Experimental demonstration of perturbative anticrossing mitigation using non-uniform driver Hamiltonians

Trevor Lanting,1,* Andrew D. King,1,† Bram Evert,1,‡ and Emile Hoskinson1,§

1D-Wave Systems Inc., Burnaby B.C.
(Dated: August 11, 2017)

Perturbative anticrossings have long been identified as a potential computational bottleneck for quantum annealing. This bottleneck can appear, for example, when a uniform transverse driver Hamiltonian is applied to each qubit. Previous theoretical research sought to alleviate such anticrossings by adjusting the transverse driver Hamiltonians on individual qubits according to a perturbative approximation. Here we apply this principle to a physical implementation of quantum annealing in a D-Wave 2000Q system. We use samples from the quantum annealing hardware and per-qubit anneal offsets to produce nonuniform driver Hamiltonians. On small instances with severe perturbative anticrossings, our algorithm yields an increase in minimum eigengaps, ground state success probabilities, and escape rates from metastable valleys. We also demonstrate that the same approach can mitigate biased sampling of degenerate ground states.

I. INTRODUCTION

Quantum annealing (QA) holds the potential to confer a computational advantage over classical methods via multiqubit tunneling and entanglement1,2. In the near term it offers a scalable alternative to circuit-model quantum computing2,3. QA is based on physical evolution of a time-dependent quantum Hamiltonian towards a classical Hamiltonian; the fact that the mid-anneal Hamiltonian may not resemble the final Hamiltonian provides both an advantage and a challenge. In some cases, QA fails with high probability when the instantaneous quantum ground state has overwhelming support from excited states in the computational basis. As the anneal progresses this support can vanish at a small-gap perturbative anticrossing, where QA fails via Landau-Zener transition.

Previously available quantum annealing systems have allowed only uniform QA, in which each qubit is initialized with the same transverse-field driver Hamiltonian and all qubits are annealed in unison. In this case small-gap perturbative anticrossings can arise in unfavorably structured inputs, even those that are not particularly hard for classical solvers4–6.

Just as this computational bottleneck can be identified via a perturbative expansion, it can be mitigated with perturbative expansion. Such approaches, in particular those that modify the transverse field on a per-qubit basis, have been proposed and simulated7. Here we experimentally demonstrate a similar approach in which qubit dynamics are tuned via anneal offsets, which allow for suppression or enhancement of a qubit’s dynamics by annealing the qubit slightly in advance of or behind other qubits, respectively. We find that both the longitudinal and the transverse perturbative corrections contribute to improved performance across a set of tailored inputs. Measurements of tunneling dynamics and spectral analysis support these findings.

II. MITIGATING PERTURBATIVE ANTICROSSINGS WITH NON-UNIFORM DRIVER HAMILTONIANS

We consider a system of qubits connected to one another with tunable longitudinal spin-spin interactions $J_{ij}$. We can write the uniform QA Hamiltonian as a time-dependent linear combination of an initial driver Hamiltonian $H_D$ and a final problem Hamiltonian $H_P$

\[
H(s) = \frac{1}{2} A(s) H_D + \frac{1}{2} B(s) H_P
\]

(1)

\[
H_D = -\sum_i \sigma_i^z
\]

(2)

\[
H_P = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z
\]

(3)

where $A(s)$ is the transverse driver energy, $B(s)$ is the problem energy, $s = t/t_f$ is a normalized annealing parameter, $\sigma_i^z, \sigma_i^z$ are Pauli matrices operating on the $i$th qubit. At the beginning of the QA algorithm, $A(0) \gg B(0)$ and at the end of the annealing algorithm $A(1) \ll B(1)$. Fig. 1 shows the annealing schedule $\{A(s), B(s)\}$ for the system used herein. All instances studied have $h_i = 0$ for each qubit.

