Simulated-quantum-annealing comparison between all-to-all connectivity schemes

Tameem Albash, Walter Vinci, and Daniel A. Lidar

Simulated Quantum Annealing Comparison between All-to-All Connectivity Schemes

Tameem Albash,[1,2,3,*] Walter Vinci,[1,2,4,*] and Daniel A. Lidar[1,2,4,5]

[1]*Department of Physics and Astronomy, University of Southern California, Los Angeles, California 90089, USA*
[2]*Center for Quantum Information Science & Technology,*
University of Southern California, Los Angeles, California 90089, USA
[3]*Information Sciences Institute, University of Southern California, Marina del Rey, California 90292, USA*
[4]*Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA*
[5]*Department of Chemistry, University of Southern California, Los Angeles, California 90089, USA*

Quantum annealing aims to exploit quantum mechanics to speed up the solution to optimization problems. Most problems exhibit complete connectivity between the logical spin variables after they are mapped to the Ising spin Hamiltonian of quantum annealing. To account for hardware constraints of current and future physical quantum annealers, methods enabling the embedding of fully connected graphs of logical spins into a constant-degree graph of physical spins are therefore essential. Here, we compare the recently proposed embedding scheme for quantum annealing with all-to-all connectivity due to Lechner, Hauke and Zoller (LHZ) [Science Advances 1 (2015)] to the commonly used minor embedding (ME) scheme. Using both simulated quantum annealing and parallel tempering simulations, we find that for a set of instances randomly chosen from a class of fully connected, random Ising problems, the ME scheme outperforms the LHZ scheme when using identical simulation parameters, despite the fault tolerance of the latter to weakly correlated spin-flip noise. This result persists even after we introduce several decoding strategies for the LHZ scheme, including a minimum-weight decoding algorithm that results in substantially improved performance over the original LHZ scheme. We explain the better performance of the ME scheme in terms of more efficient spin updates, which allows it to better tolerate the correlated spin-flip errors that arise in our model of quantum annealing. Our results leave open the question of whether the performance of the two embedding schemes can be improved using scheme-specific parameters and new error correction approaches.

I. INTRODUCTION

Many important and hard optimization problems can be mapped to finding the ground state of classical Ising models [1, 2]. The observation that such models can also be solved via quantum annealing [3–5] has spurred a great deal of recent interest in building special-purpose analog quantum devices [6, 7] with the hope of realizing quantum speedups [8]. In general, all pairs of spins in the Ising Hamiltonian can be coupled, resulting in long-range interactions, e.g., as in the Sherrington-Kirkpatrick model [9–11]. A direct physical implementation of such spin systems would require all-to-all connectivity, an impossibility for analog devices with local hardware connectivity. Instead one must embed the logical problem defined by the given Ising model into the available physical device connectivity. This embedding represents long-range logical interactions in terms of short-range physical interactions, a process which introduces unavoidable tradeoffs between energy scales, hardware resources, and accuracy [12–18].

The first embedding technique used for quantum annealing is due to Choi and involves a graph-theoretical construction known as minor embedding (ME) [12, 13]. In ME, the spins of the given Ising problem with long-range interactions, which we refer to as the logical spins, are replaced by chains of physical spins with short-range interactions implemented on the device hardware. Physical spins within a chain are induced to behave as a single logical spin using strong ferromagnetic interactions acting as energy penalties. The hardware connectivity graph should allow for the minor embedding of complete graphs, i.e., the latter should be obtainable from the former via edge contractions. A well known example is the "Chimera" graph used in the D-Wave quantum annealing devices [19].

Recently, Lechner, Hauke and Zoller (LHZ) proposed an elegant alternative embedding technique, realized on a two-dimensional triangular-shaped grid [17]. In the LHZ scheme, the relative alignment (parallel or antiparallel) of pairs of logical spins is mapped to a physical hardware spin. Both the logical local fields and the logical couplings are mapped to local fields acting on the physical spins. This is appealing, since it places the burden of implementing the logical problem entirely on the implementation of local fields, thus obviating the need for simultaneous control of local fields and couplings. Couplings between the physical spins (acting as energy penalties) must still be introduced to ensure that the mapping is consistent.

Since in the LHZ scheme the logical spins are encoded redundantly and non-locally in the physical degrees of freedom, it was suggested in Ref. [17] that the scheme includes an intrinsic fault tolerance, with some similarities to the error robustness of topological quantum memories [20]. This has been confirmed by Pastawski and Preskill (PP) under the assumption of a noise model of random spin-flips [21]. More specifically, PP pointed out that the LHZ scheme may be viewed as a classical low-density parity-check code that makes the scheme highly robust against weakly correlated spin-flip noise.

While allowing for a rigorous analytical treatment, weakly correlated random spin-flips are unfortunately not the most relevant errors in quantum annealing, the algorithm for which the LHZ scheme was proposed. In adiabatic quantum com-

puting and quantum annealing, a time-dependent Hamiltonian evolves a system from an easily prepared and trivial ground state to the non-trivial ground state of the Hamiltonian of interest. In a closed system, the adiabatic theorem provides a guarantee that the probability of dynamical excitations can be made arbitrarily small if the evolution is sufficiently slow relative to the inverse of the minimum gap [22–24]. Under the favorable assumption of weak system-environment coupling, decoherence in open-system, finite temperature quantum annealing takes place in the instantaneous energy eigenbasis [25–27]. In both the closed and open-system cases, errors may occur throughout the evolution, resulting in the generation of final-time excited states that differ from the ground state by a large number of spin-flips, that are neither random nor weakly correlated [28–31], as also recognized in Ref. [21].

Motivated by these considerations, here we study whether the ME or the LHZ scheme is preferable under a realistic noise model for quantum annealing. This is particularly pertinent since the LHZ scheme would require the development of a new quantum annealing architecture, while the ME scheme has already been experimentally implemented in numerous studies, e.g., Refs. [32–39]. We employ simulated quantum annealing (SQA) [40, 41], a quantum Monte Carlo method that iteratively updates an approximation to the instantaneous quantum Gibbs state governed by the time-dependent system Hamiltonian [Eq. (2) below]. While it is not a completely faithful description of the annealing process since it neglects unitary dynamics, it is the method of choice for large-scale simulations of stoquastic (sign-problem free) quantum annealing [42], and has been used to successfully describe experiments on the D-Wave devices [43–47]. Our SQA simulations aim to model the behavior of two quantum annealing devices implementing the same annealing schedule with identical physical parameters, but with different Hamiltonians, representing the ME and LHZ embedded problem Hamiltonians, respectively. In addition to SQA, we also use parallel tempering (PT) [48, 49], a highly efficient variant of classical Monte Carlo that enables us to focus purely on the classical Hamiltonian, independently of the quantum evolution that takes place during the annealing.

For the range of parameters and problem sizes studied, we find that the ME scheme outperforms the LHZ scheme for the majority of instances studied. Using PT simulations on the final (classical) Hamiltonian, we show that the two schemes perform similarly when the number of updates is sufficiently large. We interpret this to mean that the difference between the two schemes arises from more efficient SQA updates for the ME Hamiltonian, arising during the evolution governed by the intermediate (quantum) Hamiltonian. Our results imply that the LHZ scheme does not necessarily exhibit an intrinsic fault tolerance against a noise model that is appropriate for quantum annealing.

The rest of this paper is organized as follows. In Sec. II we describe the ME and LHZ schemes for embedded quantum annealing in more detail, as well as errors and majority vote decoding. We present our SQA and PT results in Sec. III, where we compare the ME and LHZ schemes subject to majority vote decoding. In Sec. IV we consider several other decod-

ing strategies for the LHZ scheme. In particular we describe how to map the decoding of the LHZ scheme to the decoding of a fully-connected, quadratic Sourlas code [50, 51].This allows us to define minimum-weight (MWD) and maximum-likelihood (MLD) decoding strategies, the latter being optimal in the case of errors generated by random spin-flips. We also consider the belief propagation strategy proposed in Ref. [21]. We conclude and provide a broader discussion of our results in Sec. V. Additional details are provided in the Appendix.

II. EMBEDDED QUANTUM ANNEALING

In quantum annealing one encodes the solution of a hard combinatorial optimization problem into the ground state of a classical Ising Hamiltonian:

$$H_{\mathrm{P}} = \sum_i h_i \sigma_i^z + \sum_{i<j} J_{ij} \sigma_i^z \sigma_j^z \ . \tag{1}$$

Hard problems typically result in highly frustrated spin systems with a spin-glass phase at low temperatures [52]. A quantum annealing device attempts to solve for the ground state of Eq. (1) by implementing the following time-dependent Hamiltonian

$$H(t) = A(t)H_X + B(t)H_{\mathrm{P}} \ , \qquad t \in [0, t_f] \ , \tag{2}$$

where $H_{\mathrm{X}} = -\sum_i \sigma_i^x$ is the "driver" term whose ground state is a uniform superposition in the computational basis that initializes the computation. The annealing schedule is specified by the functions $A(t)$ and $B(t)$, where typically $A(0) \gg B(0)$ and $A(t_f) \ll B(t_f)$. The most general optimization problem on N binary variables is defined on an all-to-all connectivity problem Hamiltonian H_{P} graphically represented in Fig. 1(a).

The construction of a physical device that implements the system described by Eq. (2) with all-to-all connectivity can be technically demanding because it requires the implementation of long-range interactions. While in trapped ions long-range interactions are unproblematic [53], not all physical implementations support such interactions; e.g., in most superconducting and semiconducting devices interactions are hard-wired and quasi-local [19, 54–56], and for ultracold gases in optical lattices interactions are determined by the geometry [57]. Schemes that approximate long-range interactions with short-range interactions are thus desirable for certain implementations. The ME and LHZ schemes are both designed to address this challenge, as we briefly review next.

A. Minor Embedding Scheme

The ME scheme replaces long-range interactions between logical qubits by short-range interactions between chains (or clusters) of physical qubits, where each chain represents a logical qubit, and correspondingly replaces the logical problem Hamiltonian H_{P} [Eq. (1)] by a minor-embedded Hamiltonian
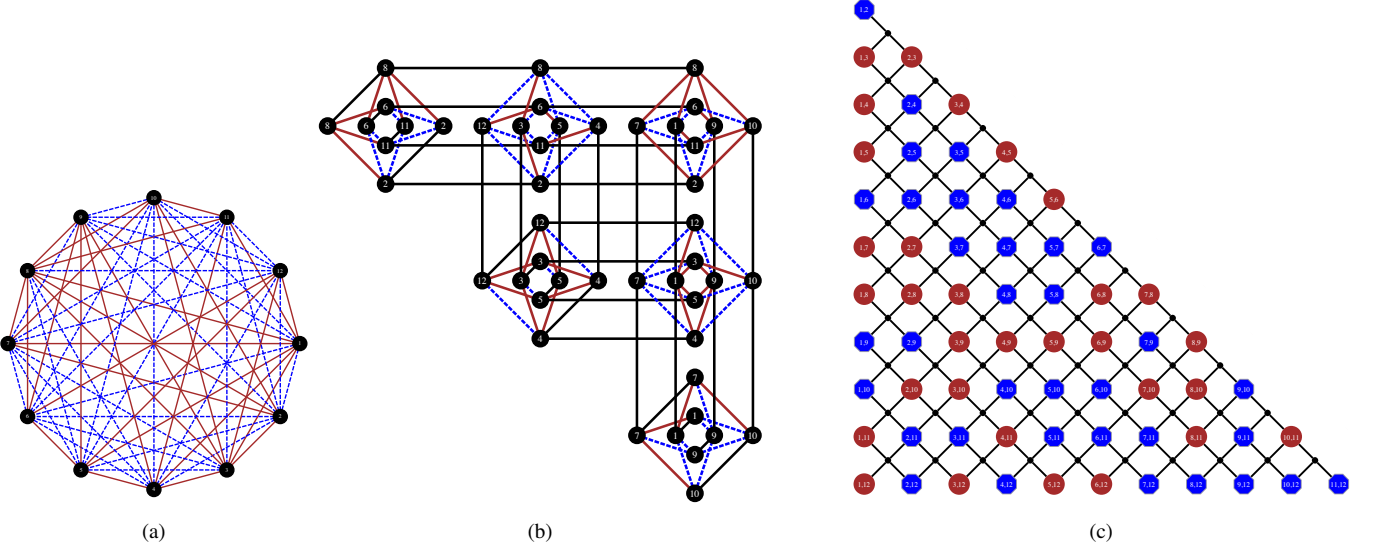
FIG. 1. **All-to-all connectivity and its representation using the ME and LHZ schemes.** (Color online) (a) A complete graph of 12 logical qubits. Brown (solid) and blue (dashed) lines correspond to types of couplings (e.g., ferromagnetic and antiferromagnetic). (b) ME of the all-to-all Hamiltonian shown in (a). Brown (solid) and blue (dashed) lines are the physical realizations of the logical couplings. Black lines are the energy penalties within the logical chains. (c) LHZ implementation of the same Hamiltonian. Red circles and blue octagons represent physical qubits with either negative (ferromagnetic) or positive (antiferromagnetic) local fields. Small black dots and lines represent three and four-body energy penalties.

