), \alpha, \beta \in \mathbb{R} \Rightarrow \alpha\mathbf{B}_1 + \beta\mathbf{B}_2 \in \mathbf{SH}(r)$,
2. $\mathbf{B}(u) \in \mathbf{SH}(r), u \in [a,b] \Rightarrow \int_a^b \mathbf{B}(u)du \in \mathbf{SH}(r)$,
3. $\mathbf{B}_1, \mathbf{B}_2 \in \mathbf{SH}(r) \Rightarrow [\mathbf{B}_1, \mathbf{B}_2] \in \mathbf{SH}(r)$.

Now we will show $\mathbf{D}_r(u) \in \mathbf{SH}(r)$, since

$$\begin{aligned}
\mathbf{D}_r(u)^\dagger &= (-it_f\mathbf{\Lambda}_r(u))^\dagger - \left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_r(u)\right)^\dagger \\
&= it_f\mathbf{\Lambda}_r(u) - \left(\left(\frac{d}{du}\mathbf{V}_r^\dagger(u)\right)\mathbf{V}_r(u)\right) \\
&= it_f\mathbf{\Lambda}_r(u) - \left(\frac{d}{du}\left(\mathbf{V}_r^\dagger(u)\mathbf{V}_r(u)\right) - \mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_r(u)\right) \\
&= it_f\mathbf{\Lambda}_r(u) + \mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_r(u) = -\mathbf{D}_r(u). \qquad (5.16)
\end{aligned}$$

Using (5.14) and (5.15), we conclude that $\mathbf{\Upsilon}_k^{(j)}(u) \in \mathbf{SH}(r)$ for $k > 1$ and $j = 1, 2, ..., k-1$, $\mathbf{\Xi}_k^{(r)}(u) \in \mathbf{SH}(r)$ for $k \geq 1$, and thus $\mathbf{\Xi}^{(r)}(u) \in \mathbf{SH}(r)$. Then we have

$$\begin{aligned}
\mathbf{U}_r^\dagger(u)\mathbf{U}_r(u) &= \exp\left\{\left[\mathbf{\Xi}^{(r)}(u)\right]^\dagger\right\}\exp\left\{\mathbf{\Xi}^{(r)}(u)\right\} \\
&= \exp\left\{-\mathbf{\Xi}^{(r)}(u)\right\}\exp\left\{\mathbf{\Xi}^{(r)}(u)\right\} = \mathbf{I}_r,
\end{aligned}$$

which proves the lemma.

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Now we show the probability lower bound (3.4). The quantum system initializes at the ground state of $\mathbf{H}(0)$—namely, $|\psi(0)\rangle$ is equal to the ground state. Thus, we have 1 for the first component of $\boldsymbol{\alpha}(0)$ and 0 for the rest of its components, which implies $\|\boldsymbol{\alpha}_r(0)\| = 1$ for any $r \geq 1$. Using (5.13) we obtain

$$\begin{aligned}
\|\boldsymbol{\alpha}_r(1)\| &\geq \|\mathbf{U}_r(1)\boldsymbol{\alpha}_r(0)\| \\
&\quad - \left\|\mathbf{U}_r(1)\int_0^1 [\mathbf{U}_r(u)]^{-1}\left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u)du\right\| \\
&= 1 - \left\|\int_0^1 [\mathbf{U}_r(u)]^{-1}\left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u)du\right\| \\
&\geq 1 - \int_0^1 \left\|\left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u)\right\| du. \qquad (5.17)
\end{aligned}$$

Note that $\|\boldsymbol{\alpha}_{r+}(u)\| \leq \|\boldsymbol{\alpha}(u)\| = 1$, $\left\|\mathbf{V}_{r+}^\dagger(u)\right\| = 1$, $\mathbf{V}_r^\dagger(u)\mathbf{V}_{r+}(u) = 0$, and

$$\begin{aligned}
\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u) &= \frac{d}{du}\left(\mathbf{V}_r^\dagger(u)\mathbf{V}_{r+}(u)\right) - \left(\frac{d}{du}\mathbf{V}_r^\dagger(u)\right)\mathbf{V}_{r+}(u) \\
&= \quad - \left(\mathbf{V}_{r+}^\dagger(u)\frac{d}{du}\mathbf{V}_r(u)\right)^\dagger.
\end{aligned}$$

