Improving nonstoquastic quantum annealing with spin-reversal transformations

Eleni Marina Lykiardopoulou [1,2,3] Alex Zucca,[1,4] Sam A. Scivier [1,5] and Mohammad H. Amin[1,4]

[1]*D-Wave Systems Inc., 3033 Beta Avenue, Burnaby, British Columbia, Canada V5G 4M9*

[2]*Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1*

[3]*TRIUMF, 4004 Wesbrook Mall, Vancouver, British Columbia, Canada V6T 2A3*

[4]*Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

[5]*School of Physics and Astronomy, University of Birmingham, Birmingham B15 2TT, United Kingdom*

Nonstoquastic Hamiltonians are hard to simulate due to the sign problem in quantum Monte Carlo simulation. It is, however, unclear whether nonstoquasticity can lead to advantage in quantum annealing. Here we show that YY interactions between the qubits make the adiabatic path during quantum annealing, and therefore the performance, dependent on spin-reversal transformations. With the right choice of spin-reversal transformation, a nonstoquastic Hamiltonian with YY interaction can outperform stoquastic Hamiltonians with similar parameters. We introduce an optimization protocol to determine the optimal transformation and discuss the effect of suboptimality.

I. INTRODUCTION

Quantum annealing (QA) [1–3] is a heuristic algorithm for finding low-energy configurations of Ising spin Hamiltonians, with applications in optimization and machine learning. Physical implementations of quantum annealers have matured to systems that include more than 5000 qubits, with increasing numbers of qubits expected in the future. The typical Hamiltonians implemented by these devices are

$$H(s) = A(s)H_D + B(s)H_P, \qquad (1)$$

$$H_D = -\frac{1}{2}\sum_i \sigma_i^x, \qquad (2)$$

$$H_P = \sum_i h_i\sigma_i^z + \sum_{i<j} J_{ij}^z\sigma_i^z\sigma_j^z, \qquad (3)$$

where $\sigma_i^{x,y,z}$ are the Pauli matrices acting on the ith qubit; H_D and H_P are known as driver and problem Hamiltonians, respectively; h_i and J_{ij} are dimensionless bias and coupling coefficients; and $s = t/t_a$ is a dimensionless annealing time parameter, with t_a being the total annealing time. The envelope functions $A(s)$ and $B(s)$ are usually fixed by the experimental implementation; an example is plotted in Fig. 1. Annealing is performed by initially letting the system relax to its ground state at $s = 0$ when $A(s) \gg B(s)$ and then evolving to a configuration in which $A(s) \ll B(s)$ at $s = 1$. Qubit states are measured at the end of annealing in the computation basis, which is defined by eigenfunctions of σ_i^z denoted by $|{\uparrow}\rangle$ and $|{\downarrow}\rangle$ with eigenvalues ± 1.

For closed systems, the adiabatic theorem [4–6] ensures that the system remains in its ground state throughout the annealing if the evolution time is long relative to a timescale that is proportional to $1/\Delta^2$, where Δ is the minimum gap between the ground state and the first excited state [7]. Reading out the N qubits then returns a configuration of $\vec{S} \equiv \{S_1, S_2, \ldots, S_N\}$, with $S_i = \pm 1$, that minimizes the problem Hamiltonian H_P. In practice, the adiabatic theorem may be violated via fast evolution or thermal excitations, resulting in a suboptimal (but maybe still acceptable) solution. In this work, we only consider closed system evolution and take the ground state as the only acceptable solution.

The existing physical implementations of QA [8] use superconducting qubits coupled via only one degree of freedom (flux), giving rise to *stoquastic* Hamiltonians [9–12], i.e., Hamiltonians with no positive or complex off-diagonal elements. Equilibrium statistics of stoquastic Hamiltonians can be simulated with quantum Monte Carlo (QMC) methods with no *sign problem* [13–15]. QMC methods may also exhibit dynamical behavior similar to QA for special stoquastic Hamiltonians [16], although this does not hold in general [17–19]. Nonstoquastic Hamiltonians, however, are not treatable by QMC methods; hence their statistical and dynamical properties are extremely hard to simulate [11,15]. They also can perform universal quantum computation [20–22] suggesting that nonstoquasticiy may be connected to quantum advantage in QA.

In order to make Hamiltonian (1) nonstoquastic, one needs to introduce interactions via other degrees of freedom, e.g., by changing the driver Hamiltonian to

$$H_D = -\frac{1}{2}\sum_i \sigma_i^x + \sum_{i<j}\left(J_{ij}^x\sigma_i^x\sigma_j^x + J_{ij}^y\sigma_i^y\sigma_j^y\right). \qquad (4)$$

We refer to the last two terms as XX and YY interactions, respectively, in contrast to the ZZ interaction in H_P. For $J_{ij}^y = 0$, the driver (4) has positive off-diagonal elements when $J_{ij}^x > 0$. This is, indeed, the regime in which most studies

have been done. It has been shown that this nonstoquastic Hamiltonian can significantly improve performance of QA, but only in special cases [23–27]. Moreover, a recent study [28] has found that the minimum gap in nonstoquastic QA generally increases by *de-signing* the Hamiltonian, i.e., making the Hamiltonian stoquastic by simply changing the sign of all positive off-diagonal elements. Therefore, whether or not nonstoquasticity can result in quantum annealing advantage over classical approaches remains an open question.