$H_{\mathrm{P}}^{\mathrm{ME}}$. Physical qubits belonging to the same logical-qubit chain are connected by strong ferromagnetic interactions (acting as energy penalties), which lowers the energy of configurations where these qubits are aligned at the end of the anneal. "Broken" chains (where not all members of a chain are aligned) do not correspond to logical states and must be corrected, e.g., by majority vote. For a more formal and complete exposition of ME see, e.g., Ref. [16].

We consider here a particular ME on "Chimera" graphs. Chimera is a degree-six graph formed from the tiling of an $L \times L$ grid of $K_{4,4}$ cells. Despite its low degree and near planarity, an $L \times L$ Chimera graph allows for the minor embedding of an all-to-all logical problem H_{P} defined on $N = 4L$ logical qubits where each logical qubit is represented by a chain of $\lceil N/4 \rceil + 1$ physical qubits. Figure 1(b) depicts a minor embedded Hamiltonian $H_{\mathrm{P}}^{\mathrm{ME}}$ of the all-to-all Hamiltonian H_{P} of Fig. 1(a), on a Chimera subgraph. Blue and red thin links correspond to physical couplers representing the logical interactions of H_{P}, while black couplers represent strong ferromagnetic interactions that couple physical qubits belonging to the same logical chain. Note in particular that while the full Chimera graph would have $8L^2 = N^2/2$ spins, the embedding of the complete graph of size N only uses $N(\lceil N/4 \rceil+1)$, i.e., about half of these spins. For a more detailed description of the ME scheme on Chimera graphs see, e.g., Refs. [12, 13].

B. Lechner-Hauke-Zoller Scheme

In the LHZ scheme, the embedded Hamiltonian $H_{\mathrm{P}}^{\mathrm{LHZ}}$ is defined on $K = \binom{N}{2}$ physical qubits corresponding to the K logical interactions of the original problem Hamiltonian H_{P}. The logical couplings $J_{ij} \equiv J_k$ are mapped to local fields applied to the physical qubits (we denote the physical qubit's value in the computational basis by $q_k \equiv q_{i,j} = \pm 1$). The value of a physical qubit in the LHZ scheme encodes the relative alignment of the corresponding pair of logical qubits: a physical qubit pointing up (down) corresponds to an aligned (anti-aligned) logical pair. Since $K > N$, the physical system includes physical states that do not correspond to logical states, similarly to what happens in the ME case. The appearance of these spurious states is suppressed by imposing a sufficient number of constraints (energy penalties). These constraints are designed to favor an even number of spin-flips around any closed loop in the logical spins, in order to induce a consistent mapping. The form of the LHZ Hamiltonian we consider here is the following [17]:

$$H_{\mathrm{P}}^{\mathrm{LHZ}} = \sum_{k \in \mathcal{V}_{\mathrm{LHZ}}} J_k \sigma_k^z + \sum_{c \in \mathcal{V}_{\mathcal{C}}} \mathcal{C}_c , \qquad (3)$$

where $\mathcal{C}_c = -\lambda \sigma_{c_u}^z \sigma_{c_d}^z \sigma_{c_l}^z \sigma_{c_r}^z$ are four-body interactions between the physical qubits. In addition, three-body interactions appear at the LHZ graph edges, as shown in Fig. 1(c). Both the three and four-body interactions can be replaced by two-body interactions, by coupling to ancilla qutrits [17]. $\mathcal{V}_{\mathrm{LHZ}}$ represents the vertex set of the LHZ physical graph depicted in Fig. 1(c) by circles, with $K = |\mathcal{V}_{\mathrm{LHZ}}|$, and where $\mathcal{V}_{\mathcal{C}}$ represents the $C = |\mathcal{V}_{\mathcal{C}}| = \binom{N-1}{2}$ constraint interactions shown as black dots. Local fields h_i can be included by representing them as couplings to an additional ancilla qubit, via $h_i \sigma_i^z \rightarrow h_i \sigma_0^z \sigma_i^z$.
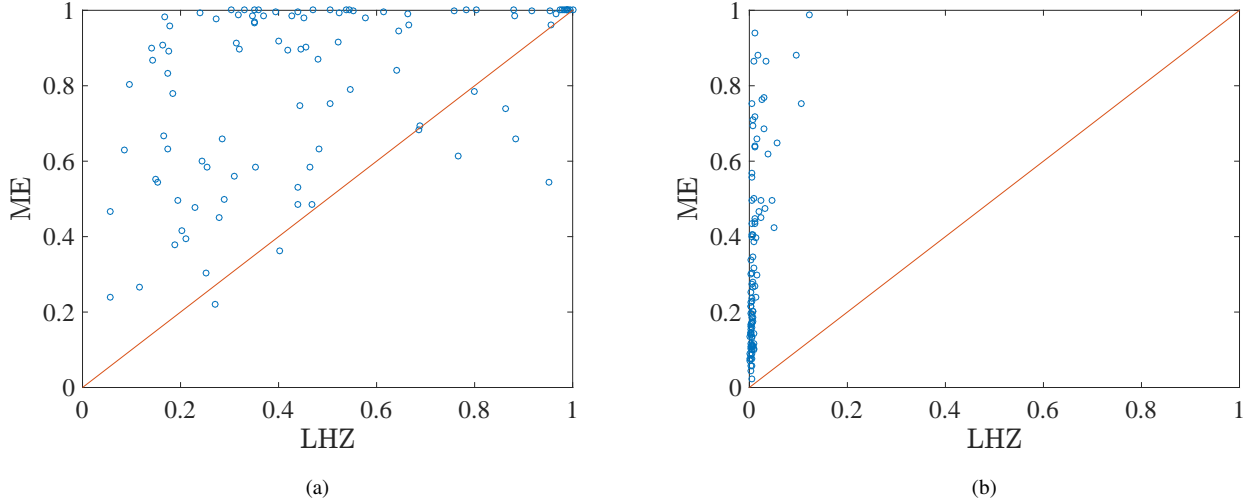
(a)

(b)

FIG. 2. **Relative performance of the ME and LHZ schemes.** (Color online) Shown are scatter plots of success probabilities for 100 random (a) K_8 and (b) K_{16} instances generated via SQA and majority vote decoding. The ME scheme exhibits substantially better performance. Majority vote decoding with the same number of votes (3 and 5 for K_8 and K_{16}, respectively) was used for both schemes, and the optimal energy penalty was used for each instance. SQA parameters: 10^4 sweeps, $\beta = 1$. Note that the effective temperature of SQA is modulated through the anneal by the annealing schedule (Fig. 11 in Appendix B). By the end of the anneal, the effective (dimensionless) temperature is ≈ 0.05, smaller than the absolute value of the smallest coupling used in our simulations, 0.1.

C. Leakage errors in embedded quantum annealing

Due to the fact that in both the ME and LHZ schemes a problem Hamiltonian defined on N logical qubits is embedded into a system with $O(N^2)$ physical qubits, the physical system has a much larger number of physical states ($\sim 2^{N^2}$) than the original logical system ($\sim 2^N$). This allows for "leakage": broken chains in the ME scheme or physical states with unsatisfied constraints in the LHZ scheme, that appear because of thermal and dynamical errors occurring during the annealing process. Such leakage states do not directly correspond to a logical state (see Fig. 10 in Appendix A for typical examples of leakage states one expects in the LHZ scheme for the random weakly correlated errors model and for finite-temperature open-system quantum annealing).

A leaked physical state must be decoded, i.e., it must be mapped back to a logical state. This decoding can be considered as partial error correction, in the sense that it recovers a logical state but it does not guarantee the recovery of the logical ground state. Next, we briefly review majority vote decoding; additional decoding schemes for the LHZ scheme will be considered in Sec. IV.

D. Majority vote decoding of the ME Scheme

A minor-embedded logical qubit i is encoded into a set of physical qubits $\{a_i\}_{a=1}^{n_A}$. Majority vote decoding (MVD) of logical qubit i consists of computing

$$\bar{q}_i = \text{sign} \sum_{a=1}^{n_A} q_{a_i} , \qquad (4)$$

where q_{a_i} represents the measured value of physical qubit a_i in the computational basis. In addition to its being used routinely in decoding minor embedded quantum annealing [58–62], MVD has been successfully used in the context of quantum annealing correction (QAC) [30, 31, 63] and hybrid minor-embedded implementations of QAC [16, 64].

MVD relies on the assumption that the decoded values $\{\bar{q}_i\}$ are the most likely to recover the logical ground state. However, this is not ensured due to the complex way errors are generated in quantum annealing. Alternatives such as energy minimization, which tends to be a better strategy for quantum annealing, have been explored as well [16], but will not be further considered here.

E. Majority vote decoding of the LHZ Scheme

In the absence of leakage, any spanning tree (i.e., a graph where there is a path connecting all vertices but any two vertices are connected by exactly one path) on the logical graph can be used to reconstruct a logical configuration from the values of the physical qubits. In the presence of leakage, however, different spanning trees may decode the same physical state to different logical states. A simple decoding technique is to perform a majority vote on n_T logical states decoded from random spanning trees [17]:

$$\bar{q}_i = \text{sign} \sum_{t=1}^{n_T} \bar{q}_{t,i} , \qquad (5)$$

where $\bar{q}_{t,i}$ represents the decoded value of logical qubit i retrieved from spanning tree t (see Appendix A for further de-
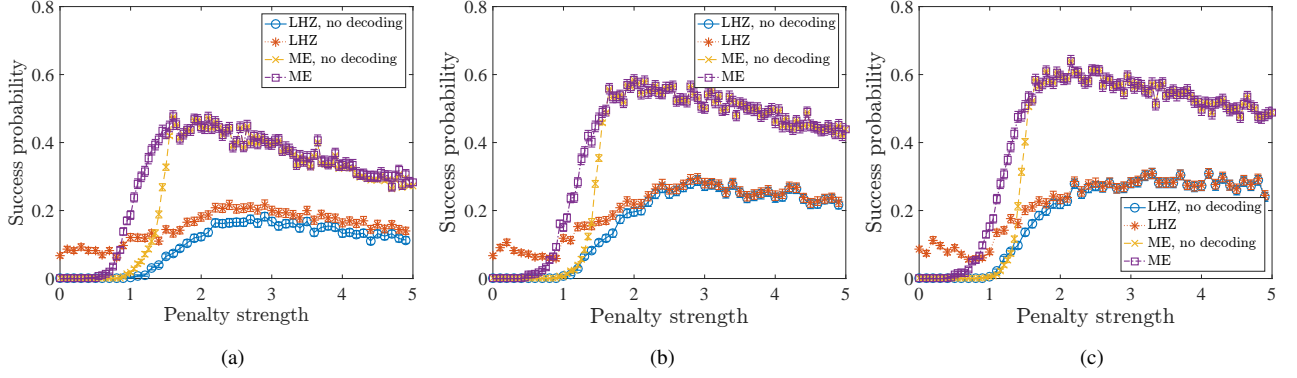
FIG. 3. **Dependence on penalty strength and the number of sweeps.** (Color online) Shown is the success probability of the ME and LHZ schemes for a representative instance generated via SQA, as a function of penalty strength, with a varying number of sweeps: (a) 10^4 sweeps, (b) 5×10^4 sweeps, and (c) 10^5 sweeps. The decoded ME scheme always outperforms the decoded LHZ scheme. The optimal penalty strength increases with the number of sweeps. Performance improves consistently for both ME and LHZ as the number of sweeps increase. The non-vanishing success probability at zero penalty for the decoded LHZ scheme is an artifact of the small size of the logical problems considered, which allows for a non-negligible probability that MVD will retrieve a logical ground state.

tails and examples).

