Demonstration of error-suppressed quantum annealing via boundary cancellation

Humberto Munoz-Bauza,1,2 Lorenzo Campos Venuti,3,1,2 and Daniel Lidar1,2,4,5

1Department of Physics & Astronomy, University of Southern California, Los Angeles, CA 90089, USA
2Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, CA 90089, USA
3Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Av., Cambridge, MA 02139, USA
4Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089, USA
5Department of Chemistry, University of Southern California, Los Angeles, CA 90089, USA

(Dated: February 8, 2023)

The boundary cancellation theorem for open systems extends the standard quantum adiabatic theorem: assuming the gap of the Liouvillian does not vanish, the distance between a state prepared by a boundary cancelling adiabatic protocol and the steady state of the system shrinks as a power of the number of vanishing time derivatives of the Hamiltonian at the end of the preparation. Here we generalize the boundary cancellation theorem so that it applies also to the case where the Liouvillian gap vanishes, and consider the effect of dynamical freezing of the evolution. We experimentally test the predictions of the boundary cancellation theorem using quantum annealing hardware, and find qualitative agreement with the predicted error suppression despite using annealing schedules that only approximate the required smooth schedules. Performance is further improved by using quantum annealing correction, and we demonstrate that the boundary cancellation protocol is significantly more robust to parameter variations than protocols which employ pausing to enhance the probability of finding the ground state.

I. INTRODUCTION

Adiabatic quantum computing [1] and quantum annealing [2] can be used to prepare ground states and thermal states of quantum systems, a central desideratum of quantum computation. The quality of ground state preparation and Gibbs sampling is central to the performance of quantum algorithms for optimization as well as machine learning problems [3]. The performance of these computational paradigms is quantified by the adiabatic theorem in its different forms, whether for closed or open quantum systems. The theorem, or family of theorems, places a bound on the adiabatic error: the distance between the state that we set out to prepare (usually the solution to a computational problem) and the state the system actually ends up in through the dynamical evolution. In general, if a spectral gap condition is satisfied, the adiabatic error is bounded by $C/t_f$ for some constant $C$, where $t_f$ is the total evolution time [4–6]. The boundary cancellation theorem (BCT) shows that an improvement is possible if the schedule from the initial to the final generator (Hamiltonian or Liouvillian) has certain properties. In the closed system case the adiabatic error is bounded by $C_k/t_{f+1}^{k+1}$ for some constant $C_k$ independent of $t_f$, if the time-dependent Hamiltonian $H(t)$ has vanishing derivatives up to order $k$ at the beginning and end of the evolution [7], a bound that can be improved to exponentially small in $t_f$ under additional assumptions [8–12]. This theorem has recently been extended to the open system setting for a general class of time-dependent Liouville operators $\mathcal{L}(t)$, where it can be used to prepare steady states of Liouvillians instead of ground states of Hamiltonians. The theorem can be succinctly stated as follows: For a time-dependent Liouvillian $\mathcal{L}(t)$ with a unique steady state $\sigma(t)$ separated by a gap at each time $t$, the adiabatic error is likewise bounded by $C_k/t_{f+1}^{k+1}$, but it is sufficient for the derivatives to vanish up to order $k$ only at the end time $t_f$ [13]. Boundary cancellation (BC) plays a significant role in the theoretical analysis of error suppression in Hamiltonian quantum computing [14, 15].

In this work we set out to demonstrate the BCT predictions in experiments using the D-Wave 2000Q (DW2KQ) quantum annealer. The DW2KQ implements the transverse field Ising model $H(t) = A(t)H_X + B(t)H_Z$, where $H_X$ and $H_Z$ are transverse and longitudinal (Ising) Hamiltonians, respectively [16, 17]. It allows limited programmability of the control schedules $A(t)$ and $B(t)$ of the system Hamiltonian, which we exploit to implement boundary cancellation protocols.

To guide our intuition we model the behavior of the D-Wave quantum annealer with the adiabatic master equation (AME) derived in [18]. This is a time-dependent Davies-like master equation [19] which has been successfully used to interpret several D-Wave experiments, e.g., [20, 21] (though not always [22]). When combining the AME for dephasing with boundary cancellation, we encounter a problem. Namely, as we explain in detail below, the Liouvillian gap vanishes at $t = t_f$, which prevents us from directly applying the BCT in the form given in Ref. [13]. To circumvent this problem, in this work we generalize the BCT and identify the conditions on the Liouvillian gap under which BC does or does not remain effective. We find that the generalized BCT plays an important role in the D-Wave implementation.

There is another significant consideration regarding the implementation of BC on D-Wave: the phenomenon of freezing. In essence, freezing refers to a significant in-
crease in all relaxation timescales well before the end of the anneal ($t < t_f$), in such a manner that the system is frozen in a state that does not correspond to the steady state of $\mathcal{L}(t_f)$ [18]. However, the frozen state is not truly static but quasi-static [23], i.e., relaxation is not fully switched off but instead proceeds on a timescale that may be much longer than practically realizable anneal times $t_f$. Such glassy-like behavior can take place even in very simple systems that do not have a glassy landscape. We give an explanation of freezing in Sec. II D below, based on the AME. Essentially, it arises due to the fact that the system Hamiltonian commutes with the system-bath coupling when $A(s) = 0$, and the fact that the schedule $A(t)$ has nearly vanished long before $t = t_f$ (see Fig. 1).

This is true in particular of currently available commercial quantum annealers manufactured by D-Wave [24], which can be effectively described by longitudinal field coupling to a bosonic bath, which commutes with the system Hamiltonian at sufficiently low transverse fields.

The state of the system is insensitive to the details of the schedules past the freezing point, so that BC becomes ineffective. This freezing mechanism plays an adverse role in demonstrating the predictions of the BCT, since the experimentally accessible anneal times are significantly shorter than the relaxation time required to reach the true steady state of $\mathcal{L}(t_f)$.

In order to overcome both issues, i.e., the gaplessness of the Liouvillian at $t = t_f$ and freezing at $t \approx 0.5 t_f$, we design the programmable DW2KQ schedule so that it flattens (and thus implements BC) at a point $t_{\text{BC}}$ before freezing, followed by a ramp to the final values of $B(t_f)$ and $A(t_f)$, upon which the system is measured in computational basis (eigenbasis of $H_Z$).

We apply the protocol to two 8-qubit gadgets: 1) a “ferromagnetic chain gadget” (FM-gadget), and 2) a “tunneling gadget” (T-gadget) with a larger tunneling barrier between the ground state and the first excited state around the minimum gap [25]. We estimate the scaling of the adiabatic error for the FM-gadget and compare the experimental results of the boundary cancellation protocol (BCP) with those predicted by open system simulations using the AME. Based on these simulations, we discuss the conditions under which the BCT-predicted scaling may be observed. We also apply the protocol to an error-corrected version of the FM-gadget using quantum annealing correction (QAC) [26]. The QAC encoding mitigates the effects of Hamiltonian programming errors [27, 28] and effectively lowers the temperature of the environment [29, 30]; we find that QAC improves the performance of the BCP.

The structure of this paper is as follows. In Sec. II we first review the BCT, and formulate a new version which accounts for the possibility of the Liouvillian gap closing at the same point as where the boundary cancellation conditions are enforced, which is relevant for our experiments. The rest of the section is devoted to a discussion of freezing, the effect of a ramp after the BC point, and anomalous heating, all of which are phenomena affecting the performance of the BCP due to their presence in our experiments. In Sec. III we describe our methods: the BCP we used in detail, the two 8-qubit gadgets used in our experiments, and encoding with quantum annealing correction. Our results are presented in Sec. IV: we start with a standard linear control schedule and then present the results for the experimental BCP implementation. We include simulation results for a higher precision version of the BCP for reference, and then report on experiments from the QAC-encoded version of one of our gadgets. Finally, we compare the performance of the BCP to that of the pausing protocol. We conclude with a discussion and outlook in Sec. V. The appendix includes background on the adiabatic master equation, proofs of various Propositions, supporting numerical results, additional information needed to reproduce our experiments and analysis, and experimental results from two other D-Wave processors.

For convenience, we include a glossary of the acronyms, terminology, and key symbols we used in Table I.

## II. THEORY

### A. Review of the BCT results of Ref. [13]

We consider a finite-dimensional system whose density matrix $\rho_{t_f}$ satisfies the master equation $d\rho_{t_f}/dt = \mathcal{L}_{t_f}(t)\rho_{t_f}(t)$, where $\mathcal{L}_{t_f}(t)$ is a time-dependent Liouvillian superoperator. If $\mathcal{L}_{t_f}(t)$ depends on the time $t$ only through the “anneal parameter” $\tau := t/t_f$, where $t_f$ is the total evolution time, then defining $\rho(\tau) := \rho_{t_f}(t)$ and $\mathcal{L}(\tau) := \mathcal{L}_{t_f}(t)$, the master equation takes the form

$$\frac{d\rho}{d\tau} = t_f\mathcal{L}(\tau)\rho(\tau).$$  

(1)

| $\eta$ | $\rho$ | $\sigma$ | $s_{BC}$ | $s_0, t_0$ | AME | BCP | BCT | $D_{GS}$ | DW2KQ | FM-gadget | NTS | $P_{GS}$ | $P^*$ | PR | QAC | T-gadget |
|-------|-------|---------|--------|----------|-----|-----|-----|-------|--------|-----------|-----|--------|-------|----|-----|--------|
| Adiabatic error scaling exponent | Solution of the master equation | Gibbs state at end of BCP | Termination point of BCP | Dynamical freezing point | Adiabatic master equation | Boundary cancellation protocol | Boundary cancellation theorem | $\text{adiabatic error}$ | D-Wave 2000Q annealer | Ferromagnetic chain gadget | Number of tries-to-solution | Measured ground state prob. | Gibbs ground state prob. | Pause-ramp protocol | Quantum annealing correction | Tunneling gadget |
| Sec. II F | Eq. (1) | Sec. II E | Sec. III A | Eq. (16) | Sec. II A, [18] | Sec. III A | Sec. II, [13] | Eq. (22) | [31, 32] | Sec. III B | Eq. (34) | Eq. (20a) | Eq. (20b) | Sec. IV E | [25] | [26] | Sec. III C, [26] | [25] |
The main result of Ref. [13] is summarized in the following proposition:

**Proposition 1** Assume that $L(\tau)$ is such that $\|L(\tau)\|_{1,1}$ is summable in $\tau \in [0,1]$. $L(\tau)$ is differentiable to order $k + 2$ in a neighborhood of $\tau = 1$, and generates a trace-preserving and hermiticity-preserving contractive semigroup, i.e., $\|e^{\tau L(\tau)}\|_{1,1} \leq 1$ for all $\tau \geq 0$, $\tau \in [0,1]$. Suppose $L(\tau)$ has a unique steady state $\sigma(\tau)$ with eigenvalue zero separated from the rest of the spectrum by a nonzero gap $\sqrt[4]{\tau} \in [0,1]$. Let $\rho(\tau)$ be the solution of Eq. (1) with initial condition $\rho(0) = \sigma(0)$. If $L(\tau)$ has vanishing derivatives at $\tau = 1$ to order $k$, i.e., $\partial^{[k]} L \Big|_{\tau=1} = 0 \ \forall j = 1, 2, \ldots, k$, then there is a constant $C_k$ independent of $t_f$ such that

$$\|\rho(1) - \sigma(1)\|_1 \leq \frac{C_k}{t_f^{k+1}}. \quad (2)$$

We note that the smoothness assumptions of Prop. 1 are more relaxed than in Ref. [13].

From here on we focus in particular on the AME, which we briefly review in Sec. II B. It was shown in [13] that, under the action of the AME Liouvillian, it is possible to enforce the BCT as required in Prop. 1 by controlling just the time-dependent system Hamiltonian $H_S$:

**Proposition 2** Assume the system evolves according to the time dependent Liouvillian in Eq. (5a). Assume further that $\partial^{[k]} H_S(t) \big|_{t=t_f} = 0$, for $j = 1, 2, \ldots, k$. Then

$$\partial^{[k]} L \big|_{t=t_f} = 0 \ \forall j = 1, 2, \ldots, k.$$

This means that rather than having to directly check that the entire Liouvillian satisfies the conditions of vanishing derivatives at the end, it suffices to check that the system Hamiltonian $H_S(t)$ satisfies the same conditions. However, the Prop. 1 requirement that the Liouvillian has a finite gap above its zero eigenvalue throughout the evolution window must be checked independently. For the finite dimensional systems that we consider here, this is equivalent to the requirement that the steady state is unique throughout the evolution, or alternatively, that there are no “level crossings” of $L(\tau)$’s zero eigenvalue throughout the evolution. Since the system Hamiltonian is naturally the controllable component of the total Hamiltonian or Liouvillian, this is also a physically sensible controllability condition.

Moreover, it was verified numerically in [13] that Eq. (2) remains valid for a more general Liouvillian in time-dependent Redfield form, even though it does not satisfy the hypothesis of Props. 1 and 2, i.e. for which the vanishing of the derivatives of the system Hamiltonian does not imply the vanishing of the derivatives of the Liouvillian. See [13] for details.

