Comparing Relaxation Mechanisms in Quantum and Classical Transverse-Field Annealing

Tameem Albash[1,2,*,‡] and Jeffrey Marshall[3,4,†,‡]

[1]*Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, New Mexico 87131, USA*
[2]*Department of Physics and Astronomy and Center for Quantum Information and Control, CQuIC, University of New Mexico, Albuquerque, New Mexico 87131, USA*
[3]*QuAIL, NASA Ames Research Center, Moffett Field, California 94035, USA*
[4]*USRA Research Institute for Advanced Computer Science, Mountain View, California 94043, USA*

Annealing schedule control provides opportunities to better understand the manner and mechanisms by which putative quantum annealers operate. By appropriately modifying the annealing schedule to include a pause (keeping the Hamiltonian fixed) for a period of time, we show that it is possible to more directly probe the dissipative dynamics of the system at intermediate points along the anneal and examine thermal relaxation rates, for example, by observing the repopulation of the ground state after the minimum spectral gap. We provide a detailed comparison of experiments from a D-Wave device, simulations of the quantum adiabatic master equation, and a classical analogue of quantum annealing, spin-vector Monte Carlo, and we observe qualitative agreement, showing that the characteristic increase in success probability when pausing is not a uniquely quantum phenomena. We find that the relaxation in our system is dominated by a single timescale, which allows us to give a simple condition for when we can expect pausing to improve the time to solution, the relevant metric for classical optimization. Finally, we also explore in simulation the role of temperature whilst pausing as a means to better distinguish quantum and classical models of quantum annealers.

I. INTRODUCTION

Over the last few years there has been a healthy debate surrounding the operation of current generation quantum annealers, such as the D-Wave family of devices [1–3]. These devices implement in hardware a realization of quantum annealing [4–8], whereby a continuous-time interpolation between two noncommuting Hamiltonians is performed. At any point along the interpolation, the low-energy spectrum of the device is approximated by a transverse-field Ising model, with the relative strength of the transverse field and Ising Hamiltonians determined by the interpolation schedule. These devices are known to be sensitive to the operating temperature [9], although the exact model for this is still under scrutiny. On the one hand, some experimental observations (see, e.g., Refs. [10,11]) are consistent with a quasistatic model [12] induced by the coupling of the quantum system to a nonzero temperature bath [13], but on the other hand, studies aiming to directly observe quantum thermal distributions have

been mixed [10,11,14–18]. Whilst these previous results clearly demonstrate that the dynamics associated with pausing are purely dissipative, whether these processes are able to improve performance in the context of optimization, where the goal is to reach or approximate the ground state of the Ising Hamiltonian, and ultimately help achieve an advantage over classical approaches is still unknown.

The attempt to measure nontrivial quantum thermal statistics on such devices utilizes a midanneal pause, whereby the Hamiltonian is held fixed such that the system can equilibrate. Given a sufficiently fast quench (*or* ramp) and readout, it should in principle be possible to observe the quantum Gibbs distribution associated with the Hamiltonian at the pause. To date, numerical studies into the pause have been conducted purely by quantum adiabatic master equation simulations [19–21], which generally agree with observations in experiments [14]. In Ref. [21], a deeper study into the theory of pausing was conducted, from the point of view of quantum thermalization. There sufficient conditions were identified such that a pause has a nontrivial influence on the output of an anneal. These results are consistent with the picture put forward in Ref. [14] whereby the competition of various

*talbash@unm.edu
†jmarshall@usra.edu
‡These authors contributed equally to this work.

timescales—relaxation, pause, annealing—determines the general dynamics during the mid to late anneal region.

These timescales are not unique to the quantum adiabatic master equation however, motivating a study to determine whether or not the ground-state statistics under pausing can be reproduced by a simple semiclassical model of the quantum annealing dynamics. In such a picture, the quantum Hamiltonian is replaced by a related classical potential [22,23], which is then explored by thermal hopping (with no quantum tunneling). Here we use spin-vector Monte Carlo (SVMC) [24] as our model, where qubits are replaced by two-dimensional rotors (with no entanglement), and the system evolves using a Metropolis-Hastings algorithm [25,26], whereby each rotor's orientation is updated at some fixed temperature while the potential is changed as an analog to the Hamiltonian changing during the quantum anneal. Such a description has been extremely successful in discriminating quantum from classical effects in such devices and to what extent entanglement determines the output statistics [24,27–29].

By adapting the standard update in SVMC, we are able to capture the general quasistatic (thermal) behavior observed in experiments on D-Wave devices, using the pause feature as a probe for the dynamics. This adaptation to the original SVMC algorithm, we believe, will be helpful in understanding thermal dynamics more generally in such devices going forward. Such a picture provides a simple intuition for the role of thermal effects during the anneal and whilst pausing, and can help determine whether pausing can be expected to provide an advantage over classical thermal methods for sampling.

Motivated by our experimental and simulation results, we derive a condition under which the time to solution (TTS) can be reduced by relaxation effects. This condition is agnostic to the underlying dynamics and depends only on the observation of a single dominant timescale determining the exponential repopulation effect associated with the pause. Analysis of this type can be useful in determining whether pausing can improve performance.

Our paper is organized as follows. In Sec. II, we provide details about the problem instance that we focus on in this study. In Sec. III, we describe our simulation methods. In Sec. IV, we give our experimental results using the D-Wave 2000Q^{TM} device and our simulation results, and in Sec. V we provide a simple condition under which we can expect pausing to improve the performance of quantum annealing in the context of classical optimization. We conclude with a discussion of our results in Sec. VI.

II. BACKGROUND

We study the standard transverse-field Ising model of the form $H(s) = A(s)H_x + B(s)H_p$, where the transverse-field Hamiltonian $H_x = -\sum_i \sigma_i^x$ is known as the driver Hamiltonian, and the Ising Hamiltonian $H_p = \sum_{i<j}$
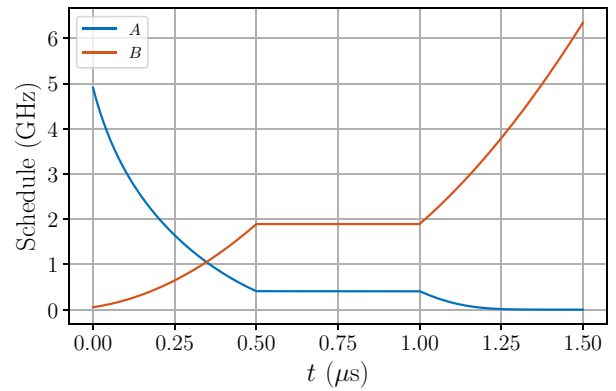

FIG. 1. Example schedule with a pause of length $t_p = 0.5$ μs inserted into an anneal with anneal time $t_a = 1$ μs. The pause occurs at $s_p = 0.5$ (i.e., midway through the interpolation). We sometimes refer to the anneal after the pause as the ramp. We use units $h = 1$.

$J_{ij}\sigma_i^z\sigma_j^z + \sum_i h_i\sigma_i^z$ is the classical "problem" Hamiltonian. The functions $A(s), B(s)$ determine the interpolation between H_x and H_p, and the dimensionless annealing parameter $s \in [0,1]$ is a function of the physical time t. For the standard anneal, it is given by $s(t) = t/t_a$, where t_a denotes the total annealing time, but in general it can be any piecewise continuous function of t.

A pause is where the value of s is held fixed for a certain period of time, or, equivalently, the values of A, B are held fixed. We use the notation s_p to denote the value of s at which a pause is inserted into the annealing schedule, and t_p the time paused, e.g., as shown in Fig. 1. For our simulations, we fix the annealing schedule to be that of the D-Wave 2000Q device [30] that was used to collect experimental data (for more information on this device, see Ref. [14]).