A. Generalizing the uniform QA algorithm

The terms $A(s)$ and $B(s)$ represent the transverse and longitudinal energy scales, which in the uniform QA algorithm do not differ from qubit to qubit. Allowing non-uniform energy scales $A_i(s)$ and $B_i(s)$ on qubit $i$, we can rewrite $H$:

\[
H(s) = -\frac{1}{2} \sum_i A_i(s) \sigma_i^z + \frac{1}{2} \sum_i B_i(s) h_i \sigma_i^z
\]

\[
+ \frac{1}{2} \sum_{i<j} \sqrt{B_i(s) B_j(s)} J_{ij} \sigma_i^z \sigma_j^z.
\]

(4)
FIG. 1. $A(s)$ and $B(s)$ versus $s$ for the QA system under the uniform QA algorithm. A global time-dependent annealing bias tunes both $A(s)$ and $B(s)$ simultaneously for all qubits throughout the QA algorithm.

B. Connecting perturbative anticrossings and degeneracy

A spin in a classical state is said to be free or floppy if flipping the spin results in another classical state that is degenerate. Floppy qubits, and associated degeneracies, have been implicated in perturbative anticrossings\cite{4,6,8}. Here we provide an argument, based on first-order perturbation calculations, to connect observed probabilities of qubit floppiness to the change in energy of degenerate excited states when a finite transverse field is present. This analysis, from which we derive our mitigation strategy, closely mirrors the approach of Dickson and Amin in Ref.\cite{7}, where the authors use a second-order perturbative expansion because the first order is zero in the problem studied.

We first consider a system with $n$ qubits with two-fold degeneracy in the second classical energy level. We can label these excited classical states $|\alpha\rangle$ and $|\beta\rangle$, each having energy $E_1$. We also suppose that $|\alpha\rangle$ and $|\beta\rangle$ differ by a single bit flip. At $s = 1 - \varepsilon$, a small transverse field $A(s)$ is present which lifts the degeneracy between these states. The new eigenenergies are given by

$$E_1' = E_1 + \lambda_k$$

where $\lambda_k$ are the two eigenvalues of the matrix

$$V = \frac{1}{2} A(s) \begin{bmatrix} \langle \alpha | H_D | \alpha \rangle & \langle \alpha | H_D | \beta \rangle \\ \langle \beta | H_D | \alpha \rangle & \langle \beta | H_D | \beta \rangle \end{bmatrix}$$

$$= \frac{1}{2} A(s) \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$  \hspace{1cm} (6)

In this simple example, $\lambda_0 = -A(s)/2$ and $\lambda_1 = A(s)/2$. To first order, the presence of $A(s)$ lifts the degeneracy, lowering the energy of the symmetric superposition of $|\alpha\rangle$ and $|\beta\rangle$ by $A(s)/2$. Since $A(s)$ grows as $s \to 0$, this produces a perturbative anticrossing between the ground state and the first excited state.

We now consider a system with $n$ qubits and $N$ degenerate excited states with energy $E_1$. When these states are connected via single bit flips, the presence of a transverse field $A(s)$ again lifts the degeneracy of these $N$ states. To first order, the lowest energy

$$E_1' = E_1 - \Delta'$$

and

$$\Delta' = \frac{1}{N} \sum_{a=1}^{N} \frac{1}{N} \sum_{b(a,b) \in B} \frac{1}{2} A(s),$$

where $B$ is the set of pairs of states $a, b$ connected by a single bit flip (see Appendix B and\cite{9}). Defining $q(a, b)$ as the qubit index associated with the single bit flip difference, we can write

$$\Delta' = \frac{1}{2} A(s) \sum_{i=1}^{n} \sum_{b(a,b) \in B} \delta(q(a,b) - i)$$

or

$$\Delta' = \frac{1}{2} A(s) \sum_{i=1}^{n} \sum_{b(a,b) \in B} \delta(q(a,b) - i),$$

where $\delta$ denotes the Dirac delta function. We define a per-qubit floppiness metric

$$F_i = \left\langle \sum_{b(a,b) \in B} \delta(q(a,b) - i) \right\rangle_N,$$

which can be interpreted as the fraction of the $N$ states in which the $i$th qubit is floppy. To first order, then, there is a relationship between qubit floppiness and the reduction in excited state energy:

$$\Delta' = \frac{1}{2} \sum_{i} A(s) F_i.$$  \hspace{1cm} (13)

Again, $A(s)$ grows as $s \to 0$, producing a perturbative anticrossing between the ground state and the first excited state.