We next review the AME for quantum annealing problems and discuss a modified form of the BCT, which is tailored to the situation we encounter in our experiments using the D-Wave annealer.

### B. Adiabatic master equation for quantum annealing

We now consider the quantum annealing problem defined by a system Hamiltonian of the form

$$H_S(s) = \frac{A(s)}{2} H_X + \frac{B(s)}{2} H_Z,$$

where

$$H_X = -\sum_i \sigma_i^x, \quad H_Z = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z. \quad (4)$$

and $\gamma \in (0,1]$ is a scaling factor that controls the energy scale of the computational problem. We refer to $A(s)$ and $B(s)$ as the physical schedules of the anneal. They are fixed properties of the quantum annealing hardware (see Fig. 1), but $s$ can be made to depend on $\tau = t/t_f$, in the adjustable way that we describe below for the DW2QK. We call the function $s(\tau) : [0,1] \mapsto [0,1]$ the control schedule.

We model the anneal according to the AME [18] with the Hamiltonian in Eq. (3). The AME is derived from the total Hamiltonian $H(t) = H_S(t) + H_B + H_{SB}$, with
\[ H_{SB} = g \sum_i A_i \otimes B_i \] where the \( A_i \) (\( B_i \)) are the dimensionless system (bath) operators, subject to an adiabatic approximation. The resulting Liouvillian is of the time-dependent Davies form

\[ \mathcal{L}(t) = -i \left[ H_S(t) + H_{LS}(t), \bullet \right] + D(t) \quad (5a) \]

\[ D(t) = i \sum_{ij\omega} \gamma_{ij}(\omega) \left( A_j(\omega) \bullet A_i^\dagger(\omega) - \frac{1}{2} \left\{ A_i^\dagger(\omega)A_j(\omega), \bullet \right\} \right). \quad (5b) \]

Here \( H_{LS}(t) \) is the Lamb-shift term, \( \gamma_{ij}(\omega) \) is the Fourier transform of the bath correlation function

\[ G_{ij}(t_1,t_2) = \langle B_i(t_1)B_j(t_2) \rangle = G_{ij}(t_1-t_2), \quad (6) \]

\[ B_i(t) = e^{iH_S t}B_i e^{-iH_S t}, \quad (X) = \text{Tr}(\rho_B X) \] where \( \rho_B \) is the initial bath state, and \( \omega(t) \) are the Bohr frequencies of \( H_S(t) \) (to simplify notation we suppress their explicit time-dependence when convenient), i.e., the differences between the instantaneous eigenenergies of \( H_S(t) \).

The Lindblad operators \( A_i(\omega) \) that appear in the Davies generator (“dissipator”) \( D \) are defined by

\[ e^{i\tau H_S(t)}A_i e^{-i\tau H_S(t)} = \sum_{\omega} e^{-i\tau \omega} A_i(\omega), \quad (7) \]

i.e.,

\[ A_i(\omega) = \sum_{E_n-E_n=\omega} \Pi_n A_i \Pi_n, \quad (8) \]

where \( H_S = \sum_n E_n \Pi_n \) and \( \Pi_n \) are the eigenprojectors of \( H_S \).

The Lamb shift is given by

\[ H_{LS} = \sum_{ij\omega} S_{ij}(\omega) A_i^\dagger(\omega)A_j(\omega), \quad (9) \]

with

\[ S_{ij}(\omega) = \int_{-\infty}^{+\infty} d\omega' \gamma_{ij}(\omega') \mathcal{P} \left( \frac{1}{\omega - \omega'} \right), \quad (10) \]

where \( \mathcal{P} \) is the Cauchy principal value. It commutes with \( H_S \).

This master equation generates a completely positive map and describes a system with a slowly varying system Hamiltonian \( H_S(t) \), weakly interacting with an infinite bath. Moreover \( \mathcal{L}(t) \) has the property that it has the Gibbs state as a steady state, i.e., \( \mathcal{L}(t) \sigma(t) = 0 \) with \( \sigma(t) = e^{-\beta H_S(t)} / Z(t), Z(t) = \text{Tr} e^{-\beta H_S(t)} \).

In our simulations we assume longitudinal field coupling, i.e., the system operators in the system-bath interaction Hamiltonian are \( A_i = \sigma_i^z \) with \( \gamma_{ij}(\omega) = \delta_{ij} \gamma(\omega) \) (independent coupling) and the bath has an Ohmic spectral density, so that

\[ \gamma(\omega) = 2\pi \eta_0 g^2 \frac{\omega}{1 - e^{-\beta \omega}}. \quad (11) \]

\[ \text{C. A modified BCT} \]

Unfortunately, we cannot directly apply Prop. 1 to annealers with \( A(t_f) = 0 \) (such as the DW2KQ) because of the following result (proven in App. A):

**Proposition 3** Assume that \( \mathcal{L}(\tau) \) is the AME Liouvillian [Eq. (5a)] with the Hamiltonian in Eq. (3) and longitudinal system-bath coupling (i.e., that commutes with \( H_Z \)). Then, at the end of the anneal when \( A(t_f) = 0 \), the zero eigenvalue of \( \mathcal{L}(t_f) \) is at least \( d \)-fold degenerate, where \( d \) is the system’s Hilbert space dimension.

Since presumably, when \( A(t) \neq 0 \) for \( t < t_f \) the Lindbladian \( \mathcal{L}(t) \) has a unique steady state \( \sigma(t) \) which corresponds to its zero eigenvalue, Prop. 3 means that the Liouvillian gap closes at \( t_f \) and we cannot use Prop. 1.

We now generalize the BC result Prop. 1 to the situation where one enforces the BC conditions at the same point as where the Liouvillian gap closes, e.g., at \( t = t_f \).

**Proposition 4** Assume that \( \| \mathcal{L}(\tau) \|_{\infty,1} \) is summable in \( \tau \in [0,1] \). \( \mathcal{L}(\tau) \) is differentiable to order \( k + 2 \) in a neighborhood of \( \tau = 1 \), and generates a trace-preserving and hermiticity-preserving contractive semigroup, i.e.,

\[ \| e^{r \mathcal{L}(\tau)} \|_{\infty,1} \leq 1 \text{ for all } r \geq 0, \tau \in [0,1]. \]

Suppose \( \mathcal{L}(\tau) \) has a unique steady state \( \sigma(\tau) \) at each time \( \tau \in [0,1] \) except at the point \( \tau_0 = 1 \) where the Liouvillian gap closes as \( \Delta \sim v(\tau_0 - 1)^{\alpha} \) for a linear schedule (i.e., in the absence of a BC schedule). Let \( \rho(\tau) \) be the solution of Eq. (1) with initial condition \( \rho(0) = \sigma(0) \) and \( \sigma(1^-) := \lim_{\tau \to 1-} \sigma(\tau) \).

Assume we enforce boundary cancellation at the end, i.e., \( \mathcal{L}(\tau) \) has vanishing derivatives at \( \tau = 1 \) to order \( k \):

\[ \frac{\partial^l_\tau \mathcal{L}(\tau)}{l!} = 0 \forall j = 1,2,\ldots,k \text{ [note that here } \mathcal{L}(\tau) \text{ is written for simplicity as } \mathcal{L}(\sigma(\tau))]. \]

Then boundary cancellation is only mildly effective in that the adiabatic error satisfies

\[ \| \rho(1) - \sigma(1^-) \|_1 \leq C \frac{\eta^2}{\tau_f^2}, \quad (12a) \]

with \( \eta = \frac{k + 1}{k\alpha + \alpha + 1} < \frac{1}{\alpha} \). \quad (12b)

The proof is given in App. B, which also discusses the relaxed smoothness assumptions of Prop. 1 relative to Ref. [13]. The fact that the gap closes for the Lindbladian modeling the D-Wave annealers was already noted in Ref. [33], and indeed Eq. (12b) with \( k = 0 \) was derived there.

For completeness we comment also on the case where the Liouvillian gap closes in the middle of the evolution and one enforces boundary cancellation at the end. We have performed numerical simulations under these conditions (see App. C), all agreeing with a scaling of the adiabatic error as in Eq. (12a), but with an exponent given by \( \eta = k + 1 \), as in Eq. (2). Analytical results will be presented elsewhere.
Summarizing, if the gap closes at the end of the evolution, implementing BC is mildly advantageous in that the best one can do is to change the exponent $\eta$ from $1/(\alpha+1)$ to $1/\alpha$ for an infinitely flat schedule ($k \to \infty$). If, instead, the gap closes in the middle of the evolution, we infer from our numerical results that is has no effect on the scaling of the adiabatic error and implementing BC results in a favorable exponent $\eta = k+1$, just as in the case where the Liouvillian gap does not close at all.

D. Freezing

As the Hamiltonian starts to commute with the system-bath coupling operators near the end of the anneal, the dynamics of the system will begin to slow down. This gives another significant consideration regarding the successful implementation of BC schedules, which is that the system state appears to stop evolving at a point $t_0 < t_f$, a phenomenon of quantum annealing called freezing. To explain freezing, we write the master equation in the instantaneous energy eigenbasis $\{|n\rangle\}$, $n = 0, 1, \ldots, d-1$ (we omit the time dependence for clarity), i.e., the basis which diagonalizes the Hamiltonian $H_S(t)$. Consider the region of the anneal where $A(t)$ is small but non-zero, and also both $B(t) > 0$. It is then natural to assume that the spectrum of $H_S(t)$ is non-degenerate. The density matrix in the energy basis is $\rho(t) = \sum_{mn} \rho_{mn}|m\rangle\langle n|$. The diagonal elements of $\rho$ in this basis evolve according to the Pauli master equation [34] (for a modern derivation see, e.g., Ref. [35]). In particular, the ground state probability $\rho_{00} = \langle 0|\rho(t)|0 \rangle$ evolves according to

$$\dot{\rho}_{00} = \sum_n (\rho_{nn} W_{0n} - \rho_{00} W_{n0})$$

$$= \sum_{n>0} W_{0n} \left( \rho_{nn} - \rho_{00} e^{-\beta(E_n - E_0)} \right),$$

where the transition rate matrix $W$ has the following matrix elements:\footnote{The matrix $W$ satisfies detailed balance, i.e., $W_{mn} e^{-\beta E_m} = W_{nm} e^{-\beta E_n}$, which follows from an analogous equation for $\gamma(\omega)$, which in turns follows from the Kubo-Martin-Schwinger (KMS) conditions on the bath correlation function.}

$$W_{mn} = \sum_{i,j} \gamma_{ij}(E_n - E_m) \langle m|\sigma^z_j|n\rangle \langle n|\sigma^z_j|m\rangle.$$  

On the basis of Eqs. (13) and (14), freezing is seen to be a consequence of the following argument. Since we are in the adiabatic regime, only the lowest levels are populated, i.e., $\rho_{nn} \approx 0$ for $n > n_A$. Eq. (13) is then replaced by

$$\dot{\rho}_{00} \approx \sum_{n=1}^{n_A} W_{0n} \left( \rho_{nn} - \rho_{00} e^{-\beta(E_n - E_0)} \right)$$

$$- \rho_{00} \sum_{n=n_A+1}^{d} W_{0n} e^{-\beta(E_n - E_0)},$$

where $d$ is the system’s Hilbert space dimension. When $A(t) = 0$, the Hamiltonian is diagonal in the computational basis. Let the Hamming distance between the ground state $|0\rangle$ and the excited states $\{|n\rangle\}_{n=1}^{n_A}$ when $A(t) = 0$ be at least $q$ (note that $q \geq 1$). Using perturbation theory around $A(t) = 0$ it can be shown that $\langle 0|\sigma^z_j|n\rangle = O(A^q)$; see App. D. Hence, at the time $t_0$ at which $A$ is sufficiently small (but non-zero) one has $W_{0n} = O(A^{2q}) \approx 0$ for $n = 1, 2, \ldots, n_A$, and we can neglect the sum in line (15a); we determine $t_0$ below. As for the term in line (15b), the transition rates between higher excited states are typically smaller, and moreover, the terms are exponentially suppressed because of the larger gaps (i.e., for sufficiently small temperatures $E_n - E_0 \gg 1/\beta$, given that $n \geq n_A + 1 \geq 2$). As a consequence, Eq. (15) becomes $\dot{\rho}_{00} \approx 0$, i.e., the ground state population is effectively frozen for $t \geq t_0$.

The location of the freezing point for the $n$th excited state is determined by the point where the relaxation time for the $n$th excited state, given by $\tau^{(n)}_{\text{rel}} = W_{0n}^{-1}$, becomes longer than the anneal time. As long as just one of these relaxation times, with $n \in \{1, \ldots, n_A\}$ is too long, the system will not thermalize. For the system to freeze, i.e., $\rho_{00} \approx 0$, we need all transitions to cease. Hence, we define the freezing point as the solution $t_0$ of $\min_{n \in \{1, \ldots, n_A\}} \tau^{(n)}_{\text{rel}} = t_0$, or:\footnote{A shortcut to deriving Eq. (16) is to interpret $1/\tau^{(n)}_{\text{rel}} = \gamma(E_n - E_0)\langle 0|V|n\rangle^2 = W_{0n}$ as a statement of Fermi’s golden rule for the transition rate, with $V = \sum_\sigma \sigma^z_j \gamma$ playing the role of the perturbation, and $\gamma$ the density of states.}

$$t_0 := \min \{ t \in [0, t_f] \text{ s.t. } \max_{n \in \{1, \ldots, n_A\}} W_{0n}(t) = 1/t_f \}. \quad (16)$$

In terms of the control schedule $s(\tau)$, we have $s_0 := s(t_0)$, where $s_0 = t_0/t_f$. In practice, to determine the freezing point $s_0$ we implement a closely related procedure described in App. E.

