

This is the accepted manuscript made available via CHORUS. The article has been published as:

Quantum annealing correction for random Ising problems

Kristen L. Pudenz, Tameem Albash, and Daniel A. Lidar

Phys. Rev. A **91**, 042302 — Published 2 April 2015

DOI: [10.1103/PhysRevA.91.042302](https://doi.org/10.1103/PhysRevA.91.042302)

Quantum annealing correction for random Ising problems

Kristen L. Pudenz,^{1,2,3} Tameem Albash,^{2,3,4} & Daniel A. Lidar^{1,2,3,4,5}

⁽¹⁾*Department of Electrical Engineering,* ⁽²⁾*Center for Quantum Information Science & Technology,*
⁽³⁾*Information Sciences Institute,* ⁽⁴⁾*Department of Physics and Astronomy,* ⁽⁵⁾*Department of Chemistry*
University of Southern California, Los Angeles, California 90089, USA.

We demonstrate that the performance of a quantum annealer on hard random Ising optimization problems can be substantially improved using quantum annealing correction (QAC). Our error correction strategy is tailored to the D-Wave Two device. We find that QAC provides a statistically significant enhancement in the performance of the device over a classical repetition code, improving as a function of problem size as well as hardness. Moreover, QAC provides a mechanism for overcoming the precision limit of the device, in addition to correcting calibration errors. Performance is robust even to missing qubits. We present evidence for a constructive role played by quantum effects in our experiments by contrasting the experimental results with the predictions of a classical model of the device. Our work demonstrates the importance of error correction in appropriately determining the performance of quantum annealers.

I. INTRODUCTION

It is widely accepted that no form of quantum information processing can be scalable without some form of quantum error prevention, suppression, or correction [1]. This applies in particular to quantum annealing [2–7] and the closely related quantum adiabatic algorithm [8], strategies designed to take advantage of quantum mechanics in solving classical optimization problems, such as finding the ground state of a disordered Ising Hamiltonian, a well-known NP-hard problem [9]. Interest in quantum annealing has piqued in recent years since commercial processors comprising hundreds of programmable superconducting flux qubits have become available to the research community [10, 11], and a lively debate has erupted concerning their quantumness [12–19] and the possibility of observing a quantum speedup [20–22], for which there exists theoretical evidence via specific examples [4, 5, 23].

While error mitigation strategies for quantum annealing and more generally, adiabatic quantum computing have been proposed [24–33] and implemented [34], much less is known compared to the relatively mature state of quantum error correction in the circuit model [1, 35]. In particular, an accuracy threshold theorem [36–38] for fault-tolerant quantum annealing remains elusive, in spite of some degree of inherent robustness of adiabatic quantum computation to thermal excitations and control errors [39–41]. Notwithstanding, we recently proposed a practical error suppression and correction strategy for quantum annealing and implemented it using a D-Wave Two (DW2) quantum annealing processor [42] on a toy problem of antiferromagnetic chains [34]. We demonstrated that this quantum annealing correction (QAC) scheme provided a substantial fidelity enhancement in the presence of thermal excitation and control errors.

Here, we experimentally study the performance of QAC on random Ising problems with quenched disorder using a DW2 processor, on up to 112 logical qubits comprising 4 physical qubits each. In contrast to anti-

ferromagnetic chains, these are hard optimization problems, of the type studied in recent benchmarking work probing for a quantum speedup [13, 20] against simulated annealing [43]. In this manner we hope to demonstrate the importance of including error correction in future quantum annealing devices, and in particular the utility of QAC in improving the performance of the current D-Wave devices. However, rather than attempting to demonstrate a speedup, which seems likely to be precluded in our setting for the reasons discussed in Ref. [21], we focus on establishing a performance improvement, i.e., an enhancement in the success probability of finding the Ising spin glass ground state, when using QAC. Moreover, we demonstrate that QAC is also effective at extending the precision range of the D-Wave device, thus in effect overcoming control errors. The price to be paid for these improvements is a reduction in the number of qubits and the degree of the qubit connectivity graph due to the use of an encoding, but such tradeoffs seem inevitable if the goal is to reach scalability of quantum information processing.

A key question raised by the performance gains of the QAC strategy is to what extent it is a form of *quantum* error correction. To address this we compare our experimental results for the ground state probability to those we compute numerically using a classical model of interacting spins. This SSSV model [16] has been very successful in reproducing the success probabilities of random Ising instances reported in Ref. [13] and has been playing a central role in the quantumness discussion concerning the D-Wave devices [16–18]. We demonstrate that for random Ising problems subject to the QAC encoding, there is a strong discrepancy between the SSSV model and the experimental results. This conclusion is robust to varying the parameters of the SSSV model, and it provides indirect evidence that quantum effects play an important role in the success of the QAC strategy in our experiments.

II. RESULTS

A. Quantum annealing and the D-Wave Two processor

These topics have been described in detail in a number of publications (see, e.g., Refs. [13, 20, 42]), so here we give just the details needed for our work.

Quantum annealing (QA) is a method for finding the ground state of an Ising spin Hamiltonian

$$H_{\text{Ising}} = \sum_{i=1}^N h_i \sigma_i^z + \sum_{i<j}^N J_{ij} \sigma_i^z \sigma_j^z. \quad (1)$$

The local fields h_i and couplings J_{ij} are given and the problem is to find the spin configuration $\{\sigma_i^z\}_{i=1}^N$ that minimizes H_{Ising} , where each spin variable $\sigma_i^z \in \pm 1$. In QA this is done by adiabatic evolution from the ground state of an initial transverse field, i.e., the time-dependent Hamiltonian is

$$H(t) = A(t)H_X + B(t)H_{\text{Ising}}, \quad (2)$$

where $H_X = \sum_{i=1}^N \sigma_i^x$ and the σ_i 's are now the standard Pauli spin-1/2 matrices acting on the i th qubit. The function $A(t)$ decreases monotonically to zero, while $B(t)$ increases monotonically from zero, with $t \in [0, t_f]$. For a closed system the adiabatic theorem guarantees that if the initial state is the ground state then the final state will be arbitrarily close to the ground state provided t_f is large enough compared to the minimum energy gap of $H(t)$ and provided the functions $A(t)$ and $B(t)$ are sufficiently smooth [44, 45].

The DW2 processor is a physical realization of the quantum annealing algorithm. The system is initialized in the thermal Gibbs state of $A(0)H_X$, which has almost its entire weight on the ground state since $kT \ll A(0)$. The idealized conditions of the adiabatic theorem can of course not be realized in a physical device such as the DW2, which operates at a finite temperature and suffers from programming control errors on the h_i and J_{ij} terms. In such an open system, thermal processes can depopulate the ground state, reducing the success probability of the algorithm (though thermal relaxation can sometimes be beneficial [46, 47]). Control errors can unintentionally cause the annealer to evolve according to the wrong Hamiltonian.

B. Quantum Annealing Correction

What can be done to mitigate the detrimental effect of thermal excitation and control errors? While a variety of theoretical proposals exist [24–33] we require one that is implementable using the DW2 device. To this end we employ two strategies that were proposed and studied in Ref. [34]. The first is a purely classical (C) repetition

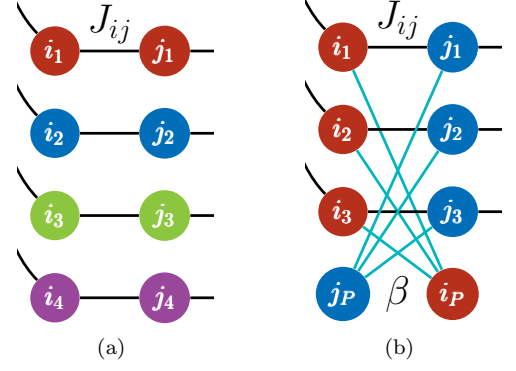


FIG. 1. (Color online) Embedding of the C and QAC strategies on the D-Wave processor. (a) C strategy, (b) QAC strategy. The basic unit cell of the D-Wave processor includes 8 qubits (shown as circles), with the qubits on the left side coupling to adjacent unit cells up and down, and those on the right side coupling left and right. Panel (a) shows how 4 parallel copies of the problem for the C strategy are embedded, each with the same problem couplings J_{ij} (black lines). Panel (b) shows how two logical qubits (red and blue) for the QAC strategy are embedded within a unit cell; the problem qubits $i_{1,2,3}$ are coupled using the same problem couplings J_{ij} (black lines), and the penalty qubit i_P couples ferromagnetically to the problem qubits with magnitude β (light blue lines).

strategy, whereby we evolve K independent copies of the problem, i.e.,

$$H_C = \sum_{i=1}^{\bar{N}} \sum_{k=1}^K h_i \sigma_{i_k}^z + \sum_{i<j}^{\bar{N}} \sum_{k=1}^K J_{ij} \sigma_{i_k}^z \sigma_{j_k}^z. \quad (3)$$

where $K\bar{N} = N$. This strategy runs the quantum annealing algorithm $K \geq 2$ times in parallel to increase its chances of finding the ground state. An example of how the C strategy is embedded on the DW2 device is shown in Fig. 1(a), where $K = 4$.

The second strategy, QAC, uses the same physical resources as the C strategy, but adds two important aspects to suppress and correct bit-flip errors: (i) we encode our problem using a $K - 1$ qubit repetition code, and (ii) supplement H_{Ising} with energy penalty terms in the form of the the stabilizer generators of the repetition code. The former allows us to correct bit-flip errors via decoding and to boost the energy scale, while the latter allows us to suppress thermal excitations. The resulting final Hamiltonian takes the form

$$H_{\text{QAC}} = \alpha \bar{H}_{\text{Ising}} - \beta H_P, \quad \alpha > 0, \beta \geq 0 \quad (4)$$

$$\bar{H}_{\text{Ising}} = \sum_{i=1}^{\bar{N}} h_i \bar{\sigma}_i^z + \sum_{i<j}^{\bar{N}} J_{ij} \bar{\sigma}_i^z \bar{\sigma}_j^z, \quad H_P = \sum_{i=1}^{\bar{N}} \bar{\sigma}_i^z \sigma_{i_P}^z,$$

where $\bar{\sigma}_i^z = \sum_{k=1}^{K-1} \sigma_{i_k}^z$, $\bar{\sigma}_i^z \bar{\sigma}_j^z = \sum_{k=1}^{K-1} \sigma_{i_k}^z \sigma_{j_k}^z$ are (scaled) logical $\bar{\sigma}_i^z$ and $\bar{\sigma}_i^z \bar{\sigma}_j^z$ operators respectively. The QAC