This suggests a simple mitigation algorithm based on producing a non-uniform per-qubit driver Hamiltonian $A_i(s)\sigma_z^i$: If an initial set of samples obtained via the “uniform” annealing algorithm is dominated by states from a highly degenerate cluster of first excited states, we can reduce $\Delta'$ by decreasing $A_i(s)$ for qubits for which $F_i$ is high relative to other qubits, based on the empirical samples returned by the hardware. If there are multiple clusters of first excited states, this algorithm based on $F_i$ should also lift the lowest energy of each of these clusters. More generally, the goal is to suppress the appearance of states that are disproportionately favored by a perturbative transverse field, and simply mitigating based on
these observed states will suppress oversampled states disproportionately.

Before describing our algorithmic implementation of this approach, we need to lay out the limitations of our \textit{in situ} per-qubit anneal control.

\section{Per-qubit anneal offsets}

Currently available QA systems do not facilitate independent control of \( A_i(s) \) and \( B_i(s) \). Instead we control \( A_i(s) \) and \( B_i(s) \) together on a per-qubit basis using \textit{anneal offsets}.

The D-Wave 2000Q system by default implements Hamiltonian (1) with networks of coupled rf-SQUID flux qubits\(^1\). See Appendix A for more details. The QA algorithm is run by adjusting a global time-dependent bias, \( \Phi \), simultaneously for all of the qubits. We define a normalized global bias, \( c \), as

\begin{equation}
    c = \frac{\Phi_{CCJJ} - \Phi_{sCCJJ}}{\Phi_{CCJJ} - \Phi_{sCCJJ}}
    \end{equation}

where \( \Phi_{sCCJJ}/\Phi_0 = -0.6457 \) and \( \Phi_{sCCJJ}/\Phi_0 = -0.7140 \) for the particular processor used in this study. Like the annealing parameter \( s \), \( c \) ranges from 0 at the beginning of the anneal to 1 at the end of the anneal. Increasing \( c \) over the course of the anneal both decreases \( A(s) \) and increases \( B(s) \) for all qubits; the ratio \( A(s)/B(s) \) is a fixed function of \( c \) determined by the macroscopic rf-SQUID flux qubit device parameters (see Appendix A). Fig. 1 shows \( A(s) \) and \( B(s) \) for the experiments discussed herein, and Fig. 2 shows \( s(c) \). Note that we choose this particular global trajectory to produce a quadratic growth of \( B(s(c)) \).

We can modify the annealing trajectory for an individual qubit through the use of a tunable \textit{in situ} static bias \( \delta_i \) that adds or subtracts from the global time-dependent annealing bias signal \( c \), giving each qubit a bias signal \( c_i = c + \delta_i \). We then define \( A_i(s) \) as \( A(s(c + \delta_i)) \) and \( B_i(s) \) as \( B(s(c + \delta_i)) \). Fig. 2 shows \( A_i(s) \) and \( B_i(s) \) for \( \delta_i = \{ -0.02, 0, 0.02 \} \). On current hardware, the tunable static bias allows an adjustment of up to \( \delta_i \sim 0.1 \). This allows one to advance or delay the annealing signal locally for each qubit.

\section{Algorithmic mitigation approach}

Following the perturbative analysis in Section II B, we employ an iterative method similar to the one described in Ref.\(^7\). At each iteration \( k \), per-qubit anneal offsets \( \delta_{i,k} \) are adjusted to slow the dynamics of qubits that are frequently floppy in the observed states. Adjustments are scaled down for each successive iteration as the iterative search is refined by combining the new and current offsets in an ansatz ratio of \( \sqrt{k} : 1 \).