III. RESULTS FOR MAJORITY VOTE DECODING

In order to compare the ME and LHZ schemes we first generated a set of 100 logical random instances on complete graphs $\{K_8, K_{16}\}$ with couplings $J_{ij} \in \{\pm 0.1, \pm 0.2, \pm 0.3, \cdots \pm 1\}$ chosen uniformly at random, and all logical local fields set to zero. For each instance, we constructed the embedded physical Hamiltonians $H_{\mathrm{P}}^{\mathrm{ME}}$ and $H_{\mathrm{P}}^{\mathrm{LHZ}}$. For each instance and embedding scheme, we ran SQA [40, 41] using the quantum annealing Hamiltonian Eq. (2) with the same annealing schedule, temperature, and number of Monte Carlo sweeps. A single sweep involves applying a Wolff cluster update along the imaginary time direction for all spatial quantum Monte Carlo slices (see Appendix B for more details on SQA).

The annealing schedule sets the effective energy scale of the Hamiltonian [Eq. (2)], and the ratio of this scale to the temperature changes during the anneal. The annealing schedule can play a crucial role in determining the performance of the adiabatic algorithm, and in principle can be optimized separately for the ME and LHZ schemes. Furthermore, the ME and LHZ schemes may benefit from controlling the constraints (the chains in the ME case, and the four-body terms in the LHZ case) using an independent annealing schedule. We do not address these issues in this work, and leave the exploration of possible improvements using these strategies for future work.

In order to understand the role of the purely classical Hamiltonians $H_{\mathrm{P}}^{\mathrm{ME}}$ and $H_{\mathrm{P}}^{\mathrm{LHZ}}$ in the success of each scheme, we used parallel tempering (PT) [48, 65]. For sufficiently long runtimes, PT samples from the classical thermal (Gibbs) state. If the temperature is sufficiently low, PT samples predominantly from the ground state, and hence it can be used as a solver to find the ground state of the classical Hamiltonian.

PT is a more efficient algorithm than simulated annealing [66] for both the sampling and solver tasks (see Appendix B for more details on PT). Here we use PT as a sampler on the ME and LHZ Hamiltonians, whereby we run PT with a long but fixed runtime, while restricting it to single-spin updates in order to keep the algorithm as close to the SQA implementation as possible for both the ME and LHZ schemes[1] At sufficiently large penalty values, we expect the leakage states to be completely decoupled from the low-lying levels in both the ME and LHZ schemes. The low energy physical states are then equivalent to the low energy logical states, which are the same for the two schemes, so at sufficiently low temperatures we expect the PT success probabilities for both schemes to coincide. As we increase the temperature but keep the runtime constant, we expect deviations between the two schemes to emerge. This allows us to compare how well each scheme performs in recovering the logical ground state from the low-energy physical states.

In this section we perform the comparison between the ME and LHZ schemes using only majority vote decoding, before considering more sophisticated decoding strategies in Sec. IV. MVD allows us to equalize the decoding effort between the two schemes.

A. Comparing ME and LHZ via simulated quantum annealing

Fig. 2 shows scatter plots comparing the ME and LHZ schemes on SQA-generated states. For both the ME and

---

[1] Cluster updates that take into account the geometry of each embedding scheme are likely to help speed up the convergence of PT to the Gibbs state for both the ME and LHZ schemes. For example, cluster updates associated to the flip of a logical qubit (a chain in the ME case and a cluster consisting of all physical qubits $\{q_{i,k}, q_{k,i}\}_k$ for logical qubit $\bar{q}_i$ in the LHZ case).
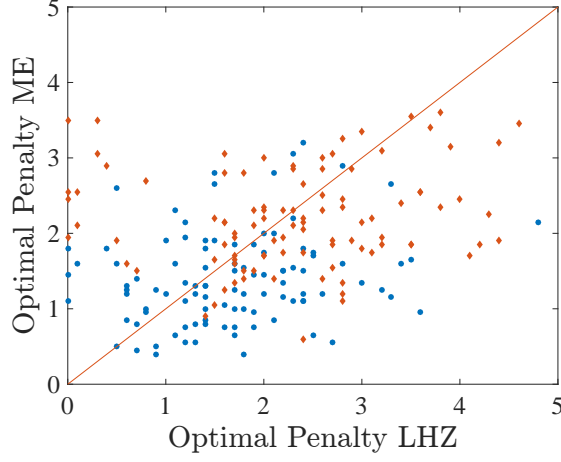
FIG. 4. **Optimal penalty strength comparison.** (Color online) Shown is a comparison of optimal penalties, i.e., the penalty values that maximize the success probability after MVD and using SQA. These correspond to the penalty strengths used for the results shown in Fig. 2. The optimal penalty is roughly equally distributed on K_8 instances (blue dots) but tends to be higher for the LHZ scheme on K_{16} instances (red diamonds). SQA parameters: 10^4 sweeps, $\beta = 1$.

LHZ cases we ran SQA simulations over a wide range of energy penalty values and, for each instance, chose the value that maximizes the success probability, which we refer to as the optimal penalty value. In both cases we used majority-vote decoding as described in Secs. II D and II E. In the ME case, logical values were obtained with a majority vote on $n_A = \lceil N/4 \rceil + 1 = 3$ and 5 physical qubits, corresponding, respectively, to the length of the chains required to minor-embed K_8 and K_{16}. Figure 2 clearly shows an advantage for the ME scheme over the LHZ scheme using majority vote decoding, with the advantage growing as the problem size increases from K_8 to K_{16}.

The number of sweeps is fixed in Fig. 2; to test the dependence on this number, Fig. 3 shows the success probability for a representative K_8 instance as a function of the penalty strength, when the number of sweeps is increased from 10^4 to 10^5. The performance of both schemes improves, but we find that the decoded ME scheme's success probability is always larger than LHZ's, except at very small penalty values; we discuss the reason for this below. In addition, the optimal penalty strength for ME appears smaller; this is studied more systematically in Fig. 4, which shows a roughly equal distribution for K_8, but it confirms that the LHZ scheme tends to require higher optimal penalties for K_{16}. This could be a disadvantage given practical limitations of quantum annealing devices.

As we demonstrate in the next subsection, the reason for the relatively poorer performance of the LHZ scheme is likely due to its less efficient spin updates in the SQA simulations. Two factors are likely responsible for this difference: first, the LHZ scheme requires asymptotically twice as many physical qubits as ME [more precisely, $N(N-1)/2 > N(\lceil N/4 \rceil + 1)$ for $N \geq 8$]; second, the four-body constraint terms may make

updates via time-like cluster spin-flips less effective. It is possible that alternative implementations of the constraints (using for example two-body interactions along with coupling to qutrits, as suggested in Ref. [17]) may improve the performance of the LHZ scheme under SQA simulation; this is left for a future study.

B. Comparing ME and LHZ via parallel tempering

We next study the probability with which the logical ground state can be retrieved from the results of PT simulations that use the classical physical Hamiltonians H_P^{LHZ} and H_P^{ME}. This analysis allows us to directly study the resiliency to leakage due to purely classical thermal errors.

Fig. 5 presents our PT simulation results, as a function of energy penalty, for representative K_8 and K_{16} instances, where we also examine the size and temperature-dependence. We confirm in Fig. 5 for the K_8 case that for sufficiently large penalties and sufficiently low temperatures, the success probability for both schemes is identical and high. Therefore, the observed difference is attributable to more efficient spin updates in the case of SQA simulations of the ME scheme.

At smaller energy penalties, thermal excitations are more likely to populate leakage states, leading to differences between the two schemes. This is instance and size-dependent, as can be seen by the differences between the two schemes in Figs. 5(a) and 5(c), but the absence of a difference in Fig. 5(b).

The typical behavior is presented in the scatter plot of Fig. 6, where we compare the success probabilities of the LHZ and ME schemes at particular values of the inverse temperature and energy penalty, for the same 100 random K_8 instances as in the SQA case of Fig. 2(a). At this relatively small penalty value, the ME scheme exhibits a small but consistent advantage in its ability to recover states that have leaked due to thermal excitations.

Differences between the two schemes are also expected if the number of updates is insufficient to reach the low-energy physical states, and hence leakage states can still be populated. In this case, even at large penalty values the physical state need not be equivalent to the logical state. This is what happens in the LHZ case for the K_{16} instance shown in Fig. 5(c), where evidently the success probability does not saturate at both the high and low temperature value. We find that increasing the number of steps in the PT simulations eliminates this feature (Fig. 14 in Appendix D). Moreover, we find that for most of the instances and penalty values, the SQA states are closer to the PT states for ME than for LHZ (Fig. 15 in Appendix D). Thus, we see once again that the LHZ scheme has greater difficulty with single spin updates than ME, in agreement with our earlier SQA results.

The different response to small and large energy penalties explains the striking step-like behavior of the success probability observed in Fig. 5 for the low temperature simulations, where PT overwhelmingly samples the physical ground state. This step-like behavior appears at the smallest value of the energy penalty such that the physical ground state is correctly mapped to the logical ground state. A scatter plot of the cor-
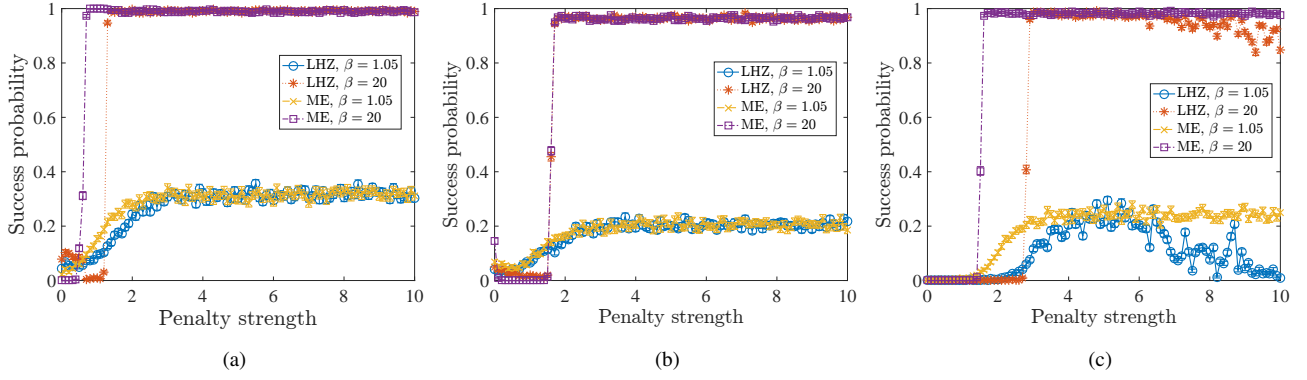
FIG. 5. **Performance of the ME and LHZ schemes using PT.** (Color online) Shown are the success probabilities after MV decoding for three representative instances generated via PT on K_8 [(a) instance 3 and (b) instance 95] and K_{16} (c) (instance 72) graphs. Results for two temperatures are shown; reducing the temperature boosts the success probabilities, as expected. The ME and LHZ schemes exhibit similar performance for the K_8 instances, though the optimal penalty value is again smaller for the ME scheme in (a). Similar results are seen for the K_{16} instances. The drop in success probability at large penalty strength for $\beta = 1.05$ in the LHZ case (all but two of our 100 instances exhibit this signature) can be understood as a signature of incomplete thermalization at the chosen parameter values. It is alleviated by increasing the number of PT swaps (see Appendix D). PT parameters are discussed in Appendix B.
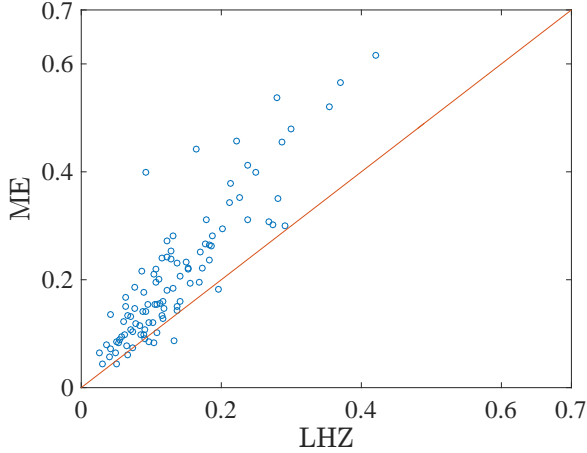