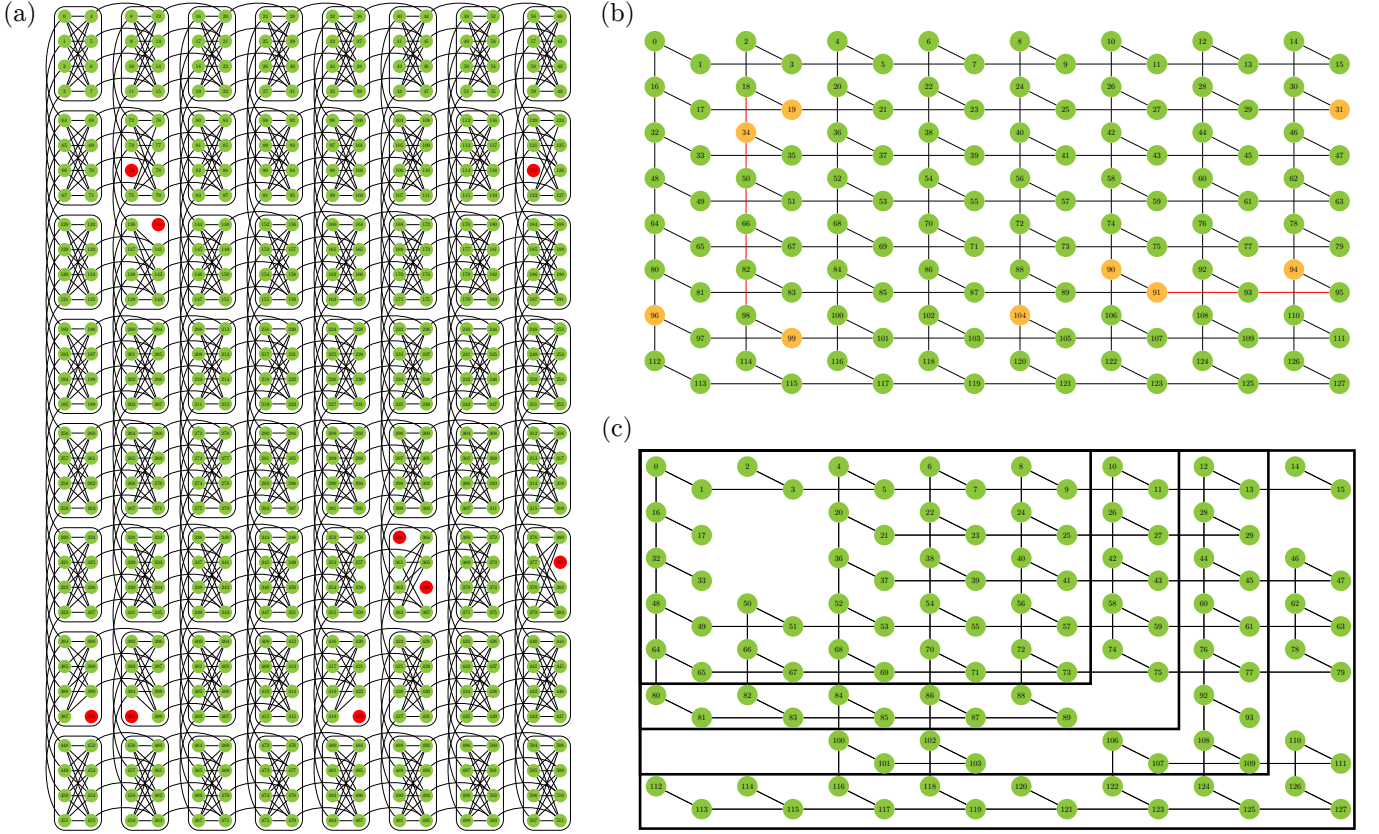


FIG. 2. (Color online) Physical and logical qubit connectivity graphs. (a) The physical connectivity graph consists of 8×8 unit cells of eight qubits (denoted by circles), connected by programmable inductive couplers (lines). The 503 green (red) circles denote functional (inactive) qubits. In the ideal case, where all qubits are functional and all couplers are present, one obtains the non-planar, degree-6 “Chimera” connectivity graph. (b) The complete degree-3 logical connectivity graph, with perfect (imperfect) logical qubits and their couplers shown in green and black (orange and red) respectively. (c) The actual logical connectivity graph. Imperfect qubits and their couplings are not shown as these were not used. We generated random problems over each of the regions shown in the rectangles of increasing sizes $\bar{N} \in \{46, 66, 86, 112\}$ in (c). These problem sizes were chosen because they consist of square blocks of unit cells, and the treewidth of a planar square lattice grows as a function of the smallest dimension of any rectangular region chosen [48].

strategy treats $K - 1$ of the K qubits comprising the i th logical qubit as “problem” qubits which encode the original problem into a repetition code, while the remaining “penalty” qubit i_P is used to implement the energy penalty. This is illustrated in Fig. 1(b) for the DW2. The parameter α is the problem scale factor, through which we can control the effective noise level on the logical Ising Hamiltonian \bar{H}_{Ising} . The penalty term H_P enforces a ferromagnetic coupling between the problem qubits and their penalty qubit, thus aligning them in agreement and forcing errors that do not commute with σ^z to pay an energy penalty. This compensates for the fact that the repetition code can only be used to detect and correct bit-flip errors. However, note that if the dominant dephasing is in the instantaneous energy eigenbasis rather than the computational basis then the adiabatic algorithm is not adversely affected by phase errors [39, 41, 49]. The parameter β is the penalty scale factor, which we optimize to maximize the success probability for each problem instance, balancing H_P against \bar{H}_{Ising} [34]. We note that

we cannot at the same time encode H_X since this would require N -body interactions, and that $\alpha, \beta \leq 1$ on the DW2, thus precluding energy penalty strategies of the type suggested in Refs. [25, 33].

If the copies are statistically independent in the C strategy and each succeeds with probability p , then the probability that at least one copy will succeed is $1 - (1 - p)^K$, which is greater than p if $0 < p < 1$ and $K \geq 2$, so C will improve performance compared to a single copy of the same problem instance. In Ref. [34] we established that the QAC strategy outperforms the C strategy for sufficiently long antiferromagnetic chains. The pertinent question, then, is whether QAC will continue to outperform C for sufficiently large and hard optimization problems.

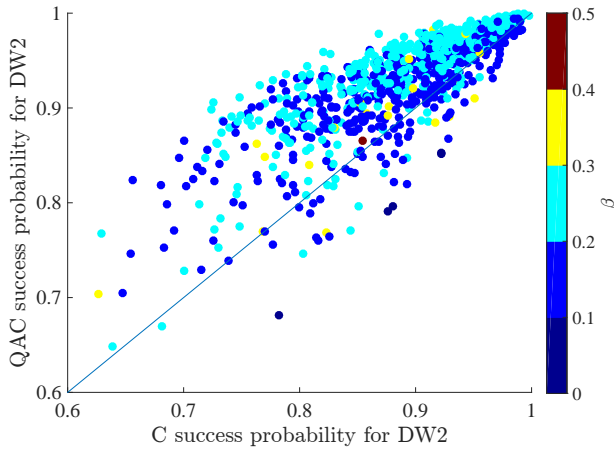


FIG. 3. (Color online) Correlation plot of the success probability for the QAC and C strategies for DW2. The case of $\bar{N} = 112$ logical qubits is shown. Above the diagonal line QAC performs better, while below it C performs better. The color scale indicates the optimal β for each instance. The QAC strategy uses both problem group decoding and logical group decoding, while the C strategy only uses problem group decoding, as discussed in Appendix A. QAC outperforms C for the overwhelming majority of instances. The few instances where C outperforms QAC are those with low-lying undecodable energy states due to logical qubits with < 3 (and weak) problem couplings to neighbors, arising from holes in the logical connectivity graph (see Appendix B).

C. Success probability for random Ising instances

Earlier work studied the performance of the D-Wave processors on random Ising problem instances of increasing size N without error correction [13, 20] by embedding these problems on the physical “Chimera” connectivity graph shown in Fig. 2(a). We now present the results of our study of similar problem instances, using the C and QAC strategies. When the Chimera graph is contracted by replacing each set of 4 physical qubits by the corresponding logical qubit we obtain the *logical* connectivity graph shown in Fig. 2(b). For each problem size \bar{N} we generated 1000 instances where all local fields $h_i = 0$ and the couplings J_{ij} were specified by drawing uniformly at random from the set $\pm\{\frac{1}{6}, \dots, \frac{5}{6}, 1\}$, slightly above the $\Delta J_{ij} \sim \frac{1}{7}$ precision limit of the DW2 [50]. These instances were embedded on the actual logical connectivity graph shown in Fig. 2(c), comprising only perfect logical qubits, defined as consisting of 4 physical qubits (in contrast, an imperfect logical qubit consists of only the 3 problem qubits, without the penalty qubit). We ran each instance 1000 times using different gauges to reduce systematic errors (see Appendix D) and counted the number of times a state with the correct logical ground state energy was found, and defined this as the success probability for that instance (see Appendix A for more details). Note that we distinguish between the logical and

physical ground states: the former is the ground state corresponding to the Ising instance defined in terms of the \bar{N} logical qubits, while the latter is the ground state over the N physical qubits. We also ran each instance with $\beta \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ and defined the optimal β as the value β_{opt} that maximized the success probability after gauge-averaging (see Appendix A for more details).

Since 9 of the physical qubits are missing in the DW2 hardware graph [see Fig. 2(a)], the maximum problem size we can study while implementing the full 4-copy C strategy is $\bar{N} = 112$. This maximal encoded graph has holes in the regular structure due to the missing physical qubits that reduce the number of couplings for some of the QAC logical qubits [see Fig. 2(c)]. Fig. 3 shows the performance of QAC *vs* C in this case, where all $\bar{N} = 112$ perfect logical qubits were used. Answering our earlier question, we find that QAC is a better strategy choice than C for the overwhelming majority of problem instances. We discuss the reason for the appearance of a small number of instances where C outperforms QAC in Appendix A.

D. Dependence on problem size

In order to quantify performance with respect to the hardness of the Ising instances for DW2, we consider the dependence on \bar{N} of the expected number of annealing runs necessary to observe a success (a logical ground state) at least once with 99% probability [13, 20]:

$$R = \frac{\log(1 - 0.99)}{\log(1 - p_S)}, \quad (5)$$

where p_S is the probability of success. For the C strategy, p_S corresponds to observing a logical ground state in at least one of the four copies (this is almost the same as running a single copy four times, as discussed in Appendix D), which we refer to as problem group decoding. For the QAC strategy, p_S corresponds to observing a logical ground state after decoding using two complementary decoding strategies, problem group decoding over the three problem qubits and logical group decoding. Details of the decoding strategies is discussed in Appendix A.

R is a proxy for the time-to-solution $t_a R$, where t_a is the annealing time, set to $20\mu s$ (the minimum possible with the DW2) in all our experiments. This annealing time is certainly suboptimal, i.e., too long to enable the extraction of meaningful scaling behavior, as discussed in detail in Ref. [20]. Therefore R should not be interpreted as reflecting the true value of the time-to-solution for the different strategies. Instead, we focus on the increasing separation between the values of this quantity for QAC and C as a function of problem size to establish the relative performance of the different strategies.