1. Choose a static offset magnitude \( \alpha \).

2. Initialize each anneal offset \( \delta_{i,0} \) to zero.

3. For iteration \( k = 1, \ldots, n_{\text{iter}} \),

   (a) Anneal \( r \) times, saving each result.

   (b) Compute \( \mu_{i,k} \), the probability of floppiness, for each qubit based on the \( r \) saved results of the current annealing run.

   (c) Adjust each anneal offset according to \( \mu_{i,k} \):

   \begin{equation}
   \delta_{i,k} := \delta_{i,k-1} + \frac{\alpha \mu_{i,k} - \delta_{i,k-1}}{1 + \sqrt{k}}.
   \end{equation}

As a simple demonstration of the method we apply one iteration of the method to the 16-qubit system studied in Ref.\(^{10}\), similar to those studied elsewhere\(^4,11\), shown in Fig. 3. This instance has a unique ground state and a metastable valley of \( 2^8 \) first excited states in which all eight outer qubits (those coupled to only one other qubit) are floppy, leading to a small-gap anticrossing and low ground state probability when run without mitigation. In the first iteration if mitigation, the outer qubits have \( \mu_{i,1} \approx 1 \) and the others have \( \mu_{i,1} \approx 0 \). Applying mitigation with \( \alpha > 0 \) increases both ground state probability and minimum eigengap, while the opposite is true when \( \alpha < 0 \).
III. IMPROVING OPTIMIZATION WITH ANNEAL OFFSETS

A. Optimization testbed

To study the effect of mitigating perturbative crossings with nonuniform driver Hamiltonians, we need small instances—small enough to exactly diagonalize Hamiltonian (1)—in which the uniform QA algorithm will be drawn into a large valley of metastable states by perturbative anticrossings as described in Section II B. We begin with a qubit connectivity graph with 24 qubits, with each qubit coupled to four others, and construct many thousands of random Ising instances by randomly assigning each $J_{ij}$ a value of +1 or −1. Fig. 4 shows an example. The median ground state probability observed on these instances using the uniform QA algorithm is over 99%. To identify a set of instances for which uniform QA fails due to perturbative anticrossings, we first discard all instances with more than three ground states or fewer than 50 first excited states. We then run all remaining instances with the uniform QA algorithm, performing $10^4$ anneals and collecting all spin configurations. Our testbed consists of those 100 instances—unique up to graph isomorphism and spin reversal—for which the processor yielded the lowest observed ground state probability. Each instance has exactly two classical ground states forming a symmetric pair: ↑↑ ⋯ ↑ and ↓↓ ⋯ ↓. Each has large valleys of first excited states connected by floppy qubits.

B. Ground state probability

Fig. 5 shows measurements of ground state probability $p_{GS}$ for the 100 testbed instances after zero, one, and four iterations of the algorithm outlined in Section II D for all 100 instances. We run with $\alpha = 0.04$ and $r_s = 3.15 \times 10^3$. The median $p_{GS}$ improves from 62% to 85% over four iterations of mitigation.

We examine the minimum eigengaps calculated for these instances—see Appendix C—and compare to the minimum eigengaps calculated after a single iteration of the mitigation algorithm applying offsets $\delta_i = \alpha \mu_i$ for both $\alpha = 0.02$ and $\alpha = -0.02$, which we term mitigation and antimitigation respectively. Fig. 6 shows the mitigated and antimitigated eigengaps plotted versus the baseline eigengap. The algorithm systematically increases the minimum eigengap for $\alpha = 0.02$ and systematically decreases the eigengap for $\alpha = -0.02$. 