FIG. 6. **Comparing leakage resiliency of the ME and LHZ schemes.** (Color online) Shown is a scatter plot for the K_8 instances. The success probability on both axes is measured relative to the PT state ($\beta = 1.05$) at a fixed intermediate value of the energy penalty ($\gamma = 1$). Decoding was performed via majority vote.
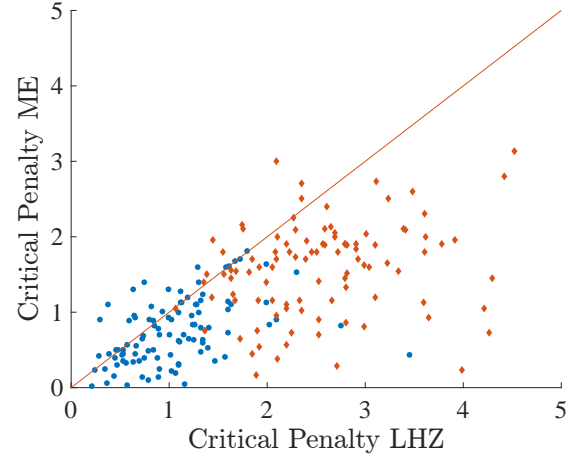

FIG. 7. **Critical penalty strength comparison.** (Color online) Shown is the penalty strength where the PT success probability after MVD is 0.5 for $\beta = 20$. There is a systematic increase in the critical penalty strength for both schemes as the problem size increases from K_8 (blue dots) to K_{16} (red diamonds). Additionally, the ME critical penalty strength is systematically lower for the K_{16} instances.

responding critical energy penalty values is shown in Fig. 7, where it is seen to be larger for the LHZ scheme, in agreement with the SQA optimal penalty results seen in Fig. 4. However, a lower bound argument shows that the value of the optimal penalty should scale at least with N in both schemes (see Appendix F).

IV. MINIMUM WEIGHT, MAXIMUM LIKELIHOOD, AND BELIEF PROPAGATION DECODING OF THE LHZ SCHEME

In an attempt to improve the performance of the LHZ scheme, in this section we present and study a decoding strategy that extends the approaches of Refs. [17, 21], and is designed to address independent or weakly correlated physical spin-flip errors. As we have already stressed, this is not an accurate model of noise in quantum annealing. However, adopting this perspective allows us to develop well-defined decoding strategies, which we in turn test on relevant distributions

generated via SQA and PT. We shall find that this leads to substantial improvements.

A. Minimum Weight Decoding

Let us denote the final states of the K physical qubits at the end of a quantum annealing run by $\Psi_{t_f} = \{s_k\}_{k=1}^K$ (with $s_k = \pm 1$). This readout amounts to a syndrome measurement $\mathcal{S}$ defined by the list of values of the four-local constraints: $\mathcal{S} = \{\zeta_c\}_{c=1}^C = \{-\lambda s_{c_u}^z s_{c_d}^z s_{c_l}^z s_{c_r}^z\}_{c=1}^C$, with $\zeta_c = \pm 1$. Given a physical state with violated constraints, the goal of minimum-weight decoding (MWD) is to find the nearest Hamming distance constraint-satisfying physical state.

Consider the ground state $\Phi = \{e_k\}_{k=1}^K$ ($e_k = \pm 1$) of the following Hamiltonian, defined on the same LHZ graph as the original optimization problem:

$$H_{\text{MWD}} = - \sum_{k \in \mathcal{V}_{\text{LHZ}}} \sigma_k^z - \sum_{c \in \mathcal{V}_{\mathcal{C}}} \zeta_c C_c , \qquad (6)$$

where it is assumed that the four-body interactions $C_c = -\lambda \sigma_{c_u}^z \sigma_{c_d}^z \sigma_{c_l}^z \sigma_{c_r}^z$ are sufficiently strong such that Φ minimizes (satisfies) all constraint terms. The MWD decoded state is then given by:

$$\Psi_{\text{MWD}} = \{e_k s_k\}_{k=1}^K , \qquad (7)$$

for the following reason: in the absence of any constraint-violating terms $\Phi = \{+1\}_{k=1}^K$, i.e., the original physical state need not be changed. When the physical state violates a constraint, a sequence of spin-flips must be performed to correct this violation. The second term in Eq. (6) forces the state Φ to undo the measured syndrome, and the first term minimizes the number of spin-flips (i.e., $e_k = -1$) in G.

To gain additional insight into the MWD problem, and to connect it to maximum likelihood decoding of Sourlas codes [50, 51], we first note that the MWD Hamiltonian (6) can be brought back to the LHZ form (3). Let us assume that we can find a set of spin-flips $G = \{g_k\}_{k=1}^K$ with $g_k = \pm 1$ (-1 means a spin-flip) that removes all unsatisfied constraints, i.e., that replaces every ζ_c by $+1$. It is easy to find at least one such G state, by performing a sequence of spin-flips that changes the sign of each constraint one by one [this is not necessarily the MWD state since it need not minimize the number of spin-flips; an example is shown in Appendix F, Fig. 19(b)]. Alternatively, one can think of G as a gauge transformation $\sigma_k^z \mapsto g_k \sigma_k^z$ such that $\zeta_c \mapsto 1$, which allow Eq. (6) to be rewritten as

$$H_{\text{MWD}}^g = - \sum_{k \in \mathcal{V}_{\text{LHZ}}} g_k \sigma_k^z + \sum_{c \in \mathcal{V}_{\mathcal{C}}} C_c . \qquad (8)$$

Equation (8) is now in the form of Eq. (3), and thus can be mapped back to an optimization over a complete (logical) graph K_N with couplings $g_{ij} = g_k$:

$$H_{\text{MWD}}^{\text{spin}} = - \sum_{i<j} g_{ij} \sigma_i^z \sigma_j^z . \qquad (9)$$

We can extract a vector $E^g = \{e_k^g\}_{k=1}^K$ from the ground state $\{s_i^g\}_{i=1}^N$ of this optimization problem, where $e_k^g = g_{ij} s_i^g s_j^g$ [i.e., e_k^g is positive or negative if the coupling corresponding to the pair (i, j) is satisfies or unsatisfied, respectively]. E^g is the minimum-weight corrected state. Taking into account the gauge transformation we have:

$$\Psi_{\text{MWD}} = \{e_k^g g_k s_k\}_{k=1}^K . \qquad (10)$$

Note that a syndrome measurement does not uniquely identify an error. In fact, there are 2^N possible G states, related by gauge transformations on the logical graph. Two different choices of G will thus define two optimization problems over a K_N that only differ by a gauge transformation, i.e., the two corresponding decoding processes are equivalent.

B. Maximum Likelihood Decoding

In Eq. (9) we have formulated the decoding process in terms of a spin system with ferromagnetic or antiferromagnetic unit interactions $g_{ij} = \pm 1$. In this formulation, the spin flip error probability P_ϵ corresponds to the probability of an antiferromagnetic interaction $g_{ij} = -1$. Errors thus corrupt the sign of the couplings of the spin system Eq. (9). This formulation of the decoding of the LHZ scheme corresponds precisely to the decoding of a Sourlas code for error correction of classical signals transmitted over a noisy channel, wherein the transmitted physical bits correspond to the couplings, while the logical bits are encoded in the ground state of an Ising model [50, 51]. As was shown in Refs. [67, 68], optimal, maximum-likelihood decoding (MLD) decoding of a Sourlas code corresponds to performing a thermal average at the Nishimori inverse temperature $\beta_N = \frac{1}{2} \ln[(1 - P_\epsilon)/P_\epsilon]$:

$$(e_k^g)_{\text{MLD}} = g_{ij}^g \text{sign}(\langle s_i^g \rangle_{\beta_N} \langle s_j^g \rangle_{\beta_N}) . \qquad (11)$$

An intuitive explanation for the need to perform a finite temperature decoding for optimal performance is that for sufficiently large error probabilities P_ϵ, the Ising Hamiltonian (9) is corrupted to such a degree that the correct state is most likely encoded in one of the excited states rather than in the ground state. These excited states are optimally explored by thermal sampling, if performed at the Nishimori temperature.

C. Hardness of MWD and MLD

MWD requires the minimization of the Hamiltonian defined in Eq. (9). This is equivalent to solving an instance of MAX-2-SAT, which is NP-hard; MLD requires thermal sampling, which is #P-hard and thus even harder [69–71]. Nevertheless, we expect the typical instance to be easy in the small P_ϵ regime. This observation can be made more precise thanks to the connection between Sourlas decoding and statistical mechanics. Specifically, the spin system defined in Eq. (9) will be in a ferromagnetic phase for sufficiently small error probabilities P_ϵ, where decoding is expected to
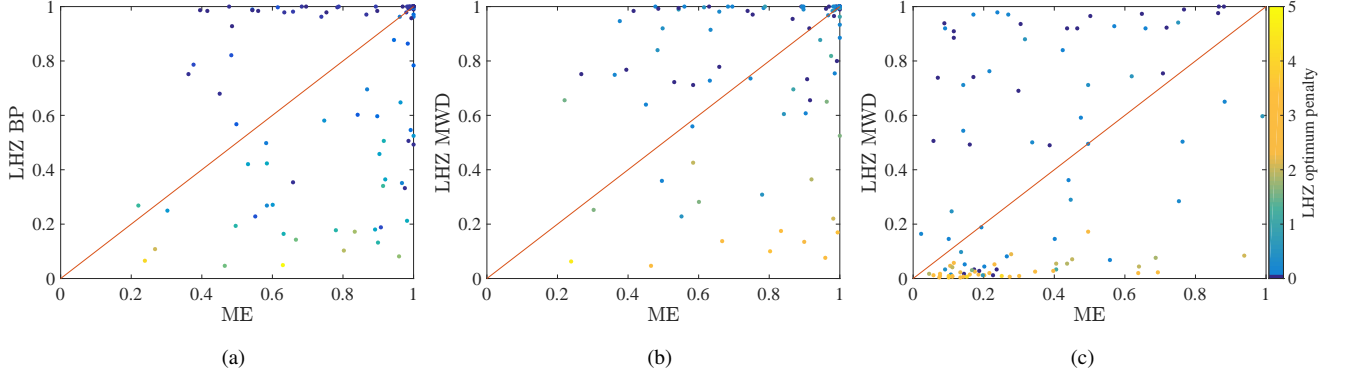
FIG. 8. **Performance of different decoding strategies for the LHZ scheme.** (Color online) Shown are scatter plots comparing success probabilities obtained at the optimal penalty strength for (a) the LHZ scheme with BP decoding (using an error rate of 0.2 and 10 iterations) and the ME scheme on the K_8 set; (b) the LHZ scheme with MWD decoding and the ME scheme on the K_8 set; (c) same as (b) but for the K_{16} set. Data was obtained with SQA simulations (10^4 sweeps, $\beta = 1$). Color denotes the energy penalty values for the LHZ scheme. Note that LHZ outperforms ME only for optimal penalty strength $\ll 1$ (blue dots).
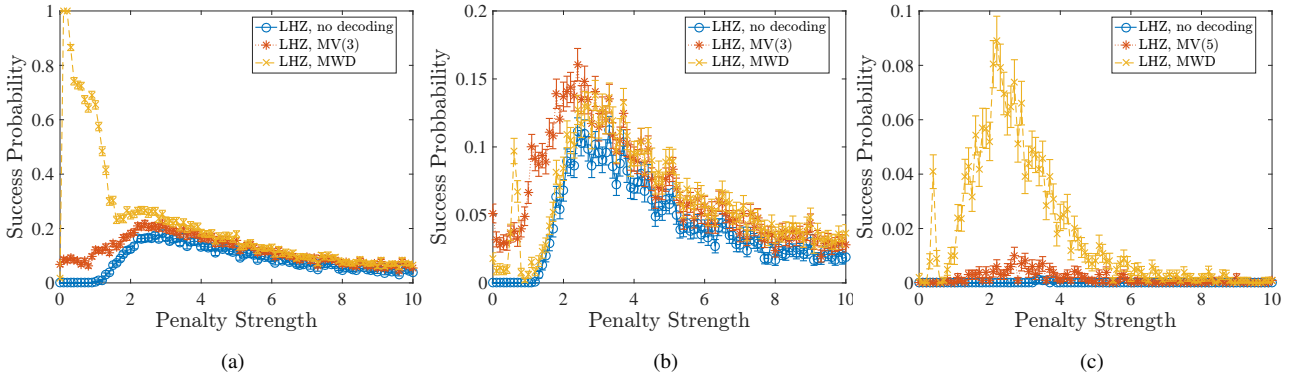

