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]

[1]*Information Sciences Institute and Department of Electrical Engineering,
University of Southern California, Los Angeles, CA 90089, USA*
[2]*Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland*
[3]*Department of Chemistry and Center for Quantum Information Science & Technology,
University of Southern California, Los Angeles, California 90089, USA*
[4]*Quantum Architectures and Computation Group, Microsoft Research, Redmond, WA 98052, USA*
[5]*Departments of Electrical Engineering, Chemistry and Physics,
and Center for Quantum Information Science & Technology,
University of Southern California, Los Angeles, California 90089, USA*
[6]*Department of Physics, University of California, Santa Barbara, CA 93106-9530, USA*

**Quantum technology is maturing to the point
where quantum devices, such as quantum commu-
nication systems, quantum random number gen-
erators and quantum simulators, may be built
with capabilities exceeding classical computers.
A quantum annealer, in particular, solves hard
optimisation 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
experiments on a 108 qubit D-Wave One device
based on superconducting flux qubits. The strong
correlations between the device and a simulated
quantum annealer, in contrast with weak correla-
tions between the device and classical annealing
or classical spin dynamics, demonstrate that the
device performs quantum annealing. We find ad-
ditional evidence for quantum annealing in the
form of small-gap avoided level crossings charac-
terizing the hard problems. To assess the com-
putational power of the device we compare it to
optimised 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 millen-
nia [1]. Mimicking this process in computer simulations
is the idea behind *simulated annealing* as an optimisa-
tion method [2], which views the cost function of an op-
timisation problem as the energy of a physical system.
Its configurations are sampled in a Monte Carlo simu-
lation using the Metropolis algorithm [3], escaping from
local minima by thermal fluctuations to find lower en-
ergy 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 optimisation problem.

The phenomenon of quantum tunneling 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 optimisation problem. Simulated quantum anneal-
ing, 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]. Further
speedup may be expected in physical quantum anneal-
ing, 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 pro-
grammable quantum device.

In this work we report on computer simulations and
experimental tests on a D-Wave One device [11] in order
to address central open questions about quantum anneal-
ers: is the device actually a quantum annealer, i.e., do
the quantum effects observed on 8 [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 dynam-
ics, 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 comput-
ers?

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

$$H_{\text{Ising}} = -\sum_{i<j} J_{ij}\sigma_i^z \sigma_j^z - \sum_i h_i \sigma_i^z, \qquad (1)$$

with N binary variables $\sigma_i^z = \pm 1$. 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 traveling salesman and satisfiability of logical formulas
[15]. 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 $O(\exp(cN^a))$. While quantum mechanics
is not expected to turn the exponential scaling into a
polynomial one, the constants c and a can be smaller on
quantum devices, potentially giving substantial speedup

over classical algorithms.

To perform quantum annealing, we map each of the Ising variables σ_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

$$H(t) = -A(t) \sum_i \sigma_i^x + B(t) H_{\text{Ising}} , \quad t \in [0, t_f] . \quad (2)$$

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 [16], 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 optimisation problems, the smallest gaps of avoided level crossings have been found to close exponentially fast with increasing problem size [17–19], suggesting an exponential scaling of the required annealing time t_f with problem size N.