FIG. 4. A randomly-generated 24-qubit instance shown on a $2 \times 2$ grid of unit cells in the D-Wave 2000Q qubit topology. Qubits are represented by circles; couplers are represented by line segments. Each nonzero coupler $J_{ij}$ is +1 (red) or −1 (blue). This instance has two classical ground states $\uparrow \uparrow \cdots \uparrow$ and $\downarrow \downarrow \cdots \downarrow$, and 424 first excited states.

FIG. 5. Algorithmic mitigation improves QA success probability. Shown are ground state probabilities $p_{GS}$ after zero, one, and four iterations of mitigation. Median $p_{GS}$ increases from 62% to 85% over four iterations.
C. Approximating orthogonal control of $H_D$ and $H_P$

Anneal offsets do not offer control of a qubit’s transverse field independent of classical energy scale. Advancing a qubit slows its dynamics by both reducing its transverse field and increasing its Ising energy scale. To determine whether or not the positive effects of mitigation can be explained by variation of classical energy scale alone, we repeat the experiment from Section III B, now compensating for variation in classical energy scale.

To balance classical energy scales at a particular point $s^*$, we find a new set of couplings $J'_{ij}$ that equalize effective energy scales for each coupler at $s^*$:

$$\sqrt{B_i(s^*)B_j(s^*)}J'_{ij} = B(s^*)J_{ij}.$$  

Following (4), applying QA to $J'$ in the place of $J$ gives us

$$H(s^*) = -\frac{1}{2} \sum_i A_i(s^*)\sigma^i_z + \frac{1}{2} B(s^*) \sum_{i<j} J_{ij}\sigma^i_z\sigma^j_z,$$

approximating independent control of $A_i$ at $s^*$. We fix $s^* = 0.3$, where $J'_{ij}$ and $J_{ij}$ differ the most. Results are shown in Fig. 7 using $\alpha = 0.04$.

Mitigation continues to provide a systematic performance improvement, but less so than what was achieved without balancing classical energy scales. This indicates that both factors contribute to improved performance in the $p_GS$ metric. When running these experiments we rescaled all Ising Hamiltonians by a factor of 0.85 to keep all couplings within the available interval $[-2, 1]$. The increase in minimum baseline $p_GS$ when downsampling the Hamiltonian is consistent with previous work\textsuperscript{6,11}.

D. Metastable valley escape rates

To further illustrate the effect of the mitigation algorithm, we study the dynamics of relaxation from an excited state during the QA process with and without mitigation. We prepare the QA hardware in a particular classical state at $s = 1$ and run the QA algorithm backwards to an intermediate $s_{\text{target}}$, where we remain for some dwell time $\tau$ before completing the QA algorithm back to $s = 1$. For a given classical state preparation, we can measure the initial escape rate $\Gamma$ out of this state at $s_{\text{target}}$.

This measurement protocol is a generalization of the macroscopic resonant tunneling (MRT) protocol described elsewhere\textsuperscript{12}. We measure $\Gamma$ for excited-state preparations for the uniform QA Hamiltonian and for the mitigated and antimitigated case. For each measurement, we choose an $s_{\text{target}}$ that corresponds to the anticrossing identified by exact diagonalization of Hamiltonian $4$. Fig. 8 shows the rate of escape from the metastable valley for the 100 instances. Mitigation sys-
tematically enhances this rate and antimitigation suppresses it.

IV. SAMPLING DEGENERATE GROUND STATES

The initial testbed of 100 instances was designed to exhibit perturbative anticrossings late in the anneal. Now we consider a second testbed of 100 24-qubit instances with highly degenerate ground states in which some ground states are observed much more frequently than others. In these instances deflection of eigenvalues late in the anneal leads to nonuniform sampling of classical states by QA—this deflection does not lead to perturbative anticrossing because the eigenvalues converge at \( s = 1 \). Again the same mitigation approach—which does not consider whether or not observed states are excited—gives both empirical and spectral improvement.