FIG. 9. **Comparison of MWD and MVD for the LHZ scheme as a function of energy penalty.** (Color online) Shown are results for two representative K_8 instances (instance 3 and 5) and a representative K_{16} instance (instance 5). (a) MWD wins by peaking at a very small energy penalty value. (b) A rare example where MVD performs better than MWD on a K_8 instance. (c) A typical K_{16} instance where MWD beats MVD by a large margin.

be easy. On the other hand, a disordered phase will arise for sufficiently large P_ϵ. A phase transition between the decodable (ordered) and undecodable (disordered) regimes is thus expected at a threshold error probability P_{th}. A mean field analysis of the problem in Eq. (9), valid in the large-N limit, reveals the presence of a phase transition between ferromagnetic and spin-glass phases at $P_{\mathrm{th}} = 1/2$ [9, 10]. Moreover, in the ferromagnetic phase the thermal average in Eq. (11) recovers the uncorrupted state perfectly at any finite temperature. In this sense, in the large-N limit MWD is equivalent to MLD.[2] On the other hand, for $P_\epsilon \geq P_{\mathrm{th}} = 1/2$ we are in the undecodable regime, in which $(e^g_k)_{\mathrm{MLD}}$ does not recover the uncorrupted state. Note that the perfect decoding achievable with MWD/MLD for $P_\epsilon \leq 1/2$ is a consequence of the fault

_____

[2] At finite N, MWD decoding is equivalent to MLD decoding in the $P_\epsilon \to 0$ ($\beta_N \to \infty$) limit only.

tolerance of the LHZ scheme when the only source of errors is random spin-flips [17, 21].

D. Weight-3 Parity Check with Belief Propagation

Given a classical error correcting code, instead of performing MLD, which is hard, one can use belief propagation (BP) as a suboptimal, but more manageable decoding technique, where decoded values are iteratively updated based on the expected error rate. It turns out that BP works very well with low-density parity check (LDPC) codes.

The relative alignment of a pair of logical spins $\bar{q}_i$ and $\bar{q}_j$ can be recovered in multiple ways. One way is to use the value of the physical spin $q_{i,j}$, but using weight-3 parity checks there are $N - 2$ additional independent ways, specifically, $q_{i,k}q_{k,j}$, with $k \neq i, j$. This particular approach was proposed by Pastawski and Preskill (PP) [21] and can be viewed as a

classical LDPC code, which can be decoded with belief propagation. We follow the same implementation of BP as used in Ref. [21]. For the case of weakly correlated errors, the probability of a decoding error is exponentially small in N [21]. PP noted that including higher-weight parity checks can yield further improvements. Indeed, to approach the performance of the MLD described in section IV B, one should include all parity checks and use finite temperature decoding. We do not consider these extensions here; they are an interesting possibility for future work.

E. Decoding Results

To understand whether the advantage that ME has over LHZ originates in the decoding step, we studied the efficacy of several other decoding strategies for the LHZ scheme. We included MWD, BP of Ref. [21], MVD over a large number of trees (100). We did not include MLD since it gives a well-defined decoding procedure only in the case of random spin-flips, where the error probability P_ϵ and the corresponding Nishimori temperature are known.

Different decoding strategies using majority vote give very similar results, as shown in more detail in Appendix C. This suggests that majority voting schemes that attempt to exploit a large number of decoding trees are not beneficial, as a consequence of the fact that the noise generated via SQA simulations does not correspond to uncorrelated random bit flips. (We explicitly show how the SQA results differ from uncorrelated noise in Appendix E.) In the latter case we would expect the decoding errors to decay exponentially with the number of trees (below a decoding threshold as discussed in the previous subsection).

We focus our attention here on MWD, which gives a substantial improvement over MVD and slightly improved performance over BP as seen in Figs. 8(a) and 8(b). Figures 8(b) and 8(c) show that after MWD the LHZ scheme becomes competitive with ME on the K_8 instances, although its performance remains worse on the K_{16} instances. We can thus not conclude that MWD (which requires substantial effort beyond MVD) is sufficient to enable the LHZ scheme to outperform ME.

A closer examination shows that when the LHZ scheme with MWD is superior to ME, it typically does so by having a peak in the success probability at almost vanishing penalty strengths [represented by the color scale of Figs. 8(b) and 8(c)]. Fig. 9(a) shows the success probability as a function of the penalty strength for an instance with the aforementioned behavior. In this situation, success arises entirely from the decoding process. This is related to the fact that MWD at zero penalty strength corresponds to finding the ground state of the following approximation of the original logical problem: $J_{ij} \to \text{sign}(J_{ij})$. To see this, note that when the energy penalties vanish, the physical ground state is trivially given by the physical configuration with all spins aligned to their local fields. This configuration is typically a leaked state. Following the steps described in Sec. IV A it is easy to check that $g_{ij} = -\text{sign}(J_{ij})$ is a possible choice for a gauge transforma-

tion which allows one to write the MWD problem Eq. (9), i.e., the approximation of the logical problem mentioned above. It turns out that for our set of K_8 instances, the ground state of this approximate problem typically corresponds to the ground state of the original problem, resulting in the observed peak in success probability. Even when this does not happen, however [as in the examples of Figs. 9(b) and 9(c)], MWD gives comparable or better results than MVD.

V. CONCLUSIONS

The path toward the achievement of "quantum supremacy" [72] using quantum annealing is fraught with many challenges, among which is the embedding problem we have focused on here. Lechner, Hauke and Zoller have proposed an embedding scheme which they claimed exhibits intrinsic fault tolerance [17]. Unfortunately, our findings do not support the fault tolerance claim in the context of a realistic error model for quantum annealing: we find that the performance of the LHZ scheme suffers under the errors generated using simulated quantum annealing, as well as (to a smaller degree) under classical thermalization. This is perhaps unsurprising, since the fault tolerance claim was based on a model of weakly correlated spin-flip errors, which does not accurately describe the errors that arise due to dynamical and thermal excitations during the course of a quantum annealing evolution (see Appendix E).

Since the earlier minor embedding scheme due to Choi [12, 13] has already found widespread use, it is important to compare the performance of the two schemes. A particularly attractive feature of the LHZ scheme is that it separates the problem of controlling local fields and couplings. In principle this appears to allow for a drastic design simplification. Still, it remains to be seen how flexible the LHZ scheme is to the occurrence of faulty qubits or couplers, a problem that has prompted the development of improvements to Choi's original ME scheme, showing that ME is easily adaptable to architectures with faults [14, 18]. Another attractive feature of the LHZ scheme is that it is amenable to a host of different decoding techniques, beyond the majority vote decoding strategy that is typically used for ME. Indeed, we have derived optimal minimum weight and maximum likelihood decoding strategies for the random bit-flip error model, and demonstrated that the former boosts the performance of the LHZ scheme for SQA. It is possible that the LHZ scheme may rival or surpass the ME scheme with even better decoding strategies, tailored to correlated spin-flip errors.

However, in several other important respects the ME scheme appears to be more attractive. First, the LHZ scheme requires a factor of two more physical qubits than ME for a given number of logical qubits. Second, the LHZ scheme requires four-body interactions (or two-body interactions with qutrits; it is an open question whether this would change our results), which can be harder to implement, while the ME scheme only requires standard, realizable two-body interactions between qubits. Third, we found that the optimal penalty strength tends to be lower for the ME scheme, which

can be important given practical constraints on the highest energy scales achievable, especially in superconducting flux qubit based quantum annealing devices. Given that the penalties are implemented differently in the two schemes, it is possible that a suitable physical implementation of the LHZ scheme achieves a higher maximum penalty strength, making this point minor. However, both schemes suffer from requiring energy penalties to grow with problem size, an issue that needs to be addressed in both schemes for scalability (see Appendix F). Finally, and perhaps most importantly, we found that the ME scheme seems to outperforms the LHZ scheme under a broad set of conditions. Namely, subject to identical simulation parameters and decoding effort, the success probabilities of the ME scheme are nearly always higher than those of the LHZ scheme for randomly generated Ising instances over complete graphs, up to the largest sizes we were able to test. We have explained this finding in terms of better spin-update properties of the ME scheme under SQA simulations.

We note several caveats regarding our results. First, the LHZ scheme's spin-update bottleneck under SQA simulations does not necessarily correspond to a performance bottleneck for an actual quantum annealing device: we used a discrete-time quantum Monte Carlo version of SQA with only time-like cluster updates, and such a model is ultimately not a complete model of a true quantum annealer. Only experiments or more detailed open-system quantum simulations can definitely address whether one scheme has a true implementation advantage over the other. However, the past success of SQA simulations in reproducing experimental quantum annealing data does suggest that our results have predictive power.

Second, our study is not comprehensive, and our conclusions are obviously limited to the set of instances we have considered. Thus, while we have observed a drop in performance for the LHZ scheme relative to the ME scheme when going from K_8 to K_{16}, our study does not suffice to ascertain whether the observed relative performance between the two embedding schemes persists for larger problem sizes.

Third, we focused on simulations with identical parameters and annealing schedule, with the choices based on present values for quantum annealing devices [47]. We expect that using lower temperatures and longer anneals will improve the performance of both schemes, but whether it changes the relative performance is unclear. Furthermore, we did not explore the possibility of separately optimizing all the parameters of both schemes to improve their respective performance. For example, one possibility is to optimize the annealing schedule for each scheme separately, as well as using a third independent annealing schedule for the constraint strength.

Finally, in our tests, we did not include the effects of errors in controlling the local fields and couplings. Given that the LHZ only needs to control the local fields precisely to implement the logical problem, while the ME scheme requires control of both local fields and couplings, it is conceivable that under inclusion of noise on the local fields and couplings, the LHZ scheme may be robust to such hardware implementation errors. This in itself would be very advantageous since such errors can dominate the performance of quantum annealing devices [73, 74].

In conclusion, there is no question that the field of quantum annealing will benefit from continued research into improved embedding methods. An important consideration that should guide such efforts is the integration of quantum error correction to achieve scalability. Our work highlights the importance of testing new embedding methods using errors models that directly match quantum annealing.