Experimental results

We performed our experiments on a D-Wave One chip, a 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 experiments. The "chimera" connectivity graph of these qubits is shown in figure 1 in the supplementary material. Instead of trying to map specific optimisation problem to the connectivity graph of the chip [20, 21], we chose random spin glass problems that can be directly realised. For each coupler J_{ij} in the device we randomly assigned a value of either $+1$ or -1, giving rise to a very 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 material. We performed experiments for problems of sizes ranging from $N = 8$ to $N = 108$. For each problem size N, we selected 1000 different *instances* by choosing 1000 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 $M = 1000$ 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 to three models: simulated classical annealing (SA), simulated quantum
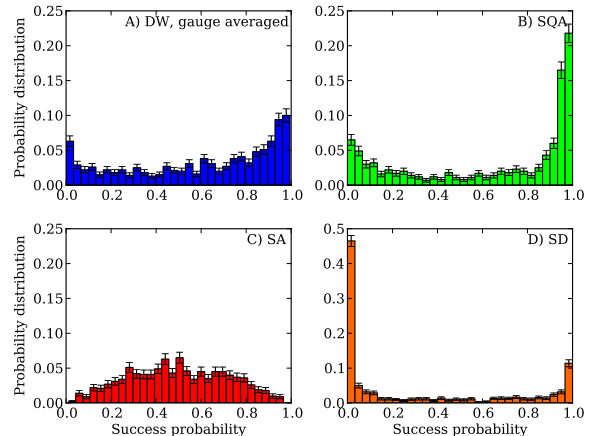

FIG. 1: **Success probability distributions.** Shown are normalized histograms $p(s) = $ (number of instances with probability s)$/K$ of the success probabilities of finding the ground states for $N = 108$ qubits and $K = 1000$ different spin glass instances. We find similar bimodal distributions for the D-Wave results (DW, panel A) and the simulated quantum annealer (SQA, panel B), and somewhat similar distributions for spin dynamics (SD, panel D). The unimodal distribution for the simulated annealer (SA, panel C) clearly does not match. The D-Wave data is taken with gauge averaging of 16 sets. Note the different vertical axis scale for D).

annealing (SQA), and classical spin dynamics (SD). We study both the distribution of the success probabilities and the correlations between the D-Wave device and the models.

For our first test, we counted for each instance the number of runs M_{GS} in which the ground state was reached, to determine the success probability as $s = M_{\text{GS}}/M$. Plots of the distribution of s are shown in figure 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 material). 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 material).

As a second test, we show in figure 2 results for the joint probability distribution $p(s, \Delta)$, which also includes the probability distribution for the final state energy Δ measured relative to the ground state. We find that the distribution for the D-Wave device (panel A) is very similar to that of the simulated quantum annealer (panel B), whereas it is quite different from that of a simulated
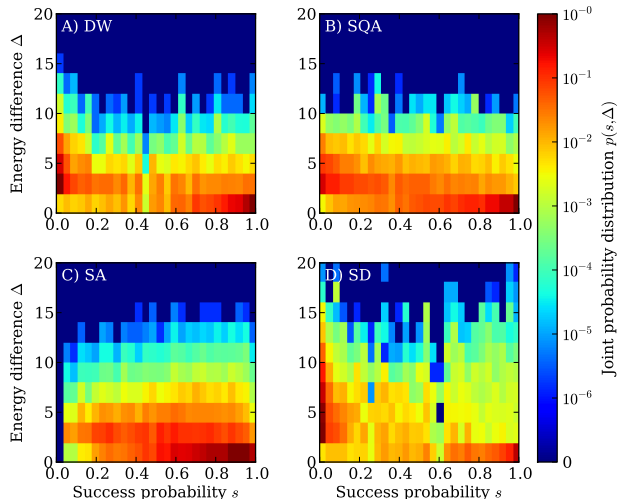
FIG. 2: **Energy-success distributions.** Shown is the joint probability distribution $p(s, \Delta)$ (colour scale) of success probability s and the final state energy Δ measured relative to the ground state. We find very similar results for the D-Wave device (panel A) and the simulated quantum annealer (panel B). The distribution for simulated classical annealing (panel C) matches poorly and for spin dynamics (panel D) matches only moderately. For the D-Wave device and SQA the hardest instances result predominantly in low-lying excited states, while easy instances result in ground states. For SA most instances concentrate around intermediate success probabilities and the ground state as well as low-lying excited states. For classical spin dynamics there is a significantly higher incidence of relatively high excited states than for DW, as well as far fewer excited states for easy instances. The histograms of figure 1, representing $p(s)$, are recovered when summing these distributions over Δ. SA distributions for different numbers of sweeps are shown in the supplementary material.

classical annealer (panel C) and spin dynamics (panel D).