We focus our study on one well-behaved instance $\mathcal{I}_{12}^0$ from Ref. [14], which we depict in Fig. 2 (the instance is part of the supplemental material there and is given
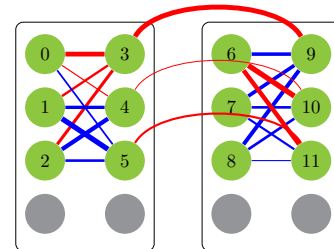

FIG. 2. Depiction of instance $\mathcal{I}_{12}^0$ used in this study. Shown are two Chimera unit cells, with the indexed nodes corresponding to the 12 qubits used in the instance. Red (blue) edges connecting nodes correspond to ferromagnetic (antiferromagnetic) Ising couplings between qubits, with the thickness of the edge in direct proportion to the magnitude of the coupling. Explicit values for the couplings are given in Appendix A.
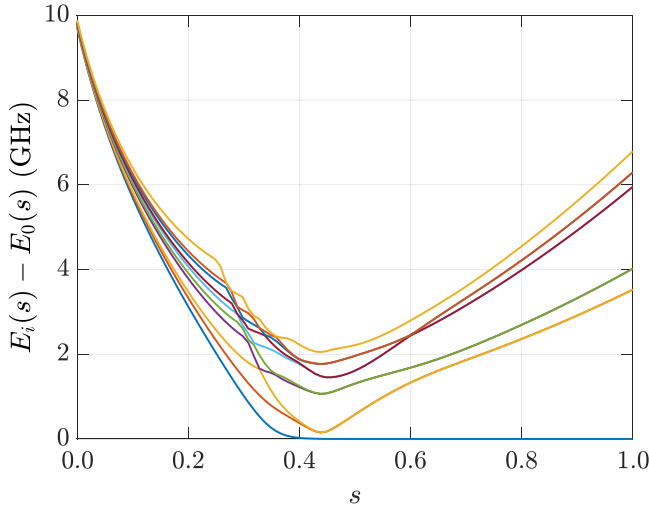
FIG. 3. Spectrum (first 20 levels) of $\mathcal{I}_{12}^0$. Minimum gap of approximately $\Delta = 0.15$ GHz at $s_\Delta = 0.44$. The specified DW operating temperature is 0.25 GHz. We use units $h = 1$.
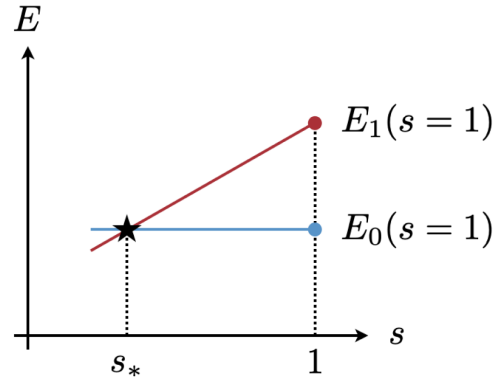

FIG. 4. Illustration of the energy level crossing predicted by first-order perturbation theory. At this order in perturbation theory, the ground-state energy $E_0(s)$ remains unchanged while the first excited state energy $E_1(s)$ is lowered as a function of the perturbative parameter $1 - s$, resulting in an energy level crossing at s_*.

here in Appendix A). It is small enough ($n = 12$ qubits) to be amenable to analytic methods and simulations. This instance has no local fields ($h_i = 0$) and thus has a doubly degenerate Ising spectrum, and the $J_{ij} \in [-1, 1]$ are chosen according to a uniform distribution. Moreover, the minimum gap is less than the thermal energy scale (Fig. 3); hence, thermal effects are expected to play a role.

The minimum gap can be understood in terms of a perturbative crossing [31]. To see this, it is useful to work in the subspace with eigenvalue 1 under the operator $P = \prod_{i=1}^n \sigma_i^x$. In the absence of dissipative dynamics, the unitary evolution would be restricted to this subspace if the initial state is the uniform superposition state. In this subspace, the Ising ground state is given by

$$|E_0(s = 1)\rangle = \frac{1}{\sqrt{2}}(|000110110000\rangle + |111001001111\rangle),$$
(1)

and we have two closely spaced ($\Delta_{21} \approx 0.0781$ in units of the Ising Hamiltonian) energy states above the ground state

$$|E_1(s = 1)\rangle = \frac{1}{\sqrt{2}}(|001010001111\rangle + |110101110000\rangle),$$
(2a)

$$|E_2(s = 1)\rangle = \frac{1}{\sqrt{2}}(|001110001111\rangle + |110001110000\rangle),$$
(2b)

where the least significant bit in the state label corresponds to the state of qubit 1. These two energy eigenstates differ in only one position, i.e., are connected by the transverse-field Hamiltonian $|E_1(s = 1)\rangle = \sigma_9^x |E_2(s =$

$1)\rangle$. For a slightly different choice of J_{ij}, these two energy states would be degenerate, and our choice of J_{ij} weakly breaks this degeneracy.

Moving away from $s = 1$ to smaller s values, which corresponds to turning on the transverse field, first-order perturbation theory predicts that the energy of the ground state $E_0(s)$ remains unchanged. In contrast, the symmetric combination of $|E_1(s = 1)\rangle$ and $|E_2(s = 1)\rangle$ is lowered in energy (the antisymmetric combination is raised in energy) at the same order in perturbation theory, and this combination is the unique first excited state away from $s = 1$ with an energy $E_1(s)$ that decreases with decreasing s. First-order perturbation theory thus predicts a crossing of the ground state and first excited state at some $s = s_* < 1$, as illustrated in Fig. 4, and this crossing is then corrected at higher order to an avoided level crossing, with the gap determined primarily by the Hamming distance between the computational basis states in Eqs. (1) and (2) (at least 3 in this case).

III. METHODS

In this work we compare experimental results of the D-Wave (DW) 2000Q device [30] located at NASA Ames to simulations of the adiabatic master equation (AME) and SVMC.

A. AME

The AME is a time-dependent Davies master equation [13,32] of Lindblad form [33] derived in the limit of weak coupling between the system and a Markovian environment. We assume independent but identical Ohmic oscillator baths for each qubit, which gives rise to a spectral

density of the form

$$\gamma(\omega) = 2\pi\kappa^2 \omega e^{-|\omega|/\omega_c}/(1 - e^{-\beta\omega}), \quad (3)$$

where $\beta = 1/k_B T$ is the inverse thermal energy scale of the bath, κ^2 is the dimensionless system-bath coupling strength squared, and we have introduced (by hand) the ultraviolet cutoff ω_c. Under these assumptions, the AME takes the form ($\hbar = 1$)

$$\frac{1}{t_a}\frac{d}{ds}\rho(s) = -i[H(s), \rho(s)]$$

$$+ \sum_{i=1}^{n}\sum_{\omega}\gamma(\omega)\left\{L_{\omega,i}(s)\rho(s)L_{\omega,i}^{\dagger}(s)\right.$$

$$\left. - \frac{1}{2}[L_{\omega,i}^{\dagger}(s)L_{\omega,i}(s), \rho(s)]\right\}, \quad (4)$$

where the index i runs over the n qubits and the index ω runs over all possible Bohr frequencies of the system Hamiltonian $H(t)$. The Lindblad operators are given by

$$L_{\omega,i}(s) = \sum_{a,b}\delta_{\omega,E_b(s)-E_a(s)}\langle E_a(s)|\sigma_i^z|E_b(s)\rangle$$

$$\times |E_a(s)\rangle\langle E_b(s)|, \quad (5)$$

where we have taken a dephasing system-bath interaction on each qubit and the $\{|E_a(t)\rangle\}_a$ are (instantaneous) energy eigenstates of $H(t)$ with eigenvalues $\{E_a(t)\}_a$.

In the context of the AME, the increase in ground-state probability associated with pausing after the minimum gap can be straightforwardly understood. At any fixed s value, the fixed point of the AME is the Gibbs state associated with $H(s)$. As the gap opens up, the ground-state population of the Gibbs state increases (ignoring pathological cases where maybe the first excited state degeneracy grows), and the dissipative dynamics restores population to the instantaneous ground state (thermal relaxation). Whether taking an adiabatic (in the open system sense [34]) anneal or pausing at a fixed s, the Gibbs state will be reached [35]; the relevant question though is how efficiently can this repopulation occur. In the AME, this is determined by the nonzero eigenvalues of the Lindbladian. These eigenvalues depend sensitively on the overlaps $\langle E_a(s)|\sigma_i^z|E_b(s)\rangle$ and $\gamma(\omega)$, and the relevant first excited state to ground-state relaxation term in the Lindblad master equation is [36]

$$\gamma_{1\to0} \propto \gamma[E_1(s) - E_0(s)]\sum_{i}|\langle E_0(s)|\sigma_i^z|E_1(s)\rangle|^2. \quad (6)$$

We further know that the ground-state population in the Gibbs state increases to 1 as $s \to 1$ (assuming that the spectral gap is large here), while the overlap

$\langle E_a(s)|\sigma^z|E_b(s)\rangle$ goes to 0. Therefore, the presence of an optimal pause location after the minimum gap might be expected as a competition of these two effects.

B. SVMC with transverse-field updates

As a brief reminder of SVMC [24], each qubit is modeled as a two-dimensional rotor with an associated angle $\theta_i \in [0, \pi]$, where $\theta_i = 0$ is aligned along the z axis, and $\theta_i = \pi/2$ along the x axis (here i labels the qubit).

In the original model, update angles are chosen uniformly randomly, i.e., θ_i' has no relation to θ_i. Updates to the spin angles $\theta_i \to \theta_i'$ are accepted according to a standard Boltzmann factor associated with the change in energy of the classical Hamiltonian

$$\mathcal{H}(s) = -A(s)\sum_{i}\sin\theta_i$$

$$+ B(s)\left(\sum_{i<j}J_{ij}\cos\theta_i\cos\theta_j + \sum_{i}h_i\cos\theta_i\right). \quad (7)$$

This Hamiltonian can be interpreted as the semiclassical potential associated with the spin-coherent path integral [22,28,37] collapsed to a plane [23].