Less studied is nonstoquasticity due to YY interaction. Recently a pair of superconducting flux qubits with a nonstoquastic Hamiltonian were implemented by coupling them via both flux and charge degrees of freedom [29]. The resulting Hamiltonian had a driver of the form (4) that was dominated by YY interaction. A nonzero J_{ij}^{y} is special because it generates positive off-diagonal elements regardless of its sign and makes adiabatic path and performance variant under *spin-reversal transformation* (SRT), defined in the next section. The goal of this paper is to systematically study the role of SRT in quantum annealing with nonstoquastic drivers.

II. SPIN-REVERSAL TRANSFORMATION

Let us define a gauge transformation:

$$|\psi\rangle \rightarrow |\psi'\rangle = U|\psi\rangle, \quad H \rightarrow H' = UHU^{\dagger}, \quad (5)$$

with the unitary operator

$$U \equiv \prod_{i} \left(\sigma_{i}^{x}\right)^{(1-\alpha_{i})/2}, \quad (6)$$

where $\vec{\alpha} \equiv \{\alpha_{1}, \alpha_{2}, \ldots, \alpha_{N}\}$ is a set of transformation parameters with $\alpha_{i} = \pm 1$. The unitary operator U flips the state of qubit i if $\alpha_{i} = -1$; otherwise, it does nothing. The sign of each term in the Hamiltonian is adjusted so that the total energy remains unchanged. The transformed Hamiltonian H' has the following parameters:

$$h_{i}' = \alpha_{i} h_{i}, \quad (7)$$

$$J_{ij}'^{x} = J_{ij}^{x}, \quad (8)$$

$$J_{ij}'^{y} = \alpha_{i} \alpha_{j} J_{ij}^{y}, \quad (9)$$

$$J_{ij}'^{z} = \alpha_{i} \alpha_{j} J_{ij}^{z}. \quad (10)$$

It has exactly the same spectrum and dynamical behavior as H, as expected for gauge transformations, and the returned solution is the transformation of the original solution:

$$S_{i}' = \alpha_{i} S_{i}. \quad (11)$$

In practice, changing the sign of J_{i}^{y} is nontrivial, at least for the physical implementation of Ref. [29]. We define SRT as transformations (7)–(11) without (9), i.e., with $J_{ij}'^{y} = J_{ij}^{y}$. Therefore, SRT only transforms the classical part of the Hamiltonian.

For $J_{ij}^{y} = 0$, SRT is a true gauge transformation and is not expected to affect the dynamics. Therefore, solving the problem with QA using H or H' should lead to exactly the same probability of success. However, if there exist systematic errors in parameter specifications of the physical Hamiltonian, the errors will not be transformed if we submit H' instead of H to the QA hardware. This means that SRT is not a true

gauge transformation at the physical level and therefore is expected to affect the probability of success. In these situations, parameter specification error can be mitigated (averaged out to some extent) by running the problem with a set of SRTs, each parametrized by a randomly selected $\vec{\alpha}$.

When $J_{ij}^{y} \neq 0$, the spectrum and the dynamical behavior of (1) change depending on the applied SRT. However, since the SRT remains a true gauge transformation of the problem Hamiltonian, the classical problem being solved stays the same while the quantum path in the Hilbert space through which the solution is reached can heavily vary for different SRTs. Specially, the minimum gap can significantly change between transformations, resulting in a huge difference in success probabilities. Our goal is to find ways to select SRTs intelligently so that the performance is improved.

The driver Hamiltonian (4) contains all the tunneling terms in Hamiltonian (1). Single-qubit tunneling is through σ_{i}^{x} operators, and the XX and YY terms contribute to two-qubit cotunneling events. The matrix elements of $\sigma_{i}^{x}\sigma_{j}^{x}$ and $\sigma_{i}^{y}\sigma_{j}^{y}$ between states with ferromagnetic (FM) and antiferromagnetic (AFM) orientations are given by

$$\langle\uparrow\downarrow|\sigma_{i}^{x}\sigma_{j}^{x}|\downarrow\uparrow\rangle = 1, \quad \langle\uparrow\uparrow|\sigma_{i}^{x}\sigma_{j}^{x}|\downarrow\downarrow\rangle = 1, \quad (12)$$

$$\langle\uparrow\downarrow|\sigma_{i}^{y}\sigma_{j}^{y}|\downarrow\uparrow\rangle = 1, \quad \langle\uparrow\uparrow|\sigma_{i}^{y}\sigma_{j}^{y}|\downarrow\downarrow\rangle = -1. \quad (13)$$

While $\sigma_{i}^{x}\sigma_{j}^{x}$ does not distinguish between FM and AFM orders, $\sigma_{i}^{y}\sigma_{j}^{y}$ has off-diagonal elements with opposite signs. To the lowest order perturbation in $A(s)/B(s) \ll 1$, the two-qubit tunneling amplitudes for FM and AFM correlations, when $h_{i} = 0, \forall i$, are given by

$$\Delta_{\text{FM}}^{h=0} = -A(s)\left(\frac{A(s)}{4|J^{z}|B(s)} - J^{x} + J^{y}\right), \quad (14)$$

$$\Delta_{\text{AFM}}^{h=0} = -A(s)\left(\frac{A(s)}{4|J^{z}|B(s)} - J^{x} - J^{y}\right). \quad (15)$$