Then

$$\begin{aligned}
&\left\|\left(\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right)\boldsymbol{\alpha}_{r+}(u)\right\| \leq \left\|\mathbf{V}_r^\dagger(u)\frac{d}{du}\mathbf{V}_{r+}(u)\right\| \|\boldsymbol{\alpha}_{r+}(u)\| \\
&= \left\|\mathbf{V}_{r+}^\dagger(u)\frac{d}{du}\mathbf{V}_r(u)\right\| \\
&\leq \left\|\mathbf{V}_{r+}^\dagger(u)\right\| \left\|\frac{d}{du}\mathbf{V}_r(u)\right\| \\
&= \left\|\frac{d}{du}\mathbf{V}_r(u)\right\|.
\end{aligned}$$

Plugging the above inequality into (5.17) we arrive at

$$\|\boldsymbol{\alpha}_r(1)\| \geq 1 - \mathcal{K}_r,$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

where

$$\mathcal{K}_r = \int_0^1 \left\| \frac{d}{du}\mathbf{V}_r(u) \right\| du.$$

Finally, the probability for the quantum annealing process to stay in a ground state at the final annealing time is equal to $\|\boldsymbol{\alpha}_\zeta(1)\|^2$, and for any $1 \le r \le \zeta$,

$$\|\boldsymbol{\alpha}_\zeta(1)\|^2 \ge \|\boldsymbol{\alpha}_r(1)\|^2 \ge [(1-\mathcal{K}_r)_+]^2 .$$

Thus, (3.4) is established.

5.3.2 Establishing the probability lower bound (3.5)

We utilize (5.7) to re-write (5.9) with $k = 0$ as the following linear differential equation system for the amplitudes $\alpha_0^{(\varpi)}(u)$ of ζ ground states:

$$\frac{d\alpha_0^{(\varpi)}(u)}{du} = -it_f\lambda_0(u)\alpha_0^{(\varpi)}(u) - \sum_{l\neq\varpi}\alpha_0^{(l)}(u)\langle\check{0}^{(\varpi)}(u)|\frac{d}{du}|\check{0}^{(l)}(u)\rangle$$
$$- \sum_{j>0}\sum_{l}\alpha_j^{(l)}(u)\frac{1}{\lambda_j(u)-\lambda_0(u)}\langle\check{0}^{(\varpi)}(u)|\frac{d\mathbf{H}(u)}{du}|\check{j}^{(l)}(u)\rangle, \qquad (5.18)$$

where $\varpi = 1,\cdots,\zeta$, the sum in the second term is taken over $l = 1,\cdots,\varpi-1,\varpi+1,\cdots,\zeta$ for ground states, and the sums in the third term are over for all excited states.

The linear differential equation system (5.18) has the solution

$$(\alpha_0^{(1)}(u),\cdots,\alpha_0^{(\zeta)}(u))' = \mathbf{U}(u)(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))' +$$
$$\mathbf{U}(u)\int_0^u [\mathbf{U}(x)]^{-1}\sum_{j>0}\alpha_j^{(l)}(x)\frac{1}{\lambda_j(x)-\lambda_0(x)}\langle\check{\mathbf{0}}(x)|\frac{d\mathbf{H}(x)}{dx}|\check{j}^{(l)}(x)\rangle dx, \qquad (5.19)$$

where $\langle\check{\mathbf{0}}(x)| = (\langle\check{0}^{(1)}(x)|,\cdots,\langle\check{0}^{(\zeta)}(x)|)'$, and $\mathbf{U}$ is a fundamental matrix for the homogeneous linear differential equation system corresponding to (5.18) with

