# Evidence for quantum annealing with more than one hundred qubits

Sergio Boixo[1], Troels F. Rønnow[2], Sergei V. Isakov[2], Zhihui Wang[3], David Wecker[4], Daniel A. Lidar[5], John M. Martinis[6] and Matthias Troyer[2]*

**Quantum technology is maturing to the point where quantum devices, such as quantum communication systems, quantum random number generators and quantum simulators may be built with capabilities exceeding classical computers. A quantum annealer, in particular, solves optimization problems by evolving a known initial configuration at non-zero temperature towards the ground state of a Hamiltonian encoding a given problem. Here, we present results from tests on a 108 qubit D-Wave One device based on superconducting flux qubits. By studying correlations we find that the device performance is inconsistent with classical annealing or that it is governed by classical spin dynamics. In contrast, we find that the device correlates well with simulated quantum annealing. We find further evidence for quantum annealing in the form of small-gap avoided level crossings characterizing the hard problems. To assess the computational power of the device we compare it against optimized classical algorithms.**

Annealing a material by slow cooling is an ancient technique to improve the properties of glasses, metals and steel that has been used for more than seven millennia[1]. Mimicking this process in computer simulations is the idea behind simulated annealing as an optimization method[2], which views the cost function of an optimization problem as the energy of a physical system. Its configurations are sampled in a Monte Carlo simulation using the Metropolis algorithm[3], escaping from local minima by thermal fluctuations to find lower energy configurations. The goal is to find the global energy minimum (or at least a close approximation) by slowly lowering the temperature and thus obtain the solution to the optimization problem.

The phenomenon of quantum tunnelling suggests that it can be more efficient to explore the state space quantum mechanically in a quantum annealer[4–6]. In simulated quantum annealing[7,8], one makes use of this effect by adding quantum fluctuations, which are slowly reduced while keeping the temperature constant and positive—ultimately ending up in a low-energy configuration of the optimization problem. Simulated quantum annealing, using a quantum Monte Carlo algorithm, has been observed to be more efficient than thermal annealing for certain spin glass models[8], although the opposite has been observed for $k$-satisfiability problems[9]. A further acceleration may be expected in physical quantum annealing, either as an experimental technique for annealing a quantum spin glass[10], or—and this is what we will focus on here—as a computational technique in a programmable device.

In this work we report on computer simulations and experimental tests on a D-Wave One device[11] to address central open questions about quantum annealers: is the device actually a quantum annealer, namely do the quantum effects observed on 8 qubits[11,12]) and 16 qubits[13] persist when scaling problems up to more than 100 qubits, or do short coherence times turn the device into a classical, thermal annealer? Which problems are easy and

which problems are hard for a quantum annealer, for a simulated classical annealer, for classical spin dynamics, and for a simulated quantum annealer? How does the effort to find the ground state scale with problem size? Does the device have advantages over classical computers?

We consider the problem of finding the ground state of an Ising spin glass model with the 'problem Hamiltonian'

$$H_{\text{Ising}} = -\sum_{i<j} J_{ij}\sigma_i^z\sigma_j^z - \sum_i h_i\sigma_i^z$$

with $N$ binary variables $\sigma_i^z = \pm 1$, where $J_{ij}$ are the couplings and $h_i$ are the local fields. This problem is non-deterministic polynomially (NP) hard[14], meaning that it is at least as hard as the hardest problems in NP, a class which includes notoriously difficult problems such as the travelling salesman and satisfiability of logical formulas. It also implies that no efficient (polynomial time) algorithm to find these ground states is known and the computational effort of all existing algorithms scales with problem size as $\mathcal{O}(\exp(cN^a))$. Although quantum mechanics is not expected to turn the exponential scaling into a polynomial scaling, the constants $c$ and $a$ can be smaller on quantum devices, potentially giving a substantial acceleration over classical algorithms.

To perform quantum annealing, we map each of the Ising variables $\sigma_i^z$ to the Pauli $z$-matrix (which defines the 'computational basis') and add a transverse magnetic field in the $x$-direction to induce quantum fluctuations, obtaining the time-dependent $N$-qubit Hamiltonian

