Unraveling reverse annealing: A study of D-Wave quantum annealers

Vrinda Mehta,[1] Hans De Raedt,[1] Kristel Michielsen,[1,2,3] and Fengping Jin[1]
[1]*Institute for Advanced Simulation, Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany*
[2]*RWTH Aachen University, 52056 Aachen, Germany*
[3]*AIDAS, 52425 Jülich, Germany*

(Received 24 February 2025; accepted 12 June 2025; published 14 July 2025)

D-Wave quantum annealers offer reverse annealing as a feature allowing them to refine solutions to optimization problems. This paper investigates the influence of key parameters, such as annealing times and reversal distance, on the behavior of reverse annealing by studying models containing up to 1000 qubits. Through the analysis of theoretical models and experimental data, we explore the interplay between quantum and classical processes in describing the empirical data. Our findings provide a deeper understanding that can better equip users to fully harness the potential of the D-Wave annealers.

DOI: 10.1103/vfxh-gjgy

I. INTRODUCTION

Quantum annealing has emerged as a promising approach for solving complex optimization problems by leveraging quantum mechanical effects [1–5]. Among the commercially available quantum annealers, D-Wave systems [6–13] are at the forefront, offering specialized hardware designed to solve optimization problems [14–27]. Some recent studies have employed these annealers to explore complex physical phenomena in systems utilizing the annealer's architecture to simulate intricate spin interactions and study emergent behaviors [11,28–37]. While a considerable amount of research has focused on the forward annealing process, recent advancements have brought reverse annealing as an intriguing variant that promises to enhance the performance of quantum annealers [38]. In reverse annealing, instead of starting in the uniform superposition state, the system is initialized in a classical state, allowing the exploration of the local energy landscape [38]. This method has been suggested to be beneficial in guiding the system toward optimal solutions, particularly in challenging problem instances [39–42].

Despite the growing interest in reverse annealing, its understanding remains a subject of active debate. The extent to which quantum effects, thermalization, and classical dynamics affect the reverse annealing results is unclear. Understanding these factors is crucial for unlocking the full potential of quantum annealers. Reverse annealing has demonstrated effectiveness in certain applications [42], but questions about the interplay between quantum coherence, thermal noise, and dissipation still need to be answered. While some recent studies have addressed the amount of coherence in the dynamics of standard quantum annealing on these annealers, they primarily focused on the collective behavior of many qubits [32–35]. In contrast, this paper investigates the detailed behavior of a small number of qubits using the reverse annealing protocol.

In our previous work [43], we explored the effects of various control parameters offered by these annealers on the sampling probabilities of the multiple solutions of hard (for quantum annealing) 2-SAT problems. However, although these results hint at a tendency of the system to relax to equilibrium, the theoretical description of the observed sampling behavior is lacking. Using models like the Bloch equations and the Lindblad master equation, among others, in the present paper, we aim to delve deeper into the mechanisms that can describe such a behavior. More specifically, we experimentally study the reverse annealing feature of D-Wave for different choices of the various control parameters and incorporate features in our theoretical models that can reproduce the D-Wave results, providing deeper insights into understanding the extent to which quantum and classical processes can describe the dynamics of reverse annealing.

The central finding of this paper is that under the reverse annealing protocol and for sufficiently long annealing times, the sampling probabilities of different states tend to converge toward those predicted by the thermal equilibrium (Gibbs) distribution. We emphasize that, although our simulation results exhibit strong agreement with the experimental data, it is premature to draw definitive conclusions about the "true" physical mechanisms underlying the system's behavior. The observed agreement does not uniquely establish the nature of the system's evolution.

The paper is organized as follows: we first provide a brief description of the methods (Sec. II) employed to investigate reverse annealing and of the specific problems that we study (Sec. III). We then proceed to give an overview of the key results produced by the quantum annealers (Sec. IV). Next, in Sec. V we present results from numerical models, designed based on the key results. This is followed by the analysis of further results from the quantum annealers (Sec. VI). Finally,

in Sec. VII we discuss the implications of these findings, highlighting potential avenues for future research.

II. METHODS

In this section we focus on the methods and approaches used to investigate the behavior of D-Wave systems. Specifically, we limit our study to the reverse annealing protocol provided by the quantum annealers and implement the equivalent protocol in our simulations. The empirical data reported in this paper are obtained from experiments on the D-Wave Advantage_5.4 system at the Jülich Supercomputing Centre.

In theory the time evolution of the D-Wave annealer is described by the Hamiltonian [6]

$$\begin{aligned}\frac{H(t)}{\hbar} &= \frac{\pi A(s)}{h} H_D + \frac{\pi B(s)}{h} H_P,\\ H_D &= -\sum_i \sigma_i^x,\\ H_P &= \sum_i h_i \sigma_i^z + \sum_{i>j} J_{i,j}\sigma_i^z\sigma_j^z,\end{aligned} \tag{1}$$

where $A(s)$ and $B(s)$ are expressed in GHz and are obtained by fitting simple functions to the annealing schedule data provided by D-Wave [6], and s is the annealing parameter (see below). In Appendix A we show the D-Wave annealing schedule data and the fitted functions used in our simulations. In reality, the time evolution under Eq. (1) is modified by interactions with external degrees of freedom, leading to decoherence and dissipation.

The idea of the reverse annealing protocol is to start in one of the low-lying excited classical states of the problem Hamiltonian ($s = 1$) and anneal backward, i.e., by decreasing the strength $B(s)$ and increasing the strength of $A(s)$ up to some reversal distance s_r. After an optional pause at s_r, the protocol then continues towards $s = 1$, as in standard quantum annealing.

The D-Wave implementation of the protocol offers several control parameters, e.g., the reverse and the forward annealing times t_{reverse} and t_{forward}, respectively, the optional waiting time t_{wait}, the value of s_r, and the choice of the initial state. The reverse annealing schedule is defined by

$$s = \begin{cases} 1 - (1-s_r)\frac{t}{t_{\text{reverse}}}, & t \leqslant t_{\text{reverse}} \\ s_r, & t_{\text{reverse}} < t \leqslant t_{\text{reverse}} + t_{\text{wait}}, \\ s_r + (1-s_r)\frac{t - t_{\text{reverse}} - t_{\text{wait}}}{t_{\text{forward}}}, & t_{\text{reverse}} + t_{\text{wait}} < t \leqslant t_{\text{end}}, \end{cases} \tag{2}$$

where $t_{\text{end}} = t_{\text{reverse}} + t_{\text{wait}} + t_{\text{forward}}$.

The effects of varying these parameters on the performance of reverse annealing in sampling the four degenerate ground states of the 2-SAT problems have been studied in [43]. In this work we focus on the behavior of the quantum annealers by changing t_{reverse}, t_{wait}, and t_{forward}, keeping the other parameters fixed. More specifically, we employ two different procedures described below. To simplify the notation, we denote by $p(t_{\text{end}})$ the probability of a particular state at time $t = t_{\text{end}}$:

Waiting time scans (WTS). In this scheme, we fix $t_{\text{reverse}} = t_{\text{forward}} = 1\,\mu\text{s}$, and vary t_{wait}. For each value of t_{end}, we then determine the respective probabilities $p(t_{\text{end}})$ of finding the relevant states of a given problem.

Annealing times scans (ATS). Fixing $t_{\text{wait}} = 0$, in this scheme we vary $t_{\text{reverse}} = t_{\text{forward}}$. As for the other scheme, we determine the probabilities $p(t_{\text{end}})$ for the relevant energy states for each value of t_{end}.

To further clarify the terminology, we refer to Fig. 1 as an example, which shows the $p(t_{\text{end}})$ for states $|\uparrow\rangle$ and $|\downarrow\rangle$ using red (squares) and blue (circles) lines, respectively.

Note that, for WTS, the system spends a larger proportion of time at a smaller $s = s_r$ compared to ATS. Therefore, for the same t_{end}, the average transverse field for WTS is larger than for ATS, which, as shown later, affects the dynamics.

In the D-Wave experiments, for each problem, we collect 4500 samples for each value of t_{end} corresponding to either changing t_{wait} in the WTS or changing $t_{\text{reverse}} = t_{\text{forward}}$ in the ATS. Furthermore, in most cases shown here, the collection of the samples is carried out by sequentially submitting a given problem ten times. In doing so, we use the minor embedding feature of the annealers, which maps the given problem to a set of physical qubits. In a few other cases, we simultaneously submit many copies of the problem to the D-Wave annealer. The values of probabilities for a given state are assigned by averaging the number of times the state is sampled over the ten sequential runs in the former case, or over the submitted number of copies in the latter.

III. PROBLEM DESCRIPTION

In order to gain a more general understanding of the behavior of D-Wave quantum annealers, it is beneficial to study their performance across different types of problems. To this end we study three kinds of problems:

One- and two-spin problems. As the first class of problem Hamiltonians H_P's we choose the simple instances of one- and two-spin problems, with fixed h_i's and $J_{i,j}$. As these problems are simple, their ground state(s), first excited state(s), and the corresponding energies and degeneracies are known. Consequently they serve as ideal problems for studying and demonstrating the behavior of the D-Wave systems. Therefore the main focus of this paper is on these problems.

For the one-spin problem, we mainly study the cases with $h_1 = 0$ or $h_1 = 0.1$. The states $|\uparrow\rangle$ and $|\downarrow\rangle$ are the two degenerate ground states of the former case. For the latter, the $|\downarrow\rangle$ state with energy -0.1 is the ground state, while the other state with an energy of 0.1 is the first excited state.

Moving to the two-spin problems, we consider three problem instances:

(i) Instance 2S1: $h_1 = h_2 = -1$ and $J_{1,2} = 0.95$. For this problem the $|\uparrow\uparrow\rangle$ state with energy -1.05 is the ground state, while the states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ with energy -0.95 are the degenerate first excited states. State $|\downarrow\downarrow\rangle$ with energy 2.95 is the second excited state.

(ii) Instance 2S2: $h_1 = h_2 = -1$ and $J_{1,2} = 1.00$. For this problem the states $|\uparrow\uparrow\rangle$, $|\uparrow\downarrow\rangle$, and $|\downarrow\uparrow\rangle$ are the threefold degenerate ground states with energy -1.00, while the state $|\downarrow\downarrow\rangle$ with energy 3.00 is the first excited state.

(iii) Instance 2S3: $h_1 = h_2 = -0.95$ and $J_{1,2} = 1.00$. For this problem the states $|\uparrow\downarrow\rangle$ and $|\downarrow\uparrow\rangle$ are twofold degenerate ground states with energy -1.00, while $|\uparrow\uparrow\rangle$ state with energy -0.90 is the first excited state. State $|\downarrow\downarrow\rangle$ with energy 2.90 is the second excited state.

