Uncertain fate of fair sampling in quantum annealing

Mario S. Könz, Guglielmo Mazzola, Andrew J. Ochoa, Helmut G. Katzgraber, and Matthias Troyer

Phys. Rev. A **100**, 030303 — Published 20 September 2019

DOI: 10.1103/PhysRevA.100.030303

Uncertain Fate of Fair Sampling in Quantum Annealing

Mario S. Könz,[1,*] Guglielmo Mazzola,[1] Andrew J. Ochoa,[2] Helmut G. Katzgraber,[2,3,4] and Matthias Troyer[1,5]

[1]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*
[2]*Department of Physics and Astronomy, Texas A&M University, College Station, Texas 77843-4242, USA*
[3]*1QB Information Technologies (1QBit), Vancouver, British Columbia, V6B 4W4, Canada*
[4]*Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501 USA*
[5]*Microsoft Quantum, Microsoft, Redmond, WA 98052, USA*
(Dated: August 12, 2019)

Recently, it was demonstrated both theoretically and experimentally on the D-Wave quantum annealer that transverse-field quantum annealing does not find all ground states with equal probability. In particular, it was proposed that more complex driver Hamiltonians beyond transverse fields might mitigate this shortcoming. Here, we investigate the mechanisms of (un)fair sampling in quantum annealing. While higher-order terms can improve the sampling for selected small problems, we present multiple counterexamples where driver Hamiltonians that go beyond transverse fields do not remove the sampling bias. Using perturbation theory we explain why this is the case. In addition, we present large-scale quantum Monte Carlo simulations for spin glasses with known degeneracy in two space dimensions and demonstrate that the fair-sampling performance of quadratic driver terms is comparable to standard transverse-field drivers. Our results suggest that quantum annealing machines are not well suited for sampling applications, unless post-processing techniques to improve the sampling are applied.

PACS numbers: 75.50.Lk, 75.40.Mg, 05.50.+q, 03.67.Lx

Quantum annealing (QA) [1–9] is a heuristic designed to harness the advantages of quantum mechanics to solve optimization problems. The performance of QA and, in particular, QA machines such as the D-Wave Systems Inc. devices are controversial to date [10–32]. Most studies have focused on finding the minimum value of a binary quadratic cost function (problem Hamiltonian), yet less on the *variety* of solutions obtained when repeating the optimization procedure multiple times. Important applications that rely on sampling, such as SAT-based probabilistic membership filters [33–36], propositional model counting and related problems [37–39], or machine learning [40, 41] rely on ideally uncorrelated states. This sought-after *fair sampling* ability of an algorithm, i.e., the ability to find (ideally all) states associated with a cost function with (ideally) the same probability, is thus of importance for a variety of applications. Moreover, the ability of an algorithm to sample ground states with similar probability is directly related its ergodicity which strongly influences the efficiency of optimization and sampling techniques.

Following small-scale studies [42], Ref. [43] recently performed systematic experiments on the D-Wave 2X annealer. The results demonstrated that quantum annealers using a transverse-field driver are biased samplers, an effect also observed in previous studies [13, 14, 44]. Matsuda *et al.* [42] conjectured that more complex drivers might alleviate this bias, something we test in this work.

Binary optimization problems can be mapped onto k-local spin Hamiltonians. Without loss of generality we study problem Hamiltonians with N degrees of freedom in a z-basis of the form

$$\mathcal{H}_{\mathrm{P}} = -\sum_{i,j=1}^{N} J_{ij}\sigma_i^{\mathrm{z}}\sigma_j^{\mathrm{z}}, \tag{1}$$

where σ_i^{z} is the z-component of the Pauli operator acting on site i. Note that local biases can also act on the variables. For such a problem Hamiltonian, in principle, a driver of the form

$$\mathcal{H}_{\mathrm{x},N} = \sum_{n=1}^{N} \Gamma^{\mathrm{x},n}[\otimes\sigma^{\mathrm{x}}]^n \tag{2}$$

would induce transitions between all states and therefore ensure a fair sampling, provided the anneal is performed slow enough. Unfortunately, such driver is hard to engineer and, at best, one can expect drivers of the form $\mathcal{H}_{\mathrm{x},2} = -\sum_j \Gamma^{\mathrm{x}}\sigma_j^{\mathrm{x}} + \sum_{j,k} K_{j,k}^{\mathrm{x}}\sigma_j^{\mathrm{x}}\sigma_k^{\mathrm{x}}$. Quantum fluctuations are induced by the driver and then reduced to sample states from the problem Hamiltonian, i.e., $\mathcal{H}(t) = (1-t/T)\mathcal{H}_{\mathrm{x},n} + (t/T)\mathcal{H}_{\mathrm{P}}$, where $t \in [0,T]$, T the annealing time, and n the order of the interactions in the driver. For an infinitely-slow anneal, the adiabatic theorem [2, 45] ensures that for $t = T$ a (ground) state of the problem Hamiltonian is reached. It is therefore desirable to know if after an infinite amount of repetitions, the process results in *all* minimizing states, i.e., fair sampling.