Consider first the results for $\alpha = 1$, in Fig. 4(a). Increasing the problem size corresponds to harder problem instances and requires more runs R , with a steeper rise

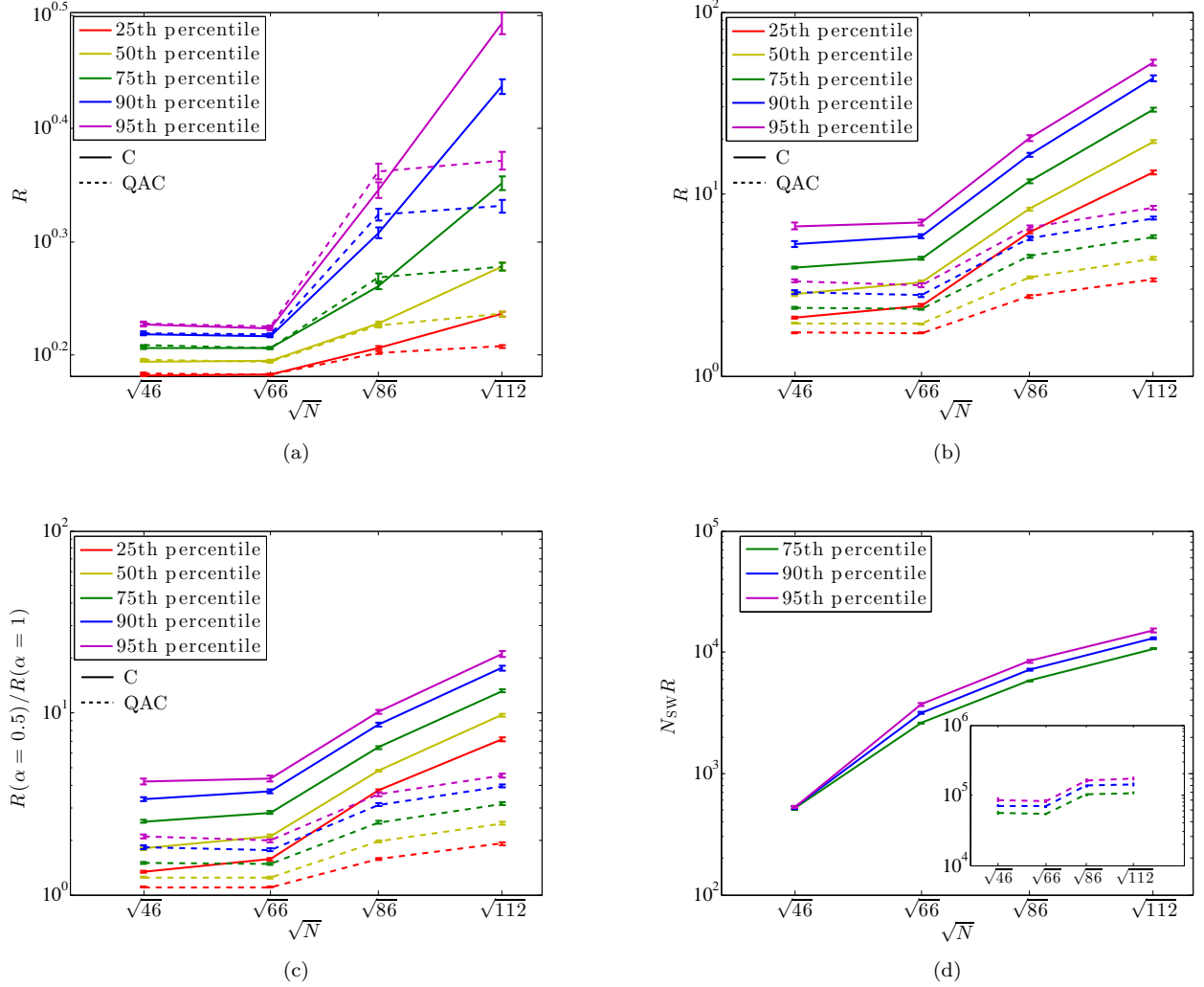


FIG. 4. (Color online) Problem size dependence of the expected number of annealing runs R for the QAC and C strategies. (a) For $\alpha = 1$. (b) For $\alpha = 0.5$. Various percentiles of problem hardness are included, as shown in the legend. The QAC strategy uses both problem group decoding and logical group decoding, while the C strategy only uses problem group decoding, as discussed in Appendix A. We observe better performance for the QAC strategy at all percentiles for the largest problem size. (c) Scaling of the ratio of R values for $\alpha = 0.5$ and $\alpha = 1$. The performance of the QAC strategy decreases by a factor of less than 3 at the hardest percentile and largest problem size, whereas the C strategy's performance decreases by a factor of close to 20. (d) Scaling of QAC for the SSSV model without coupling noise. Solid lines use an optimized number of sweeps N_{sw} for each problem instance, i.e., N_{sw} is chosen to minimize the expected number of annealing runs $N_{\text{sw}} R$, while the dashed lines in the inset show the scaling for a fixed $N_{\text{sw}} = 10^5$.

occurring for the higher percentiles. While for the three smallest problem sizes R is similar for the C and QAC strategies, when the problem size becomes sufficiently large there is a statistically significant separation for every percentile between the two strategies. The beneficial effect of QAC becomes much more prominent at $\alpha = 0.5$, as seen in Fig. 4(b), where the separation between QAC and C is apparent already at the smallest problem size. By reducing the problem energy scale α we have increased the effect of thermal noise, and by halving the values of the couplings we operate the DW2 device in a regime well below the aforementioned $\Delta J_{ij} \sim \frac{1}{7}$

precision limit. The susceptibility to programming errors is therefore higher, yet QAC continues to work. This is further visualized in Fig. 4(c), where we plot the ratio of R values for $\alpha = 0.5$ and $\alpha = 1$. The ratio rises much more rapidly as a function of problem size for the C case than for the QAC case, demonstrating the relative (as well as absolute) stability of the latter to both thermal and control errors.

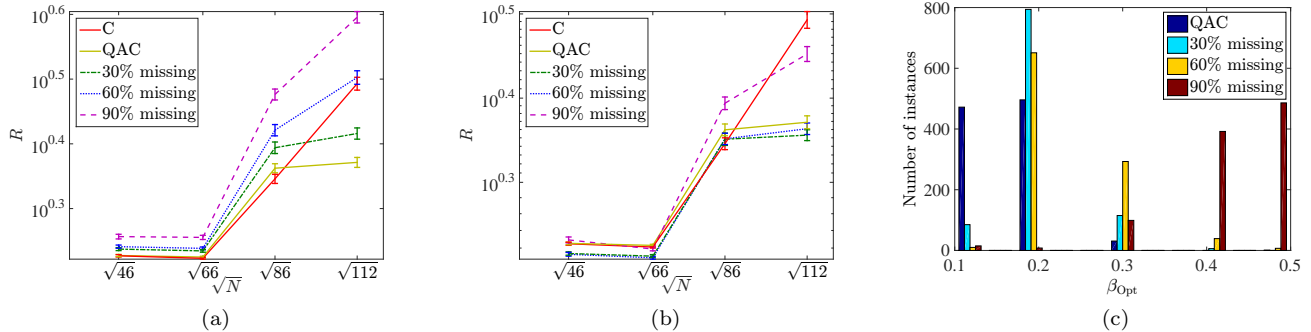


FIG. 5. (Color online) Robustness of QAC to missing penalty qubits. Effect of missing penalty qubits for $\alpha = 1$ on the 95th percentile for (a) β_{opt} held constant at its original QAC value, and (b) β optimized for the number of missing qubits. The C strategy (solid red) and QAC strategy (solid yellow) lines are the same data that was displayed in Fig. 4(a). The new dash-dotted green, dotted blue, and dashed purple series show the effects of randomly removing 30%, 60%, and 90% of the penalty qubits, respectively. For (a), QAC with 30% loss continues to outperform C, but not at the higher percentages. For (b), performance at 30% and 60% loss tracks the original QAC closely, suggesting the code is highly resilient to penalty qubit loss if the penalty magnitude is adjusted accordingly. (c) The histogram displays the optimized β values from the trial set $\{0.1, 0.2, 0.3, 0.4, 0.5\}$, for 1000 instances. The true optimal β for the perfect QAC scheme lies between 0.1 and 0.2, but for 30% of penalty qubits missing the optimum shifts to 0.2, for 60% missing between 0.2 and 0.3, and for 90% missing to the maximum allowed value.

E. Robustness of QAC to Qubit Loss

A good code should be robust, and to this end we next study the effect of imperfect logical qubits, by systematically removing penalty qubits from perfect logical qubits. We repeated our QAC experiments on the same 1000 random problem instances of each size, but removed 30%, 60%, and 90% of the penalty qubits from each instance at random, while keeping all problem couplings intact. We consider two cases: keeping β_{opt} fixed at its QAC value, as well as re-optimizing β for each new fraction of removed penalty qubits.

In Fig. 5 we show that the separation between the C and QAC strategies at the largest problem size persists even when a significant number of penalty qubits are removed. If β_{opt} is held fixed [Fig. 5(a)] QAC improves on the C strategy even when 30% of the penalty qubits are removed. When we allow β to be adjusted, up to 60% of the penalty qubits can be removed and QAC still shows an advantage over the C strategy [Fig. 5(b)]. To achieve this β_{opt} must be increased as more penalty qubits are removed [Fig. 5(c)]. With fewer penalty qubits, the remaining penalty qubits must couple more strongly to their respective problem qubits in order to maintain the benefits of the QAC strategy. At the 30% level β_{opt} is sharply peaked in Fig. 5(c), suggesting that the optimum is determined by the number of available penalty qubits, not the particular problem instance being solved. This is important because the advantages provided by QAC would be diluted if it were necessary to try several values of β to solve a relevant problem. Instead, it suffices to pick a single β value to solve a new problem.

III. DISCUSSION

We have demonstrated a substantial performance enhancement using QAC over the classical C strategy. A natural question is to what extent QAC is a truly quantum strategy. We now discuss this from two different angles: a solvable analytical model, and a comparison to the classical SSV model [16].