This model fails to capture the phenomena of freeze out that is present in current experimental quantum annealers. Freeze out occurs in transverse-field annealing with a dominant dephasing system-bath interaction when the transverse field is weak relative to the Ising Hamiltonian (i.e., "late" in the anneal where the instantaneous energy eigenstates are well approximated by the eigenstates of the Ising Hamiltonian), and the system effectively freezes with no more population dynamics occurring. In contrast, in SVMC where update angles are chosen arbitrarily, dynamics can still occur at $A = 0$. When angles are chosen near 0 or $\pm\pi$, the updates correspond to classical spin flips on the Ising Hamiltonian.

Of course, more general update procedures can be implemented, such as $\theta_i \to \theta_i + \epsilon_i(s)$, where $\epsilon_i(s)$ is associated with a random variable depending on s (and i). For example, in order to replicate the freeze-out effect, we consider updates where $\epsilon_i(s)$ is chosen randomly in the restricted range $[-\min\{1, A(s)/B(s)\}\pi, \min\{1, A(s)/B(s)\}\pi]$. For $A(s)/B(s) \gtrsim 1$ (corresponding to $s \lesssim 0.36$), the updates amount to the unrestricted updates of the original algorithm, but when $A(s) < B(s)$, the updates are restricted to be around the current angle value. This simple adaptation induces freeze out since now the transverse-field strength directly determines how large an angle update is allowed, and in the limit of $A/B \to 0$, updates stop entirely. To distinguish it from the standard SVMC, we refer to this algorithm with transverse-field-dependent updates as SVMC TF.

The dynamics during the evolution of the classical energy landscape of SVMC and SVMC TF [Eq. (7)] are

FIG. 5. Semiclassical picture of the dynamics. Left (right) corresponds to early (late) in the anneal. The green circle corresponds to the position of the system in the energy landscape.

illustrated in Fig. 5. Early in the anneal, the semiclassical potential has a single minimum, and the dynamics follows this minimum. As we approach the minimum gap, a second local minimum develops. Near the minimum gap of the quantum Hamiltonian, there are two near degenerate minima in the semiclassical potential, and the only way for the system to reach the other minimum is via thermal hopping. If this minimum is far away, corresponding to updating many spins by large angles, reaching this minimum via thermal hopping is suppressed. Crucially, SVMC is not able to tunnel to the new minimum. After the minimum gap of the quantum Hamiltonian, there is again a single global minimum in the semiclassical potential, and it is unlikely that a sizeable population has reached it. Depending on the shape of the landscape, if the state is stuck in some local minimum, it may still reach the new global minimum in this regime via thermal hopping.

In SVMC and SVMC TF, as an analogue for "time," the number of sweeps is used, where one sweep consists of checking whether each (classical) rotor in the model should be updated with a new orientation. Pausing in SVMC and SVMC TF, where the classical potential is fixed for a chosen number of rotor updates, provides more opportunities for the state to escape local minima, and will generally result in a larger population in the ground state.

C. Simulation parameters

Unless otherwise stated, in our AME and SVMC simulations, the temperature is fixed at 12 mK, the purported temperature of the DW device used in our study. In the AME simulations, we fix $\kappa^2 = 10^{-3}$ and $\omega_c = 8\pi$ GHz. We also restrict the dynamics to the lowest 16 energy eigenstates of the instantaneous Hamiltonian; we find that this is sufficient to maintain $\mathrm{Tr}\rho \approx 1$ numerically.

DW experiments and AME simulations are both performed with an annealing time of 1 μs. Each data point from the DW is computed from (at least) 10^4 samples (anneals), using (at least) five choices of gauges (for more information, see Ref. [14], where some of this data was originally presented).

In the SVMC and SVMC-TF simulations, we use a total of 10^4 sweeps for the anneal unless otherwise noted, and we perform 10^4 independent anneals per data point.

IV. RESULTS

In this section, we present our results comparing various properties of SVMC TF, the AME, and where possible, experimental results of the D-Wave annealing device. Note that the goal of this work is not to match explicit quantities through a choice of parameters. Indeed, there are many possible parameters available in each of the three systems. Rather our strategy is to make "reasonable" choices for the parameters, and study how the systems behave relative to one another. This allows us to see which dynamics are consistent in the different models, and those which may differ. When this is the case, we provide intuitive reasoning for it.

A. Varying the pause time

As a first comparison, in Fig. 6 we study the effect of changing the pause time, but keeping all other parameters fixed. The first clear observation is that both the AME and SVMC TF are broadly consistent with the experimental observation that increasing the pause time shifts the peak to a later time, and also increases its size.

This shows that the mechanism by which the ground state repopulates under a pause is certainly not a uniquely quantum effect. Rather, it relies on there being an excess population in excited states at some late time in the anneal, and then, when pausing, for any type of relaxation process to take place. In SVMC TF this corresponds to transitions from local minima in the potential at the pause location to the global minima (see Fig. 5). In SVMC TF, at a given pause point, and as we explore in more detail below, the temperature is critical to driving these transitions, since no quantum tunneling can occur. With the added restriction on the SVMC-TF updates, this adds a further constraint on how easy it is for SVMC TF to transition to the global optima, at a given pause location.

Note that in all three figures there are differences as well as similarities. We stress here again that we are searching for broad properties, and not matching precisely the curves; there may well be sets of parameters in which all three models show closer agreement.

Each model shows consistent behavior with respect to the position (and height) of the peak shifting later in the anneal with pause time t_p. This phenomenon is well understood and commented on in Ref. [14], where a longer pause allows one to pause at a later time but still repopulate the ground state effectively (where otherwise the relaxation rate is reduced). The later the system can effectively thermalize, assuming the gap opens, the larger the instantaneous ground-state (GS) population.

Another observation is that pausing at the minimum gap in the DW and SVMC TF does in fact provide a slight improvement in P_0, whereas in the AME pausing at this point is detrimental. In the AME, pausing at the minimum gap (the location where the Gibbs state has the
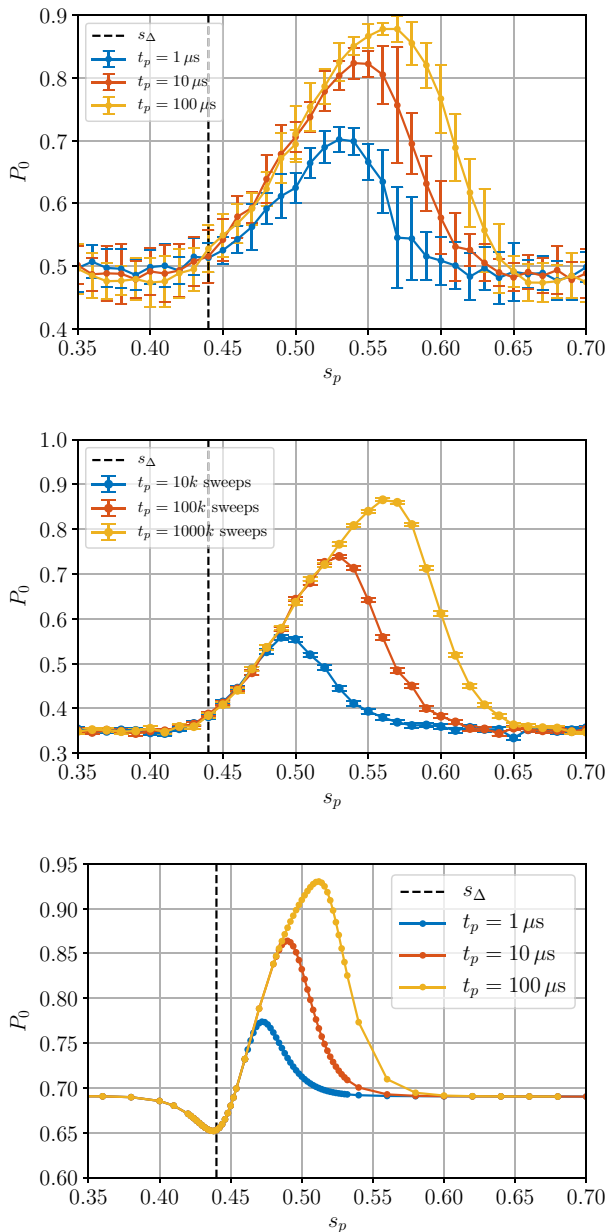
FIG. 6.   Varying the pause time for the DW (top), SVMC TF (middle), and the AME (bottom). All simulations are performed at 12 mK. SVMC TF has an anneal "time" of 10^4 sweeps, and the AME and DW of 1 μs. Note that, for the sake of comparison, the blue curves in each plot have a pause equal to the anneal time $t_a = t_p$, and the other two curves are 1 and 2 orders of magnitude larger. Error bars for the DW are 2σ, computed over at least five independent samples (each sample consisting of 10^4 anneals). Error bars in SVMC TF are 2σ over 10^4 samples. Note that the AME data is "exact." The vertical dashed line denotes the position of the minimum gap s_Δ. Note that the y scale is different in all three plots.

lowest instantaneous GS population) means that population is transferred to the excited states. If the dynamics after the anneal are too slow (e.g., ramp is too short), there is not enough time to effectively repopulate the GS. Indeed,

we see later that this dip disappears at higher temperature. (We note that, for SVMC TF, using a much smaller number of sweeps in the anneal can manifest a similar dip in P_0, but other features of the results get affected. We present these extra results in Appendix C.)