The first term in each equation describes tunneling through two single-qubit tunneling processes via σ^{x} operators. The last two terms, on the other hand, are contributions of direct two-qubit cotunneling via XX and YY interactions. Notice that with a negative J^{x} (stoquastic), the XX coupling always increases the tunneling amplitude for both FM and AFM correlations. The YY interaction with $J^{y} > 0$, however, increases (decreases) the tunneling amplitude for FM (AFM) correlated qubits, due to constructive (destructive) interference. For a pair of coupled qubits with zero bias, Eqs. (14) and (15) determine the size of the spectral gap between the ground and the first excited states. Therefore, for the same magnitude of ZZ coupling, FM coupling has a larger spectral gap than AFM coupling when $J^{y} > 0$, as experimentally demonstrated in Ref. [29]. The same argument also holds for larger clusters of strongly coupled qubits; the spectral gap is largest when couplings are maximally FM.

In problems with first-order phase transition [30], the minimum gap is typically suppressed because the ground state jumps between two states that are separated by a large hamming distance. In these cases, a large cluster of qubits needs to flip between the two crossing local minima. The cluster's tunneling amplitude at the avoided crossing determines the size of the gap. This phenomenon was experimentally
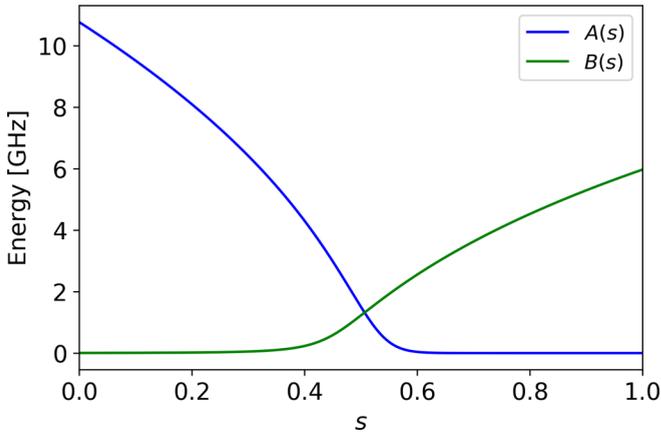
FIG. 1. The envelope functions $A(s)$ and $B(s)$ as functions of the dimensionless annealing time s.

demonstrated in Ref. [31], using a crafted 16-qubit problem with an extremely small gap. If qubits' transverse fields can be tuned individually, one can increase Δ by changing the adiabatic path either randomly [32] or algorithmically [33]. The presence of XX and/or YY interactions allows for alternative ways of changing the adiabatic path. Especially with YY interaction, every SRT introduces a new path. Therefore, by choosing the right SRTs, one can find an adiabatic path with a large Δ. This can be done either by random exploration or algorithmically. The latter is the focus of this work.

III. NUMERICAL SIMULATIONS

In this section, we explore the effect of SRT on the performance of QA assuming realistic parameters. The Hamiltonian is taken to be (1) with experimentally motivated $A(s)$ and $B(s)$ plotted in Fig. 1. All qubit couplings (XX, YY, or ZZ) are assumed to be according to the Chimera topology [34]. Based on the experimental observations of Ref. [29], we choose $J_{ij}^{y} = 0.5$ whenever they are nonzero. Also, to allow a direct comparison between stoquastic and nonstoquastic drivers, we choose $J_{ij}^{x} = -0.5$ (if nonzero). We calculate the minimum gap at each step of the evolution using exact diagonalization. We also calculate the probability of success by solving the time-dependent Schrödinger equation with annealing time $t_a = 1$ μs. We initialize the system in the the ground state of Hamiltonian (1) at $s = 0$. An approximated evolution can be obtained in the following way. Let $|k(t)\rangle$ denote instantaneous eigenstates of the Hamiltonian at time t, with instantaneous eigenvalues $E_k(t)$. At each step of the evolution, we obtain the n_{eigen} lowest energy eigenstates via exact diagonalization. We express the wave function in terms of these eigenstates: $|\psi(t)\rangle \approx \sum_{k=0}^{n_{\text{eigen}}} c_k(t)|k(t)\rangle$, where $c_k = \langle k(t)|\psi(t)\rangle$. Since the eigenstates in the truncated subspace do not form a complete set, this relation is approximate. However, for slow evolutions, the occupation probabilities of the higher energy state are extremely small and therefore the approximation is good. One may also increase n_{eigen} until there is no effect on the results. In our calculations, we use $n_{\text{eigen}} = 25$. Assuming the Hamiltonian remains constant during the small time step δt, the wave function after the time step
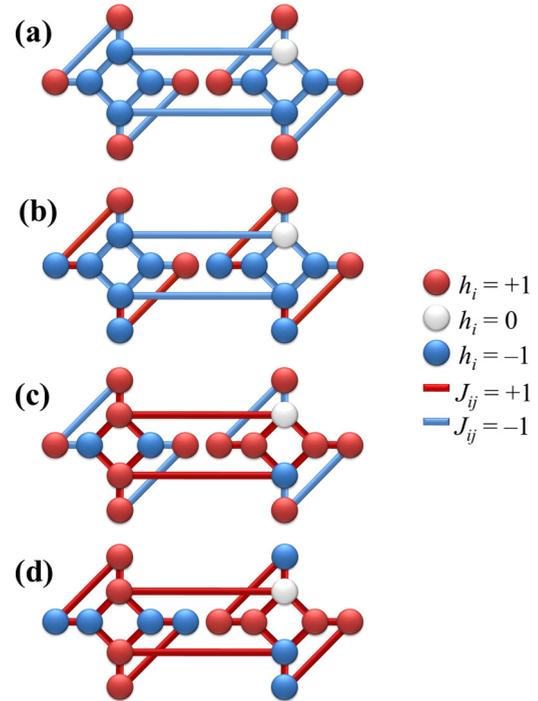
FIG. 2. (a) A crafted problem designed to have a small minimum gap. All couplers are ferromagnetic. (b)–(d) Three spin-reversal transformations of problem (a) with some or all couplers being antiferromagnetic.