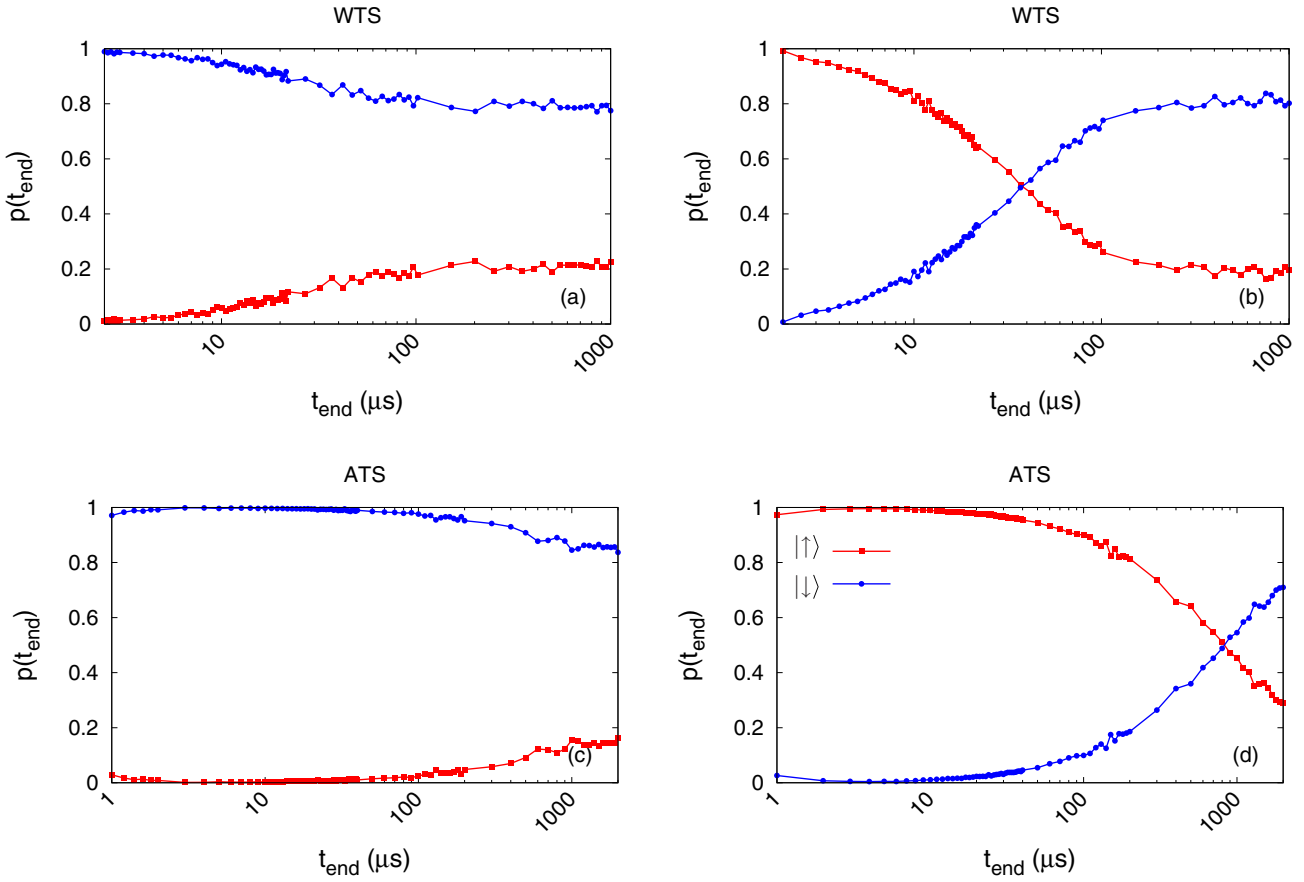

FIG. 1. (a), (b) WTS data and (c), (d) ATS data for the one-spin problem with $h_1 = 0.1$ and $s_r = 0.7$ obtained from the D-Wave annealer with (a), (c) $|\uparrow\rangle$ and (b), (d) $|\downarrow\rangle$ as initial states. The legend in panel (d) also applies to panels (a), (b), and (c).

2-SAT problems. A 2-SAT problem is defined by M clauses, each consisting of two Boolean literals (a Boolean variable x_i or its negation $\overline{x_i}$ for $i = 1, \ldots, N$) [44,45]. The goal is then to determine whether there exists an assignment to the variables x_i's that makes each clause true, and thereby the 2-SAT problem satisfiable. As the second class of problems, we use a specially constructed set of hard (for quantum annealing) 2-SAT instances with four satisfying assignments [43].

To use quantum annealing for solving these problems, we map them to the Ising Hamiltonian

$$C_{2\mathrm{SAT}} = \sum_{\alpha=1}^{M} \left(\epsilon_{(\alpha,1)} s_{i[\alpha,1]} - 1\right)\left(\epsilon_{(\alpha,2)} s_{i[\alpha,2]} - 1\right), \qquad (3)$$

where $i[\alpha, j]$ represents the variable i that is involved in the jth term of the αth clause for $i = 1, \ldots, N$, $j = 1, 2$, and $\alpha = 1, \ldots, M$. If this variable is x_i then $\epsilon_{(\alpha,j)} = 1$, whereas if it its negation $\overline{x_i}$ then $\epsilon_{(\alpha,j)} = -1$.

The resulting problem Hamiltonians have been found to have an exponentially increasing degeneracy of the first excited state with growing problem size [43]. In this paper we focus on 2-SAT problems with 6–14 variables.

Ferromagnetic spin chains. As the last class of problems, we use instances of ferromagnetic spin chains, with neighbors connected by a uniform ferromagnetic coupling $J = -0.1$. The ground state of these problems is twofold degenerate ($|\uparrow\uparrow \ldots \uparrow\rangle$ and $|\downarrow\downarrow \ldots \downarrow\rangle$) and has an energy of $J(N-1)$, where N is the number of spins in the chain.

IV. KEY RESULTS

This section showcases a few key experimental results produced by the quantum annealers which serve as the basis for the subsequent investigation.

We start by looking at the simplest case, a one-spin problem, with $h_1 = 0.1$. The results for the WTS and the ATS are shown in Fig. 1. We initialize the system in the ground state $|\downarrow\rangle$ [Figs. 1(a) and 1(c)] or first excited state $|\uparrow\rangle$ [Figs. 1(b) and 1(d)]. We note that while $p(t_{\mathrm{end}})$ for the state that the system is initialized in starts from a value close to one for small t_{end}, it systematically decreases till a certain value of t_{end}, beyond which it seems to stabilize, especially for the WTS ($p_\downarrow \approx 0.8$ for large t_{end}). The other state shows an opposite trend, i.e., the corresponding $p(t_{\mathrm{end}})$ starts from a value close to zero, increases to a certain value with increasing t_{end}, and tends to stabilize. Furthermore, it can be seen that the rate of decrease (increase) of $p(t_{\mathrm{end}})$ for the state which was (not) the initial state is faster for the WTS as compared to the ATS. The results shown here are from the D-Wave Advantage_5.4 system, but other D-Wave annealers also yield a similar systematic behavior of $p(t_{\mathrm{end}})$ (data not shown).

In Fig. 2 we show the WTS for the three instances of the two-spin problem, where we choose the state $|\uparrow\uparrow\rangle$ as the

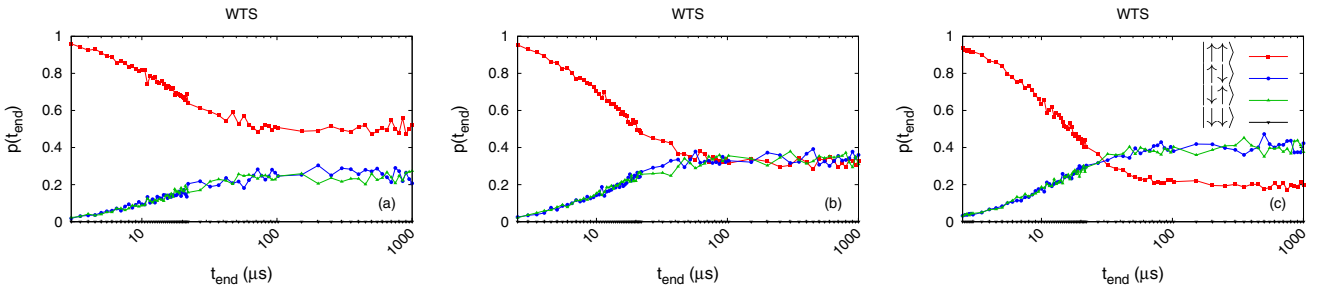

FIG. 2. WTS data for the two-spin problem instance (a) 2S1, (b) 2S2, and (c) 2S3 obtained from the D-Wave annealer with $s_r = 0.7$. The legend in panel (c) also applies to panels (a) and (b).

initial state. In this case we see that $p(t_{\text{end}})$ for the $|\downarrow\downarrow\rangle$ state remains close to zero for all values of t_{end}. For the other states, as for the one-spin cases, we see a similar systematic decrease (increase) in $p(t_{\text{end}})$ corresponding to the state in which the system was (not) initialized, followed by a stabilization of the $p(t_{\text{end}})$ values.

The results for the one- and the two-spin problems are, apart from minor details, reproducible and independent of the choice of the specific D-Wave system. This consistency points to a distinctive behavior of D-Wave quantum annealers, which is worth investigating further.

These initial observations combined provide a good motivation to compare the final probabilities $p(t_{\text{end}})$ of the various states of the problem with those resulting from the equilibrium distribution.

Conjecture. For sufficiently large t_{end}, the probabilities $p(t_{\text{end}})$, converge to the values predicted by the thermal equilibrium distribution.

According to statistical mechanics, the equilibrium distribution is given by

$$p_i^{\text{equil}} = \frac{g_i e^{-\beta E_i}}{\sum_i g_i e^{-\beta E_i}}, \tag{4}$$

where g_i and E_i are the degeneracy and the energy, respectively, of the ith level of the problem Hamiltonian, with $E_{i+1} > E_i$, and $\beta = \eta/T$ for $\eta = h\,B(s=1)/(2k_B) \times 10^9 = 0.206$ K and some effective temperature T (expressed in kelvin). Using Eq. (4) and the ground state success probability obtained from the quantum annealer for the largest value of t_{end} from the WTS of the one-spin problem with $h_1 = 0.1$ [Figs. 1(a) and 1(b)], we find $\beta = 6.93$, which corresponds to an effective temperature $T = 29.7$ mK, a value that is of a similar order as the cryogenic temperature of 15 mK, typical for the D-Wave annealers [43]. Next, using the value β obtained, we compute the ground state probabilities of the three two-spin instances according to the equilibrium distribution. For instance 2S1, this results in $p_{\uparrow\uparrow} = 0.50$, $p_{\uparrow\uparrow} = p_{\uparrow\downarrow} = p_{\downarrow\uparrow} = 0.25$, and $p_{\downarrow\downarrow} = 0$. For 2S2, we obtain $p_{\uparrow\uparrow} = p_{\uparrow\downarrow} = p_{\downarrow\uparrow} = 0.33$ and $p_{\downarrow\downarrow} = 0$, while for 2S3, Eq. (4) yields $p_{\uparrow\uparrow} = 0.20$, $p_{\uparrow\downarrow} = p_{\downarrow\uparrow} = 0.40$, and $p_{\downarrow\downarrow} = 0$. These theoretical values closely match the probabilities obtained from the D-Wave annealer for long t_{end} for these problems, thereby increasing our confidence in the conjecture that the probabilities obtained by reverse annealing on the D-Wave quantum annealers seem to relax to equilibrium probabilities for sufficiently long total annealing times.

V. SIMULATION RESULTS

As discussed in the previous section, the D-Wave results are starkly different from those of ideal quantum annealing, according to which the success probability approaches one for sufficiently long annealing times. For our numerical model to capture the observed features, we need to incorporate the elements of dissipation and decoherence. To this end, we use the Gorini-Kossakowski-Sudarshan-Lindblad master equation, which approximates the Schrödinger dynamics of the reduced density matrix for a system interacting with an environment [46–48].