On the other hand, in the model of SVMC-TF pausing near the minimum gap allows for some of the population to explore the deep welled landscape and end up in what will eventually be the global minimum (cf. Fig. 5, middle). Based on the parameters of the model therefore (determining how easy it is to leave the current global minimum), pausing even at the minimum gap can eventually lead to a larger population in the GS at $s = 1$.

The assumption of weak coupling in the derivation of the AME is violated when the minimum gap is sufficiently small. Beyond weak coupling, we can expect the broadening of the energy levels due to the interaction with the bath to be significant. The dynamics of flux qubits in the strong bath coupling regime [23] has important similarities to the dynamics of SVMC, so we can think of multiqubit SVMC as a phenomenological model for a system of flux qubits in the strong coupling regime with no entanglement. The fact that the DW exhibits behavior more akin to SVMC TF than the AME near the minimum gap is suggestive that we are observing a breakdown of the weak-coupling limit in this region on the device. Various imperfections (control errors and noise) in the device are expected to broaden the spectrum that can modify the dynamics during this region [38,39]. In this case, the minimum gap may not be a well-defined point, but more accurately modeled as a region, and this goes well beyond our AME model.

Finally, we comment that in SVMC TF, the point at which dynamics slow (here around $s_p = 0.65$) is captured fairly well. As demonstrated in Appendix D (Fig. 13), the standard SVMC continues to exhibit dynamics for significantly longer s_p values, inconsistent with experiments.

1. Probing relaxation time

The time required under the pause to reach saturation can be equated to the equilibration time. From experiments and simulations we can extract these characteristic times, which inform us about decay mechanisms involved. Using the extended pause time feature (allowing up to a 60 ms pause), we can start to see the approach to equilibrium on the DW, which was previously unreachable (e.g., as seen in Fig. 6 of Ref. [14]). In Fig. 7 we fix the pause location at the region around the peak observed in Fig. 6, and vary the time paused, while keeping all other parameters fixed.

Interestingly, at long pause times, the DW outperforms the AME simulations, which are (in principle) at the same temperature (DW reaching in excess of 95% of the GS probability). Naively, this may make one believe that the DW is in fact operating at a colder temperature than the
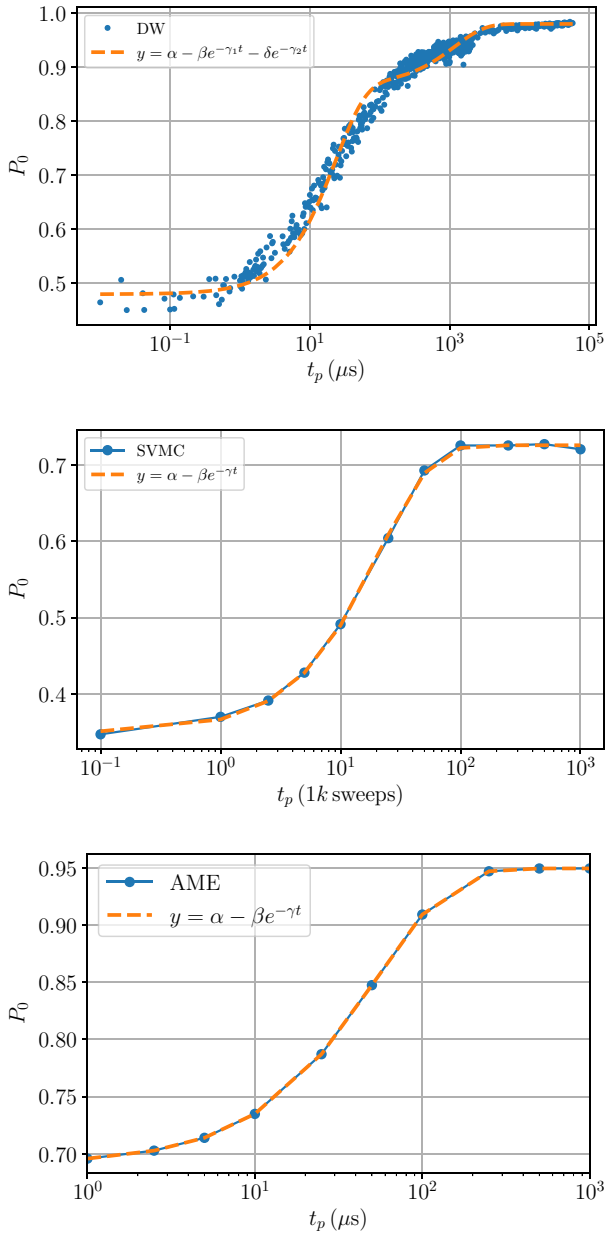
FIG. 7. Ground-state probability with pause time for the DW (top), SVMC TF (middle), and the AME (bottom). We fix the pause location in each plot to be in the region where the pause has a noticeable effect, taking $s_p = 0.57$ for the DW, and $s_p = 0.52$ for SVMC and the AME. The curve fits give excellent agreement with $(\alpha, \beta, \gamma) = (0.73, 0.38, 0.047)$ for SMVC and $(\alpha, \beta, \gamma) = (0.95, 0.26, 0.019)$ for the AME. For the DW, we had to use a function with two decay constants $(\gamma_1, \gamma_2) = (0.043, 8.5 \times 10^{-4})$. The top plot shows individual DW anneal results (from 10^4 samples), and the AME is "exact." Error bars for SVMC are 2σ over 10^4 samples.

reported 12 mK, although this would be contrary to several previous studies [9,14], but the location of the peak for the DW is at a larger s^* value, which is consistent with a larger ground-state population in the thermal state.

One other possibility is that, for increasing pause times, the device is more susceptible to an effect known as spin-bath polarization, such that outcomes of successive anneals become correlated [40]. This is caused by the appearance of local fields that bias the spin alignments and can cause them to more easily align with previously found ground states, thus resulting in a larger population.

From Fig. 7 we can extract the relevant timescales for the three models. We find that the AME and SVMC TF can be fit very well to a function depending on a single decay timescale, which in the AME corresponds to around 50 μs. This suggests that a single dominant decay channel is responsible for the repopulation effect. In the D-Wave, however, the fastest decay timescale is around half this, 25 μs, and there also exists a secondary much longer timescale of the order 1200 μs, although the curve fit here is not as accurate. Nevertheless, it is clear that there is an additional longer timescale involved (cf. Fig. 15 in Appendix E). Since this only appears at times greater than around 1 ms, this could be related to the spin-bath polarization effect.

We study the dependence of this timescale on the pause location, as shown in Fig. 8, where we plot the pause time t_p required for the success probability to reach a target value P^*, pausing at different points s_p (around the peak location in Fig. 6).

The results show a clear exponential increase in pause time required in order to obtain a particular ground-state probability at the end of the anneal. This indicates that the relaxation rate is decreasing exponentially in s after the minimum gap, and is direct evidence that "freeze out" occurs in all three models [9,12], whereby the population dynamics slow and eventually stop late in the anneal. To understand where this exponential dependence comes from in the AME, we note that $\langle E_1(s)| \sigma_i^z |E_0(s)\rangle \propto A(s)^\alpha$ for $A(s)/B(s)$ being small (see Appendix B for an argument), and since the annealing schedule $A(s)$ has an exponential form $A(s) \sim e^{-as}$ (with $a \approx 30$ for s approaching 1), it follows that the relaxation timescale in the AME will grow exponentially (see the discussion in Sec. A). It is perhaps reasonable to expect SVMC TF to replicate this given our definition of the updates, but we also find that standard SVMC exhibits similar behavior (shown in Appendix D, Fig. 14), indicating that the semiclassical landscape and the standard updates can still qualitatively capture this feature of the dissipative dynamics of the AME.