Note that, as a consequence of the perturbative argument, (for sufficiently small $A$) the rates $W_{0n}$ are decreasing functions of $A$. So if $A(t)$ is decreasing in $t$, $W_{0n}$ can be considered zero for $t \geq t^*$. A similar argument works when substituting $0 \leftrightarrow m$, and one finds that the population in the $m$th excited state, $\rho_{mm}$, is frozen, albeit with possibly different freezing points than given by Eq. (16).

Another consequence of this argument is that the phenomenon of freezing should be more pronounced (i.e., occur for smaller $t_0$ and resulting in smaller $\rho_{00}$ for $t > t_0$).
for those problems where the ground state is separated by a large Hamming distance from the excited states, i.e., a larger tunneling barrier such as the T-gadget compared to the FM-gadget, discussed in Sec. III B below. These problems are the ones which are harder to simulate with standard classical simulated annealing with single spin flip moves [36] (cluster flip moves [37] would not necessarily be similarly affected).

Finally, note that a simulation of a master equation to which the considerations above apply, such as the AME (see Sec. II B), is also expected to exhibit freezing.

E. Boundary cancellation with a ramp at the end

Since the ground state population does not change past the freezing point, no change in the schedule would be effective if performed after freezing. In view of these considerations we perform BC before freezing sets in, which — for the standard control schedule $s(\tau) = \tau$ (see Fig. 1) — happens for $s \approx 0.55$, depending on the problem. Our strategy will be the following. First, evolve from $t = 0$ to $t_f$ with a BC schedule. To avoid freezing, we ensure that $A(t_f) \neq 0$ at the end of the BC schedule. Thus, $t_f$ does not correspond to the usual total anneal time for which $A(t_f) = 0$. Right after the BC schedule ends, we perform a linear ramp\(^5\) of duration $t_r = 1 \mu s$ until the schedules reach their final values (in particular $A(t_f + t_r) = 0$), after which the system is measured in the computational basis. The state after the entire evolution can be written as $E_{\text{ramp}}E_{\text{BC}}\rho(0)$, where $E_{\text{BC}}$ ($E_{\text{ramp}}$) is the evolution through the BC schedule (ramp). However, random local fields and coupler perturbations [integrated control errors (ICE)] result in a Hamiltonian that does not behave as intended in the ideal case [38], and these errors have been well documented in the D-Wave processors [31]. The effect of such random perturbations can be controlled with error suppression and correction [26–28], which will be employed below for the ferromagnetic chain gadget. Due to this ICE effect, the measured state is better represented by

$$\rho_{\text{final}} := E_J [E_{\text{ramp}}E_{\text{BC}}\rho(0)] ,$$

where we denoted by $E_J [\bullet]$ the average over the noise on the random couplings $J_{ij}$ and fields $h_i$.

Let us now define $\rho(t_f) := E_{\text{BC}}\rho(0)$, while $\sigma(t_f)$ is the Gibbs state at the end of the BC schedule (recall that $\rho(0) = \sigma(0)$), corresponding to the Hamiltonian in Eq. (3). The adiabatic theorem in its various forms, including with boundary cancellation, provides an upper bound on the pre-ramp distance $\|\delta\|_1$, where

$$\delta := \rho(t_f) - \sigma(t_f).$$

Defining

$$\sigma_{\text{final}} := E_J [E_{\text{ramp}}\sigma(t_f)],$$

$\delta$ can be related to $\rho_{\text{final}}$ and $\sigma_{\text{final}}$ via the following bound:

$$\|\rho_{\text{final}} - \sigma_{\text{final}}\|_1 = \|E_J [E_{\text{ramp}}\delta]\|_1 \leq E_J [\|E_{\text{ramp}}\delta\|_1] \leq E_J [\|\delta\|_1].$$  \hspace{1cm} (19a)

Here Eq. (19b) follows from Jensen’s inequality [39] and the fact that every norm is convex (implying $\|E_J [x]\|_1 \leq E_J [\|x\|_1]$), while Eq. (19c) follows because $E_{\text{ramp}}$ is a completely positive and trace-preserving (CPTP) map [40] (implying $\|E_{\text{ramp}}\delta\|_1 \leq \|\delta\|_1$). Note that $\rho_{\text{final}}$ is the empirically measured state, while $\sigma_{\text{final}}$ differs from the state that we sought to prepare $\sigma(t_f)$, by the presence of the extra operations $E_J [E_{\text{ramp}}\bullet]$.

The bound (19) implies a similar bound for the ground state probabilities. Let us define

$$P_{G\text{S}} := \text{Tr} \left[ |G\text{S}\rangle\langle G\text{S}| \rho_{\text{final}} \right], \hspace{1cm} (20a)$$

$$P_{G\text{S}}^* := \text{Tr} \left[ |G\text{S}\rangle\langle G\text{S}| \sigma_{\text{final}} \right], \hspace{1cm} (20b)$$

where $|G\text{S}\rangle$ is the ground state of the Hamiltonian at the end of the anneal, i.e., the ground state of $H_Z$. Since\(^6\)

$$\text{Tr} \left[ |G\text{S}\rangle\langle G\text{S}| x \right] \leq \| |G\text{S}\rangle\langle G\text{S}| \|_\infty \| x \|_1 = \| x \|_1 ,$$  \hspace{1cm} (21)

it follows that the adiabatic error defined as

$$D_{\text{GS}} := |P_{G\text{S}} - P_{G\text{S}}^*| ,$$

satisfies

$$D_{\text{GS}} \leq \|\rho_{\text{final}} - \sigma_{\text{final}}\|_1 .$$  \hspace{1cm} (23)

The adiabatic error quantifies the difference between the ground state overlaps of the experimentally measured state ($\rho_{\text{final}}$) and the Gibbs state ($\sigma_{\text{final}}$).

F. An adiabatic error bound that combines everything

Combining Eqs. (19) and (23) with Props. 1 and 4 and our numerical evidence, we obtain

$$D_{\text{GS}} \leq \|\rho_{\text{final}} - \sigma_{\text{final}}\|_1 \leq C \frac{t_f}{t_r} ,$$  \hspace{1cm} (24)

where $C$ is now the noise averaged constant ($E_J [\bullet]$) and $\eta$ depends on the physical assumptions. Namely, $\eta = k + 1$ if the Liouvillian either has a non-zero spectral gap throughout the anneal or if BC is enforced after the Liouvillian gap has already closed somewhere along the

\(^5\) The term “quench” is used in the D-Wave documentation instead of ramp [31].

\(^6\) Here $\|y\|_\infty$ is the operator norm of $y$, i.e., its maximum singular value, which is 1 for an orthogonal projection.
anneal, while \( \eta = (k + 1)/(k\alpha + \alpha + 1) \) if the Liouvillian gap closes at the same point at which BC is enforced (Prop. 4).

Of course we may reformulate Eq. (24) as a bound on \( P_{GS} \):

\[
P_{GS}^* - \frac{C}{t_f} \leq P_{GS} \leq P_{GS}^* + \frac{C}{t_f}. \tag{25}
\]

G. Anomalous heating

Our discussion so far has assumed that the effective temperature of the system remains constant as a function of both the anneal parameter \( \tau = t/t_f \) and the total anneal time \( t_f \). However, there is evidence to suggest that the latter is in fact not the case in the D-Wave devices, i.e., the temperature is \( t_f \)-dependent. The reason is an unintentional but omnipresent high-energy photon flux that enters the D-Wave chip from higher temperature stages through cryogenic filtering, which accumulates over long anneal times and manifests as an effectively higher on-chip temperature [41]. This anomalous heating phenomenon will hinder our ability to test the BCP, since it means that in fact \( P_{GS}^* \) is a function of \( t_f \), which complicates testing the BCT prediction as summarized by Eq. (25). Indeed, for a Gibbs state \( \sigma_{\text{final}} = e^{-\beta H_S}/Z \) (with \( \beta, H_S \), and the partition function \( Z \) all evaluated at \( t = t_f \)), expanding \( H_S \) in its eigenbasis as \( H_S = \sum_{i=0} e_i [i][i] \) we readily find

\[
P_{GS}^* = \langle GS | \sigma_{\text{final}} | GS \rangle = \frac{1}{1 + \sum_i e^{-\beta(t_f)\Delta_i}}, \tag{26}
\]

where \(|0\rangle\) is the ground state and \( \Delta_i := E_i - E_0 \).

Assuming for simplicity that \( \beta(t_f) = \beta_0 + \alpha/t_f \), i.e., a temperature that depends linearly on \( t_f \) with rate \( \alpha \geq 0 \), and that \( 1 \ll \beta(t_f)\Delta_i \ll \beta(t_f)\Delta_i \) for \( i \geq 2 \), we can write this as

\[
P_{GS}^* \approx 1 - e^{-\beta(0)\Delta_1} e^{-\alpha\Delta_1/t_f}, \tag{27}
\]

i.e., the algebraic scaling with \( t_f \) of Eq. (25) becomes obscured by an exponential scaling due to \( P_{GS}^* \).

However, in reality we do not know the functional form of \( \beta(t_f) \), and the assumption \( 1 \ll \beta(t_f)\Delta_1 \ll \beta(t_f)\Delta_i \) may not hold. Therefore, in the analysis of our experimental results in Sec. IV below, we instead use an ansatz of the form

\[
P_{GS} = P_{GS}^* + \frac{C'}{t_f}, \tag{28}
\]

where \( P_{GS}^* \) is a free fitting parameter representing an averaged value of the true (unknown) \( P_{GS}^*(t_f) \), \( C' \) becomes another fitting parameter which already accounts for noise averaging as explained in Sec. II \( \Gamma \), and \( \eta' \) plays the role of the effective scaling exponent, i.e., our proxy for \( \eta \) in Eq. (25).

![FIG. 2. An example of the D-Wave boundary cancellation protocol with the \( \beta_2 \) schedule ramped at \( a_{BC} = 0.50 \), with the constructed parametric schedule (top) and the corresponding physical schedules (bottom). For the schedule shown here we chose an anneal time of \( t_f = 100\mu s \). The ramp duration is \( t_r = 1\mu s \). Contrast with the native schedule of the DW-LN processor shown in Fig. 1. The most significant impact is on the \( B(s) \) schedule, which is not natively flat, in contrast to the \( A(s) \) schedule, which natively approaches 0 in a very flat manner.](https://example.com/fig2.png)

III. METHODS

A. Boundary Cancellation Protocol (BCP)

We performed most of our experiments with the D-Wave 2000Q low noise (DW-LN) processor accessed through D-Wave Leap. We also performed additional experiments with the D-Wave 2000Q processor at the NASA Quantum Artificial Intelligence Laboratory (DWA) as well as the D-Wave Advantage (DWA) processor through D-Wave Leap. In the processor specifications, the hardware-determined schedules \( A(s) \) and \( B(s) \) are parametrized by the user-defined control schedule \( s(\tau) \). In the standard case the control is linear, i.e., \( s(\tau) = \tau \) but in general \( s(\tau) \) can be programmed in a piecewise linear manner as a function of time. The processor permits a maximum of 12 points to specify the piecewise linear function \( s(\tau) \). We take advantage of this enhanced capacity to approximate a BC schedule. The allowed range of programmable anneal times is \( t_f \in [1, 2000]\mu s \).

Even though \( A(s) \) and \( B(s) \) themselves need not satisfy the vanishing derivative requirement of the BCP, it follows from the chain rule that \( A(s(t)) \) and \( B(s(t)) \) do, as long as the control schedule \( s(t) \) satisfies this requirement and \( A(s) \) and \( B(s) \) are differentiable to the same
order as \( s(t) \). To be concrete, we take

\[
s(t) = \beta_k(t) , \quad \tau = t/t_f ,
\]

where

\[
\beta_k(t) := \frac{B_x(1, k + 1)}{B(1, k + 1)} = 1 - (1 - \tau)^{k+1}
\]

is the regularized incomplete beta function of order \( k \) \cite{13, 42}, with

\[
B_x(a, b) := \int_0^x y^{a-1}(1-y)^{b-1} dy
\]

being the incomplete beta function. As is apparent from Eq. (30b), the function \( \beta_k(t) \) has exactly \( k \) vanishing \( \tau \)-derivatives at \( \tau = 1 \), as required by the BCT. We refer to this class of control schedules simply as *beta schedules*.