A. Analytically solvable model with an optimal β value

Consider the Ising Hamiltonian on a ring with local fields:

$$H_{1D} = -h \sum_{i=1}^d \sigma_i^z - \beta \sum_{i=1}^d \sigma_i^z \sigma_{i+1}^z, \quad (6)$$

where $\sigma_{d+1}^z \equiv \sigma_1^z$. We can reinterpret this as a single qubit “problem Hamiltonian” $-h\bar{\sigma}^z$ encoded into a distance d repetition code with (scaled) logical operator $\bar{\sigma}^z = \sum_{i=1}^d \sigma_i^z$. The ground and excited states are encoded as $|\bar{0}\rangle = |000\dots 0\rangle$ and $|\bar{1}\rangle = |111\dots 1\rangle$ respectively. The spin-spin couplings $\sigma_i^z \sigma_{i+1}^z$ are the stabilizer generators of this repetition code, acting as penalty terms that energetically penalize bit flips. Replacing H_{Ising} by H_{1D} in Eq. (2) and including a transverse field H_X with annealing schedules $A(t)$ and $B(t)$ gives rise to a quantum annealing evolution designed to terminate in the $|\bar{0}\rangle$ state. The resulting solvable model is a special case of the QAC encoding, without logical spin-spin coupling. It is also known as the transverse field Ising model for a chain

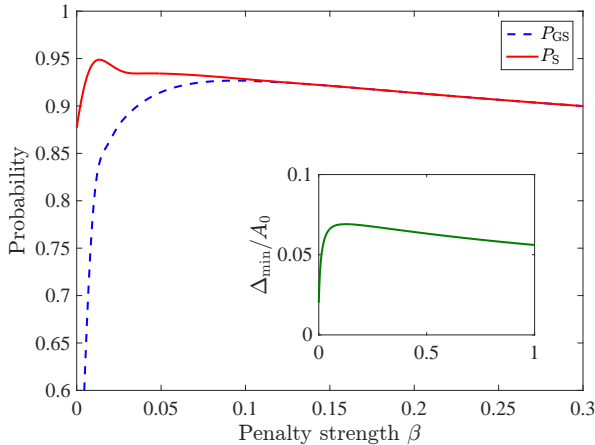


FIG. 6. (Color online) Optimal β in an open quantum system model. The main figure shows the final physical ground state probability P_{GS} and the post-decoding (majority vote with random tie-breaker) success probability P_S of the logical ground state, computed using the adiabatic master equation [49] for quantum annealing using the Hamiltonian (6) with $d = 4$, with linear annealing schedules $A(t) = A_0(1 - t/t_f)$ and $B(t) = A_0 t/t_f$. Both probabilities exhibit an optimal β value. The inset shows the numerically computed minimum gap between the ground and first excited state as a function of β . The peak in P_S and the minimum gap align as expected since a larger minimum gap means fewer transitions out of the ground state. However, the peak in P_{GS} occurs earlier, maximizing the population in all decodable states and not just the ground state. Simulation parameters are: $A_0 = 33.84$ GHz, $t_f = 100$ ns, $h = 0.01$, and system bath coupling $g^2\eta/\hbar^2 = 1.2732 \times 10^{-3}$ (see Eq. (15) in Ref. [34]).

with periodic boundary conditions, where it is commonly written as $H(t) \equiv -h_X \sum_i \sigma_i^x - h_Z \sum_i \sigma_i^z - J \sum_i \sigma_i^z \sigma_{i+1}^z$ [51], where $h_X = A(t)$, $h_Z = hB(t)$, and $J = \beta B(t)$. We can study this model analytically in various limits. Here we present the main results (details of calculations can be found in Appendix E).

The $\beta \gg h$ case: intuitively, in this case the penalty term overwhelms the problem Hamiltonian, which should be detrimental. More rigorously, in this case, corresponding to $J \gg h_Z$, it is well established that in the thermodynamic limit (i.e., $d \rightarrow \infty$), there is a critical point at $h_X = J$, and the energy gap at the critical point scales as $\Delta \sim (h_Z/J)^{8/15}$ [52, 53]. Clearly, in this limit increasing the energy penalty β makes the problem harder to solve at fixed t_f since the shrinking gap will increase the degree of non-adiabaticity as well as increase thermal excitations. Furthermore, the state $|\bar{1}\rangle$, which represents a logical error, is actually the first excited state of the Ising Hamiltonian. Any population lost to this state cannot be decoded and recovered. Having $\beta \gg h$ is thus undesirable.

The $\beta \ll h$ case: in the opposite case, where $J \ll h_Z$, the penalty term acts as a perturbation on the problem Hamiltonian. We show analytically in Appendix E that

the introduction of the penalty term increases the gap and shifts it to earlier in the evolution provided $h < \sqrt{2}$. An increased gap means that the evolution is more adiabatic and thermal excitations are suppressed at fixed t_f , suggesting that the introduction of a small β improves the adiabaticity of the evolution. Furthermore, the low energy excited states of the Ising Hamiltonian correspond to a small number of spin flips, which can be corrected via decoding.

There is thus clearly an optimal $\beta > 0$ value in the large d limit, a fact that helps to explain the observation of an optimal β value in our experiments and in Ref. [34]. To address this in a more realistic model, we performed numerical simulations for finite d using an adiabatic Markovian master equation of an open quantum system [49]. This master equation has been extensively used before in related work where it was shown to be a good model of the D-Wave device [12, 17] (it is briefly reviewed in the Methods section of Ref. [34]). We use it here to model a qubit encoded into a distance $d = 4$ classical repetition code as given by Eq. (6). As shown in Fig. 6, both the physical ground state probability P_{GS} and the logical ground state success probability P_S exhibit a peak as a function of β . The position of the P_{GS} peak coincides with the peak in the minimum energy gap, whereas P_S also depends on the decodability of the excited state spectrum, to which a significant amount of population is lost. The fact that $P_S \geq P_{GS}$ for all β shows that decoding is always beneficial.

These results of the closed and open system models thus lead us to associate the improved performance of QAC with $\beta > 0$ to both the enhancement of the energy gap and the decodability of low-energy excited states.

B. QAC for a classical model of the D-Wave device

Although we have demonstrated that the QAC strategy provides a significant performance advantage over the C strategy, and that quantum models exhibit an optimal penalty strength just as observed in our experiments, it is difficult to establish to what extent quantum effects are responsible for the overall success of QAC. In order to further address this question, we tested the efficacy of the same strategy on a classical model (SSSV) that has been successfully used [16] to reproduce the D-Wave One physical ground state probabilities on random Ising instances reported in Ref. [13]. The SSSV model replaces the qubits by classical planar rotors whose dynamics are governed by Eq. (2) with the replacements $\sigma_i^x \mapsto \sin \theta_i$ and $\sigma_i^z \mapsto \cos \theta_i$, with Monte Carlo updates for each angle θ_i . Though there is evidence that the SSSV model does not correctly capture experiments on specially designed Ising instances of up to 20 qubits [17], it remains an excellent classical model for random Ising instances on larger numbers of qubits, setting a high bar for genuine quantum effects. In particular, if the D-Wave device is well

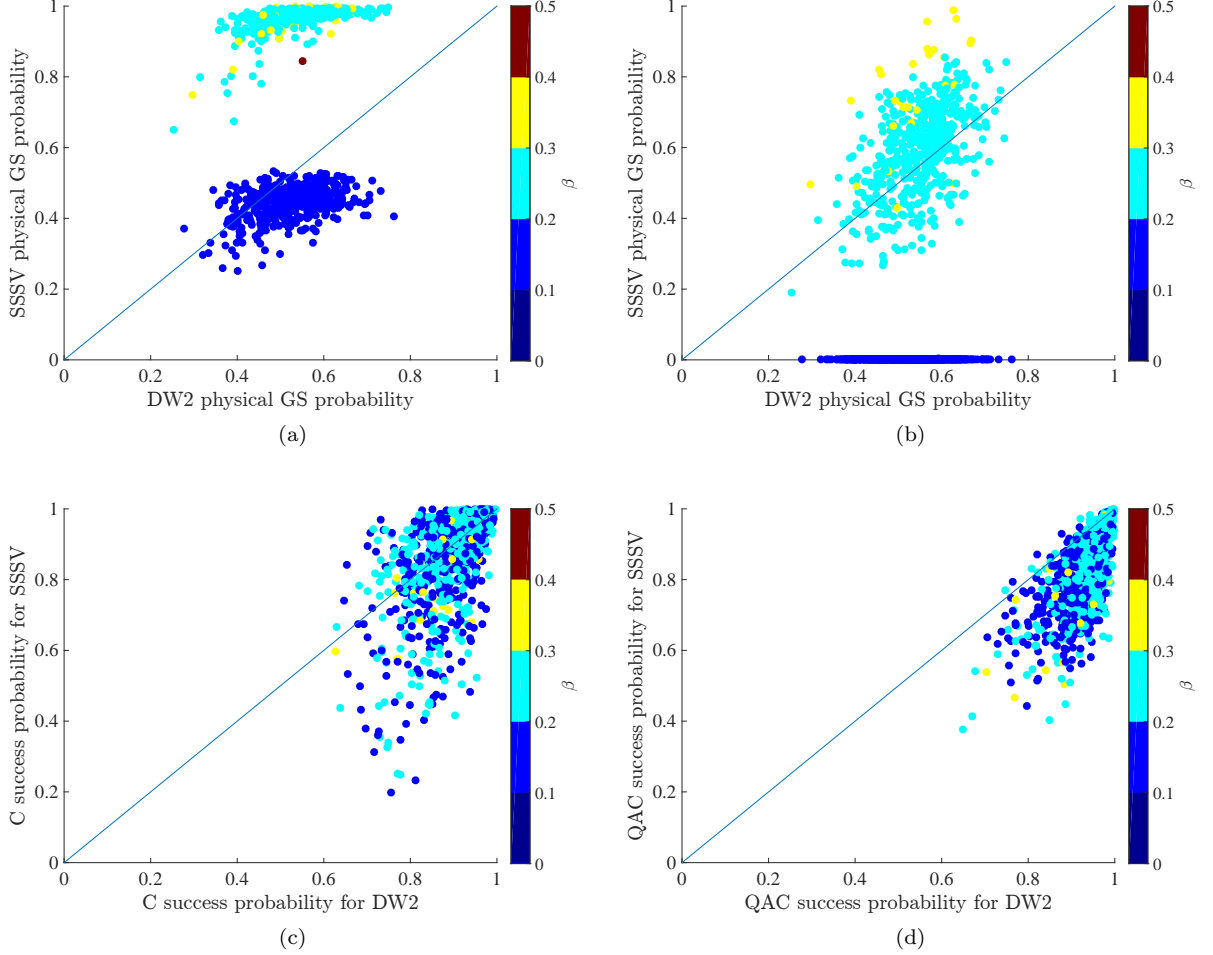


FIG. 7. (Color online) Quantumness test by comparison to the classical SSSV model. Shown are the success probability correlations for the physical or logical ground state of the DW2 *vs* SSSV results for $N = 448$ physical or $\bar{N} = 112$ logical qubits. (a) QAC, physical, without calibration noise on the couplings ($\sigma = 0$); (b) QAC, physical, with noise ($\sigma = 0.085$); (c) C, logical, with noise ($\sigma = 0.085$); (d) QAC, logical, with noise ($\sigma = 0.085$). The instances are color-coded according to their β_{opt} as used in the QAC strategy [note that in (c), there is no β but we still color-coded the instances by their QAC-optimal β values]. For QAC with $\beta_{\text{opt}} = 0$ the β_{opt} value was shifted to 0.1 to differentiate them from C. SSSV simulation parameters: temperature $T = 10.56\text{mK}$ and 1×10^5 Monte Carlo sweeps. Gaussian noise with standard deviation σ was added to the Ising couplings and local fields.

described by the SSSV model, then the QAC strategy applied to this model should give the same performance enhancement as we observed experimentally. If, on the other hand, quantum effects play an important role in the performance of the QAC strategy on the D-Wave device, then the SSSV model will not benefit from these effects.