becomes

$$|\psi(t + \delta t)\rangle \approx \sum_{k=0}^{n_{\text{eigen}}-1} e^{-iE_k(t)\delta t} c_k(t)|k(t)\rangle. \tag{16}$$

After the time step, the new instantaneous eigenstates $|k(t + \delta t)\rangle$ and eigenvalues $E_k(t + \delta t)$ are obtained by exact diagonalization and the above procedure is repeated. During the evolution, the time step δt is adjusted according to the instantaneous variation of the gap, so that the resolution of the minimum gap is adequate and the truncation and discretization error is kept under control. We also tested the configuration against the publicly available software QTIP [35,36].

A. Crafted problem

We first consider a slightly modified version of the problem studied in Ref. [31], as shown in Fig. 2(a). Parameters of the problem Hamiltonian are color coded in the figure. We divide the qubits into two groups, inner qubits (blue and white circles) and outer qubits (red circles). All couplings are FM; therefore, the two ferromagnetically oriented states $|\uparrow\uparrow \ldots \uparrow\rangle$ and $|\downarrow\downarrow \ldots \downarrow\rangle$ are energetically favored by the coupling terms. The biases in Fig. 2(a) are four positive, three negative, and one zero, making $|\downarrow\downarrow \ldots \downarrow\rangle$ the unique ground state and $|\uparrow\uparrow \ldots \uparrow\rangle$ an excited state. The outer qubits are pairwise coupled. Each outer qubit is in agreement with its applied bias in the ground state, but opposes the bias in the above excited state. If in the excited state a pair of coupled outer qubits are flipped together, two bias terms will be
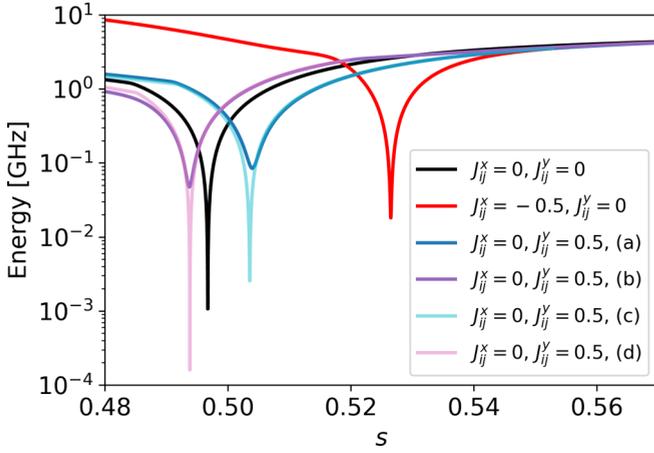
FIG. 3. The energy splitting between the ground state and the first excited state as a function of the annealing parameter s for the problems in Fig. 2 with different driver Hamiltonians.

FIG. 4. The success probabilities for the problems in Fig. 2 with different driver Hamiltonians.

satisfied but two couplers will be violated, leaving the energy unchanged. Therefore, with the existing four outer pairs there are $2^4 = 16$ degenerate excited states all connected by two-qubit flips. This degeneracy is lifted by the transverse field. Each coupled pair would lower their energy by forming an entangled state, $(|\uparrow\uparrow\rangle + |\downarrow\downarrow\rangle)/\sqrt{2}$. The lowest excited state is therefore a superposition of these 16 degenerate states. As the transverse field is increased (moving back in s), the splitting of the excited states grows until the lowest excited state crosses the ground state. The minimum gap at this avoided crossing is proportional to the tunneling amplitude between the two (localized) crossing states. Each of the 16 classical states in the superposition has a large hamming distance to the ground state (8 to 16 bit flips), resulting in a very small Δ. Figure 3 shows the energy splitting between the ground state and the first excited states for Hamiltonian (4) during the annealing according to the schedule in Fig. 1. As expected, the minimum gap for the original problem, with no YY or XX interactions (curve with minimum energy splitting at $s = 0.5$), is very small, $\Delta_0 \approx 10^{-3}$ GHz (the index 0 indicates $J_{ij}^x = J_{ij}^y = 0$ for every i and j). The gap is significantly increased when the XX interaction is turned on (curve with minimum energy splitting at $s = 0.53$ in Fig. 3).