As we discussed above, in contrast to the theoretical setup of Ref. \cite{13}, due to freezing there is no practical advantage to flattening the schedule as \( s \) approaches 1. The effectiveness of the BCP is most apparent when the flattening of the schedule occurs after the avoided level crossing of the anneal (at \( s = s^* \), but before the dynamics freeze (at \( s = s_0 \)). Thus, rather than using the standard schedules \( \{A(s), B(s)\} \) with linear \( s(t) \) that freeze out around \( s \approx 0.5 \), the beta schedule needs to be adjusted so that it terminates at an appropriate point \( s_{BC} \) corresponding to \( t = t_f \) and satisfying \( 0 < s^* < s_{BC} < s_0 < 1 \), rather than always at \( s = 1 \), when the open system dynamics are already frozen out. The anneal is then ramped to the final values \( A(t_f + t_r) = 0 \) and \( B(t_f + t_r) \). The precise construction of the control schedule is given in App. F.

Note that the piecewise linear approximation of Eq. (29) cannot exactly satisfy the requirement of vanishing derivatives at the end. However, Ref. \cite{13, Prop. 3} shows that (at least for the case \( k = 1 \)) if one tries to enforce a vanishing derivative but achieves only approximate vanishing, the overall adiabatic error is diminished nonetheless. A reasonable expectation is that the scaling will be milder than the theoretical ideal of \( 1/t_f^{k+1} \) for Liouvillians with a spectral gap (Prop. 1), i.e., will be of the form \( 1/t_f^\eta \) with \( \eta < k + 1 \) for schedules that only approximate the desired \( \beta_k \) schedule. An example of this schedule construction is shown in Fig. 2.

B. 8 Qubit Gadgets

We briefly describe the two gadgets utilized in our experiments, which are both illustrated in Fig. 3 along with their couplings and bias values in the Hamiltonian \( H_Z \) in Eq. (4). The first gadget we use is an 8-qubit ferromagnetic (FM) chain, depicted in Fig. 3 using its embedding in a single unit cell of the Chimera graph in the D-Wave 2000Q processor, which is a bipartite \( K_{4,4} \) graph. The local fields in this gadget were adjusted to break the most significant degeneracies in the ground subspace and first excited subspace that are present in a simple ferromagnetic chain without any fields. This gadget thus enforces a large Hamming distance between its ground state and any first exited state without introducing additional two-body interactions. The remaining degeneracy in the first excited state is broken by crosstalk interactions in the device, which we model and include in simulations of the Hamiltonian (see App. G). In addition, it is simple to tile and embed the QAC encoding of the FM-gadget (see below) on the Chimera graph.

The second gadget we utilize is the one introduced in Ref. \cite{25}. This gadget was thoroughly analyzed there and established to pose a small-gap tunneling barrier between the ground state and the first excited state, so we refer to it as the “T-gadget”. For finer control of the energy gap, the T-gadget is scaled to \( \gamma = 0.5 \) [recall Eq. (3)]. The ground state of both gadgets is the all-0 state, denoted by \( |0\rangle \). Both gadgets also share the all-1 state \( |1\rangle \) as their first excited state.

The spectral gaps of the gadgets are shown in Fig. 4. The most crucial distinction between them is that the FM-gadget has a minimum gap \( \Delta_{01}(s^*) \) to the first excited state that is larger than the dilution fridge operating temperature, while the opposite holds for the T-gadget. Figure 4 suggests that implementing the BCP with the termination point \( s_{BC} \) somewhere in the range of \([0.45, 0.50]\) will be the most effective for both gadgets, since \( s^* \approx 0.43 \) and \( s_0 \approx 0.5 \). Before this range, the BCP spends more time slowing down before the minimum gap, while after this range it goes through the minimum gap too quickly and only slows down going into the freezing phase.
**C. Encoding with Quantum Annealing Correction**

We supplement the BCP with quantum annealing correction (QAC) [26] and compare the scaling improvement versus QAC with a simple linear anneal. The QAC encoding of the computational Hamiltonian in (3) has the form

\[ H_{QAC} = \mathbf{P}_Z + \lambda H_P \]  

(32)

where \( \mathbf{P}_Z \) is an \( m \)-qubit repetition of the problem Hamiltonian \( H_Z \) wherein each \( \sigma^z_i \) and \( \sigma^z_i \sigma^z_j \) term is replaced by its encoded counterpart \( \bar{\sigma}^z_i = \sum_{i=1}^{d} \sigma^z_{i_d} \) and \( \bar{\sigma}^z_i \bar{\sigma}^z_j = \sum_{i=1}^{d} \sigma^z_{i_d} \sigma^z_{j_d} \), respectively, and the penalty Hamiltonian imposes an energy cost by ferromagnetic coupling to an ancilla penalty-qubit,

\[ H_P = -\sum_{i} (\sigma^z_{i_1} + \ldots + \sigma^z_{i_m}) \sigma^z_{i_p}. \]  

(33)

Here the index \( i \) is over the logical qubits, \( i_1 \ldots i_m \) are the physical qubits corresponding to the logical qubit \( i \), and \( i_p \) is the ancilla for qubit \( i \). A majority vote over the results of the measurement on the computational basis of the (odd) \( m \) physical qubits decides the spin value of logical qubit \( i \). This strategy of combining an energy penalty against errors anticommuting with \( H_P \) (the sum of the stabilizer generators of the bit-flip code) [15, 43–48], along with a bit-flip repetition code, has been demonstrated numerous times to be successful at enhancing the performance of quantum annealing, despite not being able to protect against pure dephasing errors (which commute with \( H_P \)) [29, 50–53]. We only consider \( m = 3 \) QAC encoding, which is naturally embedded onto the Chimera graph of all the DW2KQ processors. As the encoded FM-gadget can be naturally embedded with no additional minor embedding, we restrict our attention to it, and do not consider the T-gadget under QAC, which would require minor embedding. For the same reason we do not consider the more powerful nested QAC strategy [30, 54–56], which is left for future work. We estimate the optimal penalty strength \( \lambda \) for QAC by finding the value that maximizes the ground state probability of a simple linear anneal.

**IV. RESULTS**

In this section we report on our results for \( P_{GS} \), arising from both runs on the DW-LN and numerical simulations. As explained in Sec. II G, we use Eq. (28) to fit our empirical results. Our data collection procedure is described in App. H.

**A. Linear control schedule \((k = 0)\)**

We first consider the FM and T-gadgets under the native \( A(s) \) and \( B(s) \) schedules with linear control \( s(\tau) = \tau \)

![Graph of spectral gaps for FM-gadget and T-gadget](image-url)
FIG. 5. DW-LN and AME results for $P_{GS}$ with a simple linear control schedule (without a ramp) for the FM-gadget [(a) and (b)] and the T-gadget [(c) and (d)]. A non-linear fit to (28) is applied on the data and yields estimates for $\eta'$ as described in App. H. The initial constrained nonlinear fit to (28) yields values for $\eta'$ and $P^*_{GS}$ that may have high variance. A linear refit on $D_{GS}$ is done assuming $P_{GS} = P^*_{GS}$ to refine $\eta'$ with a smaller error. The values of $\eta'$ and their associated uncertainties we find in this manner are reported in the legends. Error bars on the measured ground state probabilities represent 95% confidence intervals. AME simulation parameters for both gadgets: $\eta_0 g^2 = 5.0 \times 10^{-4}$, $T = 13.5$ mK (see Sec. II B) and crosstalk strength $\chi = 0.02$ (see App. G).

and no ramp ($s_{BC} = 1$), i.e., a standard quantum annealing protocol. This means that the anneal is stopped deep inside the frozen regime, i.e., freezing is expected to have a noticeable effect.

In Fig. 5 we plot, for both the FM-gadget and the T-gadget, the measured ground state probability $P_{GS}$ for the DW-LN, along with the same quantity obtained through AME simulations. We found $P^*_{GS}$ for DW-LN by a non-linear fit to the function $C'/t_f^\eta + P_{GS}$, where $C'$ and $\eta'$ are free parameters (left column of Fig. 5). In the AME case only the last four data points were used in the fit (since otherwise $t_f$ is too small, as evidenced by the curvature observed for small $t_f$ in the right column), while in the DW-LN case we dropped the first data point for both gadgets for the same reason. In the T-gadget case we also dropped the last two data points, as they are likely corrupted by anomalous heating (see Sec. II G). This leads to the values $P^*_{GS} = 0.97$ and $P^*_{GS} = 0.78$ for the FM-gadget [Fig. 5(a)] and T-gadget [Fig. 5(c)], respectively. For the AME, we took $P_{GS}$ to be 1, as this is consistent with the BCT for a large gap and the non-linear fits did not find a smaller value. We then refined the value of $\eta'$ by a linear fit to $\log_{10}(P_{GS} - P^*_{GS})$ as a function of $\log_{10} t_f$, since we expect $P_{GS} - P^*_{GS}$ to scale as $C/t_f^\eta$ by Eq. (25). Using Eqs. (24) and (25), we interpret $\log_{10}(P_{GS} - P^*_{GS})$ as a proxy for $\log_{10} D_{GS}$, plotted in Fig. 5(b) and Fig. 5(d) for the FM-gadget and T-gadget respectively.

Several observations are in order from Fig. 5.
- The $P_{GS}$ values at the largest $t_f$ are larger for the
FIG. 6. DW-LN adiabatic distance scaling fits for the FM-gadget with BCP at (a) $s_{BC} = 0.45$, (b) $s_{BC} = 0.47$, and (c) $s_{BC} = 0.50$ past the minimum gap location of $s^* \approx 0.43$ but before freeze-out at $s_0 \approx 0.51$. (d) and (e): Adiabatic distance scaling of QAC encoded FM-gadget with BCP at $s_{BC} = 0.45$ and $s_{BC} = 0.50$ (see Sec. IV D). Symbols represent the experimentally measured data points, solid lines are fits to $1/t_f'$. The $\eta'$ values reported in the legends an well as other parameters are obtained from nonlinear fits to the measured data as explained in App. H. Error bars represent 95% confidence intervals.

- The FM gadget AME simulation has a larger $\eta'$ (0.43) than the T-gadget (0.21). This signifies a faster convergence to the adiabatic limit of $P_{GS} = 1$. This too is consistent with the larger FM-gadget gap and the larger tunneling barrier of the T-gadget. However, the opposite holds for the DW-LN case. This is consistent with the larger FM-gadget gap (see Fig. 4) and the larger tunneling barrier of the T-gadget (quantified by the ground state being separated by a larger Hamming distance from the excited states), which causes an earlier onset of freezing and hence “locks in” the smaller ground state population of the T-gadget.

- The AME data for the T-gadget is in closer agreement with the power-law bound of the adiabatic theorem than the FM-gadget. This is evidenced by the good agreement with the linear fit in Fig. 5(d), but the persistence of some curvature at large $t_f$ in Fig. 5(b). We may speculate that the FM-gadget will transition to a power-law behavior at even larger $t_f$ than we were able to simulate, though one should keep in mind that in any case, the power-law bound of the adiabatic theorem is not tight.

- For the DW-LN at large $t_f$, we observe an appreciable reduction in $P_{GS}$ [Fig. 5(c)], and a concomitant increase in $D_{GS}$ [Fig. 5(d)] in the T-gadget case, and a small positive curvature in the FM-gadget case. This suggests that an anomalous excitation (heating) mechanism is at work in the D-Wave case. This anomaly was discussed in Sec. II G. The anomalous heating phenomenon hinders our ability to test the BCP, since it means that the effective temperature is itself $t_f$-dependent and increasing. As explained above, this is why we treated $P^*_n$ as a fitting parameter.

- The DW-LN FM-gadget scaling [Fig. 5(b)] is much closer to the expected $D_{GS} \sim 1/t_f'$ than the T-gadget [Fig. 5(d)], which is obscured by anomalous heating. It is unclear why this phenomenon should more severely impact the T-gadget.

We report the values of $\eta'$ extracted from our fits in Fig. 5. To understand what value of $\eta'$ to expect, note that according to our model of the D-Wave annealer and in view of Proposition 3, the Liouvillian gap $\Delta(\tau)$ must close at the end of the anneal. Recall that $\Delta \sim v(\tau - \tau_0)^\alpha$ with $\tau_0 = 1$ (or $t = t_f$) being the end of the anneal.
We do not know the associated value of the exponent \( \alpha \); however, given that \( \mathcal{L} \) is a normal operator, if the Liouvillian depends smoothly on the schedule in a small interval around \( t_f \), a standard result [57] implies that \( \alpha \) must be an integer. On the other hand, \( \alpha = 1 \) is incompatible with the fact that the lowest eigenvalue of \( \mathcal{L} \) is zero, and that the gap must close smoothly. The value observed in the single qubit case is \( \alpha = 2 \) [33]. According to Prop. 4 and more specifically Eq. (12b) with \( k = 0 \) (due to the linear schedule), this would correspond to an exponent \( \eta = 1/3 \), which is also the scaling observed in the single qubit case [33].

However, as seen in Fig. 5(d), the AME T-gadget case has \( \eta' = 0.21(1) \), which – when equated with \( 1/(\alpha + 1) \) according to Eq. (12b) with \( k = 0 \) – yields \( \alpha = 4 \). This larger value of \( \alpha \) (and corresponding smaller value of \( \eta' \)) than was found in the single qubit case is not unexpected given the very different spectrum associated with the T-gadget.