Here we analyze the behavior of more complex drivers of the form $\mathcal{H}_{\mathrm{x},n}$ ($n > 1$) on the fair sampling abilities of QA. Following Ref. [42] we first study small systems where the Schrödinger equation can be integrated using *QuTiP* [46]. We have exhaustively analyzed all possible graphs with up to $N = 6$ with both ferromagnetic and antiferromagnetic interactions and show in Fig. 1 paradigmatic examples that illustrate different scenarios using drivers with $n \leq 2$. Even for some of these small instances, in some cases the inclusion of higher-order driver terms does not remove the bias. If we

* mkoenz@itp.phys.ethz.ch

anneal adiabatically, i.e., T large enough, the instantaneous ground states is never left, which means towards the end of annealing at $T - \lambda$ (for a small $\lambda > 0$) the system is in the ground state of $\mathcal{H}(T - \lambda)$. This observation is key to predicting the sampling probabilities for different degenerate ground states. These probabilities are given by squaring the amplitudes of the lowest eigenvector of $\mathcal{H}(T - \lambda)$, assuming for now the small contribution from the driver lifts the degeneracies. Because $\mathcal{H}(T - \lambda)$ can be viewed as $\mathcal{H}_{\rm P}$ perturbed by $\mathcal{H}_{{\rm x},n}$, we analyze fair sampling using a perturbative approach [47]. To better quantify the fair-sampling behavior of a given system, we introduce the term "*hard suppression*" (i.e., total suppression) if the sampling probability is 0 for a particular ground-state configuration at the end of the anneal and the term "*soft suppression*" if a particular state is undersampled by a certain finite fraction in comparison to other minimizing configurations. Finally, we complement these studies with quantum Monte Carlo simulations for large two-dimensional Ising spin-glass problems following Ref. [43] and discuss the effects of higher-order drivers. Our results show that QA is not well suited for sampling applications, unless post-processing techniques are implemented [48].

Perturbation theory.— In the following, we show how to determine the sampling probabilities, as well as the influence the driver has on it. In short, if we apply $\mathcal{H}_{\rm D}$ as a perturbation of strength λ to $\mathcal{H}_{\rm P}$, some degeneracies will be lifted, i.e., the perturbed ground-state-space is smaller. The ground-state-space is never left during an adiabatic anneal, hence it will not be possible to reach the entire ground-state-space of the unperturbed Hamiltonian by annealing in the generic case. This analysis holds for any driver Hamiltonian $\mathcal{H}_{\rm D}$, not just the stoquastic $\mathcal{H}_{{\rm x},n}$-type drivers we use in this work. In nondegenerate perturbation theory, the first-order corrected wave function $|n\rangle$ is given by $|n\rangle = |n^0\rangle + \lambda \sum_{m \neq n} \frac{\langle m^0|H_{\rm D}|n^0\rangle}{E_m^0 - E_n^0} |m^0\rangle$, where $|n^0\rangle$ are the eigenstates and E_n^0 the eigenvalues of the unperturbed Hamiltonian $\mathcal{H}_{\rm P}$. If states $m \neq n$ are degenerate, i.e., $E_m^0 = E_n^0$, there is a singularity. To avoid it, degenerate perturbation theory requires linear combinations $|\alpha^0\rangle$ which satisfy $\langle \alpha^0|H_{\rm D}|\beta^0\rangle \sim \delta_{\alpha,\beta}$ in every degenerate subspace. This ensures that the corrected wave function does not diverge due to singularities. We focus on the ground-state subspace, but the procedure is identical for any subspace. Given k ground-states $|n_{\rm gs}^0\rangle$ of $\mathcal{H}_{\rm P}$ with energy $E_{\rm gs}^0$, we need to form the $k \times k$ subspace matrix $V_{n,m} = \langle n_{\rm gs}^0|H_{\rm D}|m_{\rm gs}^0\rangle$. Because $H_{\rm D}$ is hermitian, V is too. Every hermitian matrix can be diagonalized by a unitary transformation ($U^{-1}VU = D$) and we find the correct linear combinations $|\alpha_{\rm gs}^0\rangle$ in the columns of U. It satisfies $\langle \alpha_{\rm gs}^0|V|\beta_{\rm gs}^0\rangle \sim \delta_{\alpha,\beta}$ since D is diagonal. The diagonal entries of D are the eigenvalues of V and also the first order energy corrections E_α^1. We need to pick the lowest eigenvalue $E_{\alpha,{\rm low}}^1$ and find the corrected ground state energy $E_{GS} = E_{\rm GS}^0 + E_{\alpha,{\rm low}}^1$. The corresponding l eigenvectors $|\alpha_{\rm gs}^0\rangle$ will now determine the sampling behavior, since the annealing state will be in their span. The following scenarios can occur:

(i) $l = 1$ – In this case $p_i = \langle n_{\rm gs}^0|\phi_{\alpha,{\rm low}}^0\rangle^2$, because there is a single state $|\phi_{\alpha,{\rm low}}^0\rangle$. If sampling is fair, it will remain fair, regardless of how much the higher energy eigenvalues of $\mathcal{H}_{\rm P}$ change during the adiabatic anneal. If certain states have $p_i = 0$, $|n_{\rm gs}^0\rangle$ will never be available at the end of the anneal.