One interesting observation is how the exponents change with the probability target P^*. In experiments on the D-Wave (top plot in Fig. 8), the higher the target, the larger the exponent. This suggests that it is relatively harder to reach a success probability of, for example, 90% compared to 70%. This is consistent with the observation of a slower decay channel, since it means that it takes relatively longer to reach the highest success probabilities. This is in contrast to the AME and SVMC TF that exhibit
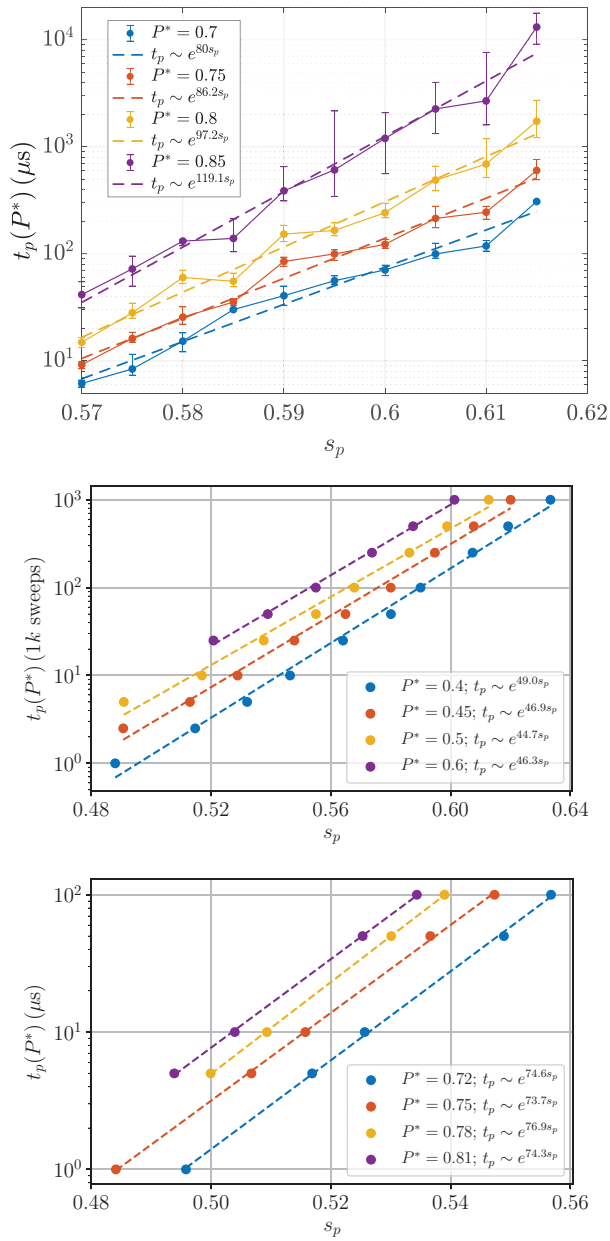
FIG. 8. Pause time required to reach target success probability P^* for the DW (top), SVMC TF (middle), and the AME (bottom). The fits (dashed lines) are from a least-square fitting, with exponents given in the legend. Note how the exponents change in the top plot compared to the bottom plots. In the AME and SVMC, the data is generated by taking curves as in Fig. 6, and finding the value on the x axis corresponding to some fixed target P^*. If it lies between data points, we connect the points on either side by a straight line and extrapolate to find the value. Since the raw DW data is noisier when pausing at a fixed location (e.g., as in Fig. 7), we do a similar analysis but take a range around $P^* \pm 0.01$. For all data points in this range, the value reported here is the median, with error bars being the interquartile range.

approximately constant exponents across the various P^*, which is again consistent with the previous observation of a single relaxation timescale.

It can be argued that the open quantum system model of the AME or SVMC TF is too simplistic to capture quantitatively the D-Wave output; this includes going beyond the weak coupling limit [39], but also unaccounted noise sources and effects, such as $1/f$ noise, which become more prevalent at the longer timescales required to reach high P^*, but also the fact that superconducting flux qubits only approximate a two-level system. If the temperature at which the device operates is in fact larger than documented (i.e., larger than 12 mK, which has also been suggested previously [14]), the population loss to higher excited states may be more severe than in the AME and SVMC TF (at the minimum gap for example).

B. Effect of temperature

One important distinction between "quantum thermalization" and "classical thermalization" is how the quantum system couples (e.g., through the z component of the spin) to its thermal environment. The mechanism for thermalization then depends heavily on the transverse field of the system as ultimately this drives the transitions between computational basis states (cf. Fermi's golden rule [39]). In particular, quantum systems can find their way out of local minima even at low temperatures when the barrier is higher than the energy scale set by the temperature. In the classical case, transitions of this type are exponentially suppressed. Thus, we expect the temperature parameter to be critical to understanding differences between the two.

Although probing the temperature in experiments is not a current feature available to D-Wave users, we can gain an understanding through simulations.

First, we note that the update procedure for SVMC is crucial here. As explained in Sec. B, using the standard
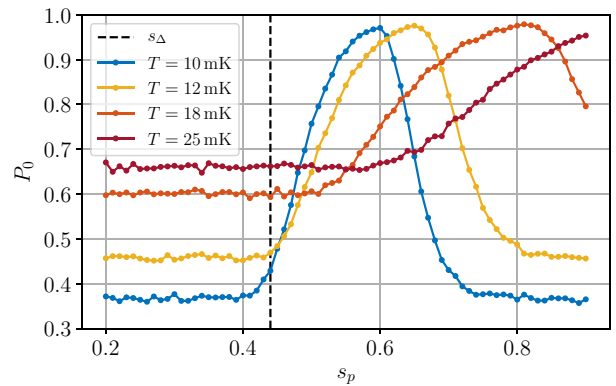

FIG. 9. SVMC dependence on temperature with random angle updates. In the standard version of SVMC with new angles chosen uniformly randomly in $[0, \pi]$ dynamics continue late in the anneal; thus, even at high temperatures, a large success probability can be achieved, assuming the pause is late enough. The anneal "time" is 10^4 sweeps, and the pause "time" is fixed at 10^6 sweeps for all curves. The location of the minimum gap s_Δ is marked with a vertical dashed line.

updates can lead to dynamics even at $s = 1$. We see this clearly in Fig. 9 where we show the success probability as a function of pause location for various temperatures. In SVMC the dynamics during a pause can be illustrated by considering the semiclassical potential that changes as a function of s. At a higher temperature, it is required to pause later in order to observe the same increase in success probability, compared to a model at a lower temperature. The reason for this is that, as $s \to 1$, valleys in the potential become deeper, and the hotter system can more easily jump out of the global minimum when the well is too shallow. Since the standard update procedure allows for one to in principle explore the entire landscape more readily, we do not see the characteristic freeze-out effect.

In contrast to this, Fig. 10 (top plot) shows a dramatic change in the behavior of SVMC TF. Here, more in line with intuition, the higher temperature system performs much worse compared to lower temperatures in the region where pausing has an effect. This is due to the freezing of the dynamics, since now the hotter system, which otherwise requires a pause late in the anneal to avoid jumping out of the global minima, has slowed dynamics in this region.

There are some intriguing differences between the dynamics of SVMC TF compared to the AME however. In the AME, as expected from the general theory, a colder temperature is always best. This is because there are fewer thermal excitations out of the ground state in the first place since the Gibbs state has a higher GS population. As such, at a given pause location, a colder temperature always arrives at a greater population in the ground state. This is especially true for pausing both earlier in the anneal before the minimum gap and later in the anneal after the minimum, where we do not expect pausing to repopulate the ground state (early in the anneal the state already has overwhelming population in the ground state, while late in the anneal, thermal relaxation is slowed down considerably). In both cases, the ground-state population is determined primarily by how much population is lost due to thermal excitations by crossing the minimum gap (we note that in the closed-system case, the ground-state population would be almost 1 for these simulation parameters), which is reflected by their equal P_0 values for low and high s_p.

In SVMC TF, while the peak GS probability increases with decreasing temperature, it is not generally true that a colder temperature results in a higher GS probability for a fixed pause location. We see this in Fig. 10 (top plot), where, e.g., when pausing after $s > 0.6$, the hotter temperature can be the most effective. This is explained again by the semiclassical picture. SVMC cannot tunnel to the new global minimum but must reach it via thermal hopping. At a given pause location, a hotter temperature may enable the system to hop out of local minima more effectively than a colder system, and if a colder system pauses too late, it is likely to become stuck in suboptimal configurations.
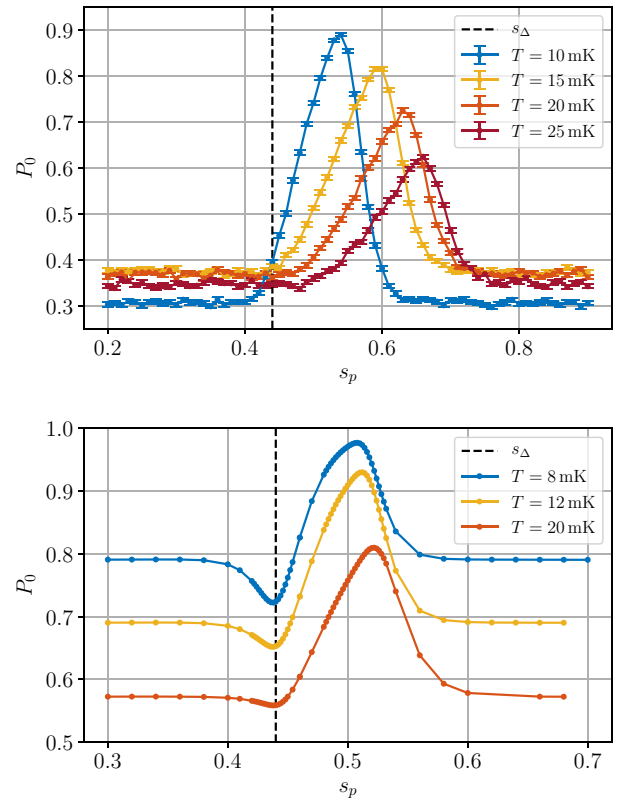

FIG. 10.   Temperature dependence for SVMC TF (top) and the AME (bottom). Plots showing dependence of temperature under the pause at s_p. In SVMC, the anneal "time" is 10^4 sweeps, and the pause "time" is fixed at 10^6 sweeps for all curves. In the AME we use an anneal of 1 μs and pause of 100 μs. The location of the minimum gap s_Δ is marked with a vertical dashed line.