To test this we numerically solved the SSSV model applied to the same set of instances as we used on the DW2, with the same β_{opt} values, and compared the probability of observing the physical ground state on each (details of the algorithm used here can be found in ref. [17]). As seen in Fig. 7(a), we find that the SSSV physical ground state probabilities separate the instances into two sets according to β_{opt} , with a higher probability for those with large β_{opt} values. This is unlike the DW2 results, where

the instances cluster irrespectively of the relevant β_{opt} .

Since it is known that the D-Wave device is susceptible to calibration noise and it is important to account for this when comparing its results to the SSSV model [17, 18], we checked the robustness of this conclusion by including a variable amount of Gaussian noise with standard deviation σ in the range $[0, 0.085]$ (the range found to be relevant in Ref. [17]) on the couplings J_{ij} , and also included local fields $h_i \sim \mathcal{N}(0, \sigma)$. As shown in Fig. 7(b), depending on the amount of calibration noise introduced, we can only correlate the SSSV results with the DW2 results for a subset of the instances. We find that with no calibration noise, we can correlate the $\beta_{\text{opt}} = 0.1$ instances, while with calibration noise, we can correlate those with $\beta_{\text{opt}} = 0.2$. Since we know that calibration

noise is present [17], the latter case is the more realistic fit. If we accept this conclusion, then Fig. 7(b) shows that the SSSV model almost never finds the physical ground state for the $\beta_{\text{opt}} = 0.1$ instances, while the DW2 still has a substantial probability of finding the physical ground state. While increasing the number of sweeps helps improve the SSSV success probabilities for all quantities presented, this conclusion is robust to varying the temperature and the number of sweeps (Monte Carlo updates per spin) in the range [10mK, 25mK] and [50k, 200k] respectively.

Furthermore, Fig. 7(c) shows that the separation by β_{opt} does not appear when using the C strategy, suggesting that the separation is an effect of the physical, rather than the logical Ising instance. Figure 7(d) shows that the separation vanishes when we decode using the QAC strategy, suggesting that the SSSV model suffers from many correctable errors for these instances. While these results of course do not amount to a proof of quantumness, they do support the notion that quantum effects play a relevant role in separating the SSSV model from the DW2 results observed in our QAC experiments.

The SSSV model is useful in another sense. In our earlier discussion of the scaling results shown in Figs. 4(a) and 4(b) we stressed that they do not exhibit the correct scaling curves for the DW2 device, since the minimal possible annealing time of $20\mu\text{s}$ is too long and is hence suboptimal [20]. To illustrate the importance of this point, we show in Fig. 4(d) the scaling of the noise-free SSSV model with and without optimizing the number of annealing sweeps. The scaling curves differ substantially, and one would be misled about the true scaling by the suboptimal curves. Finally, we note that the general shape of the SSSV no-optimization curves is similar to the experimental curves, indicating that the QAC strategy has reduced the effect of control errors on the DW2 device.

IV. CONCLUSIONS

We have demonstrated that QAC substantially enhances the performance of an experimental quantum an-

nealer, boosting its success probabilities on hard random Ising instances well beyond a classical repetition code strategy using equal hardware resources. Moreover, we have demonstrated that quantum effects appear to play an operational role in the success of the QAC strategy. These results demonstrate that the encouraging conclusions concerning the beneficial role of QAC based on anti-ferromagnetic chains reported in Ref. [34] extend to hard computational problems as well, with increasing benefit as problem instances grow in size and hardness.

While extrapolation of our DW2 scaling results to larger problem sizes would be inappropriate due to the issue of suboptimal annealing times, the improvement in performance relative to a simple classical repetition strategy validates the importance of QAC, especially for benchmarking studies. Future studies will explore harder problem instances, both larger and with a non-zero spin glass phase critical temperature [21], where the optimal annealing time will be greater than the minimum currently allowed by the D-Wave device, allowing us to extract its true scaling under QAC. Our work reinforces the importance of the inclusion of error correction in quantum annealing, with the ultimate goal of demonstrating a quantum speedup.

ACKNOWLEDGMENTS

This research was supported by the Lockheed Martin Corporation, by ARO-MURI Grant W911NF-11-1-0268, by ARO-QA Grant number W911NF-12-1-0523. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. Computation for the work described in this paper was supported by the University of Southern California's Center for High-Performance Computing (<http://hpcc.usc.edu>).

-
- [1] D. Lidar and T. Brun, eds., *Quantum Error Correction* (Cambridge University Press, Cambridge, UK, 2013).
 - [2] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, *Chemical Physics Letters* **219**, 343 (1994).
 - [3] T. Kadowaki and H. Nishimori, *Phys. Rev. E* **58**, 5355 (1998).
 - [4] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, *Science* **295**, 2427 (2002).
 - [5] S. Morita and H. Nishimori, *J. Math. Phys.* **49**, 125210 (2008).
 - [6] A. Das and B. K. Chakrabarti, *Rev. Mod. Phys.* **80**, 1061 (2008).
 - [7] V. Bapst, L. Foini, F. Krzakala, G. Semerjian, and F. Zamponi, *Physics Reports* **523**, 127 (2013).
 - [8] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, *Science* **292**, 472 (2001).
 - [9] F. Barahona, *J. Phys. A: Math. Gen* **15**, 3241 (1982).
 - [10] 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. Enderud, 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).
- [11] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, F. Cioata, I. Perminov, P. Spear, C. Enderud, C. Rich, S. Uchaikin, M. C. Thom, E. M. Chapple, J. Wang, B. Wilson, M. H. S. Amin, N. Dickson, K. Karimi, B. Macready, C. J. S. Truncik, and G. Rose, *Phys. Rev. B* **82**, 024511 (2010).
- [12] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, *Nat Commun* **4** (2013).
- [13] S. Boixo, T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, *Nat Phys* **10**, 218 (2014).
- [14] J. A. Smolin and G. Smith, *arXiv:1305.4904* (2013).
- [15] L. Wang, T. F. Ronnow, S. Boixo, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, (2013), *arXiv:1305.5837*.
- [16] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, *arXiv:1401.7087* (2014).
- [17] W. Vinci, T. Albash, A. Mishra, P. A. Warburton, and D. A. Lidar, *arXiv:1403.4228* (2014).
- [18] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, *arXiv:1404.6499* (2014).
- [19] 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, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, S. Uchaikin, A. B. Wilson, and G. Rose, *Physical Review X* **4**, 021041 (2014).
- [20] T. F. Ronnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, *Science* **345**, 420 (2014).
- [21] H. G. Katzgraber, F. Hamze, and R. S. Andrist, *Physical Review X* **4**, 021008 (2014).
- [22] D. Venturelli, S. Mandrà, S. Knysh, B. O’Gorman, R. Biswas, and V. Smelyanskiy, *arXiv:1406.7553* (2014).
- [23] R. D. Somma, D. Nagaj, and M. Kieferová, *Physical Review Letters* **109**, 050501 (2012).
- [24] A. Mizel, M. W. Mitchell, and M. L. Cohen, *Physical Review A* **63**, 040302 (2001).
- [25] S. P. Jordan, E. Farhi, and P. W. Shor, *Phys. Rev. A* **74**, 052322 (2006).
- [26] D. A. Lidar, *Phys. Rev. Lett.* **100**, 160506 (2008).
- [27] G. Quiroz and D. A. Lidar, *Phys. Rev. A* **86**, 042333 (2012).
- [28] K. C. Young, M. Sarovar, and R. Blume-Kohout, *Physical Review X* **3**, 041013 (2013).
- [29] M. Sarovar and K. C. Young, *New Journal of Physics* **15**, 125032 (2013).
- [30] K. C. Young, R. Blume-Kohout, and D. A. Lidar, *Physical Review A* **88**, 062314 (2013).
- [31] A. Ganti, U. Onunkwo, and K. Young, *Physical Review A* **89**, 042313 (2014).
- [32] A. Mizel, *arXiv:1403.7694* (2014).
- [33] A. D. Bookatz, E. Farhi, and L. Zhou, *arXiv:1407.1485* (2014).
- [34] K. L. Pudenz, T. Albash, and D. A. Lidar, *Nat Commun* **5** (2014).
- [35] F. Gaitan, *Quantum Error Correction and Fault Tolerant Quantum Computing* (Taylor & Francis Group, Boca Raton, 2008).
- [36] E. Knill, R. Laflamme, and W. H. Zurek, *Science* **279**, 342 (1998).
- [37] J. Preskill, *Quant. Inf. Comput.* **13**, 181 (2013).
- [38] D. Gottesman, *arXiv:1310.2984* (2013).
- [39] A. M. Childs, E. Farhi, and J. Preskill, *Phys. Rev. A* **65**, 012322 (2001).
- [40] M. S. Sarandy and D. A. Lidar, *Physical Review Letters* **95**, 250503 (2005).
- [41] M. H. S. Amin, D. V. Averin, and J. A. Nesteroff, *Phys. Rev. A* **79**, 022107 (2009).
- [42] P. I. Bunyk, E. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, and J. Whittaker, *arXiv:1401.5504* (2014).
- [43] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
- [44] S. Jansen, M.-B. Ruskai, and R. Seiler, *Journal of Mathematical Physics* **48**, (2007).
- [45] D. A. Lidar, A. T. Rezakhani, and A. Hamma, *Journal of Mathematical Physics* **50**, (2009).
- [46] M. H. S. Amin, P. J. Love, and C. J. S. Truncik, *Phys. Rev. Lett.* **100**, 060503 (2008).
- [47] N. G. Dickson, M. W. Johnson, M. H. Amin, R. Harris, F. Altomare, A. J. Berkley, P. Bunyk, J. Cai, E. M. Chapple, P. Chavez, F. Cioata, T. Cirip, P. deBuen, M. Drew-Brook, C. Enderud, S. Gildert, F. Hamze, J. P. Hilton, E. Hoskinson, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Lanting, T. Mahon, R. Neufeld, T. Oh, I. Perminov, C. Petroff, A. Przybysz, C. Rich, P. Spear, A. Tcaciuc, M. C. Thom, E. Tolkacheva, S. Uchaikin, J. Wang, A. B. Wilson, Z. Merali, and G. Rose, *Nat. Commun.* **4**, 1903 (2013).
- [48] N. Robertson and P. Seymour, *Journal of Combinatorial Theory, Series B* **36**, 49 (1984).
- [49] T. Albash, S. Boixo, D. A. Lidar, and P. Zanardi, *New J. of Phys.* **14**, 123016 (2012).
- [50] T. Lanting, D-Wave Inc., private communications (2013).
- [51] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, UK, 2001).
- [52] A. B. Zamolodchikov, *International Journal of Modern Physics A* **04**, 4235 (1989).
- [53] M. Henkel and H. Saleur, *Journal of Physics A: Mathematical and General* **22**, L513 (1989).
- [54] S. Noschese, L. Pasquini, and L. Reichel, *Numerical Linear Algebra with Applications* **20**, 302 (2013).