The Lindblad master equation reads [48]

$$\frac{d\rho(t)}{dt} = -\frac{i}{\hbar}[H(t), \rho(t)] + \frac{1}{2}\sum_j \gamma_j [2L_j\rho(t)L_j^\dagger - L_j^\dagger L_j \rho(t) - \rho(t)L_j^\dagger L_j], \tag{5}$$

where $\rho(t)$ is the density matrix of the system and $\gamma_j \geqslant 0$ are the damping rates corresponding to the dissipation operators L_j. In general, the operators L_j are linear combinations of the matrices that form a basis for the matrices operating on the Hilbert space of the system [48].

The question that remains is whether there exist choices of dissipation operators that can reproduce the results obtained from the quantum annealers for the one- and the two-spin problems presented in the previous section. In the following, we tackle these cases one by one.

A. One-spin problems: Bloch equations

In general, the one-spin Hamiltonian can be written as

$$H = -\tfrac{1}{2}\mathbf{B}\cdot\boldsymbol{\sigma}, \tag{6}$$

where $\mathbf{B}$ is the applied magnetic field and $\boldsymbol{\sigma} = (\sigma^x, \sigma^y, \sigma^z)$ are the Pauli matrices. Substituting $B^x = 2\pi A(s)/h$, $B^y = 0$, and $B^z = -2\pi B(s)h_1/h$ transforms Eq. (6) into Eq. (1) for a single-spin system.

According to linear algebra, the 2×2 density matrix for the state of a spin-1/2 object can be completely expressed in the basis of the three Pauli matrices plus the unity matrix $\mathbb{I}$, i.e.,

$$\rho(t) = \frac{1}{2}\left[\mathbb{I} + \sum_\alpha S^\alpha(t)\cdot\sigma^\alpha\right], \tag{7}$$

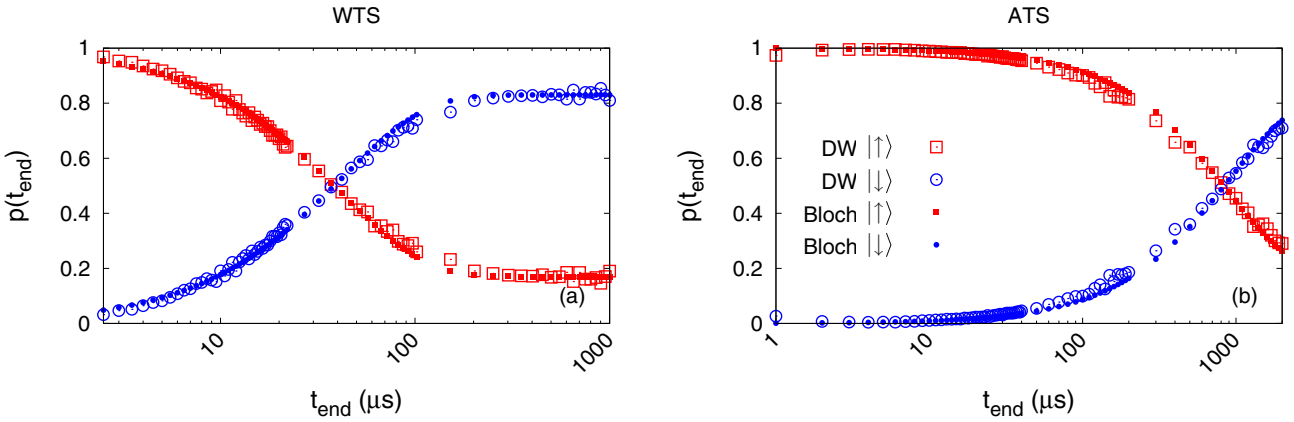

FIG. 3. Comparison of Bloch equations simulation results with those from the D-Wave annealer for (a) WTS and (b) ATS for the one-spin problem with $h_1 = 0.1$ and $s_r = 0.7$. For (a) $T_1 = 41.67\,\mu s$, $T_2 = 41.67\,\mu s$, and $M_0 = -0.66$ while for (b) $T_1 = 909.09\,\mu s$, $T_2 = 909.09\,\mu s$, and $M_0 = -0.66$. To emphasize that each data point from the simulations is obtained from an independent run we show each simulation data point as a marker. For improved legibility, in the subsequent figures, data points from our simulations are represented by lines through these points instead of by markers. The legend in panel (b) also applies to panel (a).

where $\mathbf{S(t)} = (S^x(t), S^y(t), S^z(t))$ is a vector of real numbers satisfying $\sum_\alpha (S^\alpha)^2 \leqslant 1$ for $\alpha = x, y, z$. To solve the Lindblad equation (5) for this system, we need to find the dissipation operators L_j that can produce results comparable to those obtained from the D-Wave annealers. One such choice for the dissipation operators is

$$L_1 = \sigma^+ = \frac{1}{2}(\sigma^x + i\sigma^y) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$
$$L_2 = \sigma^- = \frac{1}{2}(\sigma^x - i\sigma^y) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$
$$L_3 = \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{8}$$

As outlined in Appendix B, with this choice, the Lindblad equation (5) is equivalent to the Bloch equations given by

$$\frac{dS^x(t)}{dt} = S^y(t)B^z(t) - S^z(t)B^y(t) - \frac{S^x(t)}{T_2},$$
$$\frac{dS^y(t)}{dt} = S^z(t)B^x(t) - S^x(t)B^z(t) - \frac{S^y(t)}{T_2},$$
$$\frac{dS^z(t)}{dt} = S^x(t)B^y(t) - S^y(t)B^x(t) - \frac{S^z(t) - M_0}{T_1}, \tag{9}$$

with $T_2 = 2/(\gamma_1 + \gamma_2 + 4\gamma_3)$, $T_1 = 1/(\gamma_1 + \gamma_2)$, $M_0 = (\gamma_1 - \gamma_2)/(\gamma_1 + \gamma_2)$ denoting the transverse and longitudinal relaxation time and the equilibrium magnetization, respectively. Note that this choice of the dissipation operators yields $T_2 \leqslant 2T_1$. The numerical method used for solving Eq. (9) is discussed in Appendix D.

Figure 3 shows a comparison of the D-Wave results with those of the WTS and ATS for the one-spin problem with $h_1 = 0.1$ obtained from the simulations in which state $|\uparrow\rangle$ was chosen as the initial state. From the same figure it is evident that with the above-mentioned choice of the dissipation operators and with $T_1 = 41.67\,\mu s$, $T_2 = 41.67\,\mu s$, and $M_0 = -0.66$ for the WTS and $T_1 = 909.09\,\mu s$, $T_2 = 909.09\,\mu s$, and $M_0 = -0.66$ for the ATS, the simulation can reproduce both the trend of the probability curves and the final value of the probability for large t_{end} rather well. The fact that the relaxation times T_1 and T_2 obtained from fitting Bloch equations to the D-Wave data results in significantly different values for WTS and ATS only reflects that the underlying procedures are very different. Note that the relaxation times for a single qubit are obtained by procedures very different from the WTS or ATS.

A similar treatment of the one-spin problem with $h_1 = 0$ results in an oscillatory behavior between the two states of the system, as shown in Fig. 4. These oscillations are not present in the data from the quantum annealers and signal the coherent motion of the spin between the two degenerate states at $s = 1$. However, slightly changing the value of the field $h_1 = 0.001$ eliminates these oscillations from the simulation data and produces results that closely match those from D-Wave. As the D-Wave annealers are known to be prone to small errors in implementing the specified h and J values of a problem exactly, the absence of oscillations in the D-Wave data suggests that a minute amount of error in problem representation can lift the degeneracy between the energy levels. This suggests that, in their current state, the D-Wave annealers are incapable of implementing problems with exact degeneracies.

B. Two-spin problems: Lindblad master equation

We now turn to the two-spin problem instances discussed in the previous section. Even for such simple cases, selecting the appropriate dissipation operators demands careful deliberation. The mathematical considerations and the choice of the dissipation operators for this case are presented in Appendix C, while the other numerical aspects for the Lindblad master equation simulation are outlined in Appendix D.

In Fig. 5 we show the results obtained from the simulations of the ATS for these problems, in comparison to those from the annealers. Choosing a value of β that makes the probabilities from the equilibrium distribution (4) close to the probabilities $p(t_{end})$ obtained from D-Wave for the largest value of t_{end}, we determine ratios of the dissipation rates by setting the coherent part of the dynamics [the first term of

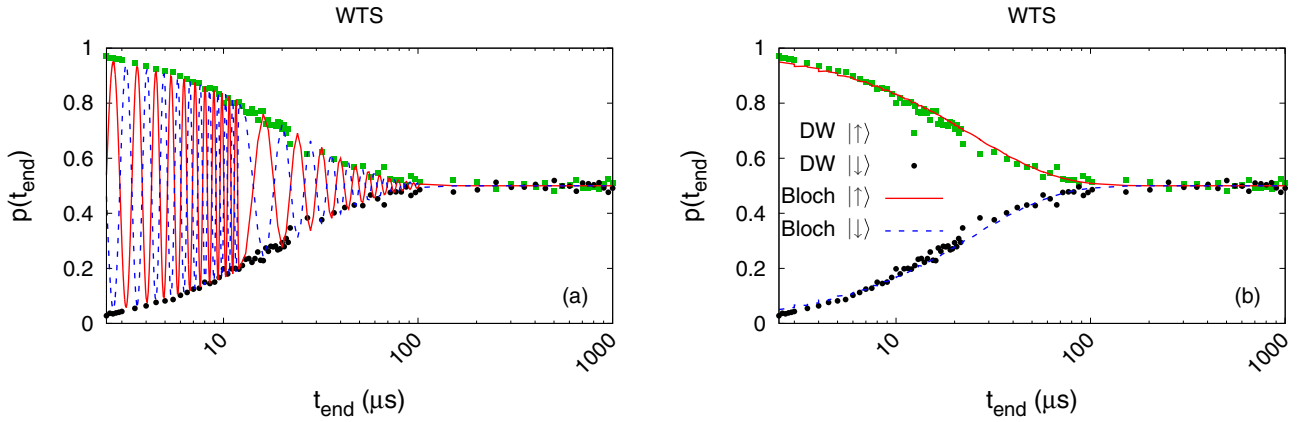

FIG. 4. Comparison of the WTS data from Bloch equation simulations (lines) for the one-spin problem with $s_r = 0.7$ and (a) $h_1 = 0$ and (b) $h_1 = 0.001$ with the D-Wave data (markers) for the same with $h_1 = 0$. For both plots $T_1 = 25\,\mu\text{s}$, $T_2 = 25\,\mu\text{s}$, and $M_0 = 0$. The legend in panel (b) also applies to panel (a).

the right-hand side of Eq. (C2)] to zero. This approach is described in more detail in Sec. V C. The resulting relations between the equilibrium probabilities and the dissipation rates are given by Eq. (13). We find that with this choice for the dissipation rates, the simulations can reproduce the behavior of the $p(t_{\text{end}})$ data produced by the D-Wave quantum annealers rather well, except for a few points corresponding to small t_{end}; see, for instance, 2S3.