(ii) $l > 1$ – Let A be the $k \times m$ matrix consisting of all l $|\phi_{\alpha,{\rm low}}^0\rangle$. If there is a vector x such that $Ax = y$ and $y_i \cdot y_i^* = 1$

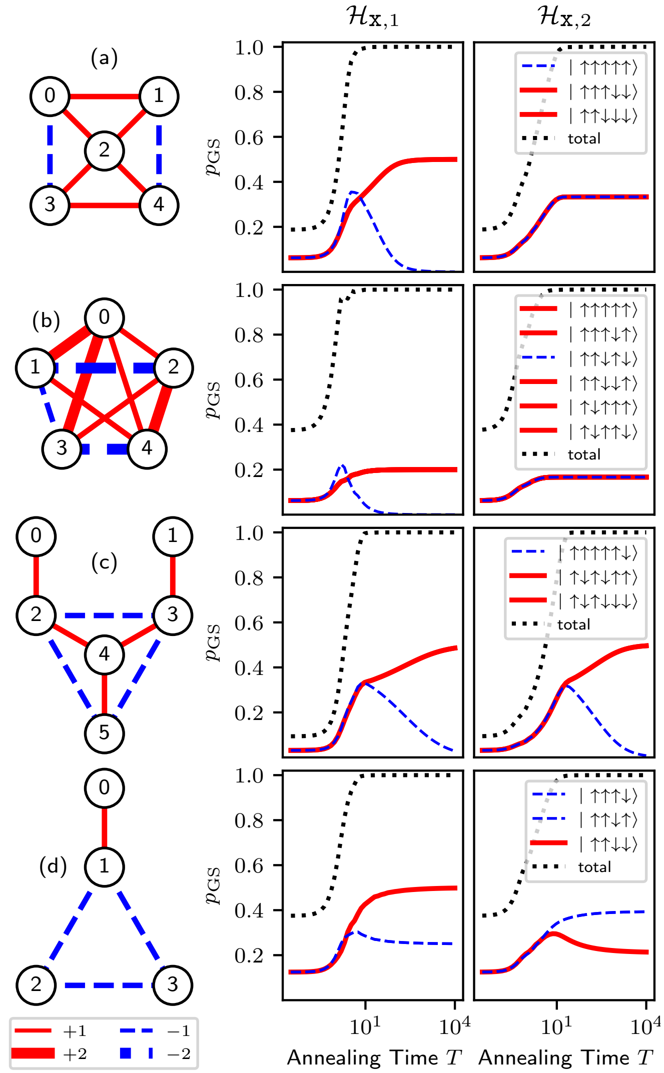

Figure 1. Toy problems with up to $N = 6$ variables and both ferromagnetic (solid lines) and antiferromagnetic (dashed lines) interactions of different strength (thickness of the lines) integrated using *QuTiP* [46]. Data for both transverse field ($\mathcal{H}_{{\rm x},1}$) and quadratic ($\mathcal{H}_{{\rm x},2}$) drivers are shown. The data show the instantaneous probability $p_{\rm GS}$ to find different states that minimize the cost function (up to spin-reversal symmetry) as a function of anneal time t. In all cases, the first spin (labeled with 0) is in the state $|\uparrow\rangle$. (a) Toy problem studied in Ref. [42] where drivers with $n = 2$ sample the states fairly. (b) Similar behavior to panel (a), however, the unfair sampling sets in earlier in the anneal. (c) Even the inclusion of $n = 2$ drivers does not remove the unfair sampling. As in (a) and (b) at least one state is suppressed. Note that a driver Hamiltonian with $n = 4$ results in fair sampling. (d) The sampling is not exponentially biased. However, one state occurs twice as likely as the others. Note that the sampling probabilities swap when going from $n = 1$ to $n = 2$.

for all i, then fair sampling is potentially possible according to first order. If there exists an i such that $y_i = 0$ for all x, then that ground-state is never found. The same argument can be made for biased sampling where there is no suppression but certain states are over-sampled.