We also mention that the behavior observed in the SVMC model is, of course, heavily dependent on the update scheme used. One can imagine different schemes where the angles are updated according to a different heuristic (i.e., other than used here with updates proportional to $A(s)/B(s)$). It may therefore be feasible that different schemes can give significantly different behavior. We leave this for future study.

V. CONDITION FOR ADVANTAGE BY RELAXATION

Based on the above observations of Sec. 1 where the dynamics are dependent on a single dominant decay rate, we provide a condition such that the time required to find the ground state is reduced. Note that in Ref. [21] conditions for which an optimal pause location exists were derived under assumptions of the AME model, i.e., conditions for which there is a pause location that will increase the final ground-state success probability. Whilst the existence of an optimal pause location is a necessary prerequisite for reducing the TTS by pausing, this alone does not guarantee an improvement in performance, which must

take into account the total time used to find a ground state (according to the TTS metric). The condition provided here is agnostic to the model by which thermalization occurs so long as there is a single dominant decay rate, and therefore applies to all three systems studied (DW, SVMC TF, AME).

Assuming a functional form as empirically demonstrated in Fig. 7, dependent on a single dominant decay rate γ (in the DW case, neglecting the long timescale; see Appendix E), we can write the ground-state probability measured as

$$P_0(s^*, t_a, t_p) = P_G(s^*, t_a) - [P_G(s^*, t_a) - P_a(t_a)]e^{-\gamma(s^*)t_p},$$
(8)

where $P_0(s^*, t_a, t_p)$ is the ground-state probability at the end of an anneal with anneal time t_a and a pause of length t_p inserted at the optimal location s^*. Here $P_a(t_a)$ is the ground-state success probability achieved under a standard anneal with anneal time t_a and no pause, and $P_G(s^*, t_a)$ the probability reached under an anneal with an *infinite* pause at location s^* (i.e., one reaching the Gibbs state), and with an anneal time t_a. We expect $P_G(s^*, t_a)$ to have minimal, if any, dependence on t_a because s^* is typically after the minimum gap, and annealing from s^* to $s = 1$ should be effectively adiabatic.

The standard definition of the TTS (T_S) is the time required to reach the ground state with 99% probability:

$$T_S(s^*, t_a, t_p) = \frac{\log(1 - 0.99)}{\log[1 - P_0(s^*, t_a, t_p)]}(t_a + t_p).$$
(9)

Note that we are assuming that the optimal pause location s^* has already been identified.

At long pause times, one expects TTS to increase as the approach to equilibrium occurs and any gain in success
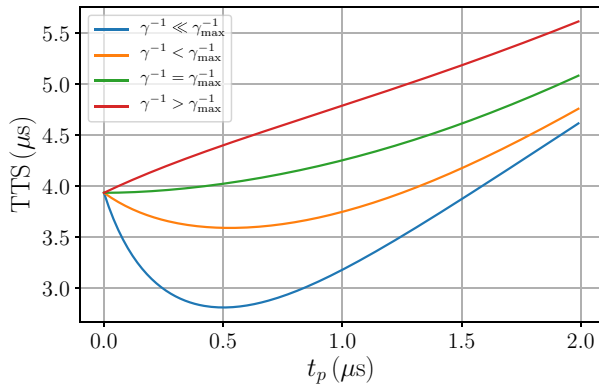

FIG. 11. Time to solution for various decay rates, as a function of the pause time. For the plot, we use parameters of the AME as found in Fig. 7 ($P_G = 0.95, P_a = 0.69, t_a = 1 \mu s$). Here $\gamma_{\max}^{-1}$ is 0.72 μs as shown in Table I. When γ^{-1} is small enough, an initial decrease in TTS can be observed. For t_p larger than shown, the curves eventually overlap and scale linearly in t_p.

TABLE I. Table of values for the three models decay time γ^{-1} (from Fig. 7), compared to the required value to observe a reduction in TTS. In all three models, γ^{-1} is too large.

Model	γ^{-1} for TTS reduction	γ^{-1} measured
DW	1.5 μs	23 μs
SVMC TF	$14k$ sweeps	$21k$ sweeps
AME	0.72 μs	54 μs

probability is over compensated by the increase in run time ($t_a + t_p$), as shown in Fig. 11. With this, we assert that the TTS should initially decrease:

$$\left. \frac{\partial T_S}{\partial t_p} \right|_{t_p=0} < 0.$$
(10)

This provides the condition

$$\gamma > \frac{-(1 - P_a)\log(1 - P_a)}{t_a(P_G - P_a)},$$
(11)

stating that the decay rate (time) needs to be large (small) enough in order to observe a decrease in TTS. This condition on γ is not simply about increasing the success probability; it reflects the *rate* the success probability needs to increase whilst pausing, in order for it to be advantageous over a standard anneal. It gives a nontrivial relationship between the annealing dynamics that give rise to P_a and the mostly static properties at s^* that determine P_G. As an illustrative example, let us consider a scenario where $P_G = 1$ and $P_a = 1 - e^{-ct_a}$, where c denotes the convergence rate to the ground state due to annealing. In this case, we find that $\gamma > c$, indicating that in order for the pause to be beneficial, the thermal relaxation rate to the ground state must be faster than the annealing convergence rate to the ground state.

In Table I we compute the maximum allowable decay time γ^{-1} in order to observe a reduction in TTS, comparing to the parameters extracted from our simulations (Fig. 7). We see that in all three models, the measured value of γ^{-1} is too large in order to observe a reduction in TTS.

In Fig. 11 we see that only for relatively short pause times can an improvement in TTS be observed. This is consistent with the results of Gonzalez Izquierdo et al. [41], who found experimentally that, for a certain set of problems requiring embedding (problem size more than 100 qubits), an improvement in TTS was observed for pause times less than around 2 μs.

Given a decay rate γ extracted from experiments and simulations, it is possible to find the optimal pause time by setting $\partial T_S / \partial t_p = 0$ (around 0.5 μs in Fig. 11).

VI. CONCLUSION

By performing a detailed analysis of a small instance with a sufficiently small gap, we have shown that the characteristic effect of improved ground-state probability observed under pausing in transverse-field annealing is certainly not unique to quantum models. Our model of open-system dynamics is the quantum AME, while our classical simulations are based on the SVMC algorithm, which, with simple modifications, can be made to exhibit many of the thermal effects such as freeze out observed with experiments on D-Wave. We find significant qualitative agreements between all three results, which is perhaps striking given how SVMC and SVMC TF do not have any detailed model of a thermal environment with which the system interacts. This casts doubt on whether the origin of an improved success probability whilst pausing is an inherently quantum one, though our analysis has been for a single problem instance so we caution against generalizations. There may be problem classes that exhibit drastically different behavior for these models and favor a quantum model under pausing. We also should not discount the possibility that the physical process is significantly faster than a simulation of the process. For example, one can view the results of Ref. [42] as a positive indication of physical thermal processes exhibiting a faster timescale than the simulated process. In our analysis, we focus on the probability of measuring the ground state of the Ising Hamiltonian, which is the relevant task for solving classical optimization problems, so we believe that identifying classical optimization problems that can take advantage of these faster physical thermal processes remains an important research direction. The three models studied here all exhibit a single dominant decay channel, causing the repopulation of the ground state, which we use to derive a simple condition [Eq. (11)] for which pausing can lead to a reduction in the time to solution. The temperature features prominently in this condition, since it determines the Gibbs ground-state probability P_G and the default annealing probability P_a [from Eq. (11)], but it also appears to be the most relevant tool to elucidate the dynamics involved in nonzero temperature annealing. We see that the two models studied here behave qualitatively differently as a function of temperature, which can serve as a useful benchmark for experimental devices [39].

The role of temperature is expected to become even more important at larger system sizes. Since we know that the Gibbs ground-state probability decreases exponentially with increasing system size at a fixed temperature [43], in order to maintain an advantage acquired with pausing with increasing system size will require the temperature to also decrease with system size. With increasing system size, it seems reasonable to expect γ^{-1} to scale exponentially for problems with exponentially small gaps [35], so it is likely that the parameter regimes under which the TTS may be improved are not practically scalable. This conclusion is not surprising given that we should only expect scalable quantum advantages in the context of fault-tolerant quantum computing. Nevertheless, we believe our analysis will be helpful going forward in determining whether or not a pause can be beneficial for quantum annealing devices of a fixed size.