As remarked above, in contrast to the T-gadget, the AME value of \( \eta' \) for the FM-gadget is at least 0.43, and is likely to be slightly higher since the largest \( t_f \) value for which we were able to run AME simulations appears to be too small for convergence. From Eq. (24) we expect \( D_{\text{GS}} \) to scale with an exponent \( \eta' \geq \eta \), where for \( k = 0 \) we have \( 1/(\alpha + 1) \leq \eta \leq 1/\alpha \) [Eq. (12b)]. Thus \( \eta' \approx 0.43 \) is consistent with these expectations for any integer \( \alpha \geq 2 \).

The DW-LN scaling is consistent with the expectation \( \eta' \geq \eta \) for the T-gadget, for which \( \eta' = 0.37 \). However, anomalous heating makes it difficult to be confident in the extracted \( \eta' \) value for this gadget. For the FM-gadget we find a good fit to \( D_{\text{GS}} \approx 1/t_f^{\eta'_{\text{F}}} \) with \( \eta' = 0.19 \), which could be explained by \( \alpha = 4 \), as in the AME T-gadget case. However, the small positive curvature in the DW-LN FM-gadget data at large \( t_f \) suggests that the true scaling may be obscured by anomalous heating here as well.

Since anomalous heating plays such a significant role in the T-gadget behavior at large \( t_f \), we abandoned it for our experiments involving BCP schedules with \( k \geq 1 \). We focus for the rest of this work on the FM-gadget.

### B. BCP schedules with \( k \geq 1 \)

We now turn to actual boundary cancellation schedules. In light of our discussion in Sec. II, a successful BCP must occur before the closing of the Liouvillian gap and freeze-out, while at the same time the Hamiltonian gap remains large enough to promote relaxation to the ground state. We thus apply the BCP at a point \( s^* < s_{\text{BC}} < s_0 \), such that \( A(s_{\text{BC}}) \neq 0 \). The BC schedule is followed by a ramp of 1 \( \mu \)s to the final values of \( A \) and \( B \).

Following the same nonlinear fit procedure as in Sec. IV A, we plot the DW-LN \( P_{\text{GS}} \) results for the FM-gadget in Fig. 6. More specifically, Fig. 6(a)-6(c) shows the results for \( s_{\text{BC}} \in \{0.45, 0.47, 0.50\} \), while Fig. 6(d)-6(e) shows analogous plots after applying QAC, which we discuss in Sec. IV D.

The first observation that emerges from Fig. 6 is that the \( \beta_0 \) schedule results in a significantly larger (i.e., better) scaling exponent than the \( k = 0 \) schedule of Sec. IV A, for which we found \( \eta' = 0.19 \). The difference between the two is that for the case shown in Fig. 6 the schedule includes a ramp at \( s = s_{\text{BC}} \). As the ramp point grows from 0.45 to 0.47 to 0.50, the value of \( \eta' \) correspondingly declines from 0.47 to 0.23 to 0.21. This confirms that stopping the BC schedule before freeze-out (at \( s_0 \approx 0.52 \)) has a significant positive impact on the performance of the BCP.

If the Liouvillian is gapped at the endpoint of the BC schedule, the theoretical expectation for the scaling of the adiabatic error is \( \eta = k + 1 \) for the \( \beta_2 \) schedule (recall Sec. II F). Instead, Figure 6 shows that the measured exponents \( \eta' \) exhibit a strong dependence on the value of \( s_{\text{BC}} \). The most favorable case is \( s_{\text{BC}} = 0.45 \) [Fig. 6(a)], for which the exponents \( \eta' \) are \{0.47 (\( k = 0 \)), 0.97 (\( k = 1 \)), 1.1 (\( k = 2 \))\}, i.e., roughly half that value for \( k = 1 \). The \( k = 2 \) exponent is statistically indistinguishable from that for \( k = 1 \). It is likely that this is due to the fact that the relatively small number of schedule interpolation points does not allow us to approximate the ideal \( \beta_2 \) schedule sufficiently well (see Sec. IV C below). Other sources of error, such as ICE and anomalous heating, may also play a role.

While this level of agreement with the BCT for the gapped Liouvillian case is far from impressive, the dependence of \( \eta' \) on \( k \) is less consistent with the BCT prediction for when the Liouvillian gap closes at \( s_{\text{BC}} \); in this case Prop. 4 for \( \alpha = 1 \) yields \( \eta = \{0.5, 0.66, 0.75\} \) for \( k = 0, 1, 2 \) with an asymptotic value of 1 for a flat schedule \( (k \to \infty) \), while using \( \alpha = 2 \) yields \( \eta = \{0.33, 0.4, 0.42\} \) for \( k = 0, 1, 2 \) and an asymptotic value of 0.5. It is thus reasonable to conclude that the predictions of Prop. 4 are less consistent with the data than those of Prop. 1, and in this sense we can interpret Fig. 6 as evidence that the Liouvillian describing the experiment does not become gapless at the end of the BC schedule.

The improvement in the scaling of the adiabatic error provided by the BC schedule is most prominent for \( s_{\text{BC}} = 0.45 \). As we increase \( s_{\text{BC}} \) to 0.47 and 0.5 the improvement gradually disappears. We interpret this as being due to the fact that the freezing point \( s_0 \approx 0.52 \) is being approached, and so any change made to the schedule can be expected to be effective only on much longer timescales \( t_f \).

We note that in contrast to the scaling behavior, which favors the smallest \( s_{\text{BC}} \) value, the asymptotic value of

---

7 One might be tempted to associate the flatness of the native \( A(s) \) schedule (see Fig. 1) with the enforcement of a BC-type schedule with a \( k \gg 0 \) value, which would in turn correspond to \( \eta = 1/\alpha \) according to Eq. (12b). However, the native \( B(s) \) schedule is not flat at all, and it dominates the end of the anneal.
FIG. 7. Adiabatic error of AME simulations of the D-Wave programmed BCP and high-precision variants (denoted HP in the legend), for the FM-gadget. Here $s_{\text{BC}} = 0.50$, using the schedule of the DW-LN processor and a crosstalk parameter of 0.02 (see App. G). The adiabatic error is evaluated relative to the Gibbs state ground state probability $\bar{P}_{\text{GS}} = 0.9690$ at $s = 0.50$.

$P_{\text{GS}}$ is highest for the largest $s_{\text{BC}}$ value. This is consistent with the the larger gap at $s_{\text{BC}} = 0.50$ [see Fig. 4(a)]. This effect is in fact more significant than the improvement due to the BCP.

For additional experiments performed using the the NASA Quantum Artificial Intelligence Laboratory as well as the D-Wave Advantage processor, see App. I. These results are in qualitative agreement with those shown in Fig. 6, and do not modify our conclusions in any significant way.

C. High precision BCP

To test the effect of the relatively low precision of the implementation of the BCP schedules (using only 12 interpolation points), we simulated the AME for an additional set of higher-precision constructions of the BC schedule, with 8 logarithmic partitions up to $s_{\text{BC}}$, rather than the 4 allotted for the D-Wave schedule (see the construction of App. F). The scaling of the adiabatic error for the FM-gadget is shown in Fig. 7. Since the adiabatic error still appears to have some concavity, the scaling exponents should be interpreted as lower bounds for the asymptotic scaling. We see that the high precision schedule brings the scaling of the adiabatic error significantly closer to the value predicted by the Prop. 1 version of the BCT within anneal times that are currently accessible to D-Wave devices: we find $\eta' = 1.66, 2.69$ for the higher precision schedules, compared to the theoretically expected values of $k + 1 = 2, 3$, respectively. The lower precision schedules result in $\eta' = 1.49, 1.74$, respectively. While the values of $\eta'$ for the high precision schedules could be refined by simulations at longer anneal times, we do not include them here to avoid floating point error effects. In any case, it would be much more difficult to experimentally verify the adiabatic error to the precision required at longer annealing times. As such, Fig. 7 is likely to be the most practical precision test of boundary cancellation to date on open systems larger than a single qubit.

We also simulated the AME with a precision of 100 evenly spaced points in the control schedule, to approximate reaching the limit of the analytically exact beta schedule. The results are indiscernible from the high-precision piecewise schedule shown in Fig. 7, indicating that the latter already implements the beta schedule at the smallest precision that is meaningful within the simulated annealing times.

We conclude that the relatively poor agreement with the Prop. 1 BCT predictions we observed in Figure 6(a)-(c), is likely to be improved in future higher precision implementations of the beta schedule.

D. QAC-Encoded FM Gadget

The scaling of the adiabatic error of the QAC-encoded FM-gadget for various penalty strengths $\lambda$ [re-
call Eq. (32)] is shown in Fig. 8, for the case of the standard linear control schedule with no ramp. We observe that from the set of values tried, the optimal penalty strength is \( \lambda = 0.07 \). Using this optimal penalty value, the ground state probabilities of BCP with QAC are shown in Fig. 6(d)-6(e). The QAC-protected \( P_{GS} \) results for both \( s_{BC} \) values shown are significantly higher than their unprotected (\( \lambda = 0 \), NP) counterparts in Fig. 6(a)-6(c), consistent with many prior results on the improved performance offered by QAC.

However, we are primarily interested here in QAC’s effect on the BCP. In this regard, at \( s_{BC} = 0.45 \), there is a notable distinction between different \( k \) values, so that QAC amplifies the effect of the BC schedule.

At \( s_{BC} = 0.50 \), QAC-protected BCP still improves over the linear anneal (\( k = 0 \)) by \( P_{GS} \) value, but the improvement is less distinct for different \( k \) values. Moreover, \( \eta' \) is smaller (at most 0.61) than that for the optimal-\( \lambda \) linear QAC anneal in Fig. 8 (for which \( \eta' = 0.69 \)), signaling that freezing has already mostly occurred before \( s = 0.50 \), and that the optimal \( \lambda \) value is \( s_{BC} \)-dependent.

Our results demonstrate that the combination of BCP and QAC as error suppression methods is more powerful in terms of the scaling of the adiabatic distance than either method alone: while \( k = 2 \) BCP has \( \eta' = 1.1 \) [Fig. 6(a)] and optimal-\( \lambda \) linear (\( k = 0 \)) QAC has \( \eta' = 0.69 \) (Fig. 8), their combination yields \( \eta' = 1.3 \) [Fig. 6(d)] and the largest \( P_{GS} \) of any other schedule for \( t_f = 12 \mu s \).

E. Comparison of the BCP to the Pause-Ramp protocol

A different protocol that attempts to exploit slowing down the anneal is the pausing protocol, which interrupts an ordinary linear anneal by a single pause [58–62]. Pausing directly uses thermal relaxation at a single point in the anneal to try to increase the ground state probability. As in the BCP, this point should also be after an avoided level crossing, but before the open system dynamics freeze [59].

We compare the BCP to a variant of the usual pausing schedule [58] referred to here as the pause-ramp (PR) schedule. While the paused protocol is typically constructed by interrupting a linear anneal with a pause (but no ramp), PR is constructed by making the first linear segment last for \( t_0 \), pausing for \( t_p \), and ramping to the end. As with the BCP, we use a ramp that is 1 \( \mu \)s long in all cases. PR is better suited for comparison with BCP than the usual pausing schedule since they differ only in their equilibrium state preparation and not in their behavior during the frozen phase.

As our metric for comparison we use not the ground state probability but rather the number of tries-to-solution (NTS) at a fixed anneal time (with 90% confidence):

\[
\text{NTS}_{90\%}(P_{GS}) = \frac{\log(1 - 0.90)}{\log(1 - P_{GS})}
\]

This metric is simply another way to study the ground state probability, interpreted as the expected number of tries needed to find the ground state at least once with 90% confidence, at the given anneal time. We use this rather than the standard time-to-solution (TTS) metric, since the latter requires identifying the optimal anneal time [25, 63, 64]. However, neither the FM-gadget nor the T-gadget exhibited an optimal anneal time in our experiments (not shown).

Figure 9 shows the NTS results for the T-gadget at various values of \( s_{BC} \). It compares the standard schedule (no BCP or pausing), the BCP protocol at \( k = 0 \) and \( k = 1 \), and the PR schedule with different initial anneal and pause times. Starting at \( s_{BC} = 0.45 \) (top left) the gap is small and thermalization is fast. BCP exhibits two distinct optimal anneal times (minima) for \( k = 0 \) and \( k = 1 \), both of which result in slightly smaller NTS than pausing. Also, the \( k = 1 \) case is optimal at smaller \( t_a \) than \( k = 0 \), which can be interpreted as as advantage of the higher order protocol. This trend persists for most of the \( s_{BC} \) values shown. Right at \( s_{BC} = 0.46 \), pausing achieves its optimal NTS curve, with a slight advantage over BCP, and the BCP of each order \( k \) is very distinct. However, the advantage of pausing is only present at \( s_{BC} = 0.46 \), and disappears already at \( s_{BC} = 0.47 \), showing that pausing performance is highly sensitive to the pause point. Since the BCP slowdown is smooth, the influence of the optimal thermalization at \( s_{BC} = 0.46 \) remains present, although the distinction between BCP orders diminishes as \( s_{BC} \) grows. Well before \( s_{BC} \) approaches the freeze-out point (\( s_0 \approx 0.51 \)), pausing becomes detrimental for the T-gadget due to excitations (witnessed by the increase in the NTS metric), while BCP mitigates the excitations until the anneal time becomes too long to avoid them.