(iii) V *is zero* – All eigenvalues $E^1_\alpha = 0$ and the sampling probabilities are determined by second-order perturbation, i.e., the probabilities depend on higher eigenvalues of $\mathcal{H}_\mathrm{P}$ (see Fig. 3).

$$\mathcal{H}_{\mathrm{x},1} = \begin{pmatrix} 0 & -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & -1 & -1 & 0 & 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & -1 & -1 & 0 \end{pmatrix}$$

$$V = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{pmatrix} \Longrightarrow k_\mathrm{low} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

Figure 2. To obtain the sampling probabilities, the ground-state eigenvectors $|g_i\rangle$ need to be known, represented here as shaded rows and columns in the matrix $\mathcal{H}_{\mathrm{x},1}$, since the solution of the diagonal $\mathcal{H}_\mathrm{P}$ is a classical one. One then needs to analyze the subspace V that is formed by restricting the driver (here $\mathcal{H}_{\mathrm{x},1}$) to space spanned by the ground-states of $\mathcal{H}_\mathrm{P}$. The lowest eigenvector(s) determine the sampling probabilities. In this example, there is one lowest eigenvector and the first ground state corresponding to the top column (first row) is suppressed (all spins up) and will never be sampled in an adiabatic anneal.

The second-order perturbation terms only play a relevant role if V is trivial. If $l > 0$, the sampling behavior is determined by V which does not depend on $\mathcal{H}_\mathrm{P}$. This means that the sampling behavior is purely a property of the driver Hamiltonian $\mathcal{H}_{\mathrm{x},n}$ and the ground-state eigenvectors of $\mathcal{H}_\mathrm{P}$. We have verified this on numerous small systems, as well as structured and random-coupling systems with direct integration and were always able to predict the sampling probabilities that correspond to the state found after the anneal.

Figure 1(a) is the example studied in Ref. [42], where $\mathcal{H}_{\mathrm{x},2}$ lead to fair sampling. There are 6 degenerate ground states, 2 of which are suppressed. With a driver of the form $\mathcal{H}_{\mathrm{x},1}$ we obtain $l > 1$, meaning that there are multiple $|\alpha^0_\mathrm{gs}\rangle$ states that determine the sampling. However, the suppressed states have $p_i = 0$. In Fig. 1(b) we show a more complex example – the smallest problem we were able to find that has $l = 1$ and one state where $p_i = 0$. It is a 12-fold degenerate system with two states fully suppressed when $\mathcal{H}_{\mathrm{x},1}$ is used as a driver. The fact that $l = 1$ could be a reason why the suppression sets in earlier during the anneal. This case is problematic for annealing schedules that are fast quenches, because there is a much smaller window during the anneal where the total ground-state probability is approximately unity and the suppressed state has not yet reached zero probability. Using $\mathcal{H}_{\mathrm{x},2}$ results in fair sampling. Figure 1(c) shows a system that has 6 ground states with 2 ground states in total suppression. Using $\mathcal{H}_{\mathrm{x},1}$ as a driver, we obtain $l = 1$ and a unique $|\alpha^0_\mathrm{gs}\rangle$ with 2 totally suppressed states with zero probability. For $\mathcal{H}_{\mathrm{x},2}$, $l > 1$ we obtain multiple $|\alpha^0_\mathrm{gs}\rangle$ states. However, two states are totally suppressed. Using $\mathcal{H}_{\mathrm{x},3}$ as a driver results in $l = 1$ and a unique $|\alpha^0_\mathrm{gs}\rangle$. However, there is a soft suppression of two ground states (not shown). Finally, using $\mathcal{H}_{\mathrm{x},4}$ we obtain $l = 1$ and fair sampling. The case shown in Fig. 1(d) reveals the undersampled states when $\mathcal{H}_{\mathrm{x},1}$ is replaced by $\mathcal{H}_{\mathrm{x},2}$. More precisely, it changes from $l > 1$ with 4 soft suppressed states to $l = 1$ with the previously 2 oversampled state now being undersampled. Using a driver $\mathcal{H}_{\mathrm{x},3}$ results in $l = 1$ and fair sampling.

Figure 3 shows a problem where by changing the strength of $J_{3,4}$ one can change the sampling bias arbitrarily. Note that changing $J_{3,4} = -1.2$ to -1.8 does shift the relative energies of the ground state and the various low exited states, but does not change their order. In terms of perturbation theory, V is trivial, and second-order perturbations dictate the behavior of the system. Because there are terms $\propto 1/(E_i - E_\mathrm{GS})$ with E_GS the ground state energy, shifting the energy levels E_i will influence the sampling.

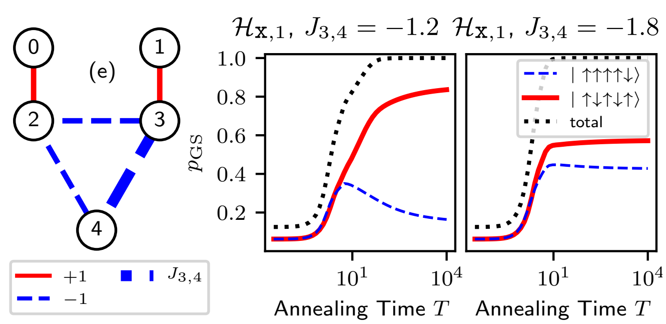

Figure 3. Toy problem with $N = 5$ variables and both ferromagnetic (solid lines) and antiferromagnetic (dashed lines) interactions of different strength (thickness of the lines) integrated using *QuTiP* [46]. Data for a driver $\mathcal{H}_{\mathrm{x},1}$. By changing $J_{3,4}$ one can change the sampling bias arbitrarily.