Recall that the ground state of the one-spin problem with $h_1 = 0$ is twofold degenerate, and that if the initial state was chosen to be $|\uparrow\rangle$, the simulation data show oscillations between the two states. For the two-spin instance 2S2, the ground state is threefold degenerate, and initializing the system in state $|\uparrow\uparrow\rangle$ also results in an oscillatory behavior of the corresponding probabilities (data not shown). However, as noted for the above-mentioned one-spin problem, slightly lifting this degeneracy by setting $h_1 = 1.001$ and $h_2 = 0.999$, eliminates the oscillations. Figure 5(b) shows simulation results for this slightly modified version of instance 2S3.

C. Larger problems: Markovian master equation

Our numerical results show that in cases of good agreement between the simulations and the D-Wave data, the absolute values of the nondiagonal elements of the corresponding density matrix are rather small (data not shown), suggesting that the contributions of the coherent part of the evolution can be ignored. Indeed, in the regime of interest, i.e., for $s \geqslant 0.7$, the effects of the transverse field produced by the σ^x terms are negligible since according to the annealing schedule given by D-Wave $A(s)/h \leqslant 0.002 \ll B(s)/h$ for $s \geqslant 0.7$. These observations suggest that it might be worthwhile to investigate the case in which we set $A(s) = 0$ and retain only the diagonal elements of the density matrix $\rho(t)$. The resulting $\rho(t)$ then commutes with Hamiltonian, and the Lindblad equations for the one- and two-spin problem reduce to

$$\frac{d}{dt}\begin{pmatrix} p_\uparrow \\ p_\downarrow \end{pmatrix} = \begin{pmatrix} -\gamma_2 & \gamma_1 \\ \gamma_2 & -\gamma_1 \end{pmatrix}\begin{pmatrix} p_\uparrow \\ p_\downarrow \end{pmatrix}, \quad p_\uparrow + p_\downarrow = 1 \tag{10}$$

and

$$\frac{d}{dt}\begin{pmatrix} p_{\uparrow\uparrow} \\ p_{\uparrow\downarrow} \\ p_{\downarrow\uparrow} \\ p_{\downarrow\downarrow} \end{pmatrix} = \begin{pmatrix} -\gamma_2-\gamma_5-\gamma_7 & \gamma_4 & \gamma_6 & \gamma_1 \\ \gamma_5 & -\gamma_4 & 0 & 0 \\ \gamma_7 & 0 & -\gamma_6 & 0 \\ \gamma_2 & 0 & 0 & -\gamma_1 \end{pmatrix}\begin{pmatrix} p_{\uparrow\uparrow} \\ p_{\uparrow\downarrow} \\ p_{\downarrow\uparrow} \\ p_{\downarrow\downarrow} \end{pmatrix},$$
$$p_{\uparrow\uparrow} + p_{\uparrow\downarrow} + p_{\downarrow\uparrow} + p_{\downarrow\downarrow} = 1, \tag{11}$$

respectively. The stationary solutions for Eqs. (10) and (11) are

$$\gamma_2 p_\uparrow = \gamma_1 p_\downarrow \tag{12}$$

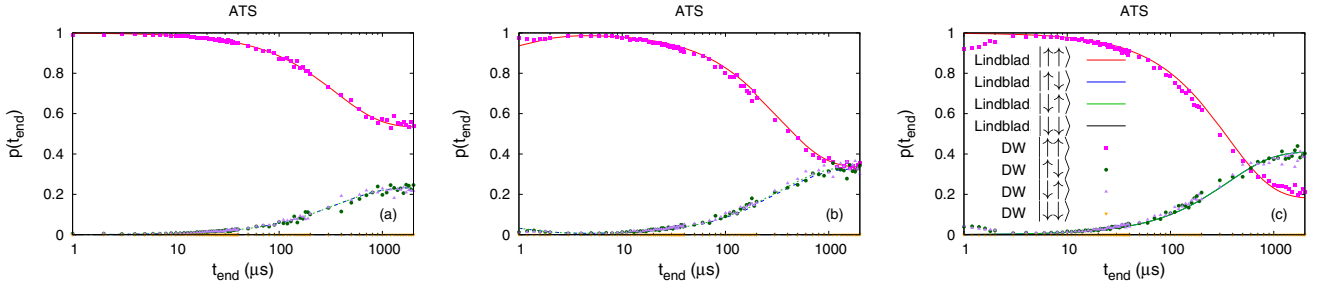

FIG. 5. Comparison of the ATS data from D-Wave (markers) with that from Lindblad master equation simulation (lines) for two-spin instances with $s_r = 0.7$ (a) 2S1, (b) 2S2, and (c) 2S3, with dissipation rates (a) $\gamma_1 = \gamma_3 = \gamma_4 = \gamma_6 = 1.5$ Hz, $\gamma_2 = 0$, $\gamma_5 = \gamma_7 = 0.6582$ Hz, (b) $\gamma_1 = \gamma_3 = \gamma_4 = \gamma_6 = 1.0$ Hz, $\gamma_2 = 0$, $\gamma_5 = \gamma_7 = 0.9837$ Hz, (c) $\gamma_1 = \gamma_3 = \gamma_4 = \gamma_6 = 0.5$ Hz, $\gamma_2 = 0$, $\gamma_5 = \gamma_7 = 1.1395$ Hz. The legend in panel (c) also applies to panels (a) and (b).

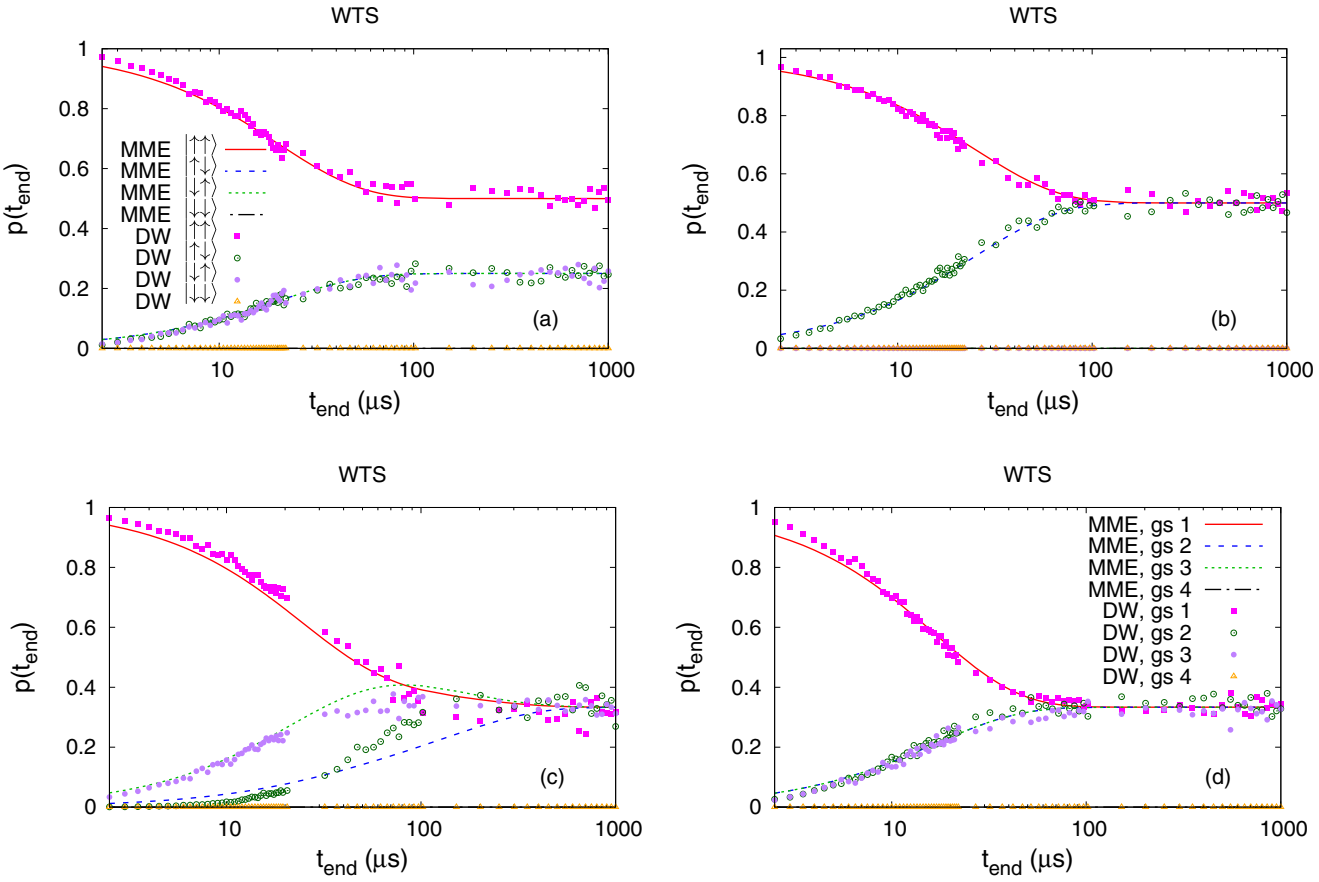

FIG. 6. Comparison of the WTS data from D-Wave (markers) with that from a Markovian master equation (MME) simulation (lines) with $s_r = 0.7$ for (a) instance 2S1, an instance of (b) six-variable, (c) 12-variable, and (d) 14-variable 2-SAT problems, the last three cases with four degenerate ground states (gs 1–4) for dissipation rates (a) $\gamma_1 = \gamma_3 = \gamma_4 = \gamma_6 = 25$ Hz, $\gamma_2 = 0$, $\gamma_5 = \gamma_7 = 12.5$ Hz, (b) $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = \gamma_6 = \gamma_7 = 20$ Hz (c) $\gamma_1 = 20$ Hz, $\gamma_2 = 0$, $\gamma_3 = \gamma_6 = \gamma_7 = 10$ Hz, $\gamma_4 = \gamma_5 = 15$ Hz, (d) $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5 = \gamma_6 = \gamma_7 = 20$ Hz. The legend in panel (d) also applies to panels (b) and (c).

and

$$\gamma_2 p_{\uparrow\uparrow} = \gamma_1 p_{\downarrow\downarrow}, \quad \gamma_5 p_{\uparrow\uparrow} = \gamma_4 p_{\downarrow\uparrow}, \quad \gamma_7 p_{\uparrow\uparrow} = \gamma_6 p_{\uparrow\downarrow}, \tag{13}$$

respectively.

Equations (10) and (11) are special cases of

$$\frac{dP(t)}{dt} = WP(t), \tag{14}$$

where $P(t) = (p_1(t), \ldots, p_N(t))^T$ is a vector of non-negative elements that sum to one, and W is a real-valued matrix. For instance, in the case of the two-spin model, $P(t)$ and W are given by the vector and matrix in Eq. (11), respectively.

The formal solution of Eq. (14) reads

$$P(t) = e^{tW} P(0). \tag{15}$$

As the columns of W add to zero, it follows immediately that $\exp(tW)$ is a Markov matrix.

For the 2-SAT problems with four satisfying assignments [43], the dimension of the relevant subspace is four, and we can still use Eq. (11) to study the relaxation processes.

The corresponding results for the two-spin problem, as well as for larger instances of 2-SAT problems with four satisfying assignments, are shown in Fig. 6. For the 2-SAT problems, this figure shows the sampling probabilities of the four known degenerate ground states of these problems (numbered in a different order than in Ref. [43]).

The good agreement between the numerical and the D-Wave data suggests that the salient features of the results produced by D-Wave's reverse annealing protocol can indeed be captured by this straightforward Markovian approach.