These results paint a picture of the two protocols as follows. The relaxation induced by the BCP depends on the properties of the Liouvillian over a neighborhood \( (s_{BC} - \epsilon, s_{BC}) \). If this neighborhood overlaps with the neighborhood of an avoided level crossing, then the BCP has an opportunity for an advantage, as the speed at which the crossing is traversed is minimized and detrimental excitations to the excited state are reduced. In contrast, PR schedules must try to discontinuously stop the anneal as close as possible to the crossing, and the time required by a paused schedule could be increased due to the longer pause time needed to recover the ground state population, unless the precise optimal pause point is quickly found.

This suggests an important practical point: optimization of the BCP can be simpler than that of PR schedules, by the need to optimize just one discrete parameter \( k \) and one continuous parameter \( s_{BC} \). Moreover, \( s_{BC} \) need not be exactly tuned, since even an approximate value can allow the BCP to closely approach its optimal performance. On the other hand, pausing requires \( s_{BC} \) to be optimized very precisely, as well as the anneal and pause times \( t_0 \) and \( t_p \).
FIG. 9. Number of tries to solution for the T-gadget, comparing the pausing protocols (empty symbols) and $\beta_0, \beta_1$ BCP (all with a ramp at $s_{BC}$), along with the linear anneal without a ramp (denoted “Lin.”; the same data for all panels), for reference. Results shown are for DW-LN, as a function of the total anneal time of each protocol. Different panels have different values of $s_{BC}$, from 0.45 to 0.55. The pausing protocol is shown for different initial anneal times $t_0$ of 1, 5, and 20 µs (empty symbols), followed by a pause of varying length $t_p$ and a 1 µs long ramp (so that the paused portion of the schedule is comparable to BCP). For BCP (filled symbols), $t_a = t_f + 1$ µs. For pausing, $t_a = t_0 + t_p + 1$ µs. Both schedules are plotted against $t_a$ rather than their natural parameters ($t_f$ and $t_p$ respectively) for direct comparison.

V. DISCUSSION AND OUTLOOK

We have extended the boundary cancellation theorem for open systems to the case where the Liouvillian gap vanishes at the end of the anneal, and derived the asymptotic scaling of the adiabatic error with the anneal time $t_f$. Armed with the corresponding theoretical expectation of the scaling of the adiabatic error for the gapped and gapless cases, we set out to test the scaling predictions and to improve the success probability of quantum annealing hardware by implementing boundary cancelling schedules. The specific functional form of these schedules induces a smooth slowdown in accordance with the boundary cancellation theorem. We experimentally tested boundary cancellation protocols for open systems and evaluated their performance and error-suppression characteristics on specifically designed 8-qubit gadgets embedded on the DW-LN annealer.

While a quantitative agreement with the theoretical predictions was not observed, we did demonstrate that as long as the protocol terminates before the onset of freezing, it can increase the ground state population in the examples studied here beyond what is achievable with simple linear anneals, and it does so with shorter anneal times. These results are in qualitative agreement with the theoretical scaling predictions of the BCT.

In conjunction with quantum annealing correction (QAC), the boundary cancellation protocol is also capable of improved adiabatic error scaling over what would be achieved with either method alone. While this does not immediately translate to a ground state solution speedup within the annealing problems studied here, we have shown that BCP-QAC is a novel error suppression strategy successfully combining two complimentary methods: the suppression of environmentally induced logical errors and the promotion of relaxation via boundary cancellation.

In contrasting the BCP with the pause-ramp protocol, we found that the BCP is significantly less sensitive to the location of the ramp point, and achieves better performance except at the exact ramp point where the pause-ramp protocol is optimal.

With the small system size of 8 qubits used in this work, it was possible to collect a large number of annealing samples as well as validate the protocol behavior against the energy spectrum and open system simulations. Future work will assess the protocol for larger system sizes and the impact it has on the scaling of time-to-solution as a function of problem size. The largest expected improvement in the protocol’s performance, based on our simulations, will arise from an increase in the number and resolution of interpolation points along the annealing schedule, which will allow a more faithful experimental implementation of the ideally smooth annealing schedules demanded by the theoretical protocol.
ACKNOWLEDGMENTS

This research is based upon work (partially) supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA) and the Defense Advanced Research Projects Agency (DARPA), via the U.S. Army Research Office contract W911NF-17-C-0050. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, DARPA, ARO, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The authors acknowledge the Center for Advanced Research Computing (CARC) at the University of Southern California for providing computing resources that have contributed to the research results reported within this publication. URL: https://carc.usc.edu.

Appendix A: Proof of degeneracy when \( A(t) = 0 \)

We assume that \( A_i = \sigma_i^z \). When \( A(t) = 0 \), \( H_S \propto H_Z \) and commutes with \( A_i \), so that Eq. (7) yields \( A_i(\omega) = \sigma_i^z \delta_{\omega,0} \). Therefore, the dissipator in Eq. (5a) becomes

\[
D|_{A(t) = 0} = \sum_{ij} g_{ij}(0) \left( \sigma_j^z \sigma_i^z \cdot \frac{1}{2} \left\{ \sigma_i^z \sigma_j^z, \cdot \right\} \right).
\]

(A1)

Now let \( |n\rangle \) be an eigenstate of \( H_Z \), where \( n = 1, \ldots, d \) and \( d \) is the system’s Hilbert space dimension. Then \( |n\rangle \langle n| \) commutes with \( \sigma_i^z \), so that \( D \langle n|\langle n| \rangle = 0 \). At the same time, \( A_i(\omega) = \sigma_i^z \delta_{\omega,0} \) implies that also the Lamb shift is diagonal in the \( \sigma^z \) basis. Together with the fact that \( |n\rangle \langle n| \) also commutes with the system Hamiltonian this implies \( L\langle n|\langle n| \rangle = 0 \), which means that the kernel of \( L \) is at least \( d \)-dimensional at \( A(t) = 0 \), i.e., that the zero eigenvalue of \( L \) is at least \( d \)-fold degenerate.

Appendix B: Proof of Proposition 4

Here we prove Proposition 4, which concerns the case where the gap closes at the end of the anneal and we enforce boundary cancellation at this point.

We denote by \( L(s) \) the Liouvillian without the BC schedule (i.e., with a linear schedule around \( s = 1 \)) while \( L(s(\tau)) \) [or \( L(\tau) \) for simplicity] indicates the Liouvillian with the BC schedule. We will use the series developed in [65] and refined in [13] together with ideas from [33, App. J]. The assumption that \( \int_0^1 ds \left\| L(s) \right\|_{1,1} < \infty \) assures, by Carathéodory’s theorem, that the solution of Eq. (1) exists and is unique in an extended sense [66], while the requirement that \( L(\tau) \) is \( k+2 \) differentiable at \( k = 1 \) implies that one can use the series Eq. (B8a) below with \( \tau = 1 \) and \( N = k \).

By assumption, we have a boundary cancelling schedule at \( s = 1 \). Such a schedule is a function \( s = s(\tau) \) where \( s(1) = 1 \) and \( s^{(j)}(1) = 0 \) for \( j = 1, 2, \ldots, k \). That is,

\[
s(\tau) = 1 - g(1 - \tau)^{k+1} + O(\delta^{k+2}), \quad \delta \equiv 1 - \tau \quad (B1)
\]

when \( \tau \to 1^- \) for some constant \( g \). Let

\[
L(s) = \sum_{j=1}^{N} \lambda_j(s) P_j(s) + D(s) \quad (B2)
\]

be the instantaneous spectral decomposition of \( L(s) \) (without the BC schedule), i.e., \( \lambda_j(s) \) are eigenvalues, \( P_j(s) \) are eigenprojectors, and \( D(s) \) is a nilpotent term. We denote by \( P(s) \) (without a subscript) the projector related to the zero eigenvalue. The assumption of contractivity implies that \( D(s) \) cannot belong to the \( \mathcal{L}(s) = 0 \) subspace. We denote by \( \sigma(\tau) \) the unique steady state of \( L(s(\tau)) \), i.e., \( L(s(\tau)) \sigma(\tau) = 0, T \sigma(\tau) = 1 \) while \( \rho(\tau) \) is the solution of Eq. (1) with the BC schedule and with the initial condition \( \rho(0) = \sigma(0) \). Note that in the main text \( \mathcal{L}(s(\tau)) \) is written for simplicity as \( \mathcal{L}(\tau) \).

We use a similar approach (and notation) as in Ref. [33]. Here \( \mathcal{E}(\tau) := \mathcal{E}(\tau,0) \) is the evolution superoperator satisfying \( \partial \mathcal{E}(\tau,\tau') = t_f L(s(\tau)) \mathcal{E}(\tau,\tau') \) and \( \mathcal{E}(\tau_3,\tau_2) \mathcal{E}(\tau_2,\tau_1) = \mathcal{E}(\tau_3,\tau_1) \) together with \( \mathcal{E}(\tau,\tau) = \mathbb{I} \), i.e.,

\[
\mathcal{E}(\tau_2,\tau_1) = \mathcal{E}^{-1}(\tau_3,\tau_2) \mathcal{E}(\tau_3,\tau_1). \quad (B3)
\]

Let \( V(\tau) \) be the adiabatic intertwiner as defined in [33] and, following the same notation therein, \( W(\tau) := P(\tau)V(\tau) = V(\tau)P(0) \). We start with Eq. (D4) of Ref. [33] which we reproduce here for clarity:

\[
\left[ \mathcal{E}^{-1}(\tau)V(\tau) - \mathbb{I} \right] P(0) = \int_0^\tau \mathcal{E}^{-1}(\tau_1) W(\tau_1) d\tau_1, \quad (B4)
\]

where the dot denotes differentiation with respect to the argument of the corresponding operator. Hence

\[
\left[ \mathcal{E}^{-1}(1)V(1) - \mathbb{I} \right] P(0) = \left( \int_0^\tau + \int_1^1 \right) \mathcal{E}^{-1}(\tau_1) W(\tau_1) d\tau_1 \quad (B5a)
\]

\[
\left( \mathcal{E}^{-1}(\tau)V(\tau) - \mathbb{I} \right) P(0) + \int_\tau^1 \mathcal{E}^{-1}(\tau_1) W(\tau_1) d\tau_1, \quad (B5b)
\]

Now apply \( \mathcal{E}(1) := \mathcal{E}(1,0) \) from the left and use Eq. (B3) repeatedly, to obtain:

\[
(V(1) - \mathcal{E}(1)) P(0) = \left( \mathcal{E}(1,\tau)V(\tau) - \mathcal{E}(1,0) \right) P(0)
\]

\[
+ \int_\tau^1 \mathcal{E}(1,\tau_1) W(\tau_1) d\tau_1 \quad (B6a)
\]

\[
\mathcal{E}(1,\tau) (V(\tau) - \mathcal{E}(\tau)) P(0) + \int_\tau^1 \mathcal{E}(1,\tau_1) W(\tau_1) d\tau_1. \quad (B6b)
\]
Now apply $\sigma(0)$ from the right. Note that $V(\tau)$ is continuous, being the solution of a differential equation. So

$$
V(1)P(0)\sigma(0) = \lim_{\tau \to 1^-} \sigma(\tau) = \sigma(1^-)
$$

and we obtain finally

$$
\sigma(1^-) - \rho(1) = \mathcal{E}(1, \tau) (\sigma(\tau) - \rho(\tau)) + \int_{\tau}^{1} \mathcal{E}(1, \tau_1) \dot{W}(\tau_1) \sigma(0) d\tau_1. 
$$

where henceforth we do not always explicitly emphasize the dependence on $\tau$ when it does not lead to a divergence. Now, the derivatives of the projectors behave as

$$
\partial_\tau P_j(s(\tau)) = \partial_\tau P_j(s) \partial_\tau s(\tau) \sim \partial_\tau P_j \delta_k.
$$

Hence, we see that when $\tau \to 1^-$, $b_1$ diverges as

$$
b_1(\tau) \sim \frac{1}{\delta_\alpha - k}.
$$

Eq. (B7) is the starting point of our analysis. Since in Eq. (B7) $\tau < 1$ and the system is gapped in $[0, \tau]$, we can use the result of [13] to obtain the behavior of $\sigma(\tau) - \rho(\tau)$. Assuming that $\mathcal{L}(\tau)$ is at least $N + 2$ times differentiable in a neighborhood of $\tau = 1$, one has the following series [13] with $\epsilon = 1/7t$ for $\tau$ in the same neighborhood:

$$
\rho(\tau) = \sigma(\tau) + \sum_{n=1}^{N} e^n b_n(\tau) + e^{N+1} r_N(\epsilon, \tau) \quad \text{(B8a)}
$$

$$
b_1(\tau) = S(\tau) \dot{P}(\tau) \sigma(\tau) = S(\tau) \dot{\sigma}(\tau) \quad \text{(B8b)}
$$

$$
b_{n+1}(\tau) = S(\tau) \dot{b}_n(\tau), \quad n = 1, 2, \ldots, N \quad \text{(B8c)}
$$

where the remainder is

$$
r_N(\epsilon, \tau) = b_{N+1}(\tau) - \mathcal{E}(\tau) b_{N+1}(0) - \int_{0}^{\tau} \mathcal{E}(\tau, \tau') \dot{b}_{N+1}(\tau') d\tau'.
$$