In the presence of YY interaction, Δ is expected to depend on SRT. Figures 2(b)–2(d) show three SRTs of Fig. 2(a). The eight inner qubits all flip at the avoided crossing; therefore, they need to be coupled ferromagnetically to allow maximum tunneling amplitude. These couplings are turned from FM to AFM in Fig. 2(c). As it is clear from Fig. 3, the size of the minimum gap for problem (c) is reduced by almost 2 orders of magnitude compared to problem (a), although its position remains almost unchanged. The outer qubits, on the other hand, determine the position of the avoided crossing. Their pairwise tunneling is what lifts the degeneracy of the 16 classical excited states and creates the avoided crossing. Therefore, turning the coupling between the outer qubits from FM to AFM should reduce the splitting of the degenerate states and push the avoided crossing back toward a smaller s, as is the case in Fig. 3 for (b) and (d) curves, where the minimum energy splitting is pushed at $s = 0.49$. Since the
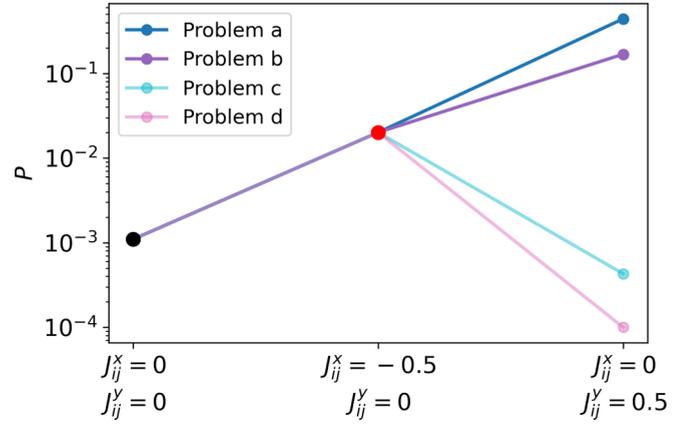
transverse field is larger earlier in the anneal, one might expect the minimum gap to be larger for problem (b) compared to that for problem (a), and likewise for problem (d) compared to problem (c). However, Fig. 3 shows the opposite behavior. This is because not only the inner qubits but also some of the outer qubits flip between the ground state and each of the 16 degenerate excited states. At a fixed transverse field, the multiqubit tunneling amplitude is largest when the outer qubits are coupled ferromagnetically. Since the transverse field is not fixed, the two effects compete with each other; coupling the outer qubits antiferromagnetically pushes the avoided crossing to a smaller s hence increasing in the transverse field, but the increase is not enough to compensate the reduction of multiqubit tunneling amplitude due to AFM coupling. As a result, the largest minimum gap happens when all couplers are FM, as they are in problem (a).

Figure 4 compares probabilities of success for the scenarios presented in Fig. 3. As expected, the success probabilities in Fig. 4 are small and correlate with the minimum gap sizes in Fig. 3. While the XX coupling enhances the performance regardless of the SRT, the YY coupling may increase or decrease the probability of success depending on the SRT. The best performance is obtained for problem (a) when YY interaction is on. Since the magnitudes of J_{ij}^x and J_{ij}^y are the same, switching from YY to XX interaction is equivalent to one of the de-signing processes proposed in Ref. [28]. Clearly from problem (a) to (b) de-signing did not improve the performance, in contrast to Ref. [28]. Therefore, although nonstoquasticity does not automatically lead to advantage in QA, with a right choice of SRT a nonstoquastic Hamiltonian with YY interaction ($J_{ij}^y = 0.5$, $J_{ij}^x = 0$) can outperform the original stoquastic Hamiltonian ($J_{ij}^y = J_{ij}^x = 0$) as well as the de-signed stoquastic Hamiltonian ($J_{ij}^y = 0$, $J_{ij}^x = -0.5$).

B. Random problems

We now investigate whether the observations in the previous example hold for random problems. We generate problems by randomly selecting J_{ij}^z from $[-1, 1]$ with 4 bits of precision (16 evenly distributed values within that range). The biases are also selected in the same way or all taken to be zero. The connectivity graph is again a Chimera graph [34], but with
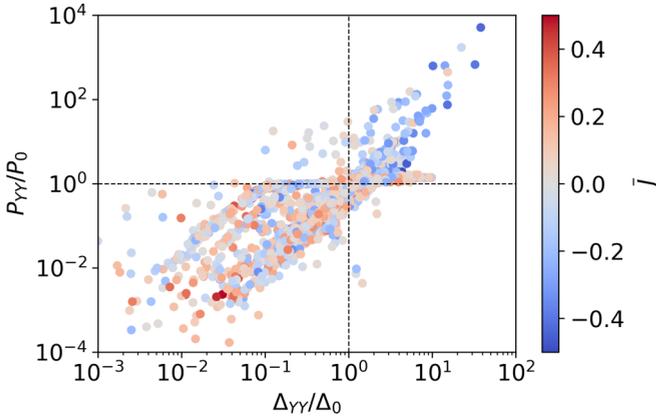
FIG. 5. Scatter plot of the relative success probability and the relative minimum gap size corresponding to the driver Hamiltonian (4) with $J_{ij}^x = 0$ and $J_{ij}^y = 0.5$. For each of the 130 random problems we have applied 50 random SRTs. The color-code shows the value of $\bar{J}$ for each instance. The majority of cases with improved performance have $\bar{J} < 0$.