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

initial condition $\mathbf{U}(0) = \mathbf{I}$. More specifically, the columns of $\mathbf{U}$ form a complete, linearly independent set of solutions for the homogeneous equation system,

$$\frac{d\alpha_0^{(\varpi)}(u)}{du} = -it_f\lambda_0(u)\alpha_0^{(\varpi)}(u) - \sum_{l\neq\varpi}\langle\check{0}^{(\varpi)}(u)|\frac{d}{du}|\check{0}^{(l)}(u)\rangle\,\alpha_0^{(l)}(u). \tag{5.20}$$

Or, in a vector-matrix form,

$$\frac{d(\alpha_0^{(1)}(u),\cdots,\alpha_0^{(\zeta)}(u))'}{du} = \mathbf{D}(u)(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))', \tag{5.21}$$

where $\mathbf{D}(u) = -it_f\lambda_0(u)\mathbf{I} - i\mathbf{\Psi}(u)$ is a matrix of size ζ, $\mathbf{\Psi}(u) = (\Psi_{\varpi l}(u))$ and $\Psi_{\varpi l}(u) = -i\langle\check{0}^{(\varpi)}(u)|\frac{d}{du}|\check{0}^{(l)}(u)\rangle$ for $l \neq \varpi$ and 0 for $l = \varpi$. Since

$$\begin{aligned} &i\Psi_{\varpi l}(u) + [i\Psi_{l\varpi}(u)]^* = \langle 0^{(\varpi)}(u)|\frac{d}{du}|0^{(l)}(u)\rangle + \left(\langle 0^{(l)}(u)|\frac{d}{du}|0^{(\varpi)}(u)\rangle\right)^* \\ &= \frac{d}{du}\langle 0^{(\varpi)}(u)|0^{(l)}(u)\rangle = 0, \end{aligned}$$

$\mathbf{\Psi}(u)$ is a real matrix. Equation (5.21) and the fundamental matrix $\mathbf{U}(u)$ correspond to equation (5.12) and fundamental matrix $\mathbf{U}_r(u)$ with $r = \zeta$ in Section 5.3.1, respectively. Hence, by Lemma 1 we have that $\mathbf{U}(u)$ is unitary, and then

$$\|\mathbf{U}(u)\| = 1. \tag{5.22}$$

As the system initializes at the ground state of $\mathbf{H}(0)$, $\|(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))\|^2 = 1$. Finally, from (5.19) and (5.22) we find

$$\begin{aligned} &\|(\alpha_0^{(1)}(1),\cdots,\alpha_0^{(\zeta)}(1))\|^2 \geq \|\mathbf{U}(1)(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))\|^2 - \\ &\left|\mathbf{U}(1)\int_0^1[\mathbf{U}(x)]^{-1}\sum_{j>0}\sum_l\frac{\alpha_j^{(l)}(x)}{\lambda_j(x)-\lambda_0(x)}\langle\check{\mathbf{0}}(x)|\frac{d\mathbf{H}(x)}{dx}|\check{j}^{(l)}(x)\rangle dx\right|^2, \end{aligned} \tag{5.23}$$

$$\|\mathbf{U}(1)(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))\|^2 = \|(\alpha_0^{(1)}(0),\cdots,\alpha_0^{(\zeta)}(0))\|^2 = 1, \tag{5.24}$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