The third test, shown in figure 3, is perhaps the most enlightening, as it 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 panel A) the correlations between two different sets of eight gauges (different embeddings of the same problem on the device, see Methods and supplementary material): no better correlations than the device with itself can be expected due to calibration errors. Panel B) shows a scatter plot of the hardness of instances for the simulated quantum annealer and the D-Wave device after gauge averaging. The 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 two are also well correlated for instances of intermediate hardness. The similarity to panel A) suggests almost perfect correlation with SQA, to within calibration uncertainties.
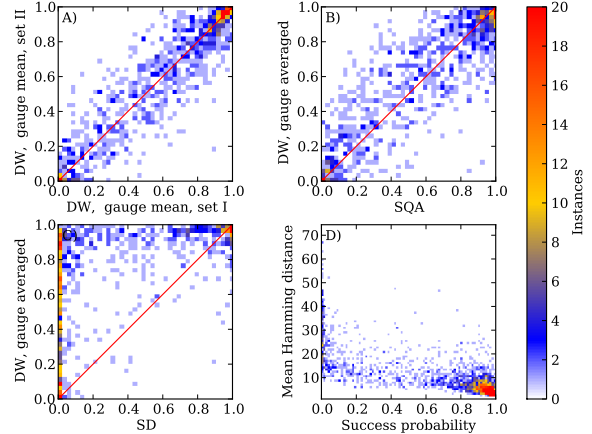

FIG. 3: **Correlations.** Panels A-C show scatter plots of correlations of the success probabilities $p(s)$ obtained from different methods. The red lines indicate perfect correlation. Panel A is for the D-Wave device between two sets of eight different gauges. This data shows the baseline imperfections in the correlations due to calibration errors in the D-Wave device. Panel B is for the simulated quantum annealer (SQA) and the D-Wave device, with the latter averaged over 16 random gauges. This correlation is nearly as good as in panel A, indicating good correlations between the two methods.. Panel C is for the classical spin dynamics and the D-Wave device, and shows poor correlation. Panel D shows the correlation between success probability and the mean Hamming distance of excited states found at the end of the annealing for $N = 108$ spin instances with local random fields. Easy (hard) instances tend to have a small (large) Hamming distance. The colour scale indicates how many of the instances are found in a pixel of the plots.

In panel C) we show the correlation between the classical spin dynamics model and the device. Some instances are easily solved by the classical mean-field dynamics, simulated quantum annealing, and the device. However, as can be expected from inspection of their respective distributions in figure 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 appear to be a correlation for instances of intermediate hardness, in contrast to the correlations seen in panel A). Similarly, there are poor correlations [22] with a classical spin dynamics model of reference [23].

The correlations between the simulated classical annealer and the D-Wave device, shown in the supplementary material, are significantly worse than between SQA and the device.

We next provide evidence for the bimodality being due to quantum effects. Our first evidence comes from the simulated quantum annealer. When lowering the temperature thermal updates are suppressed, quantum tunneling dominates thermal barrier crossing, and we observe a stronger bimodality; indeed a similar bimodal distribution arises also in an ensemble of (zero-temperature)
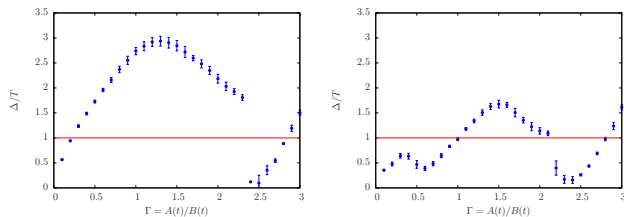
FIG. 4: **Evolution of the lowest spectral gap.** Shown in
blue are (upper bounds for) the gaps between ground state
and the lowest excited state in units of the temperature for
two typical spin glass instances of $N = 108$ spins as a function
of the ratio of transverse field to coupling, $\Gamma = A(t)/B(t)$.
Note that Γ depends on the annealing schedule details and
decreases as a function of time. In the left panel we show the
gap for an "easy" instance with success probability 98% and
on the right for a "hard" instance with success probability
8%.