$$H(t) = -A(t)\sum_i \sigma_i^x + B(t)H_{\text{Ising}}, \quad t \in [0, t_f]$$

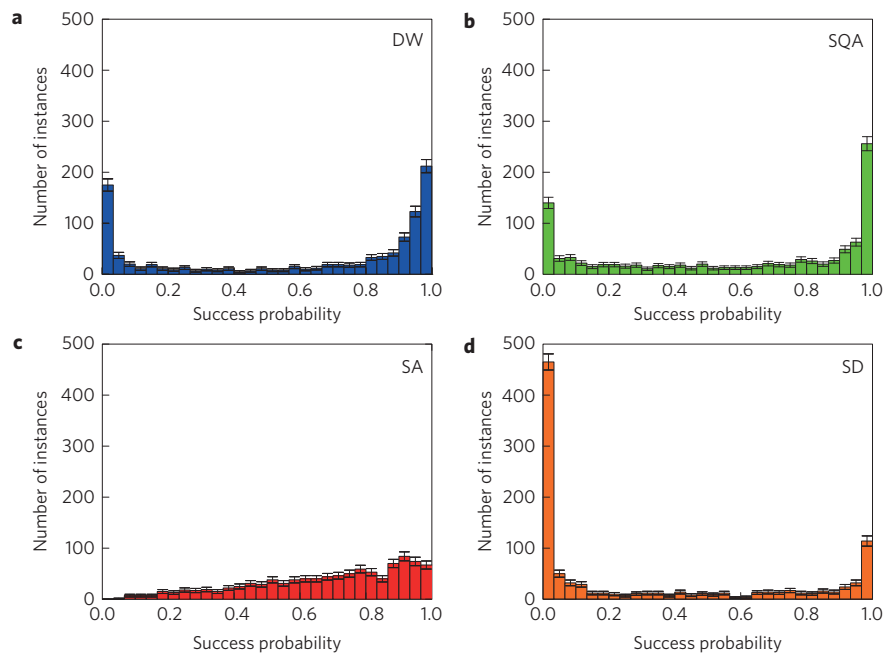where $A(t)$ and $B(t)$ are the 'annealing schedules'.

**Figure 1 | Success probability distributions.** Shown are histograms of the success probabilities of finding the ground states for $N = 108$ qubits and 1000 different spin glass instances. We find similar bimodal distributions for the D-Wave results (DW, **a**) and the simulated quantum annealer (SQA, **b**), and somewhat similar distributions for spin dynamics (SD, **d**). The unimodal distribution for the simulated annealer (SA, **c**) does not match. Error bars indicate $1\sigma$ statistical errors, but do not include systematic errors due to uncertainties in the annealing schedule and calibration errors of the D-Wave device.

Quantum annealing at positive temperature $T$ starts in the limit of a strong transverse field and weak problem Hamiltonian, $A(0) \gg \max(k_B T, B(0))$, with the system state close to the ground state of the transverse field term, the equal superposition state (in the computational basis) of all $N$ qubits. Monotonically decreasing $A(t)$ and increasing $B(t)$, the system evolves towards the ground state of the problem Hamiltonian, with $B(t_f) \gg A(t_f)$.

Unlike adiabatic quantum computing[15], which has a similar schedule but assumes fully coherent adiabatic ground state evolution at zero temperature, quantum annealing[4–6,10] is a positive temperature method involving an open quantum system coupled to a thermal bath. Nevertheless, one expects that, similar to adiabatic quantum computing, small gaps to excited states may thwart finding the ground state. In hard optimization problems, the smallest gaps of avoided level crossings have been found to close exponentially fast with increasing problem size[16–18], suggesting an exponential scaling of the required annealing time $t_f$ with problem size $N$.

## Results

We performed our tests on a D-Wave One device, comprised of superconducting flux qubits with programmable couplings (see Methods). Of the 128 qubits on the device, 108 were fully functioning and were used in our tests. The 'chimera' connectivity graph of these qubits is shown in Supplementary Fig. 1. Instead of trying to map a specific optimization problem to the connectivity graph of the chip[19,20], we chose random spin glass problems that can be directly realized. For each coupler $J_{ij}$ in the device we randomly assigned a value of either $+1$ or $-1$, giving rise to a rough energy landscape. Local fields $h_i \neq 0$ give a bias to individual spins, tending to make the problem easier to solve for annealers. We thus set all $h_i = 0$ for most data shown here and provide data with local fields in the Supplementary Methods. We performed tests for problems of sizes ranging from $N = 8$ to $N = 108$. For each problem size $N$, we selected 1,000 different instances by choosing 1,000 sets of different random couplings $J_{ij} = \pm 1$ (and for some of the data also random

fields $h_i = \pm 1$). For each of these instances, we performed $M = 1,000$ annealing runs and determined whether the system reached the ground state.

Our strategy is to discover the operating mechanism of the D-Wave device (DW) by comparing the performance of the device on many instances to three models: simulated classical annealing (SA), simulated quantum annealing (SQA) and classical spin dynamics (SD). All three models are described in detail in the Supplementary Information. We gain insight into the workings of the device by investigating the success probabilities of finding ground states for different instances, and by studying the correlations in these success probabilities between the D-Wave device and the models.

To start, we counted for each instance the number of runs $M_{GS}$ in which the ground state was reached, to determine the success probability as $s = M_{GS}/M$. Plots of the distribution of $s$ over many different instances are shown in Fig. 1, where we see that the DW results match well with SQA, moderately with SD, and poorly with SA. We find a unimodal distribution for the simulated annealer model for all schedules, temperatures and annealing times we tried, with a peak position that moves to the right as one increases $t_f$ (see Supplementary Methods). In contrast, the D-Wave device, the simulated quantum annealer and the spin dynamics model exhibit a bimodal distribution, with a clear split into easy and hard instances. Moderately increasing $t_f$ in the simulated quantum annealer makes the bimodal distribution more pronounced, as does lowering the temperature (see Supplementary Methods).