FIG. 6. Scatter plot of relative success probabilities and relative minimum gaps for SRTs obtained by minimizing Eq. (18). Color coding represents the distance from the optimum. We have only kept suboptimal solutions with $\bar{J} - \bar{J}_{opt} < 0.1$.

$N = 12$ (six qubits within each unit cell) to limit the computation time. Since most generated problems of this size are easy, we only keep the ones with small gaps: $\Delta < 0.1$ GHz. In total we generated 100 problems with random h_i and 30 problems with $h_i = 0$. Since we did not find any qualitative difference between the two cases, we combine them into a single set and present them together.

The best performance in the previous crafted example was obtained when the couplers were all FM. In random problems, however, the coupling terms are usually frustrated, meaning that no solution can satisfy them all simultaneously. Therefore, no SRT can make all the couplers FM, i.e., making a frustrated problem unfrustrated. Since qubits with strongest couplings are most likely to be correlated, it is reasonable to choose SRTs that make those couplers FM and allow the weak ones to be AFM. We define the average coupling strength for the transformed Hamiltonian as

$$\bar{J} = \frac{1}{N_J} \sum_{i,j} J_{ij}'^z, \qquad (17)$$

where N_J is the number of couplers. Clearly, $\bar{J}$ varies with SRT and the more negative it is the more ferromagnetic the couplings are.

For each of the 130 generated random problems we chose 50 SRTs by randomly selecting α_i from ± 1 with equal probability. Figure 5 shows a scattered plot of success probability versus minimum gap in the presence of YY interaction for all the 6500 instances. The probabilities and the gap values are normalized to their corresponding values in the original problem. Therefore, a value bigger than 1 means improvement. A correlation between the probabilities and the minimum gap sizes can be recognized in Fig. 5. The color-coding represents the average coupling strength $\bar{J}$. As in the previous example, adding YY interactions can improve or impair the performance depending on the SRT. However, as the color coding indicates, there is a close correlation between the performance and the sign of $\bar{J}$; most of the improved instances have FM-dominated couplers ($\bar{J} < 0$).
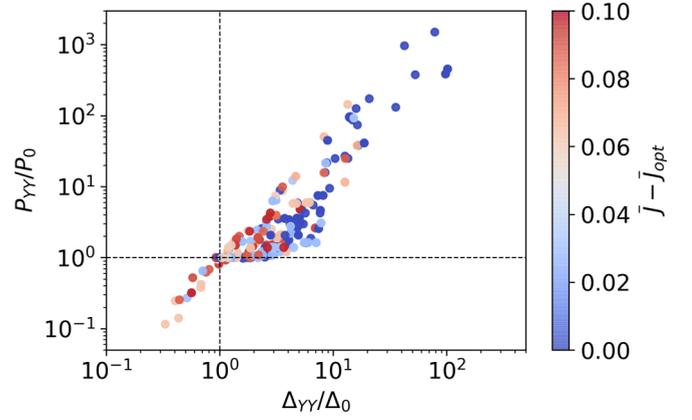
The above observation as well as the results of the previous example suggest that a SRT that minimizes $\bar{J}$ is likely to improve the performance. Using Eqs. (10) and (17), we write the objective function as

$$\bar{J}(\vec{\alpha}) = \frac{1}{N_J} \sum_{i,j} J_{ij}^z \alpha_i \alpha_j. \qquad (18)$$

The solution $\vec{\alpha}_{opt}$ that minimizes (18) defines the desired SRT. Since the number of variables is not very large ($N = 12$), we can find all global and local minima of (18) through exhaustive search. This, however, is not possible for larger problems. Since α_i is a binary variable with values ± 1 (similar to S_i), the objective function (18) itself is an Ising problem. Indeed, (18) forms the quadratic part of the problem Hamiltonian (3) and is equivalent to H_P when $h_i = 0$. Finding $\vec{\alpha}_{opt}$ is, therefore, NP-hard, i.e., as complex as minimizing the original problem Hamiltonian. However, the quantum annealer itself can be used to minimize Eq. (18). Especially for problems with $h_i = 0$, the optimal solution to the original problem H_P and the optimal SRT coincide: $\vec{S}_{opt} = \vec{\alpha}_{opt}$, Therefore, solving the problem itself with QA gives the SRT for the next run and the process can be repeated iteratively until the desired solution is reached. Moreover, suboptimal solutions to (18) also improve the performance and could be as good as, and sometimes even better than, the optimal solution, as we show below.

Figure 6 shows a scatter plot similar to Fig. 5, but instead of being random, the SRTs are obtained via optimization of Eq. (18). Colors represent the distance from the optimal $\bar{J}_{opt}$. We only kept local minima with $\bar{J} - \bar{J}_{opt} < 0.1$ to investigate what the results look like in the case where the minimization of Eq. (18) does not converge to the global but rather to a local minimum close to the global one. For each one of the 130 problems, the threshold 0.1 corresponds to 0.3%-0.6% of the total 2^{12} SRTs and is chosen according to the following argument. Ignoring frustration, the average value of $\bar{J}_{opt}$ can be calculated as

$$\langle \bar{J}_{opt} \rangle \approx \left\langle -M^{-1} \sum_{i,j} |J_{i,j}| \right\rangle, \qquad (19)$$

TABLE I. Percentages of improved instances for suboptimal and optimal solutions of Eq. (18).