Quantum Monte Carlo results.—To corroborate our results with larger systems, we perform a fair-sampling study analogous to the one done in Ref. [43] for two-dimensional Ising spin glasses on a square lattice with periodic boundary conditions. The couplers are chosen from $J_{i,j} \in \{\pm 1, \pm 2, \pm 4\}$. This ensures that degeneracies are small. The coupler-configuration space is mined for specific degeneracies as done in Ref. [43]. Figure 4 shows representative rank-ordered probabilities to find different minimizing configurations using simulated annealing (SA) [49–51], as well as transverse-field simulated quantum annealing (SQA-$\mathcal{H}_{\mathrm{x},1}$) [5, 52–56] and simulated quantum annealing with a *stoquastic* two-spins driver (SQA-$\mathcal{H}_{\mathrm{x},2}$) [57] [58]. The data are averaged over 100 disorder realizations. While the data for SA for this particular problem show a fair sampling of all minimizing configurations, neither a transverse-field $\mathcal{H}_{\mathrm{x},1}$ nor a more complex $\mathcal{H}_{\mathrm{x},2}$ driver can remove the bias. This suggests that even if QA machines with more complex $\mathcal{H}_{\mathrm{x},2}$ drivers are constructed, sampling will remain unfair unless post-processing is applied [48]. The close connection between SQA and QA performances is discussed in Refs. [53, 54].

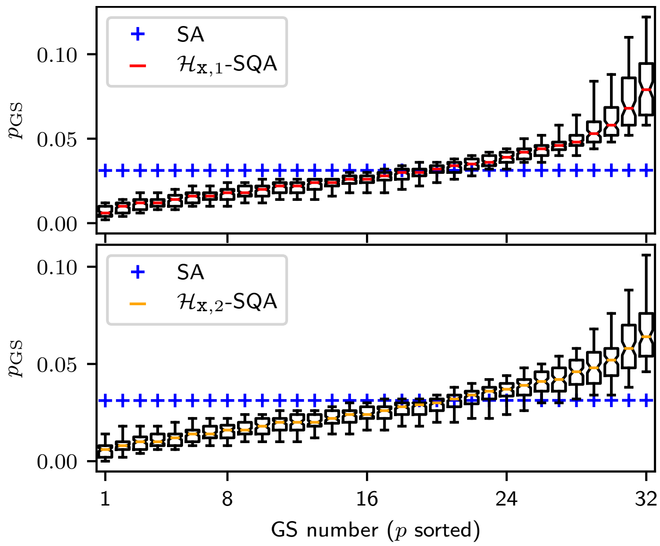

Figure 4. Rank-ordered probability p_{GS} to find different degenerate states for a two-dimensional Ising spin glass with $N = 8^2 = 64$ and a ground-state degeneracy of 32. The data are averaged over 100 disorder realizations. For each instances, 500 independent runs are performed and the probability to find a given ground-state configuration computed. While simulated annealing (SA) samples close to fair, both SQA-$\mathcal{H}_{\mathrm{x},1}$ and SQA-$\mathcal{H}_{\mathrm{x},2}$ show a clear bias in the sampling. In particular, there is no notable improvement of using a driver of the form $\mathcal{H}_{\mathrm{x},2}$ over a transverse-field driver $\mathcal{H}_{\mathrm{x},1}$. We have also simulated systems with up to $N = 12^2 = 144$ variables and ground states with up to 96-fold degeneracy obtaining similar results. Note that the bias becomes more pronounced for increasing system size N (not shown).

Effects of more complex drivers.— The following section shows that any driver (stoquastic or non-stoquastic) needs to be sufficiently dense to sample fair for generic larger systems. To predict the sampling probabilities it is sufficient to know V (except when V is trivial). V can be constructed with only the ground-state eigenvectors of $\mathcal{H}_{\mathrm{P}}$ (no eigen-energies needed) and the driver $\mathcal{H}_{\mathrm{x},n}$. This can be used to analyze different drivers—without specifying a concrete problem Hamiltonian $\mathcal{H}_{\mathrm{P}}$—by merely sampling from possible ground-state combinations. As an example consider a 2-fold degeneracy in a 5-spin system. Because we want to test the driver for all possible ground-state combinations, we can exhaustively generate all the ground-state pairs, i.e., $N(N-1)/2$, where $N = 2^5$ and check V for each one pair, to analyze the sampling behavior. For larger N, we sample instead of searching exhaustively.