A much greater insight is obtained from Fig. 2, which plots the correlation of the success probabilities between the DW data and the other models. As a reference for the best correlations we may expect, we show in Fig. 2a the correlations between two different gauges of the same problem on the device (Methods and Supplementary Methods): as a result of calibration errors no better correlations than the device with itself can be expected. Figure 2b shows a scatter plot of the hardness of instances for the simulated quantum annealer and the D-Wave device for a single gauge. The
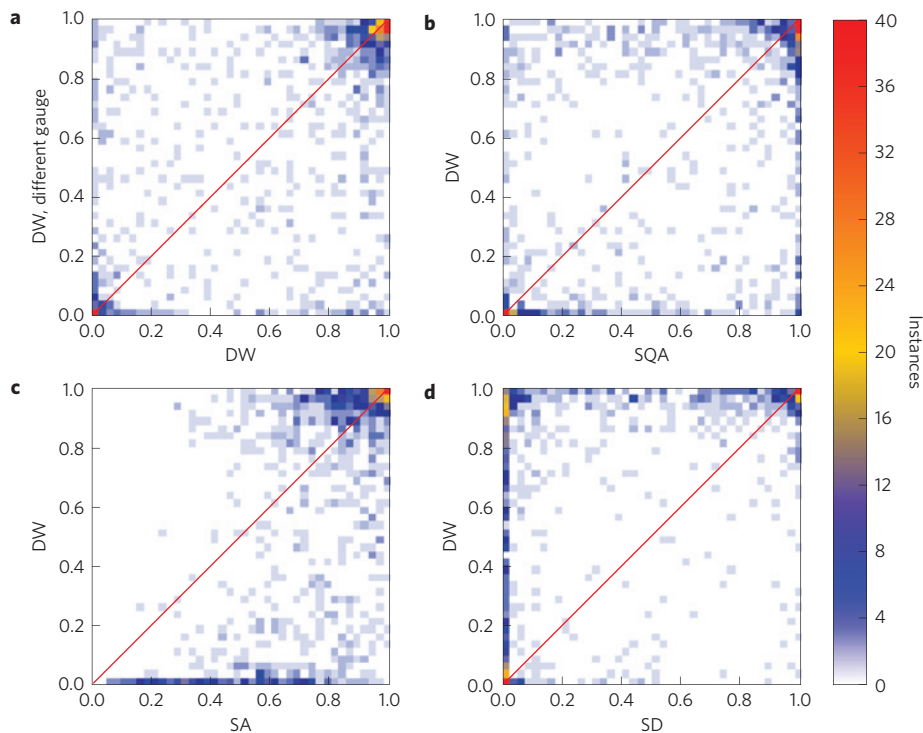
**Figure 2 | Correlations.** Scatter plots of correlations of the success probabilities $p(s)$ for different methods compared against the D-Wave device (DW). The colour scale indicates the number of instances found in a pixel. The red lines indicate perfect correlation. **a**, Correlation of the D-Wave device with itself by choosing two different gauges. This data shows the baseline imperfections in the correlation due to calibration errors in the D-Wave device. **b**, Correlation of the D-Wave device with simulated quantum annealing (SQA) using the same annealing schedule as the D-Wave device. This correlation is nearly as good as in **a**, indicating that the two methods are well correlated. **c,d**, show the much poorer correlation of the D-Wave device with simulated annealing (SA) and classical spin dynamics (SD), respectively. The case of SA is most sensitive to the annealing time and we show data for 5,000 updates per spin. Correlation plots for different annealing schedules for SA are shown in the Supplementary Methods.

high density in the lower left corner (hard for both methods) and the upper right corner (easy for both methods) confirms the similarities between the D-Wave device and a simulated quantum annealer. The similarity to Fig. 2a suggests a strong correlation with SQA, to within calibration uncertainties. To quantify the degree of correlation we performed a variant of a $\chi^2$-test of the differences between the success probabilities $s$ (Supplementary Methods for details). As expected we obtain a value of $\chi^2/M \approx 1$ between two different gauges on the D-Wave device, as the gauge-to-gauge variation was used to determine the expected error on $s$ on the device. The statistical test for Fig. 2b gives a value of $\chi^2/M \approx 1.2$, almost as good as the correlation of the D-Wave device with itself. Figure 2c,d show the correlations with a simulated classical annealer (SA) and classical mean-field spin dynamics (SD), respectively. The correlations are weaker, as can be seen both visually and by a $\chi^2$ test giving $\chi^2/M > 2.24$ for SA and $\chi^2/M \approx 9.5$ for SD. Some instances are easily solved by the classical mean-field dynamics, simulated annealing, simulated quantum annealing, and the device. However, as can be expected from inspection of their respective distributions in Fig. 1, there is no apparent correlation between the hard instances for the spin dynamics model and the success probability on the device, nor does there seem to be a correlation for instances of intermediate hardness, in contrast to the correlations seen in Fig. 2a. Similarly, there are poor correlations with a classical spin dynamics model.