Appendix A: Mapping Physical States to Logical States for LHZ

Examples of leakage errors are given in Fig. 10. The absence of unsatisfied penalties (all four-local constraints are satisfied) ensures that the physical configuration describes a legitimate logical configuration. As described in Ref. [17], a
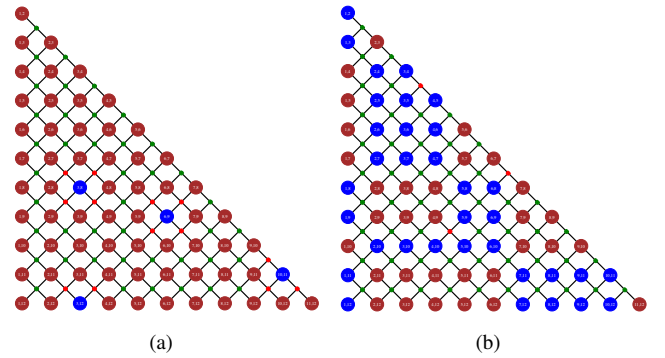

(a)   (b)

FIG. 10. **Examples of leakage errors in the LHZ scheme.** (Color online) Blue circles represent logical qubit errors, i.e., the corresponding physical qubits have been flipped from the physical state representing the logical ground state (all red circles). Because of the presence of unsatisfied constraints (red dots) the two configurations shown are leakage states, i.e., do not represent a logical state and need to be decoded. Panel (a): a small number of spin-flips, as expected for a model of weakly correlated, random spin-flips. Panel (b): a small number of constraints is unsatisfied, but a large number of spin flips have occurred, as expected for a realistic model of open-system quantum annealing.
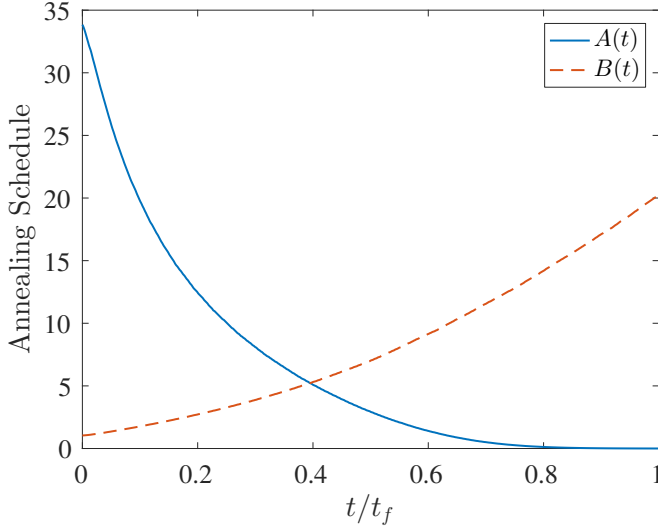
FIG. 11. **Annealing schedule used in our SQA simulations.** (Color online) The vertical axis units are arbitrary but correspond to the energy scale of our Hamiltonians.

logical qubit configuration can be reconstructed from the values of an appropriately chosen group of independent physical qubits. One possibility is to choose a group of physical qubits that corresponds to a spanning tree of the logical graph. For example, the group of physical qubits corresponding to the upper diagonal of Fig. 10(a) [or Fig. 10(b)] defines a spanning chain $q_{1,2}, q_{2,3}, \ldots, q_{N-2,N-1}, q_{N-1,N}$. To see how to perform decoding, recall that the values of the physical qubits in the LHZ scheme specify the alignment of the corresponding logical pairs. This implies that each physical configuration corresponds to a logical state only up to an overall flip of all logical qubits. We may thus choose, by convention, to fix the value of one logical qubit: $\bar{q}_1 = +1$. To decode the value of logical qubit $\bar{q}_i$, we simply read out the relative alignments of the logical pairs following the chain induced by the spanning tree which connects $\bar{q}_i$ to $\bar{q}_1$. Using the upper diagonal chain mentioned above we can decode as follows:

$$\bar{q}_i = \prod_{m=1}^{i-1} q_{m,m+1} . \tag{A1}$$

Similarly, the first column of Fig. 10(a) defines a star-like spanning tree where each qubit is only connected to $\bar{q}_1$: $q_{1,2}, q_{1,3}, \ldots, q_{1,N-1}, q_{1,N}$, thus we have:

$$\bar{q}_i = q_{1,i} . \tag{A2}$$

In the examples given in Fig. 10(a), decoding using the chain of Eq. (A1) gives the wrong state (one error hit), while decoding using the star of Eq. (A2) gives the correct state (no errors hit).
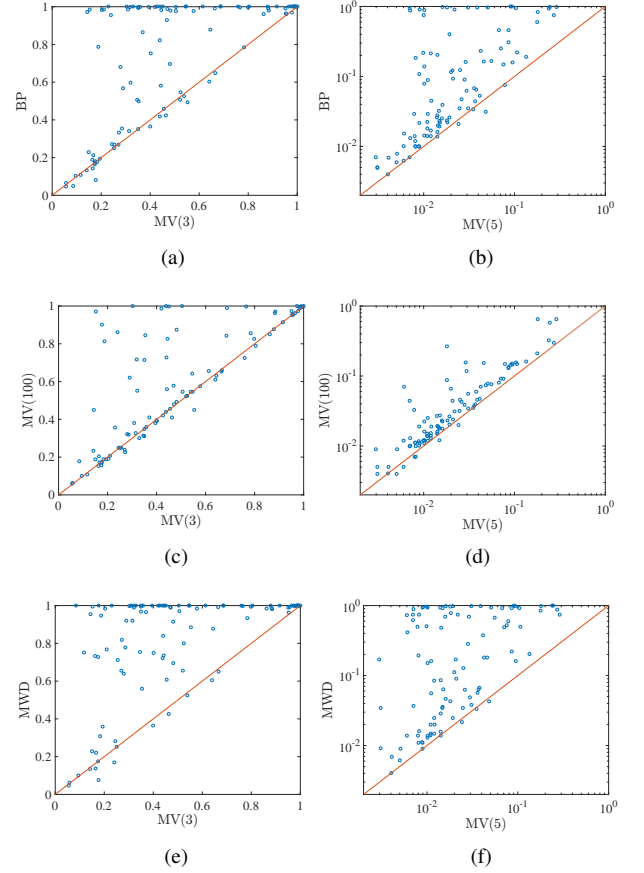

FIG. 12. **Comparison of different decoding strategies for the LHZ scheme.** (Color online) Shown are scatter plots comparing the success probabilities computed via SQA for (a,b) belief propagation of Ref. [21], (c,d) majority vote over 100 spanning trees, (e,f) MWD, all relative to majority vote decoding over three and five spanning trees for the K_8 (a,c,e) and K_{16} (b,d,f) instances respectively. The optimal penalty was used for each instance. SQA parameters: 10^4 sweeps for K_8 and 5×10^4 sweeps for the K_{16} instances, $\beta = 1$.

Appendix B: Numerical Simulations

Simulated Quantum Annealing. Our simulated quantum annealing simulations are based on a discrete-time quantum Monte Carlo algorithm [5, 40]. Details of the implementation of this algorithm have been described elsewhere [75]. Here we use a fixed Trotter slicing of $N_\tau = 64$. The annealing schedule used is shown in Fig. 11.

Parallel Tempering. Parallel tempering (also known as exchange Monte Carlo) [49] is a standard method used to thermalize spin systems. In PT simulations, replicas of the spin system at different temperatures are evolved independently for a fixed number of Metropolis updates, followed by a 'swap' operation. A swap involves comparing the energies of neighboring replicas, and swapping their temperatures with a probability p_{swap} given by:

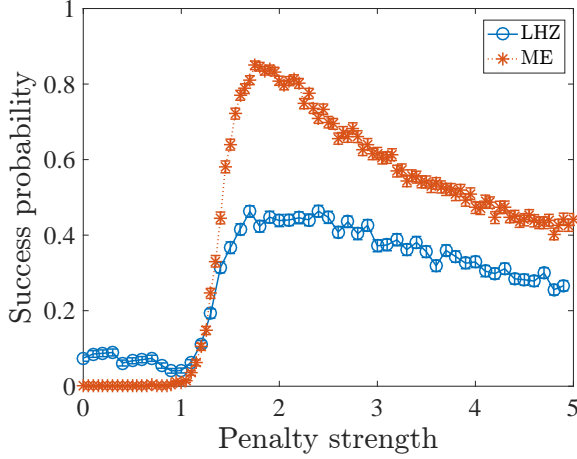$$p_{\text{swap}} = \min\{1, \exp\left[(E_{i+1} - E_i)(\beta_{i+1} - \beta_i)\right]\} \tag{B1}$$

FIG. 13. **Effect of increasing the number of sweeps and lowering the temperature.** (Color online) A comparison of the SQA results using majority vote for the K_8 instance shown in Fig. 3 for an increasing number of sweeps (10^6) and lowering the temperature ($\beta = 5$).
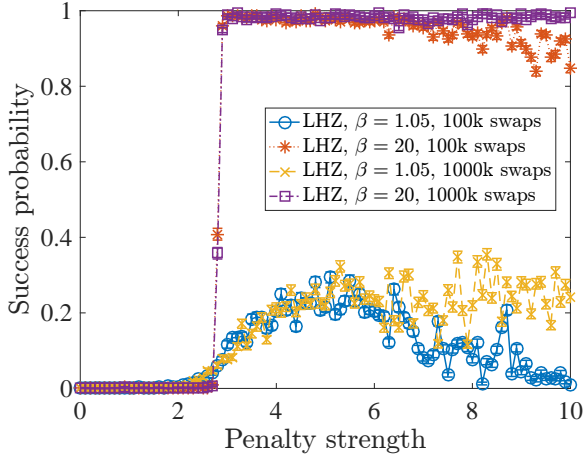

FIG. 14. **Effect of increasing the number of swaps.** (Color online) A comparison of the PT results for the K_{16} instance shown in Fig. 5 for different numbers of swaps. Increasing the number of swaps keeps the success probability more fixed at high penalty values.

where $(\beta_{i+1} - \beta_i) < 0, \forall i$. For our simulations we perform 10 sweeps of single spin Metropolis updates per swap, and we perform a total of 10^5 swaps. We use 64 different temperatures distributed according to:

$$\beta_i = \left(\frac{0.1}{20}\right)^{(i-1)/63} \beta_1 \qquad (B2)$$

with $\beta_1 = 20$ and $i = 1, \ldots, 64$.

Appendix C: Decoding Strategies

Here we provide a more direct comparison of the various decoding strategies we have tried: MWD, BP as in Ref. [21], and majority vote on a large number (100) of random spanning trees. As remarked in the main text, we did not implement MLD.

To perform MWD, we ran simulated annealing with the Hamiltonian in Eq. (9) 100 times with a linear schedule for $\beta \in [0.1, 5]$ and 1000 sweeps. We found this to be sufficient to find the ground state of the decoding Hamiltonian Eq. (9) for our relatively small problem sizes. In the BP decoding approach of Ref. [21] the alignment of a given pair of logical qubits $(\bar{q}_i, \bar{q}_j)$ is determined by iteratively updating its value using a majority vote on $N-2$ parity checks (q_{il}, q_{lj}) $(l \neq i, j)$. We then use a single random spanning tree to read out the full logical state.

The various panels of Fig. 12 present our SQA success probability results for the same 100 random K_8 and K_{16} instances considered earlier. It can be seen that the MVD scheme give comparable results, with the BP and MWD scheme giving much better results, as discussed in the main text.

Appendix D: Additional results for the LHZ scheme

1. SQA simulations for at colder temperatures and larger number of sweeps

In the main text, our SQA simulations for the K_8 instances were limited to $\beta = 1$ and 10^4 sweeps. Here we show for the same instance in Fig. 3 that increasing the number of sweeps and lowering the temperature does not necessarily improve the performance of LHZ relative to ME. We show in Fig. 13 the performance of each when using $\beta = 5$ and 10^6 sweeps. Both the ME and LHZ performance is improved, but relative performance is unchanged with ME continuing to show better results.

2. PT simulations for LHZ at higher swaps

In Sec. III A of the main text we noted that the fixed number of updates performed in the PT simulations on the LHZ embeddings was insufficient to properly thermalize the PT replicas, which led to the drop in success probability at large penalty strength in Fig. 5. In order to validate this conclusion, we increase the total number of swaps (the number of sweeps per swap remains fixed) by an order of magnitude and show the corresponding results in Fig. 14. With this increased number of swaps, the performance saturates for large penalties as expected, indicating that the behavior observed in Fig. 5 is indeed a consequence of an insufficient number of updates. This behavior is to be expected, since single spin-flip thermalization is completely frozen in the limit of very large penalty values; in such a limit, only cluster updates of physical qubits
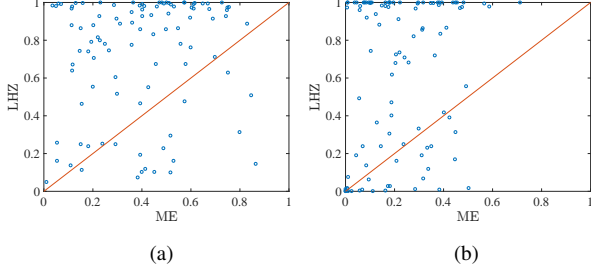
(a)   (b)

FIG. 15. **Distance test between PT and SQA.** (Color online) Scatter plot of the distance [Eq. (D2)] of the LHZ and ME states obtained via SQA from their respective PT simulation states. Panel (a): higher temperature, $\beta = 2.1$. Panel (b) lower temperature, $\beta = 4.0$. For most instances the ME scheme distance is smaller, indicating better thermalization. Results are shown for all 100 K_8 instances.