APPENDIX A: INSTANCE $\mathcal{I}_{12}^0$

The Ising parameters of instance $\mathcal{I}_{12}^0$ used in the main text are

$$J_{0,3} = -0.888\,765\,722\,269,$$
$$J_{1,3} = -0.453\,396\,499\,878,$$
$$J_{2,3} = -0.581\,810\,391\,599,$$
$$J_{0,4} = -0.222\,181\,654\,366,$$
$$J_{1,4} = 0.623\,744\,373\,452,$$
$$J_{2,4} = 0.805\,987\,681\,935,$$
$$J_{0,5} = 0.333\,955\,924\,275,$$
$$J_{1,5} = 0.995\,412\,322\,296,$$
$$J_{2,5} = 0.490\,983\,144\,977,$$
$$J_{3,9} = -0.925\,420\,427\,917,$$
$$J_{6,9} = 0.663\,343\,935\,819,$$
$$J_{7,9} = 0.687\,446\,523\,051,$$
$$J_{8,9} = 0.749\,085\,209\,325,$$

$J_{4,10} = -0.055\,994\,539\,750\,2,$

$J_{6,10} = -0.990\,358\,090\,729,$

$J_{7,10} = 0.491\,802\,375\,676,$

$J_{8,10} = 0.505\,416\,377\,921,$

$J_{5,11} = -0.400\,367\,703\,995,$

$J_{6,11} = -0.831\,748\,994\,702,$

$J_{7,11} = 0.413\,887\,841\,297,$

$J_{8,11} = 0.204\,601\,421\,856.$

APPENDIX B: ARGUMENT FOR WHY $\langle E_1(s)|\, Z\, |E_0(s)\rangle \propto A(s)^{\alpha}$

We restrict our analysis to the subspace defined by having a $+1$ eigenvalue under the operator $P = \prod_{i=1}^{n} \sigma_i^x$. We define our Hamiltonian as $H(s) = B(s)\{H_p + [A(s)/B(s)]H_x\}$, where we treat $A(s)/B(s) \equiv \lambda(s)$ as our perturbative parameter. We use $|E_i^{(0)}\rangle \equiv |E_i(s=1)\rangle$ for $i = 0, 1$ as defined in Eqs. (1) and (2) as our zeroth-order state in perturbation theory that we are expanding from, and we note that these states are eigenstates of the $Z = \sum_i \sigma_i^z$ operator. If we consider an expansion of the form (unnormalized)

$$|E_i(\lambda)\rangle = \sum_{k=0} \lambda^k |E_i^{(k)}\rangle \tag{B1}$$

then the first nonzero term that contributes in the perturbative expansion is at fourth order in λ (i.e., the Hamming distance between GS and first excited basis states at $s = 1$):

$$\begin{aligned}
\langle E_1(\lambda)|\, &Z\, |E_0(\lambda)\rangle \\
&= \lambda^4 (\langle E_1^{(0)}|\, Z\, |E_0^{(4)}\rangle \\
&\quad + \langle E_1^{(1)}|\, Z\, |E_0^{(3)}\rangle + \langle E_1^{(2)}|\, Z\, |E_0^{(2)}\rangle \\
&\quad + \langle E_1^{(3)}|\, Z\, |E_0^{(1)}\rangle + \langle E_1^{(4)}|\, Z\, |E_0^{(0)}\rangle) + O(\lambda^6). \tag{B2}
\end{aligned}$$

Therefore, we can expect that $\langle E_1(s)|\, Z\, |E_0(s)\rangle \propto A(s)^{\alpha}$.

APPENDIX C: ROLE OF THE ANNEALING TIME

For the results in the main text, we kept the annealing rate fixed. We show in Fig. 12 how the annealing rate can change the qualitative features of the DW and SVMC-TF results. For the DW, increasing the annealing rate increases the GS probability for the early and late s_p values, but does not change the location of the optimal pause point or its peak value. This is similar to the behavior observed for SVMC TF.

For SVMC TF with very few annealing updates, we observe a similar drop in the GS probability as for the AME (see Fig. 6), because with fewer sweeps during the anneal process, SVMC TF is less likely to escape the
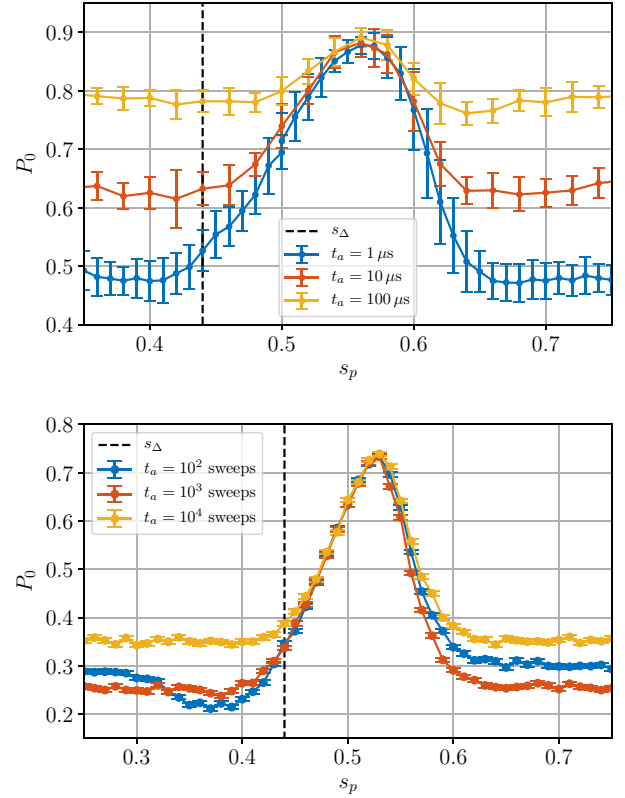

FIG. 12. Comparison of different annealing rates. Ground-state probability reached by the DW (top) and the SVMC-TF algorithm (bottom) for a fixed pause duration $100\,\mu$s for the DW, and $100k$ sweeps of SVMC TF, but a varied annealing rate.

local minimum and reach the global minimum after the pause. However, at these small annealing rates, we begin to observe differences in the GS probability for early and late pause locations.

APPENDIX D: COMPARING SVMC TO SVMC TF

Here we provide some additional comparisons between SVMC and SVMC TF. We show in Fig. 13 how the two differ in terms of changing the number of sweeps while keeping the temperature fixed. We see that SVMC does not exhibit as pronounced a suppression in the ground-state probability for large s_p as SVMC TF does, which can directly be attributed to the modification in rotor updates we discuss in the main text.

In Fig. 14 we show the pause time required for the standard SVMC algorithm to reach a target success probability (cf. Fig. 8 in the main text). As in SVMC TF, we see here an exponential increase in the time required to pause to reach the target success probability, though the scaling is more mild compared to SVMC TF, due to the fact that the dynamics slow much more in SVMC TF late in the anneal. Moreover, the dynamics persist much later in the standard SVMC (as seen by comparing the scales on the x axis).
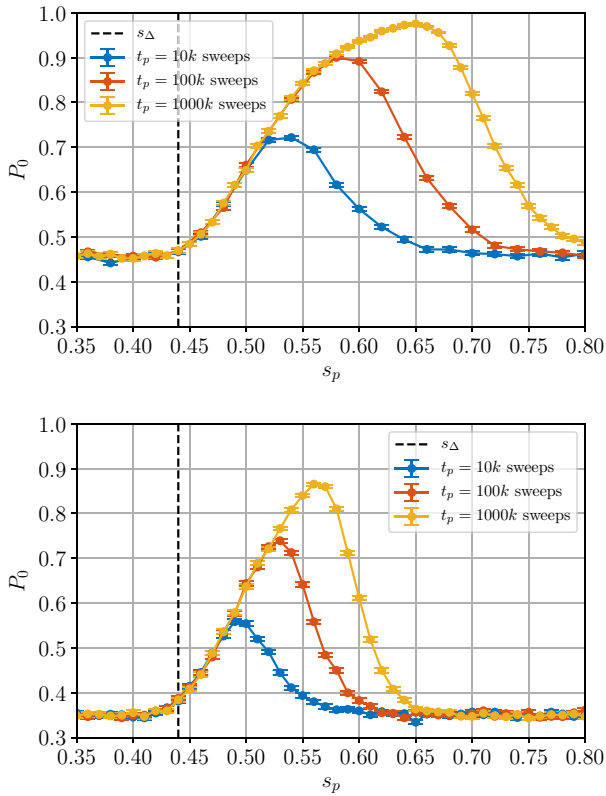
FIG. 13. Comparison between standard SVMC (top) and SVMC with transverse-field updates (bottom). These simulations are both performed at 12 mK with $10k$ sweeps in the ramp. Note that the standard SVMC continues to have updates much later in the anneal (N.B. at $s = 0.75$, the ratio $A/B < 10^{-2}$). Here we use an anneal "time" of 10^4 sweeps for all curves.