Owing to calibration errors of the device the correlation plots—including those between two different gauges on the D-Wave device—show some anti-correlated instances in the lower right and upper left corner. To reduce these calibration errors we can average the success probabilities $s$ on the device over eight gauge choices.

The resulting correlation plots in Fig. 3, show much improved correlations of the device with itself Fig. 3a. These are again comparable to the correlations of SQA with the device (Fig. 3b). Simulated annealing (Fig. 3c) does not correlate as well and classical mean-field spin dynamics (Fig. 3d) again correlate poorly. A $\chi^2$ analysis of this data, discussed in detail in the Supplementary Methods, confirms this visual impression.

We next provide evidence for the bimodality seen in Fig. 1 being related to the device performing quantum annealing. Our first evidence comes from the simulated quantum annealer. When lowering the temperature, thermal updates are suppressed, quantum tunnelling dominates thermal barrier crossing, and we observe a stronger bimodality. In contrast, thermal effects become more important as we increase the temperature, and eventually the bimodality vanishes and correlations become weaker (Supplementary Methods). To provide further evidence we picked five hard and five easy instances, then performed extensive QMC simulations to estimate the spectral gap between the ground state and the first excited state using a method similar to that described in refs 21,22. A representative result of one easy and one hard instance is shown in Fig. 4; results for the other instances are shown in the Supplementary Methods. For all instances, we found that the gap trivially closes around a ratio $\Gamma = A(t)/B(t) \approx 2.3$ of transverse field to Ising coupling, related to a global $Z_2$ spin inversion symmetry. The gap also closes towards the end of the schedule as $\Gamma \to 0$, when multiple states are expected to become degenerate ground states. Neither of these small gaps has a detrimental effect on finding the ground state, as even after choosing the wrong branch at these avoided level crossings (either by thermalization or diabatic transitions) the system still ends up in a ground state at the end of the
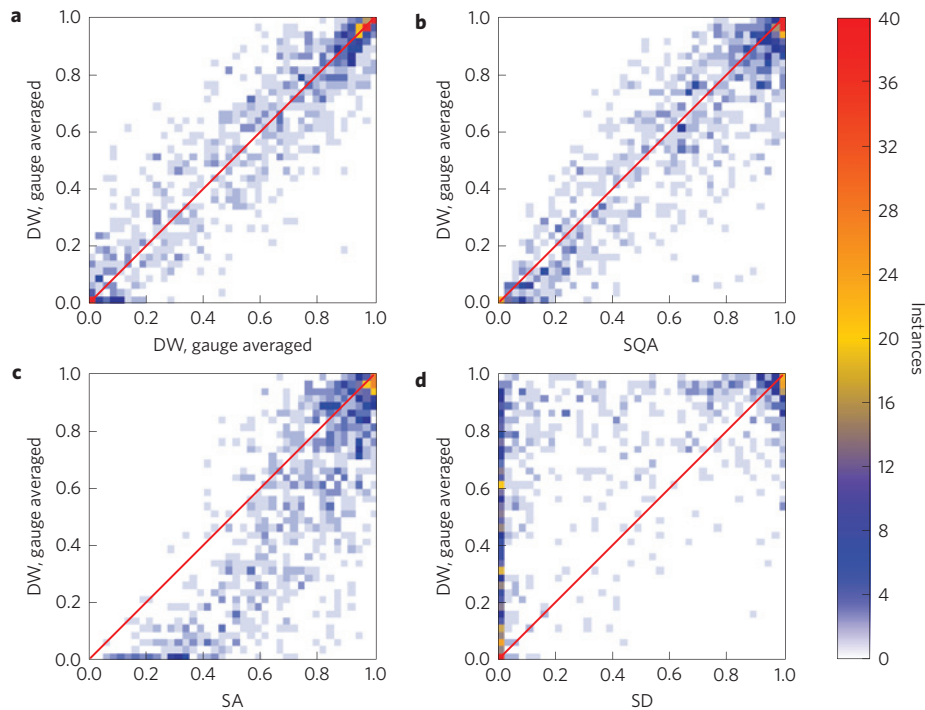
**Figure 3 | Correlations of gauge-averaged data.** Scatter plots of correlations of the success probabilities $p(s)$ obtained after averaging the success probabilities over eight different gauges of each instance on the device. The colour scale indicates the number of instances found in a pixel. The red lines indicate perfect correlation. **a**, D-Wave device (DW) between two sets of eight different gauges. This data shows the baseline imperfections in the correlations due to calibration errors in the D-Wave device. **b**, Simulated quantum annealer (SQA) using a single transverse field and the D-Wave device, with the latter averaged over 16 random gauges. This correlation is nearly as good as in **a**, indicating good correlations between the two methods. **c,d**, Poorer correlations of simulated annealing (SA) and classical spin dynamics (SD) respectively.
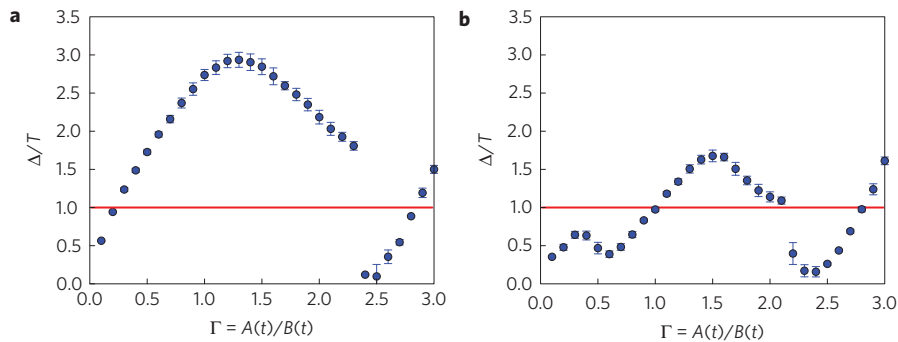