Figure 6(a) shows that this approach achieves a similar level of agreement with the D-Wave data as by solving the Lindblad master equation for the two-spin systems (see Fig. 5 for a representative example).

As discussed in [43], for the shown instances of the 2-SAT problems, the ideal quantum annealing simulations show a suppression of the sampling probabilities of one or more ground states, a behavior that could be explained on the basis of perturbation theory. While the standard forward annealing protocol on the D-Wave annealer did not show any such suppression of any of the ground states [43], the suppression of these ground states using the reverse annealing protocol, as evident in Figs. 6(b)–6(d), begs for further investigation. However, it has been observed that ground states with suppressed sampling probabilities have a large Hamming distance from other ground states [43]. This idea can be visualized by imagining the ground subspace as divided into two subspaces, one containing states that are reachable by (a sequence of)

single-spin flips starting from any of the other states, and the other containing ground states that have a Hamming distance greater than one from all the states in the first subspace. Drawing an analogy to the Metropolis Monte Carlo algorithm, at very low temperatures, only degenerate ground states separated by a Hamming distance of one can be reached when starting from one of the lowest-energy configurations. Accessing other ground states requires traversing higher energy states, which becomes more likely at higher temperatures. This scenario bears resemblance to reverse quantum annealing results from D-Wave annealers, where the system remains in a regime of small quantum fluctuations, having been initialized in one of the ground states [see Figs. 6(b)–6(d)]. A possible explanation for the suppression of these ground states in the D-Wave results could therefore be that quantum fluctuation plus noise sources within the system act like thermal fluctuations that are unable to couple the two ground subspaces characterized by the Hamming distance among the ground states.

Refocusing on the primary idea, we find that, remarkably, using the simple nonquantum Markovian description makes it possible to reproduce D-Wave results by circumventing the problem of finding the appropriate dissipation operators for the Lindblad equation for simulating these large problems, which is already a nontrivial task for the two-spin problems. We initialize the system in the first excited state, and for given values of h_1 and s_r, we perform the WTS by fixing $t_{\text{reverse}} = t_{\text{forward}} = 1\,\mu\text{s}$ for all 5000 copies of the problem.

VI. FURTHER D-WAVE EXPERIMENTS

From the results presented so far, we have seen the effects of varying the waiting time and the annealing times on the probability values $p(t_{\text{end}})$ for the different states of the problem. However, these WTS or ATS were performed for specific problems, and the reversal distance $s_r = 0.7$.

As the next step, it is interesting to investigate how the results of these scans change when either the energy gap between the ground state and the first excited state of the problem Hamiltonian, or the reversal distance in the reverse annealing protocol, is varied. To this end we first perform the WTS for the one-spin problem for various values of h_1, fixing $s_r = 0.7$. The value of the corresponding energy gap is given by

$$\Delta = 2\sqrt{A^2(s) + B^2(s)h_1^2} \tag{16}$$

at $s = s_r$. Next, setting $h_1 = 0.3$, we perform WTS for various values of the reversal distance s_r. To improve the statistics, we embed 5000 copies of the problem on the D-Wave system for each point of the scan.

The results obtained from these experiments are shown in Fig. 7, where Figs. 7(a) and 7(c) show the mean success probabilities $\langle p_0(t_{\text{end}})\rangle$ (markers) for different values of h_1 and s_r, respectively, obtained by averaging over the 5000 copies. The solid lines in these figures are obtained by fitting functions $f(t_{\text{end}}) = f_1(1 - f_2\exp(-f_3 t_{\text{end}}))$ to $\langle p_0(t_{\text{end}})\rangle$ obtained from the experiments for different values of h_1 (s_r) in Fig. 7(a) [Fig. 7(c)]. Recall that each data point in these panels has been obtained from an individual run with a fixed t_{wait}. Therefore, it is remarkable that these points fit very well to $f(t_{\text{end}})$.

Furthermore, Fig. 7(a), where s_r is fixed to 0.7, shows that for $h_1 \leqslant 1.6$, the $\langle p_0(t_{\text{end}})\rangle$'s saturate to values given by Eq. (4), in concert with the key results (Sec. IV). For $h_1 > 0.16$, the maximum annealing time (2000 μs) allowed on the D-Wave systems is too short to access the regime of equilibration. In Fig. 7(b) we present a plot of f_3 obtained from Fig. 7(a) with the energy gaps Δ calculated using Eq. (16), which fits well to the exponential function $0.12\exp(-0.06\Delta)$ for $\Delta > 1$. The exponential decrease of f_3 with Δ suggests that if the thus far observed systematic trend continues, increasing energy gaps between the ground state and the first excited states of the problems should postpone the attainment of the equilibrium probabilities to larger values of t_{end}. Such a behavior is consistent with thermal equilibration.

Next, keeping h_1 fixed at 0.3, we perform the WTS while varying the reversal distance $0.60 \leqslant s_r \leqslant 0.75$. Figure 7(c) shows $\langle p_0(t_{\text{end}})\rangle$ as a function of t_{end} for different values of s_r and suggests that the approach of the $\langle p_0(t_{\text{end}})\rangle$ to the equilibrium value given by Eq. (4) becomes slower with increasing values of s_r. In the allowed maximum annealing time limit of the D-Wave QPU of 2000 μs, this saturation can be observed only up to $s_r \leqslant 0.68$.

Figure 7(d) shows f_3 as a function of s_r. We find that f_3 fits well to $1615.79A(s_r)^{2.31}$, hinting that a stronger involvement of the transverse field leads to more fluctuations and a faster decay of $\langle p_0(t_{\text{end}})\rangle$. This observation could also explain the requirement for larger dissipation rates in our simulations for the one-spin case with $h_1 = 0.1$, shown in Fig. 3, to fit well to the D-Wave data in WTS as compared to the ATS. For the former, the system spends a larger proportion of time at a smaller $s = s_r$, resulting in a larger effective transverse field as compared to the latter where due to $t_{\text{wait}} = 0$, the effective $s \geqslant s_r$.

Next, we study the effects of varying the problem size on the WTS using ferromagnetic spin chains for $10 \leqslant N \leqslant 1000$. For this class of problems, it is straightforward to compute the equilibrium properties. This calculation shows that the ground state probability p_0^{equil} vanishes exponentially as the size of the problems increases, in concert with the corresponding D-Wave data [as shown in Fig. 10(a) in Appendix E]. To accurately estimate the ground state probability, one would therefore require an exponentially increasing number of samples as the problem size grows. Given the impracticality of this approach, a more feasible alternative is to analyze the equilibrium behavior of these problems using, for instance, the mean energy [see Fig. 10(b) in Appendix E for the absolute value of the mean energies]. In Fig. 8 we show the mean energy $\langle E(1952\,\mu\text{s})\rangle$ as a function of the problem size N. We determine β by fitting

$$\langle E\rangle = \frac{\sum_i g_i E_i e^{-\beta E_i}}{\sum_i g_i e^{-\beta E_i}} = -J(N-1)\tanh\beta J \tag{17}$$

to the empirical data for the average energy. We find that $\beta = 7.64$, which corresponds to a temperature of approximately 27 mK. This remarkably good fit strongly supports the equilibrium conjecture. In Appendix F we present the results

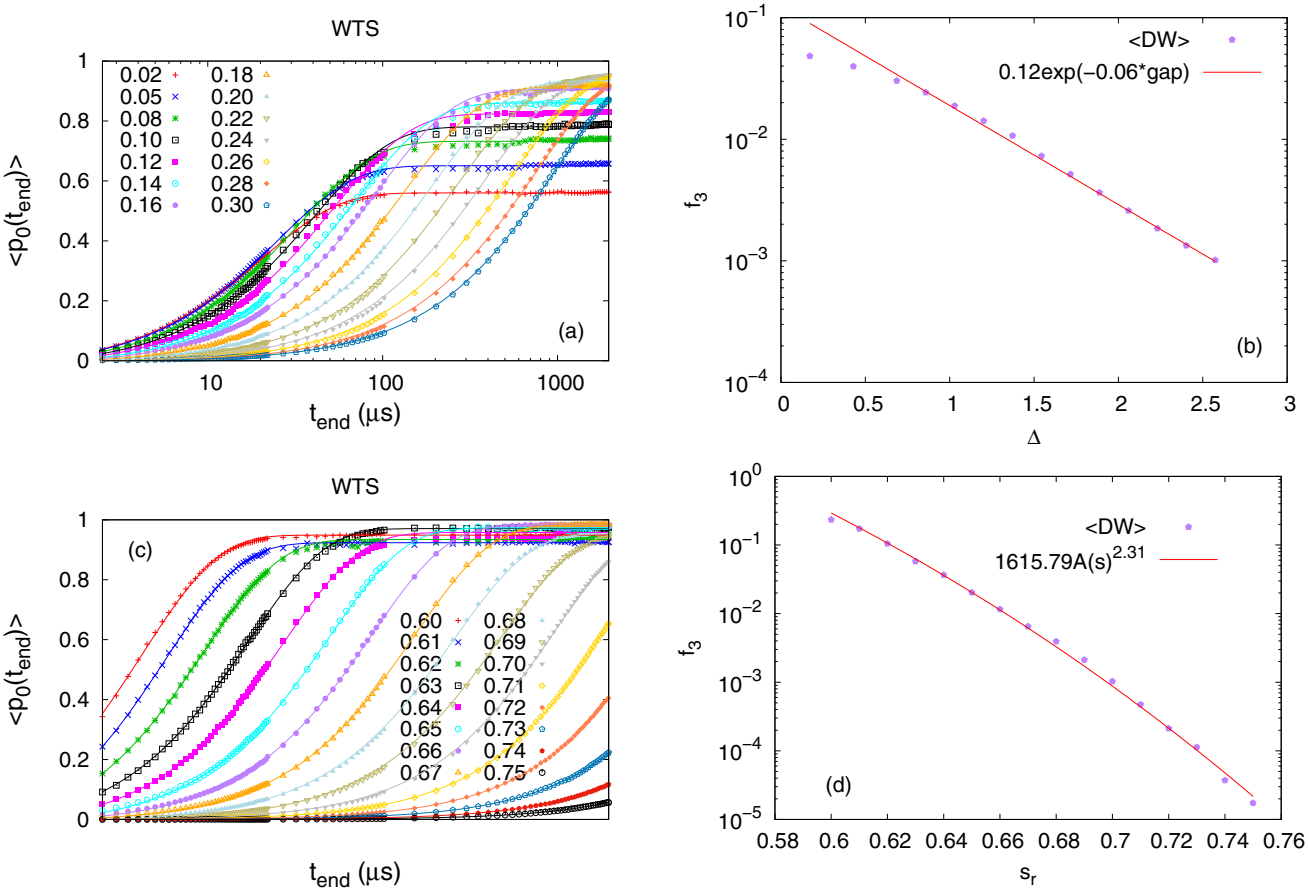

FIG. 7. (a) Data (markers) of the WTS for the mean success probability $\langle p_0(t_{\text{end}})\rangle$ obtained by averaging 5000 copies of one-spin problems with different h_1 and $s_r = 0.7$ on the D-Wave annealer with the corresponding fits (lines) to $f(t_{\text{end}}) = f_1[1 - f_2 \exp(-f_3 t_{\text{end}})]$, (b) scaling of the obtained f_3 as a function of energy gap Δ, (c) the same as (a) but for different values of reversal distance s_r and $h_1 = 0.3$, (d) the same as (b) but as a function of s_r.

of this compelling experiment using data from the Advantage_4.1 system, as a demonstration that these observations extend beyond a single D-Wave machine.