$\bar{J} - \bar{J}_{\text{opt}}$	P_{YY}/P_0	P_{YY}/P_{XX}
Suboptimal	83%	56%
Optimal	91%	82%

where the average is over the uniform distribution of J_{ij} values. With the 8-bit precision on J_{ij}, the average turns out to be $\langle \bar{J}_{\text{opt}} \rangle = -9/17$. Due to frustration, however, this is just a lower bound on the actual value. The maximum value can be obtained in a similar way without the negative sign and and gives us an average maximum of $\langle \bar{J}_{\text{max}} \rangle = 9/17$. The threshold 0.1 then corresponds to an approximation ratio, defined as $(\bar{J} - \bar{J}_{\text{opt}})(\bar{J}_{\text{max}} - \bar{J}_{\text{opt}})^{-1}$, of about 10%. As it is evident from Fig. 6, for the majority of the 130 problems, the minimum gap size is increased for the SRT corresponding to the optimal or suboptimal solutions of Eq. (18), and the probability of success is improved by up to more than 3 orders of magnitude. There were also cases where suboptimal solutions to Eq. (18) gave better probability of success than the optimal one. Table I shows the percentage of improved cases: 91% of the optimal solutions and 83% of the suboptimal solutions led to SRTs that increased the success probability.

Finally, we investigate how a nonstoquastic Hamiltonian with YY interaction compares, in terms of QA performance, to the corresponding stoquastic (de-signed) Hamiltonian with XX interaction. The magnitudes of the YY and XX interactions are the same in the two Hamiltonians ($|J_{ij}^x| = |J_{ij}^y| = 0.5$). Since for YY interaction the adiabatic path depends on the SRT, we expect the performance to be better or worse than the stoquastic case depending on the SRT. Indeed, we find that for the majority of random SRTs, stoquastic Hamiltonians yield better performance than the corresponding nonstoquastic ones, in agreement with Ref. [28]. However, when we find optimal SRTs by minimizing Eq. (18), the nonstoquastic Hamiltonians on average outperform the corresponding stoquastic ones. Figure 7 plots the relative probabilities and minimum gap sizes for the 130 random problems studied before. The percentages of improvement are also reported in Table I. Quantum annealing with the stoquastic Hamiltonian (XX interaction) is outperformed by the corresponding nonstoquastic one (YY interaction) for 82% of the problems when the optimal SRT [the global minimum of Eq. (18)] was applied and for 56% of cases when suboptimal SRTs were applied.

IV. CONCLUSIONS

We investigated the effect of a nonstoquastic Hamiltonian with YY interaction between the qubits on the performance of quantum annealing. The existence of YY interaction makes the adiabatic path and the performance strongly dependent on SRTs. Random transformations in general do not improve the performance. We found that the transformation that makes the average ZZ coupling maximally ferromagnetic is most likely to improve the performance, sometimes by several orders
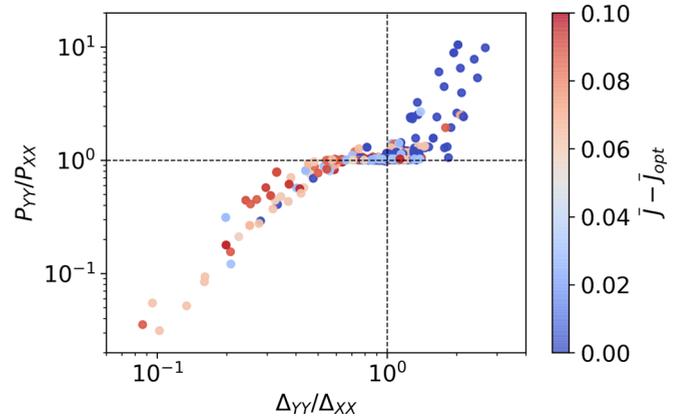
FIG. 7. Comparison between the nonstoquastic Hamiltonian with YY interaction ($J_{ij}^x = 0$ and $J_{ij}^y = 0.5$) and the (de-signed) stoquastic Hamiltonian obtained by replacing YY with XX interaction ($J_{ij}^x = -0.5$ and $J_{ij}^y = 0$). The instances are the same as in Fig. 6 with colors representing the distance from the optimum.

of magnitude. Suboptimal solutions also improve the performance, but with less frequency. We should mention that the SRT obtained by minimizing the average coupling is not the best possible SRT among all the exponentially large number of possible transformations. It is conceivable that more elegant algorithms, e.g., that use information from previously obtained solutions, lead to better transformations. One may also use machine-learning techniques to choose SRTs based on the numerical observations or experimental data once large-scale quantum annealers with nonstoquastic interactions are built.

In the physical implementation of Ref. [29], nonstoquasticity was obtained by adding capacitive coupling to magnetically coupled flux qubits. The resulting Hamiltonian had an extra XX coupling in addition to the expected YY coupling. The XX coupling was stoquastic with a magnitude that depended on ZZ interaction. Both XX and YY couplings favored FM over AFM correlation in terms of contribution to two-qubit tunneling. As a result, the dependence on SRT is expected to be stronger than that for YY coupling alone. Moreover, the addition of coupling capacitors will increase the total capacitance of each qubit, resulting in smaller tunneling amplitudes. However, only a few percent reduction is expected, not large enough to eliminate the orders of magnitude enhancement of performance observed above. More research is needed to assess the value of such capacitive interactions in practical quantum annealers at large scales.