**Figure 4 | Evolution of the lowest spectral gap.** Shown in blue are (upper bounds for) the gaps between the ground state and the lowest excited state ($\Delta$) in units of the temperature for two typical spin glass instances of $N = 108$ spins as a function of the ratio of the transverse field to the coupling, $\Gamma = A(t)/B(t)$. $\Gamma$ depends on the details of the annealing schedule and decreases as a function of time. Gaps for an 'easy' instance with success probability 98% (**a**) and for a 'hard' instance with success probability 8% (**b**). Errors bars indicate systematic uncertainties (for details see the Supplementary Methods).

annealing. The hard instances, however, typically have extra avoided level crossings with small gaps, as is seen at $\Gamma \approx 0.5$ in Fig. 4b. These extra avoided level crossings cause failures of the annealing due to transitions to higher energy states, thus making the problem 'hard'. An explanation of the origin of small gap avoided level crossings for the hardest instances is presented in the Supplementary Methods.

Investigating the excited states found by the device provides further confirmation for the 'hard' instances being due to avoided level crossings with small gaps. In Fig. 5 we show a scatter plot of the mean Hamming distance of excited states versus success probability. The Hamming distance is the number of spins that need to be flipped to reach the closest ground state. We find that for the 'easy' instances the Hamming distance is typically small. The associated

spin flips are often due to thermal errors that can easily be corrected classically, as discussed in the Supplementary Methods. The 'hard' instances, on the other hand, typically result in excited states with a large Hamming distance. This means that in these cases the device typically finds local minima far away from ground states. Many spins would need to be flipped to reach a ground state, resulting in small tunnelling matrix elements between the state found and the true ground state, and thus small gap avoided level crossings[23].

Combining all these observations we have evidence that the device's performance is consistent with quantum annealing: unlike the classical annealer and classical spin dynamics, the simulated quantum annealer and the device split instances into hard and easy cases whose success probability is strongly and positively correlated.
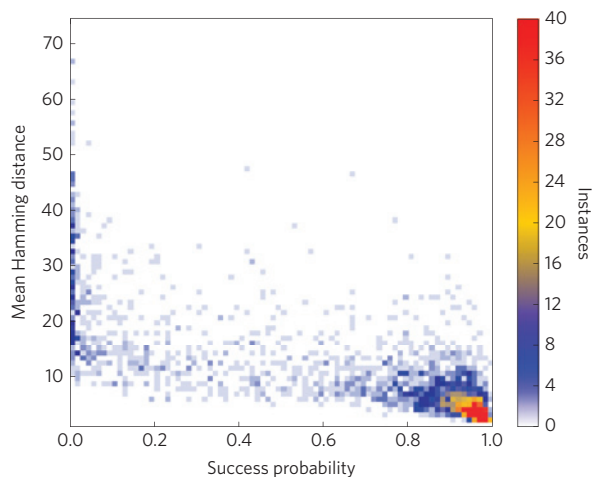
**Figure 5 | Correlation of success probability and the Hamming distance from excited states to the nearest ground state.** A scatter plot of the correlation between the success probabilities for $N = 108$ spin instances with local random fields, and the mean Hamming distance of excited states to the closest ground state for annealing runs where the device terminated in an excited state. Easy (hard) instances tend to have a small (large) Hamming distance. The colour scale indicates the number of instances found in a pixel.

The bimodality of the success probability distribution in the case of the device can be understood as having its origin in the transverse field quantum Ising model implemented by a quantum annealer, with the hard instances being such owing to small tunnelling matrix elements (and corresponding small gaps) resulting between rarely found ground states and easily found excited states during the evolution. However, our results do not rule out the possibility that there exists a classical model that is consistent with SQA, QA and the device's performance.

## Scaling

We finally investigate the scaling of the annealing effort with problem size $N$ to try to answer the question of whether quantum effects help the device to solve the hard problems faster than SQA, SA or other classical algorithms. As a first reference we investigated four exact algorithms discussed in the Supplementary Methods. An exact version of belief propagation[24] performed fastest, requiring around 60 ms for $N = 128$ and 3 min for $N = 512$ on 16 cores of an Intel Xeon CPU—comparable to previously reported values[25].