Figure 5 shows how probable it is for a random degeneracy and a ground-state combination to be sampled according to the following categories:

fair – All ground states have the same probability.

soft – At least one ground state is soft suppressed with a ratio smaller than $1:100$ (least likely vs most likely).

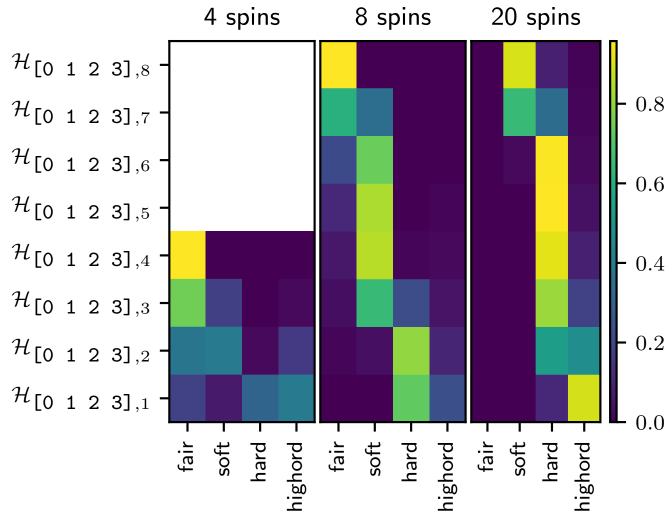

Figure 5. For each spin system all possible degeneracies are sampled with 400 random ground-state combinations. For small systems, all combinatorial possibilities of ground states can be checked, and the random sampling approximates the exact result fast. For 4 spins, fair sampling is reached for all possible problem Hamiltonians $\mathcal{H}_{\mathrm{P}}$, once the driver matrix is dense, i.e., all off-diagonal contain nonzero entries,. For example, this is the case when using $\mathcal{H}_{\mathrm{x},4}$ for a system with 4 spins. Similarly, for 8 spins the system moves from hard to soft, to fair sampling as more complex drivers are used. As the system with 20 spins illustrates, below for any $n < 7$ in $\mathcal{H}_{\mathrm{x},n}$ there is only a dependence on second-order perturbation (see Figure 3) or hard suppression in the average case.

hard – At least one ground state is soft suppressed with a ratio larger than $1:100$ or not found at all, i.e., total suppression. For better visibility in Fig. 5 we combine these two cases. However, most of the time the suppression is total.

highord – The matrix V is trivial. Higher-order perturbation will determine the sampling behavior. In the generic case of random couplings this leads to both soft or hard suppression.

In all cases and for $\mathcal{H}_{\mathrm{x},n}$ with $n \leq 8$ we use $\Gamma^{\mathrm{x},n} = 1$. Using different values for the different amplitudes leads to worse sampling, because the matrix V has multiple different entries. A random matrix has a unique eigenvector which is not parallel to $(1, 1, \ldots, 1, 1)$ in the generic case. Hence, introducing more variety into V leads to more unique (and unfair) $|\alpha^0_{\mathrm{gs}}\rangle$. How the ratio of soft to hard suppression is influenced by this was not investigated, since it is unfair in the generic case. Repeating multiple annealing runs with individually randomized $\Gamma^{\mathrm{x},n}_{i,j,\ldots}$ and averaging improves the sampling but, if not dense enough, will not be able to remove all hard suppression in a generic case.

Conclusions.— We have studied the necessary ingredients needed for quantum annealing to sample ground states fairly. From Fig. 5 we surmise that a fairly dense driver is needed to obtain fair sampling. Carefully controlling the anneal with additional parameters, for example as shown in Ref. [59] might help mitigate the bias, however, this remains to be tested. We do emphasize, however, that a $\mathcal{H}_{\mathrm{x},2}$ driver with the typical annealing modus operandi used in current hardware will not yield a fair sampling of states and performs comparably to a vanilla transverse-field driver.

ACKNOWLEDGMENTS

M. S. K. thanks Dominik Gresch for insightful discussions that accelerated the discovery process. The large scale simulated quantum annealing calculation where done using the Monch Cluster at ETH Zurich. H. G. K. would like to thank Salvatore Mandrà for multiple discussions and acknowledges the ARC Centre for Excellence in All-Sky Astrophysics in 3D (ASTRO 3D) East Coast Writing Retreat 2018 for support in preparing the manuscript. H. G. K. acknowledges support from the NSF (Grant No. DMR-1151387). H. G. K.'s research is based upon work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via Interagency Umbrella Agreement No. IA1-1198. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. We thank the Texas Advanced Computing Center (TACC) at The University of Texas at Austin for providing HPC resources (Stampede Cluster) and Texas A&M University for access to their Ada, and Lonestar clusters.

---

[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. Aepli, *Quantum annealing of a disordered magnet*, Science **284**, 779 (1999).