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

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

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

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

[5] M. H. S. Amin, Consistency of the Adiabatic Theorem, Phys. Rev. Lett. **102**, 220401 (2009).

[6] D. A. Lidar, A. T. Rezakhani, and A. Hamma, Adiabatic approximation with exponential accuracy for many-body systems and quantum computation, J. Math. Phys. **50**, 102106 (2009).

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

[8] 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 *et al.*, Quantum annealing with manufactured spins, Nature (London) **473**, 194 (2011).

[9] S. Bravyi, D. P. Divincenzo, R. Oliveira, and B. M. Terhal, The complexity of stoquastic local Hamiltonian problems, Quantum Inf. Comput. **8**, 361 (2008).

[10] S. Bravyi and B. Terhal, Complexity of stoquastic frustration-free Hamiltonians, SIAM J. Comput. **39**, 1462 (2015).

[11] M. Marvian, D. A. Lidar, and I. Hen, On the computational complexity of curing non-stoquastic Hamiltonians, Nat. Commun. **10**, 1571 (2019).

[12] J. Klassen and B. M. Terhal, Two-local qubit Hamiltonians: when are they stoquastic? Quantum **3**, 139 (2019).

[13] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Sign problem in the numerical simulation of many-electron systems, Phys. Rev. B **41**, 9301 (1990).

[14] M. Troyer and U.-J. Wiese, Computational Complexity and Fundamental Limitations to Fermionic Quantum Monte Carlo Simulations, Phys. Rev. Lett. **94**, 170201 (2005).

[15] L. Gupta and I. Hen, Elucidating the interplay between non-stoquasticity and the sign problem, Adv. Quantum Technol. **3**, 1900108 (2019).

[16] S. V. Isakov, G. Mazzola, V. N. Smelyanskiy, Z. Jiang, S. Boixo, H. Neven, and M. Troyer, Understanding Quantum Tunneling through Quantum Monte Carlo Simulations, Phys. Rev. Lett. **117**, 180402 (2016).

[17] M. B. Hastings, Obstructions to classically simulating the quantum adiabatic algorithm, Quantum Inf. Comput. **13**, 1038 (2013).

[18] M. Jarret, S. P. Jordan, and B. Lackey, Adiabatic optimization versus diffusion Monte Carlo methods, Phys. Rev. A **94**, 042318 (2016).

[19] E. Andriyash and M. H. Amin, Can quantum Monte Carlo simulate quantum annealing? arXiv:1703.09277.

[20] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, Adiabatic quantum computation is equivalent to standard quantum computation, SIAM Rev. **50**, 755 (2008).

[21] J. D. Biamonte and P. J. Love, Realizable Hamiltonians for universal adiabatic quantum computers, Phys. Rev. A **78**, 012352 (2008).

[22] A. Mizel, Proof of efficient, parallelized, universal adiabatic quantum computation, Phys. Rev. A **99**, 022311 (2019).

[23] L. Hormozi, E. W. Brown, G. Carleo, and M. Troyer, Nonstoquastic Hamiltonians and quantum annealing of an Ising Spin glass, Phys. Rev. B **95**, 184416 (2017).

[24] H. Nishimori and K. Takada, Exponential enhancement of the efficiency of quantum annealing by non-stoquastic Hamiltonians, Front. ICT **4**, 2 (2017).

[25] Y. Susa, J. F. Jadebeck, and H. Nishimori, Relation between quantum fluctuations and the performance enhancement of quantum annealing in a nonstoquastic Hamiltonian, Phys. Rev. A **95**, 042321 (2017).

[26] T. Albash, Role of nonstoquastic catalysts in quantum adiabatic optimization, Phys. Rev. A **99**, 042334 (2019).

[27] G. A. Durkin, Quantum speedup at zero temperature via coherent catalysis, Phys. Rev. A **99**, 032315 (2019).

[28] E. Crosson, T. Albash, I. Hen, and A. P. Young, De-signing Hamiltonians for quantum adiabatic optimization, Quantum **4**, 334 (2020).

[29] I. Ozfidan, C. Deng, A. Y. Smirnov, T. Lanting, R. Harris, L. Swenson, J. Whittaker, F. Altomare, M. Babcock, C. Baron, A. J. Berkley, K. Boothby, H. Christiani, P. Bunyk, C. Enderud, B. Evert, M. Hager, A. Hajda, J. Hilton, S. Huang *et al.*, Demonstration of Nonstoquastic Hamiltonian in Coupled Superconducting Flux Qubits, Phys. Rev. App. **13**, 034037 (2020).

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

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

[32] E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, H. B. Meyer, and P. Shor, Quantum adiabatic algorithms, small gaps, and different paths, Quantum Inf. Comput. **11**, 181 (2011).

[33] N. G. Dickson and M. H. Amin, Algorithmic approach to adiabatic quantum optimization, Phys. Rev. A **85**, 032303 (2012).

[34] P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz *et al.*, IEEE Trans. Appl. Supercond. **24**, 2318294 (2014).

[35] J. R. Johansson, P. D. Nation, and F. Nori, QuTiP: An open-source Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 8 (2012)

[36] J. R. Johansson, P. D. Nation, and F. Nori, QuTiP 2: A Python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 3 (2013)