VII. CONCLUSION

We have shown that with the appropriate choice of dissipation operators and rates for the Lindblad master equation, it is possible to numerically reproduce the D-Wave results for reverse annealing. We found that for sufficiently long annealing times the D-Wave quantum annealer samples states with frequencies approaching the thermal equilibrium probabilities (i.e., in concert with the Gibbs distribution). The rate at which equilibrium is attained depends on quantities like the energy gap and the reversal distance, among other factors, limiting the possibility of approaching equilibrium within the maximum annealing time admitted by the D-Wave systems. Furthermore, we have shown that by ignoring in the simulations, the coherent parts, a nonquantum Markovian master equation can also reproduce the salient features of the D-Wave data. This suggests that the coherent part of the dynamics does not play a crucial role in the reverse annealing regime in D-Wave.

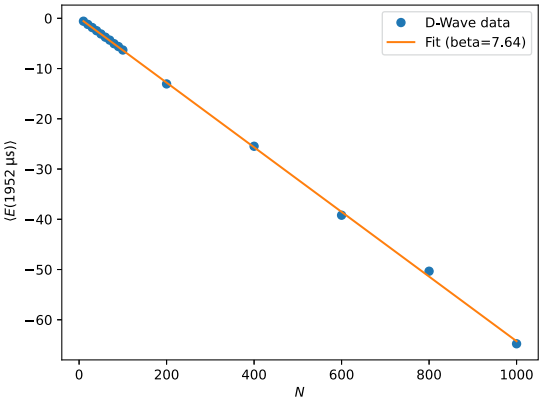

FIG. 8. Mean of empirical $E(1952\,\mu\text{s})$ data obtained from the D-Wave annealer fit to Eq. (17) with $\beta = 7.64$ corresponding to a temperature of 27 mK.

In light of these observations, it might be beneficial to develop other strategies for formulating/solving optimization problems using D-Wave quantum annealers.

Despite the strong agreement between our simulations and experimental observations, the actual physical mechanisms governing the system's dynamics remain elusive and cannot be conclusively identified based on the theoretical models analyzed in this paper.

Carrying out a similar study using standard forward annealing and the recently introduced fast annealing

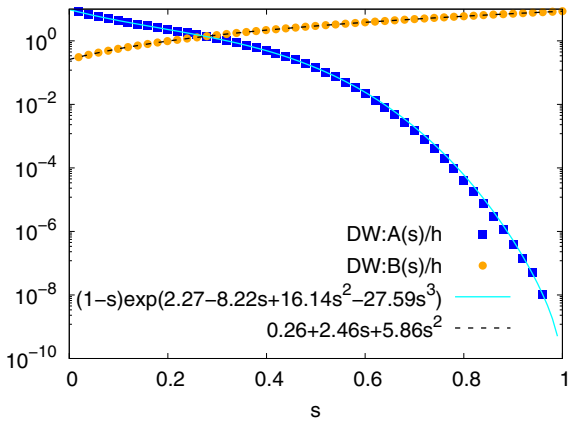

FIG. 9. Data for the annealing schedules $A(s)$ and $B(s)$ provided by D-Wave, and the functions used in the simulations.

feature of the D-Wave quantum annealers is left for future research.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS [49] at the Jülich Supercomputing Centre (JSC). V.M. acknowledges support from the project JUNIQ funded by the German Federal Ministry of Education and Research (BMBF) and the Ministry of Culture and Science of the State of North Rhine-Westphalia (MKW-NRW) and from the project EPIQ funded by MKW-NRW.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

APPENDIX A: ANNEALING SCHEDULE

As the D-Wave data for the annealing scheme are provided as tabulated values of $A(s)/h$ and $B(s)/h$ (in GHz), for numerical simulation, it is expedient to fit functions

$$A(s)/h = (1-s)\exp(A_a + A_b s + A_c s^2 + A_d s^3),$$
$$B(s)/h = B_a + B_b s + B_c s^2 \tag{A1}$$

to these annealing schedule data. The values of the parameters, obtained by fitting, are shown in Fig. 9.

APPENDIX B: MATHEMATICAL TREATMENT OF THE ONE-SPIN CASE

The most general Hamiltonian for a single-spin system is given by Eq. (6). From Eq. (7) it follows that the eigenvalues of $\rho(t)$ are $\{1 + [\sum_{k=x,y,z}(S^k)^2(t)]^{1/2}\}/2$. Therefore, we must have $\sum_{k=x,y,z}(S^k)^2(t) \leqslant 1$ for $\rho(t)$ to be a non-negative definite matrix with eigenvalues not exceeding one. From Eq. (7) it also follows that the expectation values of the spin operators are given by

$$\langle\sigma^\alpha(t)\rangle = \mathbf{Tr}\,\rho(t)\sigma^\alpha = S^\alpha(t), \quad \alpha = x, y, z, \tag{B1}$$

showing that in the case at hand, knowledge of $S^\alpha(t) = \langle\sigma^\alpha(t)\rangle$ is equivalent to the knowledge of the density matrix $\rho(t)$.

We start from Eq. (5) and derive the equation of motion for the expectation values of the spin operators. This facilitates the comparison with the Bloch equations and also helps to give meaning to the damping rates that enter Lindblad equation (5). Multiplying Eq. (5) by σ^l and computing the trace, we obtain the equations of motion of the spin expectation values [see Eq. (7)]. We have

$$\frac{dS^l(t)}{dt} = (\mathbf{S}(t)\times\mathbf{B})^l + \frac{1}{2}\sum_{j=1}^{3}\mathbf{Tr}\,\sigma^l[L_j, L_j^\dagger] + \frac{1}{4}\sum_{j,k=1}^{3}\gamma_j S^k(t)[2\mathbf{Tr}\,\sigma^l L_j\sigma^k L_j^\dagger - \mathbf{Tr}\,\sigma^l L_j^\dagger L_j\sigma^k - \mathbf{Tr}\,\sigma^l\sigma^k L_j^\dagger L_j]$$
$$= (\mathbf{S}(t)\times\mathbf{B})^l + \frac{1}{2}\sum_{j=1}^{3}\gamma_j\mathbf{Tr}\,\sigma^l[L_j, L_j^\dagger] + \frac{1}{2}\sum_{j,k=1}^{3}\gamma_j S^k(t)\mathbf{Tr}[L_j^\dagger, \sigma^l](L_j\sigma^k). \tag{B2}$$

With the specific choice of the dissipation operators given by Eq. (8), the right-hand side of Eq. (B2) can be worked out analytically, yielding the Bloch equations (9). It shall also be noted that the reduction of the Lindblad master equation (5) to Bloch equation (9) is associated with the specific choice of the dissipation operators (8). For a different choice of the dissipation operators, the Lindblad master equation might not reduce to the Bloch equations.

APPENDIX C: MATHEMATICAL TREATMENT OF THE TWO-SPIN SYSTEM

For the two-spin systems, we need to choose an appropriate basis to represent 4×4 matrices. A convenient choice is the set of 4×4 matrices constructed by taking direct products of two matrices from the set $\mathbb{I}, \sigma^x, \sigma^y, \sigma^z$, e.g., $e_1 = \mathbb{I}\otimes\mathbb{I}/2$, $e_2 = \mathbb{I}\otimes\sigma^x/2$, $e_3 = \mathbb{I}\otimes\sigma^y/2$, $e_4 = \mathbb{I}\otimes\sigma^z/2, \ldots, e_{16} = \sigma^z\otimes\sigma^z/2$. We have

$$\rho(t) = \sum_{k=1}^{16} x_k(t)\mathbf{e}_k, \tag{C1}$$

where the $x_k(t)$'s are real-valued variables. Using the properties of the $\mathbf{e}_k$'s, it follows that $\mathbf{Tr}\,\rho(t) = 1$ implies $x_1(t) = 1/2$. The equation of motion for $x_k(t)$ for $k = 2, \ldots, 16$ is found by multiplying the Lindblad equation (5) by $\mathbf{e}_k$ and computing the trace:

$$\frac{dx_k(t)}{dt} = \frac{d\mathbf{Tr}\,\rho(t)\mathbf{e}_k}{dt} = \sum_{l=1}^{16}\left[-i\mathbf{Tr}[H,\mathbf{e}_l]\mathbf{e}_k + \frac{1}{2}\sum_j \gamma_j(2\mathbf{Tr}L_j\mathbf{e}_l L_j^\dagger \mathbf{e}_k - \mathbf{Tr}L_j^\dagger \mathbf{e}_l\mathbf{e}_k - \mathbf{Tr}\,\mathbf{e}_l L_j^\dagger L_j\mathbf{e}_k)\right]x_l(t). \tag{C2}$$

Next, we have to choose the dissipation operators L's that can describe the D-Wave data well, i.e., using which we can reproduce the trend for probabilities $p(t_{\text{end}})$'s and the stationary value that they seem to approach. The minimal choice that is found to describe this behavior is

$$L_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_2 = L_1^{\mathrm{T}}, L_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad L_4 = L_3^{\mathrm{T}}, \quad L_5 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$L_6 = L_5^{\mathrm{T}}, \quad L_7 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{C3}$$

Using *Mathematica* to compute the explicit form of the equations of motion of the $x_k(t)$'s we obtain