Landau-Zener problems with a smooth distribution of
gaps. In contrast, thermal effects become more impor-
tant as we increase the temperature, and eventually the
bimodality vanishes (see supplementary material). To
provide further evidence we picked five hard and five easy
instances and performed extensive QMC simulations to
estimate the spectral gap between the ground state and
the first excited state using a method similar to that of
Refs. [24, 25]. A representative result of one easy and one
hard instance is shown in figure 4; results for the other in-
stances are shown in the supplementary material. 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 de-
generate ground states. Neither of these small gaps has
a detrimental effect on finding the ground state, since
even after choosing the wrong branch at these avoided
level crossings (either by thermalisation or diabatic tran-
sitions) the system still ends up in a ground state at the
end of the annealing. The hard instances, however, typ-
ically have additional avoided level crossings with small
gaps as is seen at $\Gamma \approx 0.5$ in the right panel of figure 4.
These additional 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 material.

Investigating the excited states found by the device
provides additional confirmation for the "hard" instances
being due to avoided level crossings with small gaps. In
panel D) of figure 3 we show a scatter plot of the mean
Hamming distance of excited states versus success prob-
ability. 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 dis-
tance is typically small. The associated spin flips are
often due to thermal errors that can easily be corrected

classically as discussed in the supplementary material.
The "hard" instances on the other hand typically result
in excited states with a large Hamming distance. This
means that there the device typically finds local minima
far away from ground states. Many spins would need to
be flipped to reach a ground state, which results in small
tunneling matrix elements between the state found and
the true ground state and thus small gap avoided level
crossings [26].

Combining all these observations we have strong ev-
idence for quantum behaviour in the device: unlike the
classical annealer and classical spin dynamics, the sim-
ulated quantum annealer and the device split instances
into hard and easy ones whose success probability is
strongly and positively correlated. The same holds for
the joint energy-success probability distributions. The
bimodality of the success probability distribution in the
case of the device can be understood as being due to
quantum effects, with the hard instances being such due
to small tunneling matrix elements (and corresponding
small gaps) resulting between rarely found ground states
and easily found excited states during the evolution.

Scaling
We finally investigate the scaling of the annealing effort
with problem size N. As a first reference we investigated
four exact algorithms discussed in the supplementary ma-
terial. An an exact version of belief propagation [27] per-
formed fastest, requiring around 60 ms for $N = 128$ and
3 minutes for $N = 512$ on 16 cores of an Intel Xeon CPU,
comparable to the timings reported in reference [28].

Since the tree width of the chimera graph scales as
$\sqrt{N}$ [29], exact solvers making use of the graph structure
scale asymptotically no worse than $\exp(c\sqrt{N})$ and sim-
ilar scaling is observed also for the simulated annealers
discussed below.

For the D-Wave device (and the simulated annealers)
we only take into account the intrinsic annealing time
and not 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 rep-
etitions 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 $R_p = \log(1-p)/\log(1-s)$, with $p = 0.99$.