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

[5] G. Santoro, E. Martoňák, R. Tosatti, and R. Car, *Theory of quantum annealing of an Ising spin glass*, Science **295**, 2427 (2002).

[6] A. Das and B. K. Chakrabarti, *Quantum Annealing and Related Optimization Methods* (Edited by A. Das and B.K. Chakrabarti, Lecture Notes in Physics 679, Berlin: Springer, 2005).

[7] G. E. Santoro and E. Tosatti, *TOPICAL REVIEW: Optimization using quantum mechanics: quantum annealing through adiabatic evolution*, J. Phys. A **39**, R393 (2006).

[8] A. Das and B. K. Chakrabarti, *Quantum Annealing and Analog Quantum Computation*, Rev. Mod. Phys. **80**, 1061 (2008).

[9] S. Morita and H. Nishimori, *Mathematical Foundation of Quantum Annealing*, J. Math. Phys. **49**, 125210 (2008).

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

[11] K. L. Pudenz, T. Albash, and D. A. Lidar, *Error-corrected quantum annealing with hundreds of qubits*, Nat. Commun. **5**, 3243 (2014).

[12] G. Smith and J. Smolin, *Putting "Quantumness" to the Test*, Physics **6**, 105 (2013).

[13] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, *Experimental signature of programmable quantum annealing*, Nat. Commun. **4**, 2067 (2013).

[14] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, *Reexamining classical and quantum models for the D-Wave One processor*, Eur. Phys. J. Spec. Top. **224**, 111 (2015).

[15] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, *Defining and detecting quantum speedup*, Science **345**, 420 (2014).

[16] H. G. Katzgraber, F. Hamze, and R. S. Andrist, *Glassy Chimeras Could Be Blind to Quantum Speedup: Designing Better Benchmarks for Quantum Annealing Machines*, Phys. Rev. X **4**, 021008 (2014).

[17] T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, et al., *Entanglement in a quantum annealing processor*, Phys. Rev. X **4**, 021041 (2014).

[18] S. Santra, G. Quiroz, G. Ver Steeg, and D. A. Lidar, *Max 2-SAT with up to 108 qubits*, New J. Phys. **16**, 045006 (2014).

[19] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, *How "Quantum" is the D-Wave Machine?* (2014), (arXiv:1401.7087).

[20] W. Vinci, T. Albash, A. Mishra, P. A. Warburton, and D. A. Lidar, *Distinguishing classical and quantum models for the D-Wave device* (2014), (arXiv:1403.4228).

[21] S. Boixo, T. F. Rønnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, *Evidence for quantum annealing with more than one hundred qubits*, Nat. Phys. **10**, 218 (2014).

[22] T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, *Consistency Tests of Classical and Quantum Models for a Quantum Device*, Phys. Rev. A **91**, 042314 (2015).

[23] H. G. Katzgraber, F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, *Seeking Quantum Speedup Through Spin Glasses: The Good, the Bad, and the Ugly*, Phys. Rev. X **5**, 031026 (2015).

[24] V. Martin-Mayor and I. Hen, *Unraveling Quantum Annealers using Classical Hardness*, Nature Scientific Reports **5**, 15324 (2015).

[25] K. L. Pudenz, T. Albash, and D. A. Lidar, *Quantum Annealing Correction for Random Ising Problems*, Phys. Rev. A **91**, 042302 (2015).

[26] I. Hen, J. Job, T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, *Probing for quantum speedup in spin-glass problems with planted solutions*, Phys. Rev. A **92**, 042325 (2015).

[27] D. Venturelli, S. Mandrà, S. Knysh, B. O'Gorman, R. Biswas, and V. Smelyanskiy, *Quantum Optimization of Fully Connected Spin Glasses*, Phys. Rev. X **5**, 031040 (2015).

[28] W. Vinci, T. Albash, G. Paz-Silva, I. Hen, and D. A. Lidar, *Quantum annealing correction with minor embedding*, Phys. Rev. A **92**, 042310 (2015).

[29] Z. Zhu, A. J. Ochoa, F. Hamze, S. Schnabel, and H. 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).

[30] S. Mandrà, Z. Zhu, W. Wang, A. Perdomo-Ortiz, and H. G. Katzgraber, *Strengths and weaknesses of weak-strong cluster problems: A detailed overview of state-of-the-art classical heuristics versus quantum approaches*, Phys. Rev. A **94**, 022337 (2016).

[31] S. Mandrà and H. G. Katzgraber, *The pitfalls of planar spin-glass benchmarks: Raising the bar for quantum annealers (again)*, Quantum Sci. Technol. **2**, 038501 (2017).

[32] S. Mandrà and H. G. Katzgraber, *A deceptive step towards quantum speedup detection* (2017), (arxiv:1711.01368).