Appendix A: Methods

1. Experiment details

Our experiments were performed on the DW2 “Vesuvius” processor at the Information Sciences Institute of the University of Southern California. The device has been described in detail elsewhere [11, 42]. The D-Wave processors are organized into unit cells consisting of eight superconducting flux qubits arranged in a complete, balanced bipartite graph, with each side of the graph connecting to a neighboring unit cell, as seen in Fig. 2(a). The annealing schedule is shown in Fig. 8.

We ran a single copy of each problem instance 1000 times, and repeated this for each gauge. A gauge is a

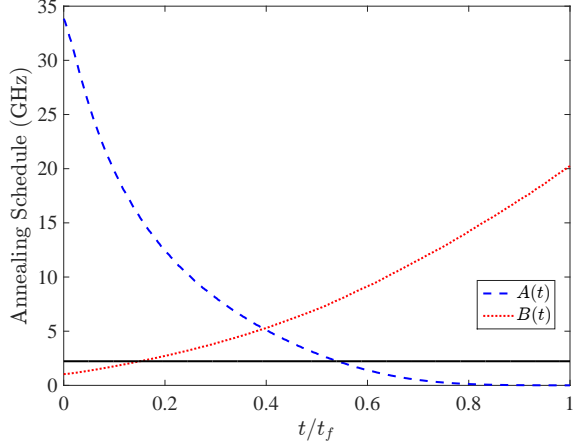


FIG. 8. (Color online) DW2 annealing schedule. The functions A and B are the ones appearing in Eq. (2). The solid horizontal black line is the operating temperature (17mK).

transformation of the couplings and fields which leaves the Ising spectrum invariant: pick $a_i = \pm 1$ at random for each spin variable i and map $J_{ij} \mapsto a_i a_j J_{ij}$, $h_i \mapsto a_i h_i$ along with $\sigma_i^z \mapsto a_i \sigma_i^z$ [13]. For each problem instance we selected the lowest energy found among all runs and gauges, and declared this to be the ground state energy. We are confident that every problem instance was solved correctly at least once since we never observed a lower energy using either the C or QAC strategies. We then implemented the C and QAC strategies as described in the main text.

2. Data analysis methodology

The following method was used to generate the number of annealing runs (R) data with their associated error bars, shown in Figs. 4 and 5. Each instance i for $\bar{N} = 112$ (86, 66, 46) was run for $G = 16$ (8) gauges, and each gauge was run $M = 1000$ times. For each gauge g , the success probability $p_{i,g}$ for QAC or C between runs may be correlated, so a binning test was performed to determine the uncertainty in $p_{i,g}$ due to the possible correlations as well as the finite number of samples. The length M sequence of successes/failures was binned into B subsequences of length $L = M/B$. The success probability $p_{i,g,l}$ of each sequence l was calculated, and the error associated with this binning was determined by:

$$\Delta p_{i,g}(L) = \frac{1}{\sqrt{L}} \sqrt{\frac{\sum_{l=1}^L (p_{i,g,l} - p_{i,g})^2}{L-1}} \quad (\text{A1})$$

If $\Delta p_{i,g}(L)$ converges as B is increased, then the converged value is taken to be the error $\Delta p_{i,g}$ associated with the success probability $p_{i,g}$. If a gauge did not yield a converged value for the error, then this gauge was discarded.

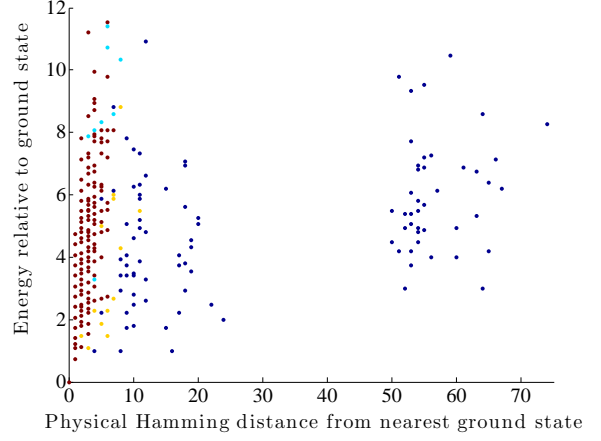


FIG. 9. (Color online) Decodability of observed states. Shown are the energy and Hamming distance relative to the ground state of all observed states in 1000 annealing cycles of one problem instance with $\bar{N} = 112$ that lies near the 95th percentile in terms of time to solution. States colored red are decodable by both logical group and problem group decoding, and tend to be low in Hamming distance. States colored light blue (yellow) are decodable by logical (problem) group only, and occupy higher Hamming distances, with logical group decodable states generally higher in energy than problem group decodable states. Dark blue states are undecodable, and cluster in groups which represent sets of logical qubits flipping together (see Appendix B).

Next, we determined the gauge-averaged number of repetitions. We performed 1000 bootstraps over the gauges, where for each gauge in each bootstrap, we calculated the number of repetitions using $\mathcal{N}(p_{i,g}, \Delta p_{i,g})$ instead of simply $p_{i,g}$. For each bootstrap we calculated the mean number of repetitions, then took the mean over the bootstraps to determine the gauge-averaged number of repetitions R_i , with the standard deviation over the bootstraps giving the error ΔR_i .

Finally for Figs. 4 and 5, we performed a bootstrap over the instances at a given \bar{N} , where for each instance in the bootstrap we used $\mathcal{N}(R_i, \Delta R_i)$. For each bootstrap we calculated the different percentiles for number of repetitions. Then the mean of the percentiles over the bootstraps is the number of repetitions at size \bar{N} with error bar given by the standard deviation of the percentiles over the bootstraps.

3. Decoding Strategies

To recover the correct solution to the encoded problem as part of the QAC strategy, we used two post-readout classical decoding methods. The methods are complementary and each has complexity linear in problem size. The first method, *logical group decoding*, is the standard scheme for decoding a repetition code. It consists of taking a majority vote over the three problem qubits within

each logical qubit, which yields a single value for each logical qubit in the original problem we seek to solve. The second method, *problem group decoding*, is equivalent to the interpretation of the results of the C method but applied only to the three copies of the problem embodied by the problem qubits, disregarding the penalty qubits. Further details of the decoding methods is discussed in Appendix B.

To illustrate the role of decoding, Fig. 9 shows all the states that were observed in 1000 annealing cycles of a particular problem instance. We observe that successful decoding depends strongly on the errors having low Hamming weight, but that a significant tolerance to excitations is permissible. This illustrates that under QAC the standard adiabatic criterion of remaining in the ground state is relaxed and replaced by maintaining correctness.

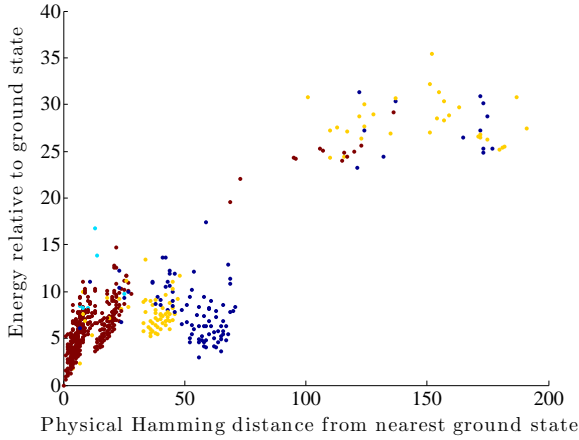


FIG. 10. Decodability of observed states for a problem instance where C outperforms QAC. States are colored by decodability as in Fig. 9. Note the large cluster of undecodable (dark blue) states at relatively low energy but at high Hamming distance from the ground state. Here $\beta_{\text{opt}} = 0.1$.

4. Explanation for cases where C outperforms QAC

As shown in Fig. 3, there is a small subset of instances where the C strategy outperforms the QAC strategy. Of this subset, only 5 instances have $\beta_{\text{opt}} = 0$. For these instances C outperforms QAC simply because it actually contains four copies of the encoded problem *vs* three copies for QAC. For the remainder of the instances in this subset, having $\beta_{\text{opt}} > 0$, the QAC strategy improves the success probability relative to simply using three copies of the encoded problem.

Why then does QAC not provide a sufficient improvement to outperform C in these cases? As seen in Fig. 2(c), there are holes in the logical connectivity graph due to the fact that we used only perfect logical qubits in our experiments. Qubits around these holes have fewer cou-

plings and therefore the energy penalty is lower for violated couplings than for qubits in the intact areas of the graph. This is exacerbated when these couplings are weak, causing the energy cost for a logical error to be small especially if the physical qubits are tied together by a finite β_{opt} , as in the case of the QAC strategy. This leads to a pronounced increase in undecodable, low-energy eigenstates, as depicted in Fig. 10. It turns out that this is typical, i.e., instances where C outperforms QAC are dominated by undecodable, low-energy excited states, arising from the flipping of large clusters of logical qubits that are weakly tied to the rest. This can be compared to what happens in the more common case, where QAC outperforms C, as shown in Fig. 9.

Appendix B: Error types and decodability

In order to more closely examine the effects of the QAC encoding and the capabilities of the problem group and logical group decoding methods, it is instructive to consider a representative problem instance. Fig. 9 of the main text shows all of the states that were observed in 1000 annealing cycles of an $N = 112$ problem instance that lies near the 95th percentile in terms of time to solution, colored by the decoding method they yielded to. Looking deeper, Fig. 11 shows sample states from the decodability categories of Fig. 9 of the main text, illuminating the error mechanisms for which each type of decoding is particularly suited.

Overall, we see that both decoding methods succeed for low Hamming weight thermal errors. While logical group decoding performs well for excited states consisting of a large number of random errors, problem group decoding is better equipped to address errors that are correlated through the problem coupling terms present in the QAC Hamiltonian. Although logical (rather than physical) errors are not expected to yield to any trivial-complexity postprocessing decoding method like the ones examined here, it is possible that a strategy involving some kind of local classical optimization may allow even these states to be recovered.

Appendix C: Robustness to physical qubit loss

In the main text we presented results demonstrating the robustness of QAC to physical qubit loss. Specifically, we investigated the case where the missing physical qubits are few enough to be embedded as the penalty qubits within logical qubits, so that all problem couplings would remain intact. To this end, we performed quantum annealing on the 1000 random problem instances of each size again, but this time we removed 30%, 60%, then finally 90% of the penalty qubits from each instance at random.