In figure 5A) we show the scaling of the typical (me-
dian) 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 prob-
lems for a fraction of problems. Focusing on the median
we see only a modest increase in total annealing time
from $5\,\mu s$ to $15\,\mu s$, corresponding to three repetitions of
the annealing. While an extrapolation of the observed
experimental scaling is tempting, this will not yield the
true asymptotic scaling. The reason is that the total an-
nealing time depends sensitively on the choice of t_f and
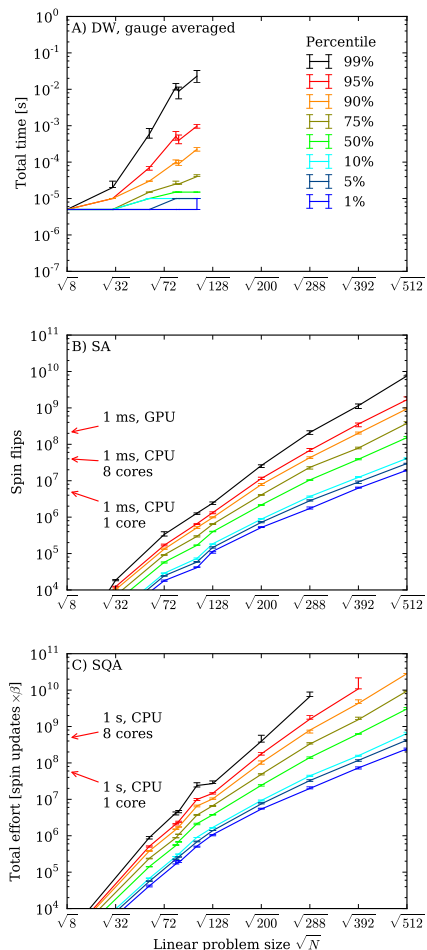for the device the minimal time of $t_f = 5\,\mu s$ turns out to

FIG. 5: **Scaling with problem size.** Shown is the scaling of the total effort to find the ground state with 99% probability for A) the D-Wave device, B) the simulated annealer and C) the simulated quantum annealer. 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 1ms or 1s on the reference machines.

be suboptimal (see supplementary material).

For the simulated classical and quantum annealer, on the other hand, we can calculate the optimal annealing time and plot its scaling in figure 5B) and C). 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 minimise 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 a total annealing time of $4.3\,\mu s$ and $0.8\,\mu s$ respectively, slightly faster than the D-Wave device, while 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 (see supplementary material); perhaps surprisingly, $N = 200$ is still a "small" problem. As a consequence, both the annealing time and the 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 few milliseconds, which is faster than first benchmarks reported for a next-generation 512 qubit D-Wave device [28]. 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.

Conclusions

Our experiments have demonstrated that quantum annealing with more than one hundred qubits takes place in the D-Wave One device, despite limited qubit coherence times. The key evidence is the correlation between the success probabilities on the device and a simulated quantum annealer, where the hard instances are characterised by avoided level crossings with small gaps. Sensitivity to these small gaps of the quantum model demonstrates that the device has sufficient ground state quantum coherence to realise quantum annealing of a transverse field Ising model. Considering the pure annealing time, the performance for typical (median) instances matches that of a highly optimised classical annealing code on a high-end Intel CPU.

While for 108 spins a majority of optimisation problems is still relatively easy, it should be possible to address the open question of quantum speedup on future devices with more qubits, by comparing the scaling results of the simulated classical and quantum annealers to experiments (see supplementary material). 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 will increase from milliseconds for 512 variables to minutes for 2048 variables, and months for 4096 variables. A quantum annealer showing better scaling than classical algorithms for these problem sizes would be an exciting breakthrough, validating the potential of quantum information processing to outperform its classical counterpart.

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 [30–32]. After programming the couplings, the device was cooled for $2.5\,s$, and then 1000 annealing runs were performed using an annealing time of $t_f = 5\,\mu s$. Annealing is performed at a temperature of $0.4\,\mathrm{GHz}$, with an initial transverse field starting at $A(0) \approx 5\,\mathrm{GHz}$, going to zero during the annealing, while the couplings and local fields are ramped up from near zero to about $B(t_f) \approx 5\,\mathrm{GHz}$ at the end of the schedule. Details of the schedule and results for longer annealing times are provided in the supplementary material.

Simulated annealing was performed using the Metropolis algorithm with local spin flips with codes optimised 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. The *simulated quantum annealing* 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 the supplementary material for details).