With the optimization testbed, our performance metric was ground state probability \( p_{GS} \). Here, with the sampling testbed, our performance metric is the expected number of samples required to see every ground state (up to antipodal symmetry) given by the empirical probabilities of each ground state, and normalized on a per-instance basis by the same expected value given perfectly uniform sampling of all ground states. This metric is based on the coupon collector problem, and we denote it \( S_{CC} \). Fig. 9 shows the results analogous to Fig. 5. As before we see a systematic improvement in performance.

To measure spectral bias among classical ground states, we look at the support of the mid-anneal instantaneous ground state in the computational basis, specifically those basis vectors corresponding to classical ground states. Where from an empirical perspective we can view classical ground state probabilities output by QA, from a spectral perspective we can view probabilities of classical ground states in the zero-temperature quantum Boltzmann distribution at some fixed point \( s \) in the anneal. In Fig. 10 we show this data for \( s = 0.3 \), contrasting the choice of \( s = 1 \) in Ref.\textsuperscript{15}; if we increase \( s \) to 0.4 the results look qualitatively similar but with larger \( S_{CC} \).

V. OUTSTANDING QUESTIONS

This work demonstrates a simple approach to a computational bottleneck in quantum annealing. To move towards a more general demonstration there are some outstanding questions that need to be addressed:

- **Orthogonal control:** Currently available QA systems do not allow independent adjustment of \( A(s) \) and \( B(s) \); there is distortion of the classical Hamiltonian when \( \delta_i \neq 0 \). This distortion may be helpful or harmful; we attempt to compensate for it in Section III C and show that our approximation to transverse-only mitigation is successful for the instances studied. The success of this initial demonstration highlights the need for future QA hardware to allow for at least partially independent adjustment of \( A(s) \) and \( B(s) \).

- **Higher-order expansion:** We have used a first-order perturbative expansion to estimate the contribution of a small transverse field; this estimate is zero in the absence of floppy qubits. In other problem classes such as Maximum Stable Set\textsuperscript{7}, disordered Ising spin glasses\textsuperscript{18}, or binary Ising spin glasses on graphs with odd connectivity\textsuperscript{16,19}, higher-order perturbative expansion is required because the first order term is zero. The corresponding second-order algorithm has been shown to work in simulation\textsuperscript{7} but remains to be studied in greater detail.

- **Multiple avoided crossings:** The instances studied in this work have a relatively simple spectrum with a single avoided crossing. The mitigation
algorithm effectively lifts this avoided crossing. In more complicated instances there may be multiple avoided crossings at different points during the QA algorithm. A single iteration of the mitigation algorithm may alleviate multiple avoided crossings but if the number of avoided crossings increases quickly with problem size or first-order perturbation is insufficient to resolve all (or any) bottlenecks, the algorithm may fail to boost ground state probability.

- Offset granularity in large systems: For this work we focused on small instances with clear avoided crossings and eigenspectra that are tractable to calculate throughout the QA algorithm. At this problem scale the granularity of available anneal offsets—$\delta_i$ values are quantized at steps of approximately 0.002—is not a barrier to success. Larger systems, in addition to facing the attendant issues mentioned already, may require finer control of anneal offsets.

- Computational advantage: This mitigation technique must ultimately confer a computational advantage in an algorithmic application in order to motivate practical use. Our choice of 24-qubit instances allowed us to study the eigenspectra of the input instances, but strong evidence for a computational advantage would require the study of larger and more difficult inputs.

VI. CONCLUSIONS

We have experimentally demonstrated an algorithm for mitigating perturbative anticrossings in quantum annealing. The algorithm uses anneal offsets to slow the dynamics of floppy qubits, reducing the associated splitting of energy levels that causes perturbative anticrossings. This results in an improvement of optimization performance in 100 small spin-glass instances designed to exhibit a response to this first-order perturbative mitigation. Exact calculations of eigenspectra throughout the anneal confirm that this algorithm increases the eigen-gap and antimitigating reduces the eigengap. Dynamics measurements confirm that under mitigation, relaxation out of excited states is enhanced; under antimitigation the opposite holds.