Figure 12 shows additional percentiles (median and 75th) supplementing Fig. 5 of the main text (95th). We

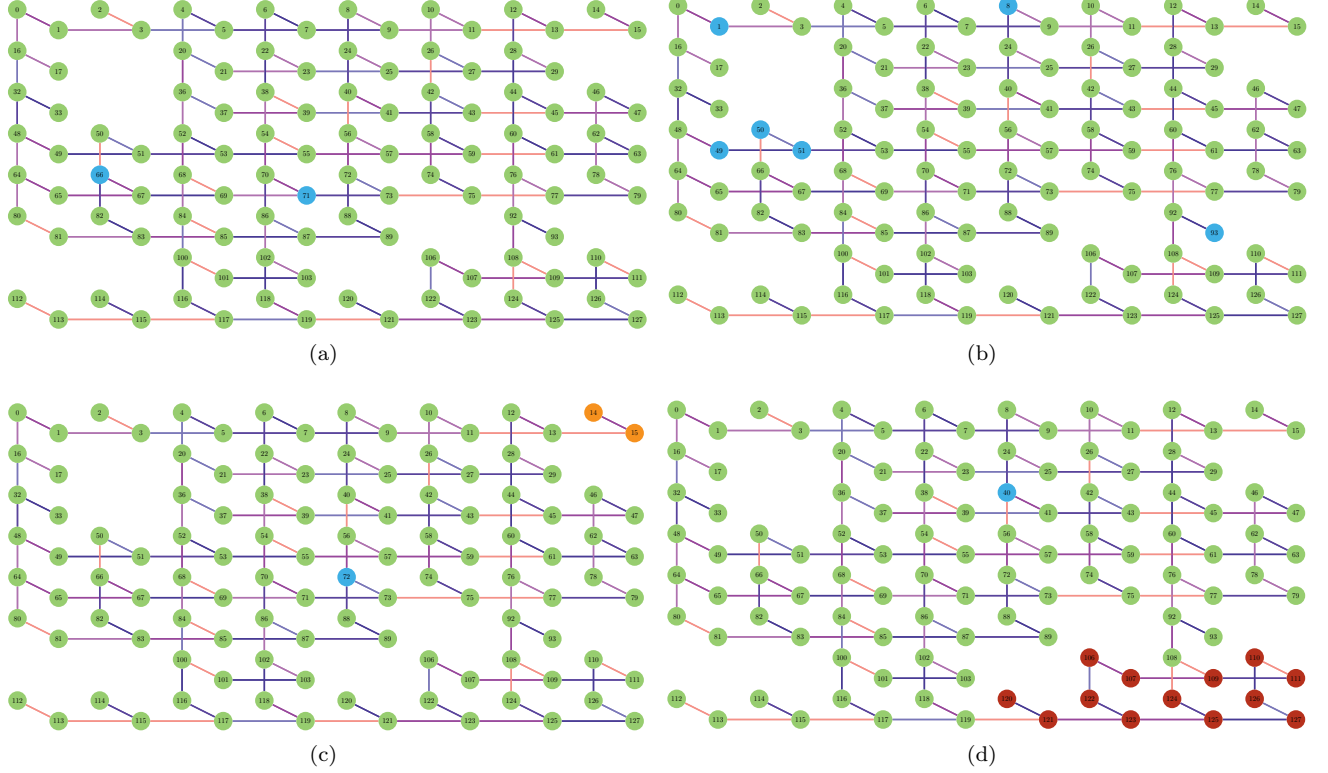


FIG. 11. (Color online) Examples of error types and decodability. All states shown are experimentally observed examples. Green circles are correct logical qubits, i.e., with all their physical qubits aligned with the ground state; blue, orange and red circles are logical qubits with, respectively, one, two, and three bit flip errors. The magnitude (though not the sign) of the problem couplings is color-coded: pink lines indicate $|J_{ij}| = \frac{1}{6}$, and the shade of the line darkens through gradations to indigo for $|J_{ij}| = 1$. (a) A state decodable by both logical group decoding and problem group decoding. In this typical example we observe two logical qubits with a single bit flip error each (blue circles) among otherwise correct logical qubits (green circles). The state is logical group decodable because single bit flips are majority-vote correctable. It is also problem group decodable because there are only two bit flips, which is not enough to ruin all three copies of the problem embedded within the QAC scheme. (b) A logical group decodable state. The 6 single bit flips are all decodable via majority vote so the state is logical group decodable, but the state is not problem group decodable because each problem group is corrupted by at least one bit flip. (c) A problem group decodable state. This state has two logical qubits in the upper right hand corner (orange circles) that are loosely coupled to the rest of the problem (pink line, $|J_{ij}| = \frac{1}{6}$) and which each have two physical problem qubits flipped from the ground state values. The problem qubits that flipped are correlated between the two logical qubits; they belong to the same copy of the problem because the problem coupling is strong between counterpart problem qubits (purple line). This leaves one copy of the problem fully intact, and the state can be decoded using problem group decoding. (d) An undecodable state. There is a cluster of logical qubit flips (red circles) in the lower right hand corner. That region is loosely coupled to the rest of the problem; all links going out of it are weak (pink) couplings. This means that the state with these logical qubits flipped together is a low-lying final excited state of the logical Ising problem, which has been suppressed via the repetition energy scale enhancement portion of QAC but is still observable in the problem instance's output statistics. This state belongs to the cluster observable near Hamming weight 20 in Fig. 9 in the main text.

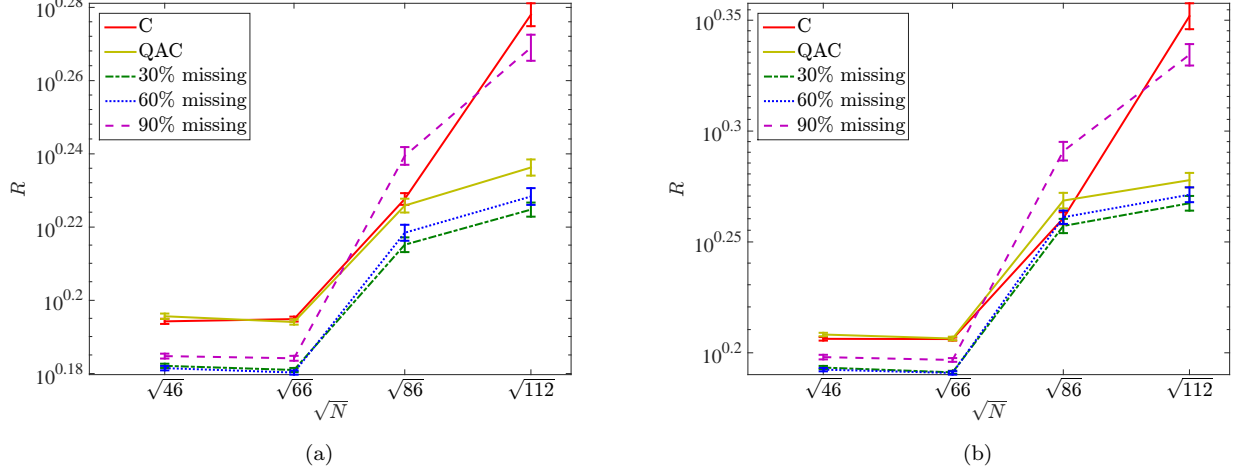


FIG. 12. (Color online) Additional percentiles supplementing Fig. 5 of the main text showing the effect of missing penalty qubits. Effect of missing penalty qubits for $\alpha = 1$ on the 50th and 75th percentiles for β optimized for the number of missing qubits. The C strategy (solid red) and QAC strategy (solid yellow) lines are the same data that was displayed in Fig. 4(a). The dash-dotted green, dotted blue, and dashed purple series show the effects of randomly removing 30%, 60%, and 90% of the penalty qubits, respectively. The performance at 30% and 60% loss tracks the original QAC closely, suggesting the code is highly resilient to penalty qubit loss if the penalty magnitude is adjusted accordingly.

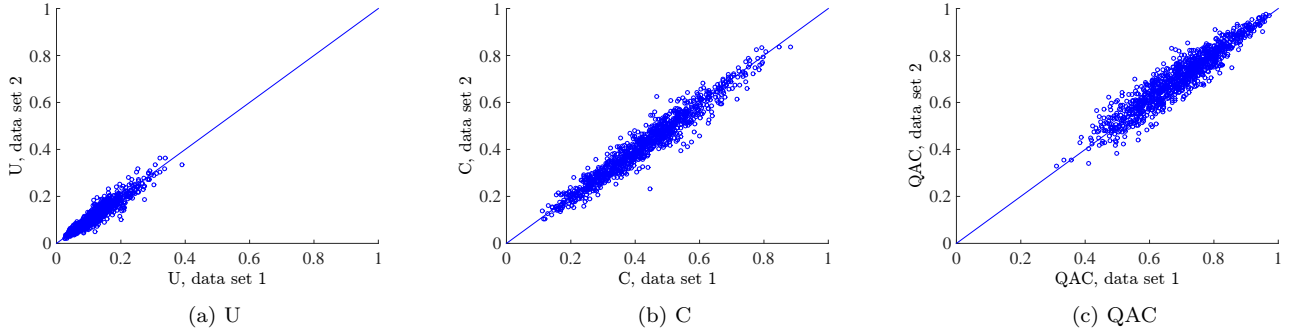


FIG. 13. Correlation of different $\alpha = 0.5$, $\bar{N} = 86$ data sets. (a) The U case, i.e., a single copy of the problem without any error correction. (b) The C strategy. (c) The QAC strategy. In all cases the correlation is excellent, with the corresponding Pearson correlation coefficients being $\rho = 0.942$ for U, $\rho = 0.969$ for C, and $\rho = 0.941$ for QAC.

observe that the separation between the C and QAC strategies at $\bar{N} = 112$ persists even when up to 60% of the penalty qubits are removed.

Appendix D: Correlation tests

Since our data collection lasted several weeks, we checked the stability of our results by performing two separate sets of experiments for $\alpha = 0.5$ and $\bar{N} = 86$, separated by 25 days. We computed the Pearson correlation coefficient, $\rho_{XY} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y}$ (the covariance of the two variables X and Y divided by the product of their standard deviations) for X and Y being the first and second data set, respectively. The results are shown

in Fig. 13, confirming that our data collection procedure was stable over time.

In a separate test we checked the correlation between the C success probability obtained by actually running 4 copies of perfect logical qubits in parallel and the theoretical C success probability obtained by running a single copy, and using it in the binomial expression for the success probability of 4 independent copies, i.e., $1 - (1 - p_{U,i})^4$, where $p_{U,i}$ is the single copy success probability for instance i . In more detail, to account for gauge averaging the procedure we used is the following. Let $p_{U,g,i}$ denote the experimentally observed success probability within the U case for a given instance i and gauge g . Then $p_{U,i} = \frac{1}{G} \sum_{g=1}^G 1 - (1 - p_{U,g,i})^4$ is the theoretical value of C from the U case, where we gauge-average the

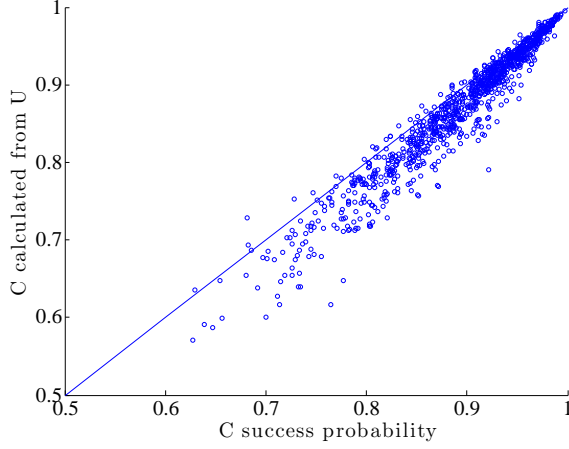


FIG. 14. Correlation between actual C and theoretical C. The plot shows the probabilities for 1000 random Ising instances with $\bar{N} = 112$ perfect logical qubits, computed using the C strategy and from a single unprotected copy (U), i.e., $p_{C,i}$ vs $p_{U,i}$ for each instance i . A strong correlation is observed, with Pearson correlation coefficient of 0.968.

success probabilities. The result is shown in Fig. 14 and the correlation is high, confirming that the 4 copies in the actual C strategy are to a large degree independent. Deviations are likely due to residual cross-talk effects [17]. In the main text we used $p_{C,i}$ in our comparisons of the C and QAC strategy, rather than $p_{U,i}$. Since as can be seen in Fig. 14 for most instances $p_{C,i} > p_{U,i}$, this is a more stringent test of the QAC strategy. The procedure for computing $p_{C,i}$ is the same as outlined in the Methods section for the gauge-averaged number of repetitions R_i .