$$\begin{aligned}
x_1 &= \tfrac{1}{2},\\
\partial_t x_2 &= \tfrac{1}{4}[8Bh_1x_3 + 8BJx_{15} + (\gamma_1 - \gamma_2 - \gamma_4 - \gamma_5 + \gamma_6 - \gamma_7)x_{14} - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_2],\\
\partial_t x_3 &= \tfrac{1}{4}[-8(Ax_4 + Bh_1x_2 + BJx_{14}) + (\gamma_1 - \gamma_2 - \gamma_4 - \gamma_5 + \gamma_6 - \gamma_7)x_{15} - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_3],\\
\partial_t x_4 &= \tfrac{1}{4}[8Ax_3 + \gamma_1 - \gamma_2 + \gamma_4 - \gamma_5 + 2(\gamma_1 - \gamma_2 - \gamma_4 - \gamma_5)x_{16} - 2(\gamma_1 + \gamma_2 - \gamma_4 + \gamma_5)x_7 - 2(\gamma_1 + \gamma_2 + \gamma_4 + \gamma_5)x_4],\\
\partial_t x_5 &= \tfrac{1}{4}[8Bh_2x_6 + 8BJx_{13} + (\gamma_1 - \gamma_2 + \gamma_4 - \gamma_5 - \gamma_6 - \gamma_7)x_{10} - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_5],\\
\partial_t x_6 &= \tfrac{1}{4}[-8(Ax_7 + Bh_2x_5 + BJx_{10}) + (\gamma_1 - \gamma_2 + \gamma_4 - \gamma_5 - \gamma_6 - \gamma_7)x_{13} - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_6],\\
\partial_t x_7 &= \tfrac{1}{4}[8Ax_6 - \gamma_2 + \gamma_6 - \gamma_7 + \gamma_1(-2x_4 - 2x_7 + 2x_{16} + 1) - 2(\gamma_2 - \gamma_6 + \gamma_7)x_4 - 2(\gamma_2 + \gamma_6 + \gamma_7)(x_7 + x_{16})],\\
\partial_t x_8 &= \tfrac{1}{4}[8Bh_1x_9 + 8Bh_2x_{11} + (\gamma_1 + \gamma_2 - \gamma_4 + \gamma_5 - \gamma_6 + \gamma_7)x_{12} - (\gamma_1 + \gamma_2 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_8],\\
\partial_t x_9 &= \tfrac{1}{4}[-8Ax_{10} - 8Bh_1x_8 + 8Bh_2x_{12} - (\gamma_1 + \gamma_2 - \gamma_4 + \gamma_5 - \gamma_6 + \gamma_7)x_{11} - (\gamma_1 + \gamma_2 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_9],\\
\partial_t x_{10} &= \tfrac{1}{4}[8Ax_9 + 8Bh_2x_{13} + 8BJx_6 + (\gamma_1 - \gamma_2 + \gamma_4 - \gamma_5 - \gamma_6 - \gamma_7)x_5 - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{10}],\\
\partial_t x_{11} &= \tfrac{1}{4}[-8Ax_{14} - 8Bh_2x_8 + 8Bh_1x_{12} - (\gamma_1 + \gamma_2 - \gamma_4 + \gamma_5 - \gamma_6 + \gamma_7)x_9 - (\gamma_1 + \gamma_2 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{11}],\\
\partial_t x_{12} &= \tfrac{1}{4}[-8(A(x_{13} + x_{15}) + Bh_2x_9 + Bh_1x_{11}) + (\gamma_1 + \gamma_2 - \gamma_4 + \gamma_5 - \gamma_6 + \gamma_7)x_8 - (\gamma_1 + \gamma_2 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{12}],\\
\partial_t x_{13} &= \tfrac{1}{4}[-8(A(x_{16} - x_{12}) + Bh_2x_{10} + BJx_5) + (\gamma_1 - \gamma_2 + \gamma_4 - \gamma_5 - \gamma_6 - \gamma_7)x_6 - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{13}],\\
\partial_t x_{14} &= \tfrac{1}{4}[8Ax_{11} + 8Bh_1x_{15} + 8BJx_3 + (\gamma_1 - \gamma_2 - \gamma_4 - \gamma_5 + \gamma_6 - \gamma_7)x_2 - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{14}],\\
\partial_t x_{15} &= \tfrac{1}{4}[-8(-Ax_{12} + Ax_{16} + Bh_1x_{14} + BJx_2) + (\gamma_1 - \gamma_2 - \gamma_4 - \gamma_5 + \gamma_6 - \gamma_7)x_3 - (\gamma_1 + \gamma_2 + 8\gamma_3 + \gamma_4 + \gamma_5 + \gamma_6 + \gamma_7)x_{15}],\\
\partial_t x_{16} &= \tfrac{1}{4}[8A(x_{13} + x_{15}) - \gamma_5 + \gamma_6 - \gamma_7 + \gamma_4(-2x_4 + 2x_7 - 2x_{16} + 1) - 2(\gamma_5 - \gamma_6 + \gamma_7)x_4 - 2(\gamma_5 + \gamma_6 + \gamma_7)(x_7 + x_{16})].
\end{aligned} \tag{C4}$$

Since $\mathbf{Tr}\,\mathbf{e}_1 = 2$, and $\mathbf{Tr}\,\mathbf{e}_j = 0$ for $j > 1$, from Eq. (C1) it follows that $x_1 = 1/2$ implies that the trace of the density matrix is always equal to one.

APPENDIX D: NUMERICAL SOLUTION

In matrix-vector notation, differential equations (B2) and (C2) take the form

$$\frac{dx}{dt} = [\mathbf{C}(t) + \mathbf{D}]\mathbf{x}(t) + \mathbf{y}, \tag{D1}$$

where $\mathbf{C}(t)$ is a time-dependent matrix describing the coherent motion of the spin(s), $\mathbf{D}$ accounts for the decoherence and dissipation, and the vector $\mathbf{y}$ is a time-independent source term. In the case of the single-spin $\mathbf{C}(t)$ is a 3×3 skew-Hermitian matrix, $\mathbf{D}$ is a 3×3 nonpositive diagonal matrix, and the value of $y_3 = M_0/T_1$ determines the stationary value of the longitudinal component of the spin. For the two-spin system, excluding the trivial equation for $x_1(t)$, $\mathbf{C}(t)$ is a 15×15 skew-Hermitian matrix, and $\mathbf{D}$ is a 15×15 matrix with no obvious symmetry properties, and the vector $\mathbf{y}$ with 15 elements determines the stationary values of the elements of the density matrix.

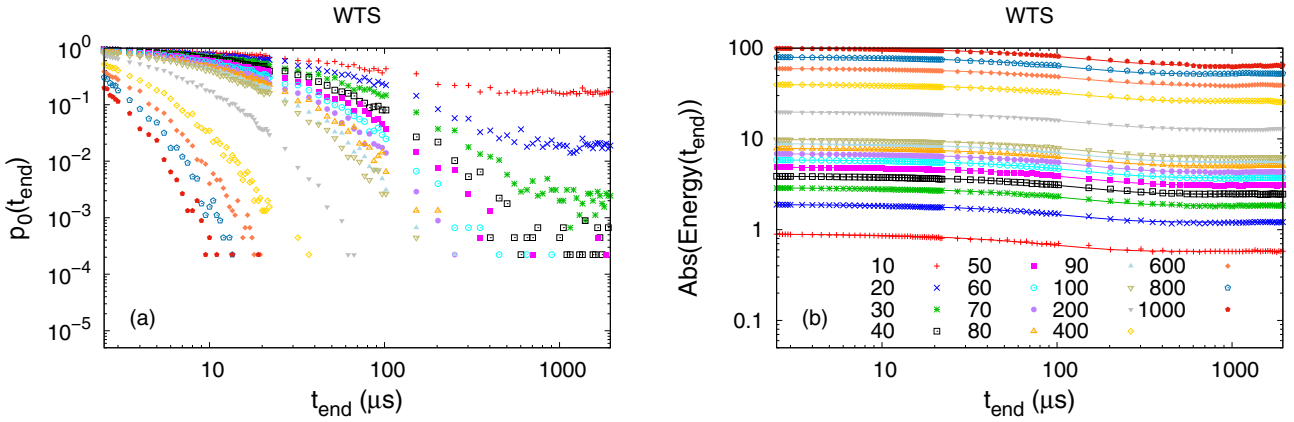

FIG. 10. Empirical data for the WTS for (a) success probability $p_0(t_{\text{end}})$ and (b) absolute values of the mean energy for ferromagnetic Ising spin chains of size $10 \leqslant N \leqslant 1000$ and $s_r = 0.7$. Lines in (b) serve as a guide to the eye. The legend in panel (b) also applies to panel (a).

Regarding $\mathbf{C}(t)$ to be piecewise constant within time intervals of duration τ, i.e., assuming $\mathbf{C}(t) = \mathbf{C}_n$ for $n\tau \leqslant t < (n+1)\tau$, we obtain

$$\mathbf{x}((n+1)\tau) = e^{\tau(\mathbf{C}_n+\mathbf{D})}\mathbf{x}(n\tau) + \int_0^\tau e^{(\tau-\lambda)(\mathbf{C}_n+\mathbf{D})}\mathbf{y}d\lambda \quad \text{(D2a)}$$

$$= e^{\tau(\mathbf{C}_n+\mathbf{D})}\mathbf{x}(n\tau) + (\mathbf{C}_n+\mathbf{D})^{-1}(e^{\tau(\mathbf{C}_n+\mathbf{D})} - \mathbb{I})\mathbf{y}. \quad \text{(D2b)}$$

We employ two different algorithms for solving Eq. (D2) numerically. The first method is numerical diagonalization. Since the dimension of the involved matrices here is rather small, we can compute the left-hand side of Eq. (D2b) through repeated numerical diagonalization of $\mathbf{M} = \mathbf{C}_n + \mathbf{D}$ for successive n, where here and in the following we suppress the subscript n for notational simplicity. As $\mathbf{M}$ is a general, real-valued matrix, we have

$$\mathbf{MR} = \mathbf{R}\Lambda, \quad \mathbf{M} = \mathbf{R}\Lambda\mathbf{R}^{-1}, \quad \text{(D3)}$$

where Λ is a diagonal matrix with complex-valued eigenvalues λ_j of $\mathbf{M}$ on the diagonal, $\mathbf{R}$ is the matrix of the eigenvectors of $\mathbf{M}$ as its columns, and $\mathbf{R}^{-1}$ is the inverse of $\mathbf{R}$, if it exists. If the latter is true, we have

$$\mathbf{x}((n+1)\tau) = \mathbf{R}e^{\tau\Lambda}\mathbf{R}^{-1}\mathbf{x}(n\tau) + \mathbf{R}\Lambda^{-1}(e^{\tau\Lambda} - \mathbb{I})\mathbf{R}^{-1}\mathbf{y}, \quad \text{(D4)}$$

where the matrices appearing in Eq. (D4) are obtained by matrix diagonalization and inversion, which is feasible for these problems, given their small size. Clearly, this numerically exact method can be used only if the inverse of $\mathbf{R}$ exists.

The second alternative to solve Eq. (D2) is to make use of a product-formula algorithm based on the decomposition of matrix exponentials. We start by approximating the contribution of the source term, represented by the last term in Eq. (D2a). Approximating the integral in Eq. (D2a) by a two-point trapezium rule we obtain

$$\mathbf{x}[(n+1)\tau] = e^{\tau(\mathbf{C}_n+\mathbf{D})}\mathbf{x}(n\tau) + \frac{\tau}{2}(1 + e^{\tau(\mathbf{C}_n+\mathbf{D})})\mathbf{y}, \quad \text{(D5)}$$

which is a second-order accurate approximation in time step τ. To keep the algorithm second-order accurate, we use a second-order accurate algorithm for computing the exponential $\exp(\tau(\mathbf{C}_n + \mathbf{D}))$.

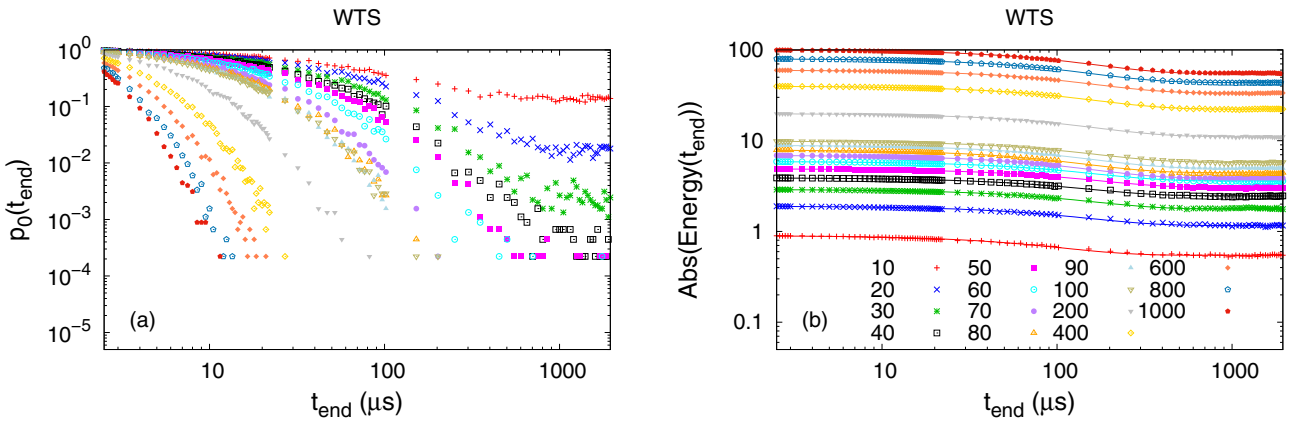

FIG. 11. Same as Fig. 10 except that the D-Wave Advantage_4.1 system was used to produce the data. The legend in panel (b) also applies to panel (a).