[33] S. A. Weaver, K. J. Ray, V. W. Marek, A. J. Mayer, and A. K. Walker, *Satisfiability-based set membership filters*, Journal on Satisfiability, Boolean Modeling and Computation (JSAT) **8**, 129 (2014).

[34] T. J. Schaefer, in *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing* (ACM, New York, NY, USA, 1978), STOC '78, p. 216.

[35] A. Douglass, A. D. King, and J. Raymond, *Constructing SAT Filters with a Quantum Annealer*, in *Theory and Applications of Satisfiability Testing – SAT 2015* (Springer, Austin TX, 2015), pp. 104–120.

[36] D. Herr, M. Troyer, M. Azinović, B. Heim, and E. Brown, *Assessment of quantum annealing for the construction of satisfiability filters*, SciPost Physics **2**, 013 (2017).

[37] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani, *Random generation of combinatorial structures from a uniform distribution*, Theoretical Computer Science **43**, 169 (1986).

[38] C. P. Gomes, A. Sabharwal, and B. Selman, *Model counting*, in *Handbook of Satisfiability*, edited by A. Biere, M. Heule, H. van Maaren, and T. Walsch (IOS Press, 2008).

[39] P. Gopalan, A. Klivans, R. Meka, D. Stefankovic, S. Vempala, and E. Vigoda, in *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on* (IEEE, Palm Springs CA, 2011), p. 817.

[40] G. E. Hinton, *Training Products of Experts by Minimizing Contrastive Divergence*, Neural Comput. **14**, 1771 (2002).

[41] S. M. A. Eslami, N. Heess, C. K. I. Williams, and J. Winn, *The shape Boltzmann machine: A strong model of object shape*, Int. J. of Computer Vision **107**, 155 (2014).

[42] Y. Matsuda, H. Nishimori, and H. G. Katzgraber, *Ground-state statistics from annealing algorithms: quantum versus classical approaches*, New J. Phys. **11**, 073021 (2009).

[43] S. Mandrà, Z. Zhu, and H. G. Katzgraber, *Exponentially Biased Ground-State Sampling of Quantum Annealing Machines with Transverse-Field Driving Hamiltonians*, Phys. Rev. Lett. **118**, 070502 (2017).

[44] A. D. King, E. Hoskinson, T. Lanting, E. Andriyash, and M. H. Amin, *Degeneracy, degree, and heavy tails in quantum annealing*, Phys. Rev. A **93**, 052320 (2016).

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

[46] J. R. Johansson, P. D. Nation, and F. Nori, *QuTiP 2: A Python framework for the dynamics of open quantum systems*, Comp. Phys. Comm. **184**, 1234 (2013).

[47] T. Lanting, A. D. King, B. Evert, and E. Hoskinson, *Experimental demonstration of perturbative anticrossing mitigation using nonuniform driver Hamiltonians*, Phys. Rev. A **96**, 042322 (2017).

[48] A. J. Ochoa, D. C. Jacob, S. Mandrà, and H. G. Katzgraber, *Feeding the Multitude: A Polynomial-time Algorithm to Improve Sampling* (2018), (arXiv:1801.07681).

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

[50] S. V. Isakov, I. N. Zintchenko, T. F. Rønnow, and M. Troyer, *Optimized simulated annealing for Ising spin glasses*, Comput. Phys. Commun. **192**, 265 (2015), (see also ancillary material to arxiv:cond-mat/1401.1084).

[51] J. J. Moreno, H. G. Katzgraber, and A. K. Hartmann, *Finding low-temperature states with parallel tempering, simulated annealing and simple Monte Carlo*, Int. J. Mod. Phys. C **14**, 285 (2003).

[52] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, *Quantum versus classical annealing of Ising spin glasses*, Science **348**, 215 (2015).

[53] 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).

[54] G. Mazzola, V. N. Smelyanskiy, and M. Troyer, *Quantum Monte Carlo tunneling from quantum chemistry to quantum annealing*, Phys. Rev. B **96**, 134305 (2017).

[55] M. Könz, B. Heim, and M. Troyer, in preparation (2018).

[56] Simulation Parameters for SQA-$\mathcal{H}_{\mathrm{x},1}$: $H = 10$, $N_{\mathrm{trotter}} = 1024$, $T = 0.001$, steps $= 1000$.

[57] Simulation Parameters for SQA-$\mathcal{H}_{\mathrm{x},2}$: $H = H_{\mathrm{x}} = 2$, $N_{\mathrm{trotter}} = 64$, $T = 0.08$, equilibrate $= 10000$, steps $= 2000000$.

[58] G. Mazzola and M. Troyer, *Quantum Monte Carlo annealing with multi-spin dynamics*, J. Stat. Mech. P53105 (2017).

[59] Y. Susa, Y. Yamashiro, M. Yamamoto, and N. Nishimori, *Exponential Speedup of Quantum Annealing by Inhomogeneous Driving of the Transverse Field*, J. Phys. Soc. Jpn. **87**, 023002 (2018).