Because the tree width of the chimera graph scales as $\sqrt{N}$[26], exact solvers making use of the graph structure scale asymptotically no worse than $\exp(c\sqrt{N})$ and similar scaling is observed also for the simulated annealers discussed below.

For the D-Wave device (and the simulated annealers) we take into account only the intrinsic annealing time and not any overhead from programming the couplers and readout of the results. We calculate the total annealing time $Rt_f$, defined as the product of the annealing time $t_f$ of one annealing run multiplied by the number of repetitions $R$ needed to find the ground state at least once with 99% probability. From the success probability $s$ of a given percentile we calculate the required number of repetitions $R_p = \lceil \log(1-p)/\log(1-s) \rceil$, with $p = 0.99$.

In Fig. 6a we show the scaling of the typical (median) instance as well as various percentiles of hardness on the D-Wave device. The rapid increase of the higher percentiles is due to calibration issues that cause problems for a fraction of problem instances. Focusing on the median we see only a modest increase in total annealing time from 5 to 15 μs, corresponding to three repetitions of the annealing.
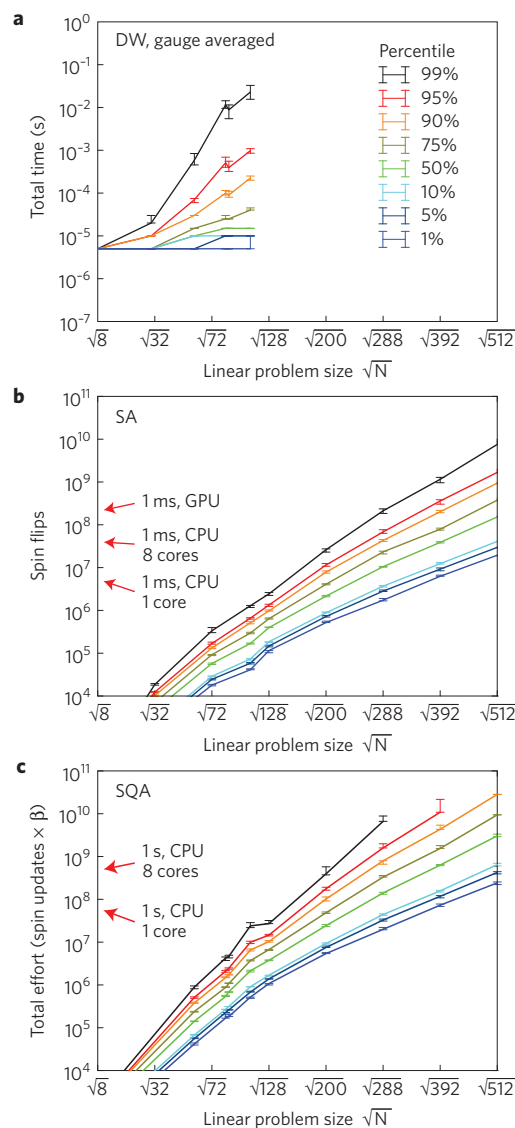


**Figure 6 | Scaling with problem size.** Scaling of the total effort to find the ground state with 99% probability for the D-Wave device (**a**), the simulated annealer (**b**) and the simulated quantum annealer (**c**). The individual lines show the scaling of the various percentiles, from the 1% easiest instances (0.01 percentile) to the 1% hardest instances (0.99 percentile). For the simulated annealers the vertical axis shows the total effort in number of spin updates for the simulated classical and quantum annealer. Arrows mark the number of spin updates that can be done in 1 ms or 1 s on the reference machines, as indicated. Error bars indicate 1σ statistical errors obtained by a bootstrap analysis.

Although an extrapolation of the observed scaling is tempting, this will not yield the true asymptotic scaling. The reason is that the total annealing time depends sensitively on the choice of $t_f$ and for the device the minimal time of $t_f = 5\,\mu s$ turns out to be suboptimal (Supplementary Methods).

For the simulated classical and quantum annealer, on the other hand, we can calculate the optimal annealing time and plot the scaling in Fig. 6b and c, respectively. The effort here is measured in the number of spin updates, defined as $RNN_{updates}$, where $N_{updates}$ is the optimal number of updates per spin to minimize the total effort. Indicated by an arrow is the number of spin updates that can be performed in a millisecond on an 8-core Intel Xeon E5-2670 CPU and on an Nvidia K20X GPU. At $N = 108$ this yields total

annealing times of 4.3 µs and 0.8 µs respectively, slightly faster than the D-Wave device, whereas the simulated quantum annealer is substantially slower. Classical spin dynamics is not competitive (as we showed above, it suffers from an abundance of hard instances) and was not considered.