For the single-spin system, the simplest choice for the decomposition is

$$e^{\tau(\mathbf{C}_n+\mathbf{D})} \approx e^{\tau\mathbf{A}_2/2} e^{\tau\mathbf{A}_1(n)} e^{\tau\mathbf{A}_2/2}, \tag{D6}$$

where

$$\mathbf{A}_1(n) = \begin{pmatrix} 0 & B^z(n) & -B^y(n) \\ -B^z(n) & 0 & B^x(n) \\ B^y(n) & -B^x(n) & 0 \end{pmatrix},$$
$$\mathbf{A}_2 = \begin{pmatrix} -1/T_2 & 0 & 0 \\ 0 & -1/T_2 & 0 \\ 0 & 0 & -1/T_1 \end{pmatrix}. \tag{D7}$$

Moving on to the two-spin system, we decompose the matrix exponential in three components, such that

$$e^{\tau(\mathbf{C}_n+\mathbf{D})} \approx e^{\tau\mathbf{A}_1(n)/2} e^{\tau\mathbf{A}_2(n)/2} e^{\tau\mathbf{A}_3} e^{\tau\mathbf{A}_2(n)/2} e^{\tau\mathbf{A}_1(n)/2}, \tag{D8}$$

with $\mathbf{A}_1(n)$ being the part of $\mathbf{C}_n$ with $B = 0$ in Eq. (1) and $\mathbf{A}_2(n)$ being the one with $A = 0$ in Eq. (1), and $\mathbf{A}_3 = \mathbf{D}$. While the matrix exponentials of $\mathbf{A}_1(n)$ and $\mathbf{A}_2(n)$ can be calculated analytically, we compute $\exp(\tau\mathbf{A}_3)$ by numerical diagonalization (once). The product formula approach can be applied when the inverse of $\mathbf{R}$ does not exist and, in practice, mainly provides an independent check on the numerical results.

APPENDIX E: D-WAVE RESULTS FOR FERROMAGNETIC SPIN CHAINS

Figure 10(a) shows empirical data of the ground state probabilities obtained from the WTS for various problem sizes up to $N = 1000$, averaged over 4500 samples. With increasing problem size, the $p_0(2000\,\mu\text{s})$'s saturate at decreasingly lower values, up to $N \leqslant 30$. For larger N, the required number of samples is prohibitive to make definite statements.

Figure 10(b) shows empirical data of the absolute value of the mean energies obtained from the WTS for various problem sizes up to $N = 1000$, averaged over 4500 samples. As discussed in Sec. VI, the values at $t_{\text{end}} = 2000\,\mu\text{s}$ are in excellent agreement with the corresponding equilibrium values of the one-dimensional ferromagnetic Ising chain.

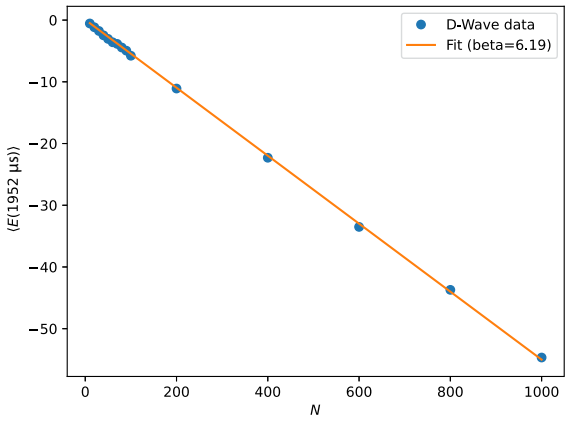

FIG. 12. Same as Fig. 8 except that the D-Wave Advantage_4.1 system was used to produce the data.

APPENDIX F: D-WAVE RESULTS USING THE ADVANTAGE_4.1 SYSTEM

While the main text focuses on results obtained from the Advantage_5.4 system, the trends and conclusions discussed are observed to hold more broadly across other D-Wave systems as well. As an illustration, we present results from an experiment on the Advantage_4.1 system using ferromagnetic spin chains of up to 1000 spins. In line with the findings discussed in Sec. VI, Fig. 11(a) shows the ground state probability p_0 as a function t_{end} for various N, while Fig. 11(b) demonstrates relaxation and eventual saturation of the mean energy with increasing annealing time.

Figure 12 shows the scaling of the mean energy $\langle E(1952\,\mu\text{s})\rangle$ as a function of system size. The solid line represents a fit of Eq. (17) to the D-Wave data, yielding a fitting parameter $\beta = 6.19$, which corresponds to a temperature of approximately 33 mK.

[1] B. Apolloni, C. Carvalho, and D. de Falco, Stoch. Process. Their Appl. **33**, 233 (1989).
[2] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, Chem. Phys. Lett. **219**, 343 (1994).
[3] T. Kadowaki and H. Nishimori, Phys. Rev. E **58**, 5355 (1998).
[4] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv:quant-ph/0001106.
[5] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).
[6] D-Wave Systems Inc., D-Wave Leap Quantum Cloud Service, https://docs.dwavesys.com/docs/latest/index.html.
[7] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh *et al.*, Phys. Rev. B **82**, 024511 (2010).
[8] M. W. Johnson, P. Bunyk, F. Maibaum, E. Tolkacheva, A. J. Berkley, E. M. Chapple, R. Harris, J. Johansson, T. Lanting, I. Perminov *et al.*, Supercond. Sci. Technol. **23**, 065004 (2010).
[9] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk *et al.*, Nature (London) **473**, 194 (2011).
[10] K. Karimi, N. G. Dickson, F. Hamze, M. H. S. Amin, M. Drew-Brook, F. A. Chudak, P. I. Bunyk, W. G. Macready, and G. Rose, Quantum Inf. Process. **11**, 77 (2012).
[11] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, Nat. Phys. **10**, 218 (2014).
[12] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, Science **345**, 420 (2014).
[13] T. Albash and D. A. Lidar, Rev. Mod. Phys. **90**, 015002 (2018).
[14] N. G. Dickson and M. H. S. Amin, Phys. Rev. A **85**, 032303 (2012).
[15] A. Lucas, Front. Phys. **2**, 5 (2014).
[16] D. Venturelli, S. Mandrà, S. Knysh, B. O'Gorman, R. Biswas, and V. Smelyanskiy, Phys. Rev. X **5**, 031040 (2015).

[17] Z. Bian, F. Chudak, R. B. Israel, B. Lackey, W. G. Macready, and A. Roy, Front. ICT **3**, 196706 (2016).
[18] H. Ushijima-Mwesigwa, C. F. A. Negre, and S. M. Mniszewski, in *Proceedings of the Second International Workshop on Post Moore's Era Supercomputing* (Association for Computing Machinery, New York, 2017), pp. 22–29.
[19] F. Neukart, G. Compostella, C. Seidel, D. Von Dollen, S. Yarkoni, and B. Parney, Front. ICT **4**, 29 (2017).
[20] D. O'Malley, V. V. Vesselinov, B. S. Alexandrov, and L. B. Alexandrov, PLoS ONE **13**, e0206653 (2018).
[21] H. N. Djidjev, G. Chapuis, G. Hahn, and G. Rizk, arXiv:1801.08653.
[22] T. Stollenwerk, E. Lobe, and M. Jung, in *Quantum Technology and Optimization Problems*, edited by S. Feld and C. Linnhoff-Popien (Springer International Publishing, Cham, 2019), pp. 99–110.
[23] M. Willsch, D. Willsch, F. Jin, H. De Raedt, and K. Michielsen, Quantum Inf. Process. **19**, 197 (2020).
[24] C. D. G. Calaza, D. Willsch, and K. Michielsen, Quantum Info. Proc. **20**, 1 (2021).
[25] D. Willsch, M. Willsch, C. D. Gonzalez Calaza, F. Jin, H. De Raedt, M. Svensson, and K. Michielsen, Quantum Info. Proc. **21**, 141 (2022).
[26] J. A. Montanez-Barrera, P. van den Heuvel, D. Willsch, and K. Michielsen, in *Proceedings of the 2023 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE, New York, 2023), Vol. 1, pp. 535–542.
[27] J. A. Montanez-Barrera, D. Willsch, A. Maldonado-Romo, and K. Michielsen, Quantum Sci. Technol. **9**, 025022 (2024).
[28] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Nat. Commun. **4**, 2067 (2013).
[29] K. L. Pudenz, T. Albash, and D. A. Lidar, Nat. Commun. **5**, 3243 (2014).
[30] T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk *et al.*, Phys. Rev. X **4**, 021041 (2014).
[31] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. Amin, A. Smirnov, M. Mohseni, and H. Neven, arXiv:1411.4036.
[32] Y. Bando, Y. Susa, H. Oshiyama, N. Shibata, M. Ohzeki, F. J. Gómez-Ruiz, D. A. Lidar, S. Suzuki, A. Del Campo, and H. Nishimori, Phys. Rev. Res. **2**, 033369 (2020).
[33] A. D. King, J. Raymond, T. Lanting, R. Harris, A. Zucca, F. Altomare, A. J. Berkley, K. Boothby, S. Ejtemaee, C. Enderud *et al.*, Nature (London) **617**, 61 (2023).
[34] A. Lopez-Bezanilla, J. Raymond, K. Boothby, J. Carrasquilla, C. Nisoli, and A. D. King, Nat. Commun. **14**, 1105 (2023).
[35] A. Lopez-Bezanilla, A. D. King, C. Nisoli, and A. Saxena, Nat. Commun. **15**, 589 (2024).
[36] A. D. King, A. Nocera, M. M. Rams, J. Dziarmaga, R. Wiersema, W. Bernoudy, J. Raymond, N. Kaushal, N. Heinsdorf, R. Harris *et al.*, Science **388**, 199 (2025).
[37] J. Vodeb, J. Desaules, A. Hallam, A. Rava, G. Humar, D. Willsch, M. Willsch, F. Jin, K. Michielsen, and Z. Papić, Nat. Phys. (2025).
[38] D-Wave Systems, Reverse quantum annealing for local refinement of solutions, Technical Report, D-Wave Whitepaper Series, 14-1018A-A (D-Wave Systems, 2017).
[39] D. Venturelli and A. Kondratyev, Quantum Mach. Intell. **1**, 17 (2019).
[40] G. Passarelli, K.-W. Yip, D. A. Lidar, H. Nishimori, and P. Lucignano, Phys. Rev. A **101**, 022331 (2020).
[41] L. Rocutto, C. Destri, and E. Prati, Adv. Quantum Technol. **4**, 2000133 (2021).
[42] M. S. Jattana, Quantum Rep. **6**, 452 (2024).
[43] V. Mehta, H. De Raedt, K. Michielsen, and F. Jin, Phys. Rev. A **112**, 012405 (2025).
[44] V. Mehta, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. A **104**, 032421 (2021).
[45] V. Mehta, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. A **105**, 062406 (2022).
[46] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).
[47] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, J. Math. Phys. **17**, 821 (1976).
[48] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2002).
[49] Jülich Supercomputing Centre, J. Large-Scale Res. Facilities **5**, A135 (2019).