This work is the first experimental demonstration that a key computational bottleneck for quantum annealing—perturbative anticrossings—can be removed or mitigated with a targeted change in driving transverse Hamiltonian. Extending these ideas to more complex input instances and more sophisticated mitigation algorithms will be the subject of future research.

VII. ACKNOWLEDGMENTS

We thank Richard Harris, Evgeny Andriyash, Mohammad Amin, Jack Raymond, and others at D-Wave for helpful conversations on this work and related topics.

---

* tlaning@dwavesys.com
† aking@dwavesys.com
‡ bevert@dwavesys.com
§ ehoskinson@dwavesys.com

1 T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, C. Folling, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, R. O’Meara, I. Oppee, C. Rich, M. C. Thom, E. Tolkacheva, S. Uchaikin, A. B. Wilson, and G. Rose, Physical Review X 4, 021041 (2014).
2 V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babush, V. Smelyanskiy, J. Martinis, and H. Neven, Physical Review X 6, 031015 (2016).
3 E. Andriyash and M. H. Amin, “Can quantum Monte Carlo simulate quantum annealing?” (2017), arXiv:1703.09277.
4 S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, Nature Communications 4, 2067 (2013).
5 D. S. Steiger, T. F. Rønnow, and M. Troyer, Physical Review Letters 115, 230501 (2015).
6 A. D. King, E. Hoskinson, T. Lanting, E. Andriyash, and M. H. Amin, Physical Review A 93, 052320 (2016).
7 N. G. Dickson and M. H. Amin, Physical Review A 85, 032303 (2012).
8 S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, Nature Physics 10, 218 (2014).
9 M. H. S. Amin and V. Choi, Phys. Rev. A 80, 062326 (2009), arXiv:0904.1387 [quant-ph].
10 N. Dickson et al., Nat. Commun. 4, 1903 (2013).
11 T. Albash, W. Vinci, A. Mishra, P. a. Warburton, and D. A. Lidar, Physical Review A 91, 042314 (2015).
12 R. Harris, M. W. Johnson, S. Han, A. J. Berkley, J. Johansson, P. Bunyk, E. Ladizinsky, S. V. Isakov, C. Rich, M. C. Thom, E. Tolkacheva, S. Uchaikin, A. B. Wilson, and G. Rose, Physical Review X 4, 021041 (2014).
13 Y. Matsuda, H. Nishimori, and H. G. Katzgraber, New Journal of Physics 11, 073021 (2009).
14 S. Mandrà, Z. Zhu, and H. G. Katzgraber, Physical Review Letters 118, 070502 (2017).
15 B. H. Zhang, G. Wagenbreth, V. Martin-Mayor, and I. Hen, Scientific Reports 7, 1044 (2017).
16 P. Flajolet, D. Gardy, and L. Thimonier, Discrete Applied Mathematics 39, 207 (1992).
17 J. King, S. Yarkoni, J. Raymond, I. Ozfidan, A. D. King, M. M. Nevisi, J. P. Hilton, and C. C. McGeoch, “Quantum Annealing amid Local Ruggedness and Global Frustration,” (2017), arXiv:1701.04579.
18 H. G. Katzgraber, F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, Phys. Rev. X 5, 031026 (2015).
19 S. Knysh, Nature Communications 7, 12370 (2016).
20 R. Harris, J. Johansson, A. J. Berkley, M. W. Johnson, T. Lanting, S. Han, P. Bunyk, E. Ladizinsky, T. Oh,
I. Perminov, E. Tolkacheva, S. Uchaikin, E. M. Chapple, C. Enderud, C. Rich, M. Thom, J. Wang, B. Wilson, and G. Rose, Phys. Rev. B 81, 134510 (2010).

R. Harris, T. Lanting, A. J. Berkley, J. Johansonn, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han, Phys. Rev. B 80, 052506 (2009).

A. D. King and C. C. McGeoch, arXiv preprint arXiv:1410.2628 (2014).

O. Perron, Mathematische Annalen 64, 248 (1907).

E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, “Quantum Computation by Adiabatic Evolution,” (2000), arXiv:0001106 [quant-ph].