and

$$\begin{aligned}
&\left\| \mathbf{U}(u) \int_0^u [\mathbf{U}(x)]^{-1} \sum_{j>0} \sum_l \frac{\alpha_j^{(l)}(x)}{\lambda_j(x) - \lambda_0(x)} \langle \check{\mathbf{0}}(x) | \frac{d\mathbf{H}(x)}{dx} | \check{j}^{(l)}(x) \rangle dx \right\|^2 \\
&= \left\| \mathbf{U}(u) [\mathbf{U}(x_*)]^{-1} \sum_{j>0} \sum_l \frac{\alpha_j^{(l)}(x_*)}{\lambda_j(x_*) - \lambda_0(x_*)} \langle \check{\mathbf{0}}(x_*) | \frac{d\mathbf{H}(x_*)}{dx} | \check{j}^{(l)}(x_*) \rangle \right\|^2 \\
&\le \sum_{\varpi=1}^{\zeta} \left| \sum_{j>0} \sum_l \alpha_j^{(l)}(x_*) \frac{1}{\lambda_j(x_*) - \lambda_0(x_*)} \langle \check{0}^{(\varpi)}(x_*) | \frac{d\mathbf{H}(x_*)}{dx} | \check{j}^{(l)}(x_*) \rangle \right|^2 \\
&\le \sum_{\varpi=1}^{\zeta} \left| \sum_{j>0} \sum_l \alpha_j^{(l)}(x_*) \frac{1}{\lambda_j(x_*) - \lambda_0(x_*)} \langle \check{0}^{(\varpi)}(x_*) | \frac{d\mathbf{H}(x_*)}{dx} | \check{j}^{(l)}(x_*) \rangle \right|^2 \\
&\le \sum_{\varpi=1}^{\zeta} \left| \sum_{j>0} \sum_l \alpha_j^{(l)}(x_*) \frac{1}{\lambda_j(x_*) - \lambda_0(x_*)} \langle \check{0}^{(\varpi)}(x_*) | \frac{d\mathbf{H}(x_*)}{dx} | \check{j}^{(l)}(x_*) \rangle \right|^2 \\
&\le \aleph_0 \zeta \left[\sum_{j>0} \sum_l |\alpha_j^{(l)}(x_*)| \right]^2 \\
&\le \aleph_0 \zeta 2^d \sum_{j>0} \sum_l \left| \alpha_j^{(l)}(x_*) \right|^2 \\
&\le \aleph_0 \zeta 2^d (1 - p_0) \le \aleph_0 \zeta 2^d, \qquad (5.25)
\end{aligned}$$

where in the above eight arrays, the second line is from the mean value theorem and $x_* \in (0, 1)$, the third and fifth lines are due to the spectral norm and (5.22), respectively, $\aleph_0$ is defined to be equal to

$$\max \left\{ \left| \frac{1}{\lambda_j(u) - \lambda_0(u)} \langle \check{0}^{(\varpi)}(u) | \frac{d\mathbf{H}(u)}{du} | \check{j}^{(l)}(u) \rangle \right|^2, j, l \ge 1, \varpi \le \zeta, u \in [0, 1] \right\},$$

and the last two lines are, respectively, from the Cauchy-Schwartz inequality and the facts that

$$p_0 = \min_{0 \le x \le 1} \sum_{\varpi=1}^{\zeta} |\alpha_0^{(\varpi)}(x)|^2 \text{ and } \sum_{j,l} |\alpha_j^{(l)}(x)|^2 = 1.$$

Statistica Sinica: Preprint
doi:10.5705/ss.202021.0228

Applying the Cauchy-Schwartz inequality and using unit norms of $|\check{0}^{(\varpi)}(u)\rangle$ and $|\check{j}^{(l)}(u)\rangle$ and the spectral norm definition, we obtain

$$\aleph_0 \leq \max_{u\in[0,1]} \left\{ [\lambda_1(u) - \lambda_0(u)]^{-2} \left\| \frac{d\mathbf{H}(u))}{du} \right\|^2 \right\} = \aleph, \tag{5.26}$$

where $\aleph$ is defined in (3.7).

Combining (5.23)-(5.26), we conclude

$$\|(\alpha_0^{(1)}(1), \cdots, \alpha_0^{(\zeta)}(1))\|^2 \geq 1 - 2^d \zeta \aleph.$$

The probability of the quantum annealing process staying in the ground states at the final annealing time is equal to $\|(\alpha_0^{(1)}(1), \cdots, \alpha_0^{(\zeta)}(1))\|^2$, and thus we establish the probability lower bound (3.5).