(B9)

and $S(\tau)$ is the reduced resolvent, defined as:

$$
S(\tau) := \lim_{z \to 0} \left( \mathbb{I} - P(\tau) \right) \left( \mathcal{L}(\tau) - z \right)^{-1} \left( \mathbb{I} - P(\rho) \right).
$$

(B10)

We will keep $N$ free and only fix it at the end when needed. The series (B8a) cannot be used at $\tau = 1$, because the assumption of a gapped Liouvillian does not hold there. Hence we must keep $\tau < 1$ but investigate the kind of divergences that arise when $\tau \to 1^-$. The other assumption is that, without the BC schedule, the Liouvillian gap closes at $s = 1$ in such a way that $\lambda_1(s) \approx v(1-s)^{n}$ as $s \to 1^-$. Therefore, using Eqs. (B1) and (B2), close to $s = 1$ the Liouvillian with the BC schedule behaves as

$$
\mathcal{L}(s(\tau)) \approx \tilde{v}(1 - \tau)^{\delta} P_j(s(\tau)) + \sum_{j=1}^{n} \lambda_j(s(\tau)) P_j(s(\tau)) + D(s(\tau)),
$$

with $\tilde{v} = v g^{\alpha}$ and $\alpha = \alpha(k+1)$. Accordingly, the most diverging term of the reduced resolvent $S(\tau)$ behaves as

$$
S(\tau) \sim \frac{P_j}{\delta^{\alpha}},
$$

(B12)

where $\delta$ is at least $N + 2$ times differentiable in a neighborhood of $\tau = 1$, and the system is gapped in $[0, \tau]$. We can use the result of [13] to obtain the behavior of $\sigma(\tau) - \rho(\tau)$. Assuming that $\mathcal{L}(\tau)$ is at least $N + 2$ times differentiable in a neighborhood of $\tau = 1$, one has the following series

$$
\rho(\tau) = \sigma(\tau) + \sum_{n=1}^{N} e^n b_n(\tau) + e^{N+1} r_N(\epsilon, \tau) \quad \text{(B8a)}
$$

$$
b_1(\tau) = S(\tau) \dot{P}(\tau) \sigma(\tau) = S(\tau) \dot{\sigma}(\tau) \quad \text{(B8b)}
$$

$$
b_{n+1}(\tau) = S(\tau) \dot{b}_n(\tau), \quad n = 1, 2, \ldots, N \quad \text{(B8c)}
$$

Next, we use $\|\mathcal{E}(\tau)\| \leq \|\mathcal{E}(\tau)\|_{1,1} \|x\|_{1,1}$ for a superoperator $\mathcal{E}$ and operator $x$, and the fact that since $\mathcal{E}(\tau_2, \tau_1)$ is a completely positive trace preserving map for $\tau_2 \geq \tau_1$, we have

$$
\|\mathcal{E}(\tau_2, \tau_1)\|_{1,1} = 1, \quad \tau_2 \geq \tau_1.
$$

(B18)

Therefore

$$
\int_{0}^{\tau} \|\mathcal{E}(\tau, \tau') \dot{b}_{N+1}(\tau')\|_{1,1} d\tau' \leq \int_{0}^{\tau} \|\dot{b}_{N+1}(\tau')\|_{1,1} d\tau'.
$$

(B19)

Now using $b_n(\tau) \sim (1 - \tau)^{-\beta_n}$ again we obtain $\|\dot{b}_{N+1}(\tau')\|_{1,1} \leq (1 - \tau')^{-\beta_{N+1}}$, which integrated from 0 to $\tau$ implies that

$$
\int_{0}^{\tau} \|\mathcal{E}(\tau, \tau') \dot{b}_{N+1}(\tau')\|_{1,1} d\tau' \leq \frac{B}{\delta^{\beta_{N+1}}},
$$

(B20)

for some positive constant $B$. All in all, we obtain for the remainder

$$
\|r_N(\epsilon, \tau)\| \leq \frac{A_{N+1}}{\delta^{\beta_{N+1}}} + \|b_{N+1}(0)\| + \frac{B}{\delta^{\beta_{N+1}}},
$$

(B21)

Here $\|b_{N+1}(0)\|$ is just another constant, independent of $\delta$. Combining these results we obtain from Eq. (B8a), after redefining the constants:

$$
\|\rho(\tau) - \sigma(\tau)\| \leq \sum_{n=1}^{N+1} e^n \frac{A_n}{\delta^{\beta_n}} + e^{N+1} A_{N+2}.
$$

(B22)

---

8 Here, and analogously in the following, the symbol ‘$\sim$’ in Eq. (B12) is intended to mean $\lim_{\tau \to 1^-} \|S(\tau)\|_1^{\delta} \|P_j(1)\|_1^{\delta}$.

This implies that one can find a positive constant $C$ such that $\|S(\tau)\|_1^{\delta} \leq C/\delta^{\beta}$ for sufficiently small $\delta$. This fact will be repeatedly used in the following.
It remains to estimate the second term on the RHS of (B7). Using the mean value theorem, \( W = PW \) [33, Eq. (C1)], and \( W(\tau)\sigma(0) = \sigma(\tau) \), we have:

\[
\int_\tau^1 \mathcal{E}(1, \tau_1)W(\tau_1)\sigma(0)d\tau_1 = \delta \mathcal{E}(1, \hat{\tau})\hat{P}(\hat{\tau})\sigma(\hat{\tau}),
\]

where \( \hat{\tau} \in [\tau, 1] \), and we used \( \delta = 1 - \tau \). Hence, when taking the norm of Eq. (B23), using Eqs. (B13) and (B18), we obtain, for some positive constant \( C \):

\[
\left\| \int_\tau^1 \mathcal{E}(1, \tau_1)W(\tau_1)\sigma(0)d\tau_1 \right\| \leq C\delta^{k+1}. \tag{B24}
\]

Using Eqs. (B22) and (B24), we finally obtain after taking the norm of Eq. (B7):

\[
\left\| \rho(1) - \sigma(1^-) \right\| \leq \sum_{n=1}^{N+1} \epsilon^n A_n \delta^{\alpha n} + \epsilon^{N+1} A_{N+2} + C\delta^{k+1} \tag{B25}
\]

\[=: f(\delta, \epsilon). \]

In this derivation \( \tau = 1 - \delta \) is sufficiently close to one but otherwise free. Thus, we can now minimize \( f \) as a function of \( \delta \). Differentiating \( f \) we obtain:

\[
\partial_\delta f = - \sum_{n=1}^{N+1} \epsilon^n \frac{\beta_n A_n}{\delta^{\alpha n+1}} + (k + 1)C\delta^k = 0 \tag{B26a}
\]

\[\rightarrow \delta^{-k} \partial_\delta f = - \sum_{n=1}^{N+1} \epsilon^n \frac{\beta_n A_n}{\delta^{(n+\alpha+1)}} + (k + 1)C = 0. \tag{B26b}
\]

We now make the ansatz \( \delta = \nu e^{\gamma} \). We wish to find a \( \gamma \) such that the solution \( \nu \) of Eq. (B26b) does not depend on \( \epsilon \). This is satisfied if all the summands in Eq. (B26b) do not scale with \( \epsilon \), i.e., are constant, which holds for \( n - \gamma(n\alpha + 1) = 0 \). This means that \( \gamma = 1/(\alpha k + \alpha + 1) \). The value of \( \nu \) can be found by solving

\[
\sum_{n=1}^{N+1} \frac{\beta_n A_n}{\nu(n\alpha k + \alpha + 1)} = (k + 1)C. \tag{B27}
\]

Picking \( N = 1 \) in Eq. (B27) one can show that the resulting quadratic equation in \( x := \nu^{-(\alpha k + \alpha + 1)} \) always has a positive solution, which means that Eq. (B27) always has a real solution \( \nu \).

Substituting \( \delta = \nu e^{\gamma} \) with \( \gamma = 1/(\alpha k + \alpha + 1) \) into Eq. (B25), we see that all terms scale as \( \epsilon^\eta \) with the exponent

\[
\eta = \frac{k + 1}{\alpha k + \alpha + 1}. \tag{B28}
\]

In other words:

\[
\|\rho(1) - \sigma(1^-)\| \leq C\epsilon^\eta, \tag{B29}
\]

as reported in the main text. See Figure 10 for numerical simulations confirming Eqs. (B28) and (B29).

---

**FIG. 10.** Test of the bound given in Proposition 4, for the case of a single qubit subject to evolution under the AME. The Hamiltonian is \( H(\tau) = \omega_x (1 - \beta_x(\tau))\sigma^x + \omega_y \beta_y(\tau)\sigma^y \) with BC schedule \( \beta_y(\tau) = B_y(1, k+1)/B_1(1, k+1) \) [B is the incomplete beta function, Eq. (31)] which has \( k \) vanishing derivatives at \( \tau = 1 \) and is linear at \( \tau = 0 \). The system-bath coupling is via \( \sigma^z \), which assures that the Liouvillian is gapless at \( \tau = 1 \) with a gap-closing exponent \( \alpha = 2 \) [33]. The bath spectral density is Ohmic [Eq. (11)]; the parameters are \( g = 1, \beta = 1, \omega_x = \omega_y = -1/2 \) in arbitrary units. Dots are results of solving for the evolution while the continuous lines are linear fits. The theoretical expectation given by Eq. (B28) is \( \eta = (0.333, 0.4, 0.428, 0.444) \) for \( k = (0, 1, 2, 3) \), respectively, and is an excellent agreement with the the last column of the legend (obtained via a linear fit).

---

**Appendix C: Gap closing in the middle: Numerical check**

Here we present numerical simulations for the case where the Liouvillian gap closes in the middle of the evolution and one performs boundary cancellation at the end, as discussed in Sec. II.C of the main text.

We first consider the following Hamiltonian:

\[
H(\tau) = \sigma^y + (\tau - 1)^{k+1}X + (\tau - 1)^{k+2}Y \tag{C1a}
\]

\[X := (-1)^k \left[ \sigma^x - 2^{k+2}\sigma^z + (2^{k+2} - 1)\sigma^y \right] \tag{C1b}
\]

\[Y := 2(-1)^k \left[ \sigma^x - 2^{k+1}\sigma^z + (2^{k+1} - 1)\sigma^y \right]. \tag{C1c}
\]

One can check that \( H(0) = \sigma^x, H(1/2) = \sigma^z, H(1) = \sigma^y \) and that the first \( k \) derivatives of \( H(\tau) \) at \( \tau = 1 \) are zero. Hence the Liouvillian gap closes with exponent \( \alpha = 2 \) at \( \tau = 1/2 \) and there is a BC schedule at the end. Numerical results confirming the expected scaling \( \eta = k + 1 \) are presented in Fig. 11(a).

Another possibility is to use the following Hamiltonian, which performs the traditional interpolation from a transverse to a longitudinal field:

\[
H(\tau) = s(\tau) \sigma^y + \left[ 1 - s(\tau) \right] \sigma^z \tag{C2a}
\]

\[s(\tau) := \left( \frac{2 \left[ B_1(1,k+1)/B_1(1,k+1) - 1 \right]}{k+1} \right)^2. \tag{C2b}
\]
The schedule in Eq. (C2b) starts with $\sigma^{x}$ at $s(0) = 1$ and ends with $\sigma^{z}$ at $s(1) = 1$, where it has $k$ vanishing derivatives. The Liouvillian gap closes with $\alpha = 4$ at some point $s^{\alpha}_{k}$. Numerical results again confirm the expected scaling $\eta = k + 1$, as shown in Fig. 11(b).

All our numerical results (we have also performed numerics for up to 4 qubit Hamiltonians, not shown here) predict the following scaling for the adiabatic error in this setting:

$$\|\rho(1) - \sigma(1)\|_{1} \leq \frac{C}{t_{f}^{\eta}}, \quad (C3a)$$

with $\eta = k + 1$. \quad (C3b)

In other words, the adiabatic error is insensitive to the closing of the gap in the middle of the evolution. Note that this result is different from the analogous one in the closed system setting where the adiabatic error scales with an exponent $\eta = 1/(\alpha + 1)$, where $\alpha$ in this case is the exponent that determines the closing of the Hamiltonian (not Liouvillian) gap.

**Appendix D: Proof that $\langle 0 | \sigma_{j}^{z} | n \rangle = O(A^{\eta})$**

In the main text we stated that $\langle 0 | \sigma_{j}^{z} | n \rangle = O(A^{\eta})$, and hence that the transition matrix elements $W_{ab} = O(A^{2\eta})$ are suppressed for $n \geq 1$. Generalizing beyond the scenario in the main text, we now allow for degeneracy in the eigenstates of $H_{Z}$.