E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science 292, 472 (2001).

V. N. Smelyanskiy, U. V. Toussaint, and D. A. Timucin, “Simulations of the adiabatic quantum optimization for the Set Partition Problem,” (2001), arXiv:0112143 [quant-ph].

Appendix A: QA apparatus and methods

We acquired the data reported in this manuscript with a D-Wave 2000Q system in Burnaby, British Columbia. Each qubit is a compound-compound Josephson-junction rf SQUID20. The total critical current of the rf-SQUID junctions is 2.56 μA, the rf-SQUID body inductance is 282 pH, and the rf-SQUID capacitance is 98 fF. Each qubit is a compound-compound Josephson-junction 22.

We run the QA algorithm by adjusting the external bias on the compound-compound Josephson-junction loop \( \Phi_{CCJJ}^s / \Phi_0 = -0.6457 \) at \( s = 0 \) to \( \Phi_{CCJJ}^s / \Phi_0 = -0.7140 \) at \( s = 1 \) according to the curve shown in Fig. 1. We ran the QA algorithm with \( t_f = 20 \mu s \) at a processor temperature of 12.7 ± 0.5mK. Of the 2048 qubits physically present, 2033 were operational. This allowed us to embed 63 disjoint copies of a 24-qubit system on the processor, each subject to an independent spin reversal transformation22.

Appendix B: Perturbation calculation details

Consider a system with \( N \) degenerate eigenstates with energy \( E_1 \) connected to one another via single bit flips. The presence of a transverse field \( A(s) \) lifts this degeneracy by an amount \( \Delta \). To first order in degenerate perturbation theory, \( \Delta \) is the smallest eigenvalue of the matrix \( V_{\alpha,\beta} = \frac{1}{2} A(s) \langle \alpha | H_D | \beta \rangle \) where \( | \alpha \rangle \) and \( | \beta \rangle \) denote specific degenerate eigenstates in the set of \( N \).

\( V_{\alpha,\beta} \) is the negative adjacency matrix for states in this excited state manifold. For simplicity, we first assume that each state is connected to the same number of other states via single bit flips, i.e. the connectivity of these states is a regular graph. It is trivial, then, to check that the \( N \) element vector

\[
| v \rangle = \frac{1}{\sqrt{N}} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{B1}
\]

is an eigenvector of \( V_{\alpha,\beta} \) with eigenvalue

\[
\lambda = -\frac{1}{N} \sum_{a=1}^{N} \sum_{b(a,b) \in B} \frac{1}{2} A(s) \tag{B2}
\]

where \( B \) is the set of pairs of states \( a, b \) connected by a single bit flip.

Because all elements \( V_{\alpha,\beta} \leq 0 \), this eigenvalue corresponds to the minimum eigenvalue23. Thus, \( \Delta' = -\lambda \) and

\[
E_1' = E_1 - \Delta'. \tag{B3}
\]

For more complicated connection topologies of degenerate first excited states, that is, non-regular graphs, Eq. B2 is still correct to first order and a good approximation as long as the connectivity graph is approximately regular.
Appendix C: Using third eigengap

Spectral analysis of quantum annealing typically considers the first eigengap\textsuperscript{24–26}, since this value governs guarantees of adiabaticity in a closed system. Since the 100 instances considered here have two antipodal classical ground states, this gap $E_1 - E_0$ goes to zero at the end of the anneal, so analyzing its minimum is meaningless. Instead we consider the third eigengap $E_3 - E_0$. The minimum third eigengap gives a consistent representation of the tunneling between ground-state and excited-state manifolds when the instantaneous ground state transitions from being mainly supported by classical excited states to being mainly supported by classical ground states. Fig. 11 shows three exemplary instances under antimitigation, no mitigation, and mitigation, where the first and second gaps do not necessarily represent the desired tunneling event.