(a)   (b)

FIG. 16. **Average error rate.** Shown is a histogram of the average error rate for the 100 instances for the (a) K_8 and (b) K_{16} instances for SQA with $\beta = 1$ and 10^4 sweeps. The error rate is calculated as follows. For the 1000 states generated for each instance and each penalty value, we find the average minimum distance in terms of spin flips to the nearest degenerate ground state. The distance divided by the total number of spins gives the average error rate for a given instance and penalty value. The data in the histogram uses the error rate at the optimal penalty from the data shown in Fig. 2.

corresponding to a logical spin-flips would be efficient update moves (these are moves we do not perform in our simulations).

3. Distance between the SQA and PT states

In order to quantify how close the ME and LHZ states are to the PT states, we use the following distance measure [76]. For each instance i we define the probability distribution function $p_i(E)$ of finding a state with physical energy E:

$$p_i(E) = \frac{1}{N_E} \sum_{n=0}^{N_E} \delta_{E_n, E} , \qquad \text{(D1)}$$

where E_n is the energy of the nth excited state and N_E is the total number of energy levels observed for the given instance. We then compute the total variation distance

$$\mathcal{D}(p, q) = \frac{1}{2} \sum_x |p(x) - q(x)| \qquad \text{(D2)}$$

for a given instance i between the probability distributions for SQA and PT, i.e., we let $p = p_i^{\text{SQA}}$ and $q = p_i^{\text{PT}}$. We resort to this distance measure because of the relatively small number of states (10^3) we have for each scheme, which prevents us from reliably computing, e.g., the trace-norm distance between states. Figure 15 is a scatter plot of the distances comparing the ME and LHZ schemes, evaluated at the penalty that maximizes the success probability (after MV decoding) for each respectively, at two inverse temperatures of PT. For the majority of the instances, the ME states tend to be closer to the PT states than the LHZ scheme, in support of our assertion that the ME scheme thermalizes more easily than the LHZ scheme.
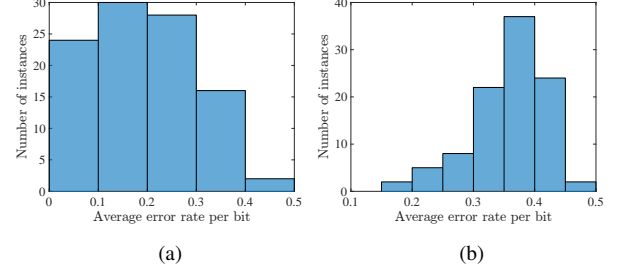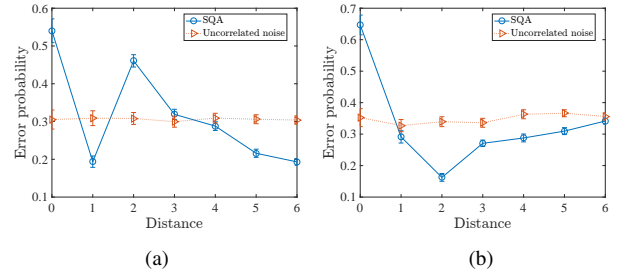

(a)   (b)

FIG. 17. **Error probability as a function of distance from the corner of the LHZ graph.** (Color online) Shown is the probability that a spin at distance d from the top corner of the LHZ graph flips subject to uncorrelated noise and SQA-generated noise, for two K_8 instances. (a) K_8: instance 3, $\gamma = 2.9$. (b) K_8: instance 95, $\gamma = 2.4$. The uncorrelated case is flat as expected, while the SQA case has a non-trivial distance dependence.

Appendix E: SQA noise is not uncorrelated

Our motivation for using SQA as a noise model is that it mimics thermal noise effects we expect from a finite temperature quantum annealer. Here we demonstrate how SQA noise differs from uncorrelated noise on the instances we have studied. First, Fig. 16 shows the average error rate for our instances in the SQA simulations. For all instances, the average error rates are below 0.5. Thus, our SQA simulations generate noise in a regime that is favorable from the perspective of decoding the LHZ scheme if in addition the noise were uncorrelated [21].

However, SQA generates noise that is not uncorrelated. To show this, we contrast it with uncorrelated noise and perform the following test. Starting at the top corner of the LHZ graph for the K_8 case [the spin labeled $1, 2$ in Fig. 1(c)], we calculate the average error rate for the spins at distances 0 to 6 (the largest for the K_8 case) from this corner spin; see the caption of Fig. 16 for how the error rate is calculated. Note that
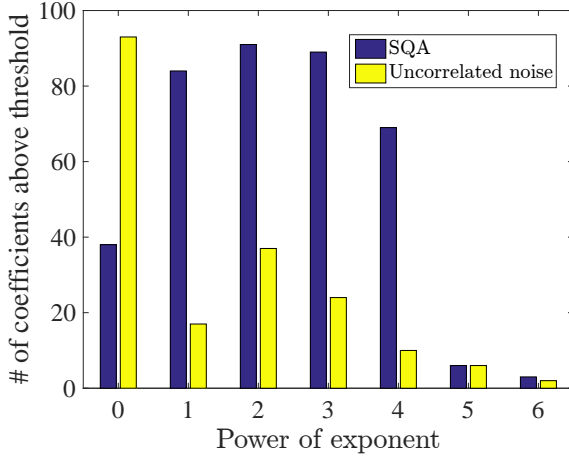
FIG. 18. **Distance dependence of error probability across all K_8 instances.** (Color online) Shown is a histogram of the values of the coefficients of a degree-6 polynomial fit to the curves generated as in Fig. 17, but for all 100 K_8 instances. We define a cutoff based on the average (over distance) error rate minus twice the standard deviation of the error rate of the uncorrelated noise. For each coefficient of the polynomial, if the coefficient is above this threshold, we count it, otherwise we do not. The histogram shown is the result of this binning process.

This is true for the ME scheme [58], but also for the LHZ scheme. We do not have a general lower bound for the strength of the penalties in the LHZ scheme, but simple argu-
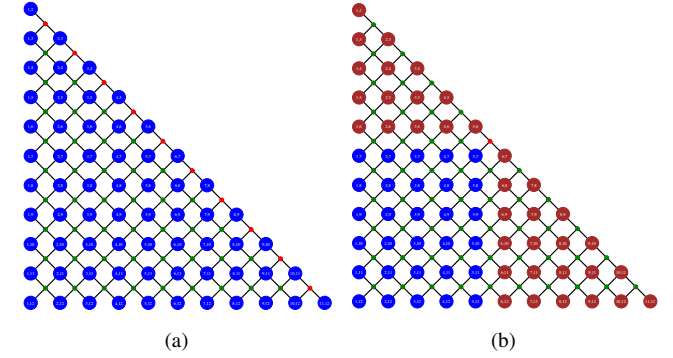

FIG. 19. **Configurations used to derive a lower bound on the scaling of the energy penalty strength.** (Color online) (a) Antiferromagnetic case with $J_{ij} = 1$. Colors represent alignment to the local fields: all the $\sim N^2$ can be aligned (blue) to the physical local field by violating only N constraints. (b) Random couplings case with $J_{ij} = \pm 1$. Colors represent physical flips (blue) from the logical ground state: a domain of $\sim N^2$ physical qubits can be flipped by violating only one constraint. To prevent these two configurations from being the physical ground state, the strength of the energy penalties should scale in both cases with N.

there are $d + 1$ spins at distance d. For uncorrelated noise we expect there not to be any dependence on distance. We show in Fig. 17 the behavior of SQA-generated noise for the two instances shown in Figs. 5(a) and 5(b), relative to uncorrelated noise, generated by randomly flipping spins with a probability equal to the averaged SQA error rate per spin (over all distances), to generate 10^3 uncorrelated error states on the physical state associated with the corresponding logical ground state of the LHZ scheme.

Clearly, the two cases behave very differently: as expected, the uncorrelated noise is effectively flat, while the SQA-generated noise has a non-trivial dependence on distance. This distance behavior is strongly instance and penalty strength dependent. To demonstrate that this difference occurs across all instances, we fit both the uncorrelated and SQA results to a degree-6 polynomial, as shown in Fig. 18. Ideally, the uncorrelated noise case will have its entire weight on the constant (x^0) term. We see that, as expected, the uncorrelated noise has most of its weight on the low powers of the polynomial, whereas the SQA-generated noise has most of its weight on the higher powers.

Appendix F: Scaling of the Strength of Energy Penalties

An important limitation of the embedded approach is that the strength of the energy penalties grows with problem size.

ments suggest that in both schemes, the strength of the energy penalties must grow linearly with the number of logical qubits N. To see this in the case of the LHZ scheme, consider, e.g., a completely antiferromagnetic K_N. Roughly half of the couplings are frustrated in the logical ground state. This should be mapped to roughly half the physical qubits in the LHZ scheme being down and the other half being up. However, for sufficiently weak constraints, the LHZ ground state is a physical configuration like that shown in Fig. 19(a) where all physical qubits are pointing down, i.e., aligned with the physical local field. The energy from aligning with the local fields is of order $-N^2$, while the number of violated constraints is of order N. This implies that the strength of the constraints must grow at least with N in order for the physical ground state to faithfully represent the logical ground state. Consider now a K_N with random couplings $J_{ij} = \pm 1$. In this case we expect the logical ground state to be highly frustrated, with roughly half of the corresponding physical qubits not aligned to the local fields. Flipping a domain of order N^2 physical qubits like that shown in Fig. 19(b) will typically result in a change in the energy of the physical configuration of the order of $\sqrt{N^2} = N$. To avoid the possibility of such domain-flips to lower the energy of the physical configuration, the constraints must grow again at least with N.

[1] F Barahona, "On the computational complexity of Ising spin glass models," J. Phys. A: Math. Gen **15**, 3241 (1982).

[2] A. Lucas, "Ising formulations of many NP problems," Front. Phys. **2**, 5 (2014).

[3] Tadashi Kadowaki and Hidetoshi Nishimori, "Quantum annealing in the transverse Ising model," Phys. Rev. E **58**, 5355 (1998).

[4] Edward Farhi, Jeffrey Goldstone, Sam Gutmann, Joshua Lapan, Andrew Lundgren, and Daniel Preda, "A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem," Science **292**, 472–475 (2001).

[5] Giuseppe E. Santoro, Roman Martoňák, Erio Tosatti, and Roberto Car, "Theory of quantum annealing of an Ising spin glass," Science **295**, 2427–2430 (2002).

[6] W. M. Kaminsky and S. Lloyd, "Scalable Architecture for Adiabatic Quantum Computing of NP-Hard Problems," in *Quantum Computing and Quantum Bits in Mesoscopic Systems*, edited by A.A.J. Leggett, B. Ruggiero, and P. Silvestrini (Kluwer Academic/Plenum Publ., 2004) arXiv:quant-ph/0211152.

[7] 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, "Quantum annealing with manufactured spins," Nature **473**, 194–198 (2011).

[8] Troels F. Rønnow, Zhihui Wang, Joshua Job, Sergio Boixo, Sergei V. Isakov, David Wecker, John M. Martinis, Daniel A. Lidar, and Matthias Troyer, "Defining and detecting quantum speedup," Science **345**, 420–424 (2014).

[9] David Sherrington and Scott Kirkpatrick, "Solvable model of a spin-glass," Physical Review Letters **35**, 1792–1796 (1975).

[10] Scott Kirkpatrick and David Sherrington, "Infinite-ranged models of spin-glasses," Physical Review B **17**, 4384–4403 (1978).

[11] P. Ray, B. K. Chakrabarti, and Arunava Chakrabarti, "Sherrington-kirkpatrick model in a transverse field: Absence of replica symmetry breaking due to quantum fluctuations," Phys. Rev. B **39**, 11828–11832 (1989).

[12] Vicky Choi, "Minor-embedding in adiabatic quantum computation: I. The parameter setting problem," Quant. Inf. Proc. **7**, 193–209 (2008).

[13] Vicky Choi, "Minor-embedding in adiabatic quantum computation: II. Minor-universal graph design," Quant. Inf. Proc. **10**, 343–353 (2011).