Increasing the problem size up to $N = 512$ spins our simulated quantum annealer shows that the fraction of easy instances drops rapidly (Supplementary Methods); perhaps surprisingly, $N = 200$ is still a 'small' problem. As a consequence, both the annealing time and number of repetitions need to be increased and the total effort grows exponentially both in the simulated quantum annealer and in the simulated classical annealer. We find an increase of the median effort by about three orders of magnitude when increasing the problem size from $N = 108$ to $N = 512$ spins. We note that even then the simulated annealer finds solutions in a few milliseconds, which is faster than the first benchmarks reported for a next-generation 512 qubit D-Wave device[25]. We also observe that the simulated quantum annealer scales slightly worse than the simulated classical annealer for our problems. Investigating for which class of problems simulated quantum annealing is better than simulated classical annealing is an important open question to be addressed in future work.

Although for 108 spins a majority of optimization problems are still relatively easy, it will be possible to address the open question of quantum acceleration on future devices with more qubits, by comparing the scaling results of the simulated classical and quantum annealers (Supplementary Methods) to such devices. Going to even larger problem sizes we soon approach the limits of classical computers. Optimistically extrapolating, using the scaling observed in our simulations, the median time to find the best solution for our test problem using simulated annealing will increase from milliseconds for 512 variables to minutes for 2,048 variables, and months for 4,096 variables. It will be interesting to explore whether a quantum annealer might exhibit better scaling than classical algorithms for these problem sizes—and hence quantum acceleration—or whether it can be efficiently matched, or even surpassed by classical algorithms such as simulated annealing, simulated quantum annealing, mean-field approaches, or other algorithms that exploit the structure of the chimera graph.

After completion of this work we learned of recent results which indicate that both our DW and SQA data correlate well with a mean-field version of SQA (S. W. Shin, J. A. Smolin, G. Smith and U. Vazirani, personal communications). This suggests that although the device's performance is consistent with quantum annealing, it operates in a temperature regime where, for most random Ising spin glass instances, a quantum annealer may have an effective semi-classical description.

## Methods

**Quantum annealing** was performed on the D-Wave One Rainer chip installed at the Information Sciences Institute of the University of Southern California. The device has been described in extensive detail elsewhere[27–29]. After programming the couplings, the device was cooled for 2.5 s, and then 1,000 annealing runs were performed using an annealing time of $t_f = 5$ µs. Annealing is performed at a temperature of 0.4 GHz, with an initial transverse field starting at $A(0) \approx 5$ GHz, going to zero during the annealing, while the couplings and local fields are ramped up from near zero to about $B(t_f) \approx 5$ GHz at the end of the schedule. Details of the schedule and results for longer annealing times are provided in the Supplementary Methods.

**Simulated annealing** was performed using the Metropolis algorithm with local spin flips with codes optimized for the ±1 couplings used as test problems. A total of $N_{updates}$ flips per spin were attempted, increasing the inverse temperature $\beta = 1/k_B T$ linearly in time from 0.1 to 3.

**Simulated quantum annealing** is a quantum Monte Carlo algorithm following the same annealing schedule as a quantum annealer, but using Monte Carlo dynamics instead of unitary evolution of the quantum system. Simulations were performed in both discrete and continuous time path integral quantum Monte Carlo simulations, with cluster updates along the imaginary time direction to account for the transverse field, combined with Metropolis rejection sampling for the Ising interactions (see Supplementary Methods for details).

**The classical spin dynamics model** replaces the quantum spins by $O(3)$ classical unit vectors $\mathbf{M}_i$, where the sign of the $z$-component of each spin is the value of the Ising variable. The spins are propagated via the equations of motion $\partial \mathbf{M}_i / \partial t = \mathbf{H}_i(t) \times \mathbf{M}_i$, where the time-dependent field $\mathbf{H}_i(t)$ acting on spin $i$ is the sum of a decaying transverse field (along the $x$ direction) and a growing coupling term (along the $z$ direction): $\mathbf{H}_i(t) \equiv (1 - t/t_f) h \hat{\mathbf{e}}_x - (t/t_f) \sum_j J_{ij} M_j^z \hat{\mathbf{e}}_z$. The initial condition is to have all spins perturbed slightly from alignment along the $x$ direction: $(-\sqrt{1 - \delta_i^2 - \eta_i^2}, \delta_i, \eta_i)$, where $|\delta_i|, |\eta_i| < 0.1$.

**Gauge averaging** was performed on the device by using gauge symmetries to obtain a new model with the same spectrum. This was achieved by picking a gauge factor $a_i = \pm 1$ for each qubit, and transforming the couplings as $J_{ij} \rightarrow a_i a_j J_{ij}$ and $h_i \rightarrow a_i h_i$. Success probabilities $s_g$ obtained from $G$ gauge choices were arithmetically averaged for the correlation plots and as $\bar{s} = \prod_{g=1}^{G} (1 - s_g)^{1/G}$ for the scaling of total effort (see Supplementary Methods for a derivation).

**The ground state energies** were obtained using exact optimization algorithms, an exact version of belief propagation using bucket sort[24] and a related optimized divide-and-conquer algorithm described in the Supplementary Methods.