APPENDIX E: RELAXATION TIMESCALES IN THE DW

In Fig. 15 we plot the data of Fig. 7 (top plot) with only a single timescale for the curve fit. We see that it does not
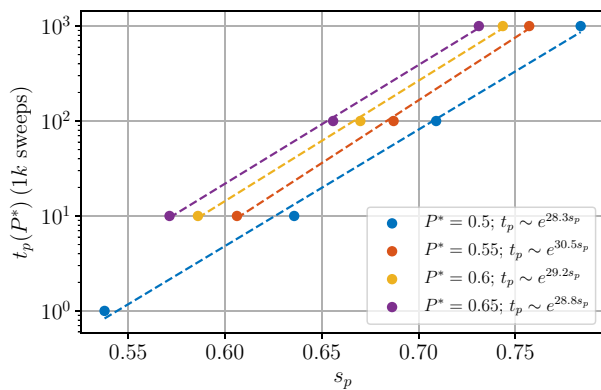

FIG. 14. Standard update SVMC pause time to target. Pause time required for the standard SVMC algorithm to reach target success probability P^*. As in SVMC TF, the time required increases exponentially in s_p (location of the pause).
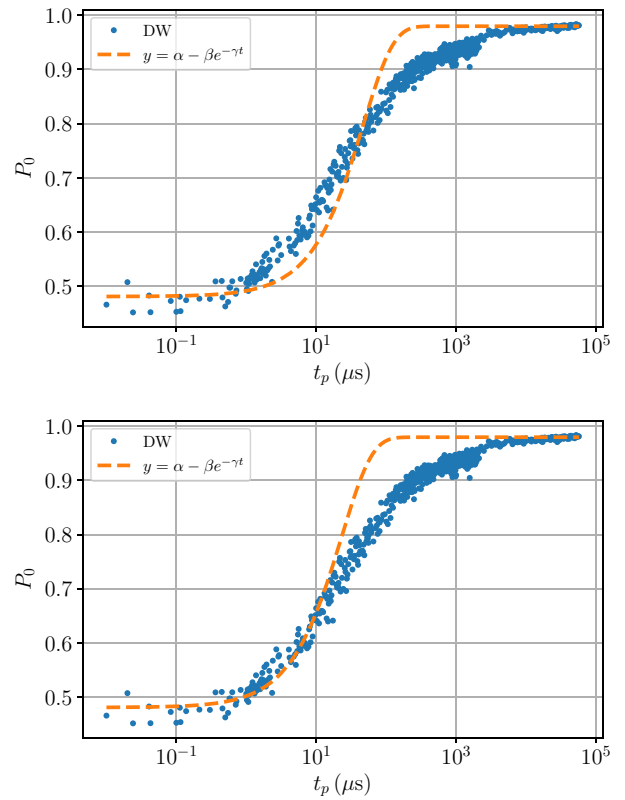
FIG. 15. Single timescale fit for the DW under the pause. Top: fixing $\alpha = P_0(t_p \to \infty)$ and $\beta = \alpha - P_0(t_p = 0)$, we obtain $\gamma^{-1} = 47\ \mu$s from a nonlinear least-squares fitting procedure. Bottom: using the same functional form, but replacing the single timescale $1/\gamma$ by the dominant timescale from the top plot of Fig. 7.

describe the data accurately, in both cases where γ is a free parameter, or where we fix γ with the dominant value (0.043) found from the two timescale fits in the main text [top plot of Fig. 7]. Note however that in the latter case, for small values of $t_p \lesssim 20\ \mu$s, the fit matches the data reasonably well, which justifies the approximation for our single timescale analysis in the main text (Sec. V). Moreover, the addition of extra decay parameters does not identify other timescales (i.e., two unique ones are identified).

[1] M. W. Johnson *et al.*, Quantum annealing with manufactured spins, Nature **473**, 194 (2011).

[2] R. Harris *et al.*, Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor, Phys. Rev. B **82**, 024511 (2010).

[3] P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. Przybysz, and J. Whittaker, Architectural considerations in the design of a superconducting quantum annealing processor, IEEE Trans. Appl. Supercond. **24**, 1 (2014).

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

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

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

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

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

[9] J. Marshall, E. G. Rieffel, and I. Hen, Thermalization, Freeze-Out, and Noise: Deciphering Experimental Quantum Annealers, Phys. Rev. Appl. **8,** 064025 (2017).

[10] A. D. King *et al.*, Observation of topological phenomena in a programmable lattice of 1800 qubits, Nature **560,** 456 (2018).

[11] R. Harris *et al.*, Phase transitions in a programmable quantum spin glass simulator, Science **361,** 162 (2018).

[12] M. H. Amin, Searching for quantum speedup in quasistatic quantum annealers, Phys. Rev. A **92,** 052323 (2015).

[13] T. Albash, S. Boixo, D. A. Lidar, and P. Zanardi, Quantum adiabatic Markovian master equations, New J. Phys. **14,** 123016 (2012).

[14] J. Marshall, D. Venturelli, I. Hen, and E. G. Rieffel, Power of Pausing: Advancing Understanding of Thermalization in Experimental Quantum Annealers, Phys. Rev. Appl. **11,** 044083 (2019).

[15] Z. Gonzalez Izquierdo, T. Albash, and I. Hen, Testing a quantum annealer as a quantum thermal sampler, arXiv:2003.00361 (2020).

[16] T. Lanting *et al.*, Entanglement in a Quantum Annealing Processor, Phys. Rev. X **4,** 021041 (2014).

[17] P. Kairys, A. D. King, I. Ozfidan, K. Boothby, J. Raymond, A. Banerjee, and T. S. Humble, Simulating the Shastry-Sutherland Ising Model using Quantum Annealing, arXiv:2003.01019.

[18] A. D. King, C. Nisoli, E. D. Dahl, G. Poulin-Lamarre, and A. Lopez-Bezanilla, Quantum Artificial Spin Ice, arXiv:2007.10555 (2020).

[19] G. Passarelli, V. Cataudella, and P. Lucignano, Improving quantum annealing of the ferromagnetic *p*-spin model through pausing, Phys. Rev. B **100,** 024302 (2019).

[20] G. Passarelli, K.-W. Yip, D. A. Lidar, H. Nishimori, and P. Lucignano, Reverse quantum annealing of the *p*-spin model with relaxation, Phys. Rev. A **101,** 022331 (2020).

[21] H. Chen and D. A. Lidar, Why and When Pausing is Beneficial in Quantum Annealing, Phys. Rev. Appl. **14,** 014100 (2020).

[22] J. R. Klauder, Path integrals and stationary-phase approximations, Phys. Rev. D **19,** 2349 (1979).

[23] P. J. D. Crowley and A. G. Green, Anisotropic landau-lifshitz-gilbert models of dissipation in qubits, Phys. Rev. A **94,** 062106 (2016).

[24] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, How 'Quantum' is the D-Wave Machine? arXiv:1401.7087 (2014).

[25] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, J. Chem. Phys. **21,** 1087 (1953).

[26] W. K. Hastings, Monte Carlo sampling methods using Markov chains and their applications, Biometrika **57,** 97 (1970).

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

[28] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, Reexamining classical and quantum models for the d-wave one processor, Eur. Phys. J. Spec. Top. **224,** 111 (2015).

[29] T. Albash, I. Hen, F. M. Spedalieri, and D. A. Lidar, Reexamination of the evidence for entanglement in a quantum annealer, Phys. Rev. A **92,** 062328 (2015).

[30] *The D-Wave 2000Q Quantum Computer Technology Overview*, Tech. Rep. (D-Wave Systems Inc., 2018).

[31] M. H. S. Amin and V. Choi, First-order quantum phase transition in adiabatic quantum computation, Phys. Rev. A **80,** 062326 (2009).

[32] E. B. Davies and H. Spohn, Open quantum systems with time-dependent hamiltonians and their linear response, J. Stat. Phys. **19,** 511 (1978).

[33] G. Lindblad, On the generators of quantum dynamical semigroups, Comm. Math. Phys. **48,** 119 (1976).

[34] L. Campos Venuti, T. Albash, D. A. Lidar, and P. Zanardi, Adiabaticity in open quantum systems, Phys. Rev. A **93,** 032118 (2016).

[35] L. Campos Venuti, T. Albash, M. Marvian, D. Lidar, and P. Zanardi, Relaxation versus adiabatic quantum steady-state preparation, Phys. Rev. A **95,** 042302 (2017).

[36] T. Albash and D. A. Lidar, Decoherence in adiabatic quantum computation, Phys. Rev. A **91,** 062320 (2015).

[37] S. Muthukrishnan, T. Albash, and D. A. Lidar, Tunneling and Speedup in Quantum Optimization for Permutation-Symmetric Problems, Phys. Rev. X **6,** 031010 (2016).

[38] D. Roberts, L. Cincio, A. Saxena, A. Petukhov, and S. Knysh, Noise amplification at spin-glass bottlenecks of quantum annealing: A solvable model, Phys. Rev. A **101,** 042317 (2020).

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

[40] Spin-Bath Polarization Effect, https://docs.dwavesys.com/docs/latest/c_qpu_2.html (accessed: 19 August 2020).

[41] Z. Gonzalez Izquierdo, S. Grabbe, S. Hadfield, J. Marshall, Z. Wang, and E. Rieffel, Ferromagnetically shifting the power of pausing, arXiv:2006.08526 (2020).

[42] A. D. King *et al.*, Scaling advantage in quantum simulation of geometrically frustrated magnets, arXiv:1911.03446 (2019).

[43] T. Albash, V. Martin-Mayor, and I. Hen, Temperature Scaling law for Quantum Annealing Optimizers, Phys. Rev. Lett. **119,** 110502 (2017).