Appendix E: Energy gap enhancement in a solvable transverse field Ising model on a ring

Here we provide details supporting the discussion of the Ising Hamiltonian on a ring with local fields presented in the main text. We first briefly review the model.

Let \mathcal{P}_d be the Pauli group over d qubits, and consider an Abelian subgroup $\mathcal{S} \in \mathcal{P}_d$ with $d - 1$ generators g_i given by:

$$g_i = \sigma_i^z \sigma_{i+1}^z, i \in \{1, \dots, d - 1\}. \quad (\text{E1})$$

This defines a distance d repetition stabilizer code that detects all bit flips σ_i^x since they anti-commute with at least one of the stabilizer generators in Eq. (E1). The code comprises a single logical qubit with codespace spanned by $|0\rangle = |0 \dots 0\rangle$ and $|1\rangle = |1 \dots 1\rangle$. We can choose the logical operators as $\bar{X} = \otimes_{i=1}^d \sigma_i^x$ and $\bar{Z} = \frac{1}{d} \sum_{i=1}^d \sigma_i^z$.

Consider a single qubit Hamiltonian $H_{\text{Ising}} = -h\sigma^z$ (with ground state denoted by $|0\rangle$ and excited state given

by $|1\rangle$) and encode it using the above repetition code. Rather than replacing σ^z by \bar{Z} we use $\bar{\sigma}^z \equiv d\bar{Z}$:

$$\bar{H}_{\text{Ising}} = -h \sum_{i=1}^d \sigma_i^z. \quad (\text{E2})$$

We can add to this Hamiltonian a penalty which is the sum over the generators g_i and their product $\prod_{i=1}^{d-1} g_i = \sigma_1^z \sigma_d^z$. With the convention $\sigma_{d+1}^z = \sigma_1^z$ the penalty term is then

$$H_P = -\beta \sum_{i=1}^d \sigma_i^z \sigma_{i+1}^z, \quad (\text{E3})$$

which describes a one-dimensional Ising chain with periodic boundary conditions, and penalizes all single bit flip errors. We use this as the final Hamiltonian of our quantum annealing algorithm. The total time-dependent Hamiltonian is that of a transverse field Ising model on a ring with a local field:

$$H(s) = -A(s) \sum_{i=1}^d \sigma_i^x + B(s) (\bar{H}_{\text{Ising}} + H_P). \quad (\text{E4})$$

For convenience we use the dimensionless time variable $s = t/t_f$. Note that $A(s)$ and $B(s)$ have dimensions of energy while h and β are both dimensionless. Also note that in order to encode the entire quantum annealing algorithm [25] we would have had to replace the transverse field term by \bar{X} . However, \bar{X} operator is a d -weight operator, which is not physically available.

We proceed to analyze $H(s)$ by treating the penalty term as a perturbation that is switched on at the end of the annealing evolution, with time going backward. Thus both β and s are considered small parameters.

Unperturbed Hamiltonian

Consider first the $d = 1$ case (single qubit Hamiltonian)

$$H(s) = -A_0(1-s)\sigma^x - A_0 s h \sigma^z, \quad (\text{E5})$$

where for simplicity we have additionally set $A(s) = A_0(1-s)$ and $B(s) = A_0 s$. This can be easily diagonalized to yield the (dimensionless) eigenenergies

$$\varepsilon_{\pm}/A_0 = \pm \sqrt{(1-s)^2 + h^2 s^2} \equiv \pm \lambda(s), \quad (\text{E6})$$

with respective orthonormal eigenstates

$$|\varepsilon_+(s)\rangle = \frac{1}{c_+(s)} \left[\frac{hs - \lambda(s)}{1-s} |0\rangle + |1\rangle \right], \quad (\text{E7a})$$

$$|\varepsilon_-(s)\rangle = \frac{1}{c_-(s)} [(hs + \lambda(s)) |0\rangle + (1-s) |1\rangle], \quad (\text{E7b})$$

where $c_{\pm}(s)$ are normalization constants. When we make d copies of this system

$$H(s) = -A_0(1-s) \sum_{i=1}^d \sigma_i^x - A_0 s h \sum_{i=1}^d \sigma_i^z, \quad (\text{E8})$$

the ground state is given by $|\varepsilon_{\text{GS}}(s)\rangle = \otimes_{i=1}^d |\varepsilon_-(s)\rangle$ with energy $\varepsilon_{\text{GS}}^{(0)}(s) = -A_0 d \lambda(s)$, and the d -fold degenerate first excited states are given by $|\varepsilon_k(s)\rangle = \otimes_{i=1}^{k-1} |\varepsilon_-(s)\rangle_i \otimes |\varepsilon_+(s)\rangle \otimes_{i=k+1}^d |\varepsilon_-(s)\rangle_i$, i.e., k labels which qubit is in the excited state. These first excited states have energy $-(d-2)A_0 \lambda(s)$, so the unperturbed gap is $\Delta^{(0)}(s) = 2A_0 \lambda(s)$. This gap is minimized at

$$s_{\min}^{(0)} = \frac{1}{1+h^2}, \quad (\text{E9})$$

where

$$\Delta_{\min}^{(0)} \equiv \Delta^{(0)}(s_{\min}^{(0)}) = 2A_0 \frac{|h|}{\sqrt{1+h^2}}. \quad (\text{E10})$$

Perturbation

We now introduce the penalty term $H_P = -\beta s V$ as a perturbation in βs , with $V = \sum_{i=1}^d \sigma_i^z \sigma_{i+1}^z$. From first order perturbation theory, the corrected dimensionless ground state energy is

$$\begin{aligned} \varepsilon_{\text{GS}}(s)/A_0 &= \varepsilon_{\text{GS}}^{(0)}(s)/A_0 - \beta s \langle \varepsilon_{\text{GS}}(s) | V | \varepsilon_{\text{GS}}(s) \rangle \\ &= -d\lambda(s) - d\beta s \langle \varepsilon_-(s) | \sigma^z | \varepsilon_-(s) \rangle^2. \end{aligned} \quad (\text{E11})$$

For the first excited states, we employ first order degenerate perturbation theory, whereby we need to calculate the projected perturbation PVP , where $P = \sum_k |\varepsilon_k\rangle \langle \varepsilon_k|$ projects on the ground state, i.e., the matrix elements $\sum_i \langle \varepsilon_k | \sigma_i^z \sigma_{i+1}^z | \varepsilon_{k'} \rangle$. The only non-zero matrix elements are

$$\begin{aligned} \sum_i \langle \varepsilon_k | \sigma_i^z \sigma_{i+1}^z | \varepsilon_k \rangle \\ = (d-2) \langle \varepsilon_- | \sigma^z | \varepsilon_- \rangle^2 + 2 \langle \varepsilon_+ | \sigma^z | \varepsilon_+ \rangle \langle \varepsilon_- | \sigma^z | \varepsilon_- \rangle \equiv a, \end{aligned} \quad (\text{E12a})$$

$$\begin{aligned} \sum_i \langle \varepsilon_k | \sigma_i^z \sigma_{i+1}^z | \varepsilon_{k+1} \rangle &= \sum_i \langle \varepsilon_{k-1} | \sigma_i^z \sigma_{i+1}^z | \varepsilon_k \rangle \\ &= \langle \varepsilon_- | \sigma^z | \varepsilon_+ \rangle^2 \equiv b. \end{aligned} \quad (\text{E12b})$$

This defines a translationally invariant $d \times d$ tridiagonal matrix (with corner terms):

$$\begin{pmatrix} a & b & 0 & 0 & \dots & b \\ b & a & b & 0 & \dots & 0 \\ 0 & b & a & b & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \\ b & 0 & \dots & 0 & b & a \end{pmatrix}. \quad (\text{E13})$$

The eigenvalues of this matrix can be found analytically and are given by [54]:

$$\lambda_n = a + 2|b| \cos\left(\frac{2\pi n}{d}\right), \quad n = 0, \dots, d-1. \quad (\text{E14})$$

Thus the corrected dimensionless first excited state energy is $\varepsilon_1(s)/A_0 = -(d-2)\lambda(s) - \beta s \max_n \lambda_n = -(d-2)\lambda(s) - \beta s(a + 2|b|)$. Therefore, to first order in perturbation theory the dimensionless gap $\Delta(s)/A_0 = [\varepsilon_1(s) - \varepsilon_{\text{GS}}(s)]/A_0$ to the first excited state is given, after some algebra, by:

$$\Delta(s)/A_0 = 2\lambda(s) - 2\beta s \frac{(1-s)^2 - 2(hs)^2}{\lambda(s)^2}. \quad (\text{E15a})$$

To find the minimum of this function we expand s_{\min} (the location of the minimum) to first order in β , keeping in mind that we are interested in the end of the evolution (near $s = 0$) and the limit where $\beta \ll h$. Thus, writing $s_{\min} = s_0 + s_1\beta + O(\beta^2)$ we find that $d\Delta(s)/ds = 0$ to first order in β at

$$s_{\min} = s_{\min}^{(0)} - \beta s_1 + O(\beta^2), \quad (\text{E16a})$$

$$s_1 = \frac{|h|(8-h^2)}{(1+h^2)^{5/2}}, \quad (\text{E16b})$$

where $s_{\min}^{(0)}$ is the unperturbed value. At this s_{\min} we have

$$\Delta_{\min} \equiv \Delta(s_{\min}) = \Delta_{\min}^{(0)} + \Delta_{\min}^{(1)} + O(\beta^2), \quad (\text{E17a})$$

$$\Delta_{\min}^{(1)} = A_0 \beta \frac{2(2-h^2)}{(1+h^2)^2}, \quad (\text{E17b})$$

where likewise $\Delta_{\min}^{(0)}$ is the unperturbed value. Since $\Delta_{\min}^{(1)} > 0$ if $h < \sqrt{2}$, and likewise then $s_1 > 0$, we find that for sufficiently small h a perturbative β has the effect of both increasing the minimum gap as well as shifting it to earlier in the evolution.