We assume the transverse field is not zero and investigate the behavior of the rates when $A \to 0$. For this value of $A(t) \neq 0$, let $H_{S} = \sum_{n} E_{n}\tilde{P}_{n}$ be the spectral decomposition of the system Hamiltonian, where $\tilde{P}_{n}$ are the (possibly degenerate) spectral projectors. The Bohr frequencies have the form $\omega = E_{b} - E_{a}$ and we can label the jump operators $A_{j}(\omega)$ as

$$A_{ab,j}(t) = \tilde{P}_{a}(t)\sigma_{j}^{z}\tilde{P}_{b}(t). \quad (D1)$$

We omit the time dependence from here on.

Let us consider the operators $A_{ab,j}$ perturbatively around $A(t) = 0$. We denote by $P_{a}$ the (unperturbed) spectral projectors at $A(t) = 0$, which are diagonal in the computational basis. By assumption, the degeneracy of $\tilde{P}_{a}$ does not change along the anneal, so that the $\tilde{P}_{a}$ are smoothly connected with the $P_{a}$. We use the Kato perturbation theory formalism [57], and denote the perturbation by $V := H_{X}$. Moreover, for level $a$, we define the reduced resolvent as

$$R_{a} = \sum_{n\neq a} \frac{P_{n}}{E_{n} - E_{a}}. \quad (D2)$$

For fixed $a \neq b$ we expand the jump operators in powers of the transverse field:

$$A_{ab,j} = \tilde{P}_{a}\sigma_{j}^{z}\tilde{P}_{b} := \sum_{n=0} A^{n}C_{n}. \quad (D3)$$

At zeroth order we have (since $a \neq b$)

$$C_{0} = P_{a}\sigma_{j}^{z}P_{b} = P_{a}\sigma_{b}\sigma_{j}^{z} = 0. \quad (D4)$$

At first order, perturbation theory gives

$$C_{1} = -P_{a}\sigma_{j}^{z}(P_{b}V R_{b} + R_{b}V P_{a})$$

$$= -(P_{a}V R_{b} + R_{b}V P_{a})\sigma_{j}^{z}P_{b} \quad (D5a)$$

$$= -P_{a}\sigma_{j}^{z}R_{b}V P_{b} - P_{a}V R_{b}\sigma_{j}^{z}P_{b} \quad (D5b)$$

$$= -\sigma_{j}^{z}R_{b}P_{a}V P_{b} - P_{a}V P_{b}R_{a}\sigma_{j}^{z}. \quad (D5c)$$

Now recall that $V = H_{X} = 2S_{z} = (S_{+} + S_{-})$ where $S_{x}$ ($S_{z}$) is the total spin-$x$ operator (ladder operator) and hence $V$ changes the Hamming distance by one. As a consequence, $P_{a}V P_{b} = 0$ if $\text{HD}(a, b) \neq 1$ where we denoted by $\text{HD}(a, b)$ the smallest Hamming distance among the
representative states of the projectors $P_a$ and $P_b$. I.e., if 
$P_a = \sum_{\mu=1}^{d_a} |\mu a\rangle \langle \mu a|$, $P_b = \sum_{\nu=1}^{d_b} |\nu b\rangle \langle \nu b|$ for some Ising 
states $|\mu a\rangle$, $|\nu b\rangle$,

$$\text{HD}(a,b) := \min_{\mu,\nu} \text{HD}(|\mu a\rangle,|\nu b\rangle).$$

At second order we obtain

$$C_2 = (P_a VR a + R_a VP a) \sigma_j^z (P_b VR b + R_b VP b) + 
P_a \sigma_j^z \left( R_b VP a VR b + R_b VR b VP b - R_b^2 VP a VP b \right) + 
\left( P_a VR a VR a + R_a VP a VR a - P_a VP a VR a^2 \right) \sigma_j^z P_b. $$

The formulas become increasingly complicated, but we 
see that at order $k$, $C_k$ is sum of terms of the form 
$O^{(k)}_1 O^{(k)}_2 \cdots O^{(k)}_{k+1}$ where there are $k V$’s and we 
denote by $O^{(j)}_i$ any operator diagonal in the same basis as 
$\sigma^z$. Two of the $O^{(j)}_i$ ($i = 1,2,\ldots,k+1$) comprise $P_a$ and 
$P_b$. Let $d_a$ be the number of $V$’s sandwiched between 
$P_a$ and $P_b$. Then this term is zero if $\text{HD}(a,b) > d_a$. 
Since at order $k$, $d_a$ is necessarily $\leq k$, we see that $C_k$ 
is zero if $\text{HD}(a,b) > k$. Alternatively, note that the ef-
fact of each $O^{(j)}$ on computational basis states is simply 
that of creating a multiplicative constant but otherwise it 
leaves computational basis states unchanged. So, apart 
from a multiplicative constant each term in $C_k$ has the form 
$|\mu a\rangle \langle \mu a| V_{d a} |\nu b\rangle \langle \nu b|$. Summarizing, if $\text{HD}(a,b) = q$, 
$A_{ab,\nu} = O(A^q)$. Since in the Pauli matrix $W_{ab}$ there 
appear two Lindblad operators per Bohr frequency ($a,b$ 
pair) we have just shown that $W_{ab} = O(A^3)$. In par-
|icular, this is true if the level $a = 0$ and $b = n$ are 
non-degenerate as it is the case in the main text.

Figure 12 shows that this prediction is borne out in 
numerical simulations.

Appendix E: Determination of the Freezing Point

The resulting transition rates between energy eigen-
states in the AME are shown in Fig. 14 for both the 
FM and the T-gadgets. Since the magnitude of available 
annealing times are limited to $10^3 \mu s$, we expect that 
a transition will become too rare to observe once it drops 
below $1/\tau_f \approx 10^{-8}$ GHz. With the majority of the pop-
ulation in the ground state and first excited state, we 
therefore expect freezing to occur when $\Gamma_{10}$ goes below 
this value. This intersection determines the freeze-out 
points highlighted in Fig. 4.

Appendix F: Schedule Construction

The annealing schedule of the DW-LN used in our main 
experiments is shown in Fig. 1. All simulations are 
based on this schedule and its operating temperature at the 
time of sample collection (13.5 mK).
FIG. 13. Ground state probabilities and fits for the BCP applied to the FM gadget using (a) DW2KQ at NASA QuAIL and (b) D-Wave Advantage at the appropriate values of $s_{\text{BC}}$ where the fits are most distinguishable.

FIG. 14. Thermal transition rates of the adiabatic master equation (with $\eta_0 g^2 = 5.0 \times 10^{-4}$) for some pairs of energy eigenstates $a \rightarrow b$ as the DW-LN schedule progresses linearly for the FM-gadget (top) and the T-gadget (bottom). The horizontal dotted line denotes the rate $1/t_f$, which is on the order of magnitude of the smallest rate that can be reliably observed within an anneal time of $t_f = 10^3$ µs. Below this line, thermal relaxation events rapidly become extremely rare, hence the onset of frozen dynamics during the anneal. The Hamming distances between the final computational states corresponding to the eigenstate transitions are also listed. When an excited state is only a small Hamming distance away from the ground state, the transition freezes later in the anneal since only a much smaller number of qubits need to tunnel through. This is the case for the $2 \rightarrow 1$ and $3 \rightarrow 1$ transitions in the FM-gadget, and the $3 \rightarrow 0$ transition in the T-gadget.

Appendix G: Crosstalk

It is known that, due to the difficulty of hardware fabrication of the D-Wave processors, programming the couplings and field to have certain values induces non-zero fields and couplings also on other links of the graph, despite designs meant to minimize such crosstalk [67]. We use the ferromagnetic crosstalk model defined by the following transformations [20]:

$$h_i \rightarrow h_i - \chi \sum_{k \neq i} J_{ik} h_k \quad \text{(G1a)}$$

$$J_{ij} \rightarrow J_{ij} + \chi \sum_{k \neq i,j} J_{ik} J_{jk} \quad \text{(G1b)}$$

where $\chi$ is the crosstalk strength. We assume a strength of $\chi = 0.02$ throughout. We have checked that optimizing the value of $\chi$ further does not substantially affect our numerical results. This crosstalk has the effect of breaking the degeneracy of the first and second excited...
states shown in Fig. 4.

We note that the transformation (G1) commutes with gauge (or spin-reversal) transformations of the Hamiltonian [68, 69], so population shifts due to crosstalk do not depend on the gauge.

Appendix H: Data Collection and Fitting

The probabilities at a given anneal time are obtained from a sample of 20 submissions (10 for QAC), where each submission programs 5 random global gauges of the gadget repeated over 31 unit cells, and samples each gauge 200 times (see Ref. [69] for a review of the gauging procedure and other best practices). To mitigate time-fluctuating systematic biases, submissions for the same anneal time are not consecutively submitted (anneal times are cycled for every schedule). Estimates for the ground state probability are taken from the average of the ground state probabilities of the 20 submissions. Error bars for each measured ground state probability are derived from twice the standard deviation over all submissions. Either by non-linear fits on the transformed data, or by linear fits on the log-log transformed data assuming a fixed $\bar{\chi}$, we find the constant $C$ and exponent $\eta$ such that the $\chi^2$ with $D_{GS}$ is minimized. Nonlinear fits are performed as a function of $1/\eta$. The estimates of $\eta^*$ for the D-Wave data are obtained by performing fits on 100 bootstrap samples of the collected ground state probabilities. We use the median values of the parameter distributions as the estimators for the parameters as they are robust to the outliers in bootstrapping that occur. The error estimate is derived from half of the range of the 90% confidence interval of the parameter distribution. The non-linear fits are evaluated using the Mathematica function NMinimize[NonlinearModelFit] with the method NMinimize[NelderMead]. The non-linear $\chi^2$ objective was weighed by the variances of the ground state probabilities.

Appendix I: Results for alternative D-Wave QPUs

We performed similar BCP quantum annealing experiments with the DW2KQ at NASA QuAIL (DW-NA) and with the D-Wave Advantage (DWA) annealers, while the main results used the low-noise DW2KQ processor through D-Wave Advantage (DWA) when $s_{BC}$ is in the appropriate region for each annealer. Furthermore, DWA appears to exhibit more significant time-dependent deviations, as can be noted in the $\beta_2$ schedule at $t_f = 431$ in Fig. 13.

---

[1] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, arXiv:quant-ph/0001106 (2000), arXiv:quant-ph/0001106.
[2] T. Kadowaki and H. Nishimori, Physical Review E 58, 5355 (1998).
[3] M. H. Amin, E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, Physical Review X 8, 021050 (2018).
[4] T. Kato, Journal of the Physical Society of Japan 5, 435 (1950).
[5] S. Jansen, M.-B. Ruskai, and R. Seiler, Journal of Mathematical Physics 48, 102111 (2007).
[6] A. Joye, Commun. Math. Phys. 275, 139 (2007).
[7] L. M. Garrido and F. J. Sancho, Physica 28, 553 (1962).
[8] G. Nenciu, Commun. Math. Phys. 152, 479 (1993).
[9] G. A. Hagedorn and A. Joye, Journal of Mathematical Analysis and Applications 267, 235 (2002).
[10] D. A. Lidar, A. T. Rezakhani, and A. Hamma, Journal of Mathematical Physics 50, 102106 (2009).
[11] N. Wiebe and N. S. Babcock, New Journal of Physics 14, 013024 (2012).
[12] Y. Ge, A. Molnár, and J. I. Cirac, Physical Review Letters 116, 080503 (2016).
[13] L. Campos Venuti and D. A. Lidar, Physical Review A 98, 022315 (2018).
[14] T. Albash and D. A. Lidar, Physical Review A 91, 062320 (2015).
[15] D. A. Lidar, Physical Review A 100, 022326 (2019).
[16] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Endreud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose, Nature 473, 194 (2011).
[17] D-Wave Systems Inc., “The D-Wave 2000Q Quantum Computer Technology Overview.” (2018).
[18] T. Albash, S. Boixo, D. A. Lidar, and P. Zanardi, New Journal of Physics 14, 123016 (2012).
[19] E. B. Davies, Comm. Math. Phys. 39, 91 (1974).
[20] T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, Physical Review A 91, 042314 (2015).
[21] T. Albash, I. Hen, F. M. Spedalieri, and D. A. Lidar, Physical Review A 92, 062328 (2015).
[22] Y. Bando, K.-W. Yip, H. Chen, D. A. Lidar, and H. Nishimori, Physical Review Applied 17, 054033 (2022).
[23] M. H. Amin, Physical Review A 92, 052323 (2015).
[24] K. Boothby, P. Bunyk, J. Raymond, and A. Roy, “Next-generation topology of d-wave quantum processors,” (2020), arXiv:2003.00133 [quant-ph].
[25] T. Albash and D. A. Lidar, Physical Review X 8, 031016 (2018).
[26] K. L. Pudenz, T. Albash, and D. A. Lidar, Nature Communications 5, 3243 (2014).