[14] Christine Klymko, Blair D. Sullivan, and Travis S. Humble, "Adiabatic quantum programming: minor embedding with hard faults," Quant. Inf. Proc. **13**, 709–729 (2014).

[15] Jun Cai, William G. Macready, and Aidan Roy, "A practical heuristic for finding graph minors," arXiv:1406.2741 (2014).

[16] Walter Vinci, Tameem Albash, Gerardo Paz-Silva, Itay Hen, and Daniel A. Lidar, "Quantum annealing correction with minor embedding," Phys. Rev. A **92**, 042310– (2015).

[17] Wolfgang Lechner, Philipp Hauke, and Peter Zoller, "A quantum annealing architecture with all-to-all connectivity from local interactions," Science Advances **1** (2015), 10.1126/sciadv.1500838.

[18] Tomas Boothby, Andrew D. King, and Aidan Roy, "Fast clique minor generation in chimera qubit connectivity graphs," arXiv:1507.04774 (2015).

[19] P. I Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, AJ. Berkley, R. Harris, J. P. Hilton, T. Lanting, AJ. Przybysz, and J. Whittaker, "Architectural considerations in the design of a superconducting quantum annealing processor," IEEE Transactions on Applied Superconductivity **24**, 1–10 (Aug. 2014).

[20] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill, "Topological quantum memory," Journal of Mathematical Physics **43**, 4452–4505 (2002).

[21] Fernando Pastawski and John Preskill, "Error correction for encoded quantum annealing," Phys. Rev. A **93**, 052325 (2016).

[22] T. Kato, "On the adiabatic theorem of quantum mechanics," J. Phys. Soc. Jap. **5**, 435 (1950).

[23] Sabine Jansen, Mary-Beth Ruskai, and Ruedi Seiler, "Bounds for the adiabatic approximation with applications to quantum computation," J. Math. Phys. **48**, – (2007).

[24] Daniel A. Lidar, Ali T. Rezakhani, and Alioscia Hamma, "Adiabatic approximation with exponential accuracy for many-body systems and quantum computation," J. Math. Phys. **50**, – (2009).

[25] Andrew M. Childs, Edward Farhi, and John Preskill, "Robustness of adiabatic quantum computation," Phys. Rev. A **65**, 012322 (2001).

[26] M. H. S. Amin, Dmitri V. Averin, and James A. Nesteroff, "Decoherence in adiabatic quantum computation," Phys. Rev. A **79**, 022107 (2009).

[27] Tameem Albash and Daniel A. Lidar, "Decoherence in adiabatic quantum computation," Phys. Rev. A **91**, 062320– (2015).

[28] Kevin C. Young, Mohan Sarovar, and Robin Blume-Kohout, "Error suppression and error correction in adiabatic quantum computation: Techniques and challenges," Phys. Rev. X **3**, 041013– (2013).

[29] Sergio Boixo, Troels F. Ronnow, Sergei V. Isakov, Zhihui Wang, David Wecker, Daniel A. Lidar, John M. Martinis, and Matthias Troyer, "Evidence for quantum annealing with more than one hundred qubits," Nat. Phys. **10**, 218–224 (2014).

[30] Kristen L Pudenz, Tameem Albash, and Daniel A Lidar, "Error-corrected quantum annealing with hundreds of qubits," Nat. Commun. **5**, 3243 (2014).

[31] Kristen L. Pudenz, Tameem Albash, and Daniel A. Lidar, "Quantum annealing correction for random Ising problems," Phys. Rev. A **91**, 042302 (2015).

[32] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik, "Finding low-energy conformations of lattice protein models by quantum annealing," Sci. Rep. **2**, 571 (2012).

[33] Zhengbing Bian, Fabian Chudak, William G. Macready, Lane Clark, and Frank Gaitan, "Experimental determination of ramsey numbers," Physical Review Letters **111**, 130505– (2013).

[34] A. Perdomo-Ortiz, J. Fluegemann, S. Narasimhan, R. Biswas, and V.N. Smelyanskiy, "A quantum annealing approach for fault detection and diagnosis of graph-based systems," The European Physical Journal Special Topics **224**, 131–148 (2015).

[35] Immanuel Trummer and Christoph Koch, "Multiple query optimization on the d-wave 2x adiabatic quantum computer," arXiv:1510.06437 (2015).

[36] Daryl P Nazareth and Jason D Spaans, "First application of quantum annealing to imrt beamlet intensity optimization," Physics in Medicine and Biology **60**, 4137 (2015).

[37] Gili Rosenberg, Poya Haghnegahdar, Phil Goddard, Peter Carr, Kesheng Wu, and Marcos López de Prado, "Solving the optimal trading trajectory problem using a quantum annealer," arXiv:1508.06182 (2015).

[38] Walter Vinci, Klas Markström, Sergio Boixo, Aidan Roy, Federico M Spedalieri, Paul A Warburton, and Simone Severini, "Hearing the shape of the ising model with a programmable superconducting-flux annealer," Scientific reports **4** (2014).

[39] Steven H. Adachi and Maxwell P. Henderson, "Application of quantum annealing to training of deep neural networks," arXiv:1510.06356 (2015).

[40] Roman Martoňák, Giuseppe E. Santoro, and Erio Tosatti, "Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model," Phys. Rev. B **66**, 094203 (2002).

[41] Bettina Heim, Troels F. Rønnow, Sergei V. Isakov, and Matthias Troyer, "Quantum versus classical annealing of ising spin glasses," Science **348**, 215–217 (2015).

[42] Sergey Bravyi and Matthew Hastings, "On complexity of the quantum ising model," arXiv:1410.0703 (2014).

[43] Sergio Boixo, Tameem Albash, Federico M. Spedalieri, Nicholas Chancellor, and Daniel A. Lidar, "Experimental signature of programmable quantum annealing," Nat. Commun. **4**, 2067 (2013).

[44] 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, "Thermally assisted quantum annealing of a 16-qubit problem," Nat. Commun. **4**, 1903 (2013).

[45] T. Lanting, A. J. Przybysz, A. Yu. 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, "Entanglement in a quantum annealing processor," Phys. Rev. X **4**, 021041– (2014).

[46] Sergio Boixo, Vadim N. Smelyanskiy, Alireza Shabani, Sergei V. Isakov, Mark Dykman, Vasil S. Denchev, Mohammad H. Amin, Anatoly Yu Smirnov, Masoud Mohseni, and Hartmut Neven, "Computational multiqubit tunnelling in programmable quantum annealers," Nat Commun **7** (2016).

[47] Tameem Albash, Walter Vinci, Anurag Mishra, Paul A. Warburton, and Daniel A. Lidar, "Consistency tests of classical and quantum models for a quantum annealer," Phys. Rev. A **91**, 042314– (2015).

[48] C. J. Geyer, "Parallel tempering," in *Computing Science and Statistics Proceedings of the 23rd Symposium on the Interface*, edited by E. M. Keramidas (American Statistical Association, New York, 1991) p. 156.

[49] Koji Hukushima and Koji Nemoto, "Exchange monte carlo method and application to spin glass simulations," Journal of the Physical Society of Japan **65**, 1604–1608 (1996).

[50] Nicolas Sourlas, "Spin-glass models as error-correcting codes," Nature **339**, 693–695 (1989).

[51] N. Sourlas, "Spin glasses, error-correcting codes and finite-temperature decoding," EPL (Europhysics Letters) **25**, 159 (1994).

[52] Giorgio Parisi, "Constraint optimization and statistical mechanics," arXiv preprint cs/0312011 (2003).

[53] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C. C. J. Wang, J. K. Freericks, and C. Monroe, "Emergence and frustration of magnetism with variable-range interactions in a quantum simulator," Science **340**, 583–587 (2013).

[54] R. Barends, A. Shabani, L. Lamata, J. Kelly, A. Mezzacapo, U. Las Heras, R. Babbush, A. G. Fowler, B. Campbell, Yu Chen, Z. Chen, B. Chiaro, A. Dunsworth, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, E. Solano, H. Neven, and John M. Martinis, "Digitized adiabatic quantum computing with a superconducting circuit," arXiv:1511.03316 (2015).

[55] Kevin Eng, Thaddeus D. Ladd, Aaron Smith, Matthew G. Borselli, Andrey A. Kiselev, Bryan H. Fong, Kevin S. Holabird, Thomas M. Hazard, Biqin Huang, Peter W. Deelman, Ivan Milosavljevic, Adele E. Schmitz, Richard S. Ross, Mark F. Gyure, and Andrew T. Hunter, "Isotopically enhanced triple-quantum-dot qubit," Science Advances **1** (2015).

[56] L. Casparis, T. W. Larsen, M. S. Olsen, F. Kuemmeth, P. Krogstrup, J. Nygård, K. D. Petersson, and C. M. Marcus, "Gatemon benchmarking and two-qubit operation," arXiv:1512.09195v1 (2015).

[57] Immanuel Bloch, "Quantum coherence and entanglement with ultracold atoms in optical lattices," Nature **453**, 1016–1022 (2008).

[58] Davide Venturelli, Salvatore Mandrà, Sergey Knysh, Bryan O'Gorman, Rupak Biswas, and Vadim Smelyanskiy, "Quantum optimization of fully connected spin glasses," Phys. Rev. X **5**, 031040– (2015).

[59] Davide Venturelli, Dominic J. J. Marchand, and Galo Rojo, "Quantum annealing implementation of job-shop scheduling," arXiv:1506.08479 (2015).

[60] Kenneth M. Zick, Omar Shehab, and Matthew French, "Experimental quantum annealing: case study involving the graph isomorphism problem," Scientific Reports **5**, 11168 EP – (2015).

[61] Travis S Humble, Alex J McCaskey, Ryan S Bennink, Jay Jay Billings, EF D Azevedo, Blair D Sullivan, Christine F Klymko, and Hadayat Seddiqi, "An integrated programming and development environment for adiabatic quantum optimization," Computational Science & Discovery **7**, 015006 (2014).

[62] Andrew D. King and Catherine C. McGeoch, "Algorithm engineering for a quantum annealing platform," arXiv:1410.2628 (2014).

[63] Anurag Mishra, Tameem Albash, and Daniel A. Lidar, "Performance of two different quantum annealing correction codes," Quant. Inf. Proc. **15**, 609–636 (2015).

[64] Walter Vinci, Tameem Albash, and Daniel A. Lidar, "Nested quantum annealing correction," arXiv:1511.07084 (2015).

[65] Helmut G Katzgraber, Simon Trebst, David A Huse, and Matthias Troyer, "Feedback-optimized parallel tempering monte carlo," J. Stat. Mech. **2006**, P03018 (2006).

[66] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," Science **220**, 671–680 (1983).

[67] Pál Ruján, "Finite temperature error-correcting codes," Phys. Rev. Lett. **70**, 2968–2971 (1993).

[68] Hidetoshi Nishimori, "Optimum decoding temperature for error-correcting codes," Journal of the Physical Society of Japan **62**, 2973–2975 (1993).

[69] C.H. Papadimitriou, *Computational Complexity* (Addison Wesley Longman, Reading, Massachusetts, 1995).

[70] Leslie G. Valiant, "The complexity of enumeration and reliability problems," SIAM Journal on Computing **8**, 410–421 (1979).

[71] Vilhelm Dahllöf, Peter Jonsson, and Magnus Wahlström, "Counting models for 2sat and 3sat formulae," Theoretical Computer Science **332**, 265 – 291 (2005).

[72] John Preskill, "Quantum computing and the entanglement frontier," arXiv:1203.5813 (2012).

[73] Victor Martin-Mayor and Itay Hen, "Unraveling quantum annealers using classical hardness," Scientific Reports **5**, 15324 EP – (2015).

[74] Zheng Zhu, Andrew J. Ochoa, Stefan Schnabel, Firas Hamze, and Helmut G. Katzgraber, "Best-case performance of quantum annealers on native spin-glass benchmarks: How chaos can affect success probabilities," Phys. Rev. A **93**, 012317 (2016).

[75] Siddharth Muthukrishnan, Tameem Albash, and Daniel A. Lidar, "Tunneling and speedup in quantum optimization for permutation-symmetric problems," arXiv:1511.03910 (2015).

[76] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, "Re-examining classical and quantum models for the d-wave one processor," Eur. Phys. J. Spec. Top. **224**, 111–129 (2015).