## References

1. Muhly, J. in *The beginning of the use of metals and alloys* (ed. Maddin, R.) 2–20 (MIT Press, 1988).
2. Kirkpatrick, S., Gelatt, C. D. & Vecchi, M. P. Optimization by simulated annealing. *Science* **220,** 671–680 (1983).
3. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21,** 1087–1092 (1953).
4. Ray, P., Chakrabarti, B. K. & Chakrabarti, A. Sherrington–Kirkpatrick model in a transverse field: Absence of replica symmetry breaking due to quantum fluctuations. *Phys. Rev. B* **39,** 11828–11832 (1989).
5. Finnila, A., Gomez, M., Sebenik, C., Stenson, C. & Doll, J. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.* **219,** 343–348 (1994).
6. Kadowaki, T. & Nishimori, H. Quantum annealing in the transverse Ising model. *Phys. Rev. E* **58,** 5355–5363 (1998).
7. Martoňák, R., Santoro, G. E. & Tosatti, E. Quantum annealing by the path-integral Monte Carlo method: The two-dimensional random Ising model. *Phys. Rev. B* **66,** 094203 (2002).
8. Santoro, G. E., Martoňák, R., Tosatti, E. & Car, R. Theory of quantum annealing of an Ising spin glass. *Science* **295,** 2427–2430 (2002).
9. Battaglia, D. A., Santoro, G. E. & Tosatti, E. Optimization by quantum annealing: Lessons from hard satisfiability problems. *Phys. Rev. E* **71,** 066707 (2005).
10. Brooke, J., Bitko, D. F. T., Rosenbaum, & Aeppli, G. Quantum annealing of a disordered magnet. *Science* **284,** 779–781 (1999).
11. Johnson, M. W. *et al.* Quantum annealing with manufactured spins. *Nature* **473,** 194–198 (2011).
12. Boixo, S., Albash, T., Spedalieri, F. M., Chancellor, N. & Lidar, D. A. Experimental signature of programmable quantum annealing. *Nature Commun.* **4,** 2067 (2013).
13. Dickson, N. G. *et al.* Thermally assisted quantum annealing of a 16-qubit problem. *Nature Commun.* **4,** 1903 (2013).
14. Barahona, F. On the computational complexity of Ising spin glass models. *J. Phys. A* **15,** 3241–3253 (1982).
15. Farhi, E. *et al.* A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science* **292,** 472–475 (2001).
16. Jörg, T., Krzakala, F., Semerjian, G. & Zamponi, F. First-order transitions and the performance of quantum algorithms in random optimization problems. *Phys. Rev. Lett.* **104,** 207206 (2010).
17. Hen, I. & Young, A. P. Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems. *Phys. Rev. E* **84,** 061152 (2011).
18. Farhi, E. *et al.* Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs. *Phys. Rev. A* **86,** 052334 (2012).
19. Bian, Z., Chudak, F., Macready, W. G., Clark, L. & Gaitan, F. Experimental determination of Ramsey numbers. *Phys. Rev. Lett.* **111,** 130505 (2013).

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

21. Kashurnikov, V. A., Prokof'ev, N. V., Svistunov, B. V. & Troyer, M. Quantum spin chains in a magnetic field. *Phys. Rev. B* **59,** 1162–1167 (1999).

22. Young, A. P., Knysh, S. & Smelyanskiy, V. N. Size dependence of the minimum excitation gap in the quantum adiabatic algorithm. *Phys. Rev. Lett.* **101,** 170503 (2008).

23. Altshuler, B., Krovi, H. & Roland, J. Anderson localization makes adiabatic quantum optimization fail. *Proc. Natl Acad. Sci. USA* **107,** 12446–12450 (2010).

24. Dechter, R. Bucket elimination: A unifying framework for reasoning. *Artif. Intell.* **113,** 41–85 (1999).

25. McGeoch, C. C. & Wang, C. *Proc. 2013 ACM Conf. Comput. Frontiers* (ACM, 2013).

26. Choi, V. Minor-embedding in adiabatic quantum computation: II. Minor-universal graph design. *Quant. Inform. Process.* **10,** 343–353 (2011).

27. Harris, R. *et al.* Experimental demonstration of a robust and scalable flux qubit. *Phys. Rev. B* **81,** 134510 (2010).

28. Harris, R. *et al.* Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor. *Phys. Rev. B* **82,** 024511 (2010).

29. Berkley, A. J. *et al.* A scalable readout system for a superconducting adiabatic quantum optimization system. *Supercond. Sci. Tech.* **23,** 105014 (2010).

## Author contributions

M.T., J.M.M. and D.A.L. designed the tests and wrote the manuscript, with input from all other authors. S.B. and Z.W. performed the tests on D-Wave One. S.V.I., T.F.R. and M.T. wrote the simulated classical and quantum annealing codes and T.F.R., S.V.I., M.T. and D.W. performed the simulations. S.B. and T.F.R. wrote the bucket sort code and the divide-and-conquer codes. T.F.R., S.V.I., M.T., S.B., Z.W. and D.A.L. evaluated the data. All authors contributed to the discussion and presentation of the results.

## Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to M. T.

## Competing financial interests

The authors declare no competing financial interests.