The classical spin dynamics model replaces the quantum spins by $O(3)$ classical unit vectors $\vec{M}_i$, 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 $\frac{\partial \vec{M}_i}{\partial t} = \vec{H}_i(t) \times \vec{M}_i$, where the time-dependent field $\vec{H}_i(t)$ acting on spin i is a sum of a decaying transverse field (along the x direction) and a growing coupling term (along z): $\vec{H}_i(t) \equiv (1 - t/t_f) h \hat{e}_x - (t/t_f) \sum_j J_{ij} M_i^z \hat{e}_z$. 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 $J_{ij} \to a_i a_j J_{ij}$ and $h_i \to a_i h_i$. 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 material for a derivation).

The ground state energies were obtained using exact optimisation algorithms, an exact version of belief propagation using Bucket Sort [27] and a related optimised divide-and-conquer algorithm described in the supplementary material.

Author Contributions MT, JM and DL designed the experiments and wrote the manuscript, with input from all other authors. SB and ZW performed the experiments on D-Wave One. SI, TR and MT wrote the simulated classical and quantum annealing codes and TR, SI, MT and DW performed the simulations. SB and TR wrote the bucket sort code and divide and conquer codes. TR, SI, MT, SB, ZW and DL evaluated the data. All authors contributed to the discussion and the presentation of the results.

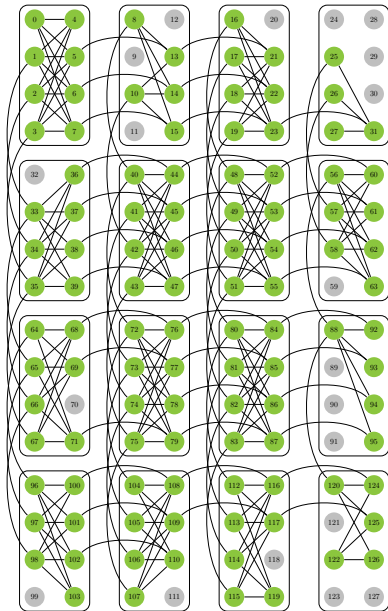# Supplementary material for "Quantum annealing with more than one hundred qubits"

FIG. 1: **Qubits and couplers in the D-Wave device.** The D-Wave One Rainer chip consists of 4×4 unit cells of eight qubits, connected by programmable inductive couplers as shown by lines.

I. OVERVIEW

Here we provide additional details in support of the main text. Section II shows details of the chimera graph used in our study and the choice of graphs for our simulations. Section III expands upon the algorithms employed in our study. Section IV presents additional success probability histograms for different numbers of qubits and for instances with magnetic fields, explains the origin of easy and hard instances, and explains how the final state can be improved via a simple error reduction scheme. Section V presents further correlation plots and provide more details on gauge averaging. Section VI gives details on how we determined the scaling plots and how quantum speedup can be detected on future devices. Finally, section VII explains how the spectral gaps were calculated by quantum Monte Carlo (QMC) simulations.

II. THE CHIMERA GRAPH OF THE D-WAVE DEVICE.

The qubits and couplers in the D-Wave device can be thought of as the vertices and edges, respectively, of a bipartite graph, called the "chimera graph", as shown in figure 1. This graph is built from unit cells containing eight qubits each. Within each unit cell the qubits and couplers realise a complete bipartite graph $K_{4,4}$ where each of the four qubits on the left is coupled to all of the four on the right and vice versa. Each qubit on the left is furthermore coupled to the corresponding qubit in the unit cell above and below, while each of the ones on the right is horizontally coupled to the corresponding qubits in the unit cells to the left and right (with appropriate modifications for the boundary qubits). Of the 128 qubits in the device, the 108 working qubits used in the experiments are shown in green, and the couplers between them are marked as black lines.

For our scaling analysis we follow the standard procedure for scaling of finite dimensional models by considering the chimera graph as an $L \times L$ square lattice with an eight-site unit cell and open boundary conditions. The sizes we typically used in our numerical simulations are $L = 1, \ldots, 8$ corresponding to $N = 8L^2 = 8, 32, 72, 128, 200, 288, 392$ or 512 spins. For the simulated annealers and exact solvers on sizes of 128 and above we used a perfect chimera graph. For sizes below 128 where we compare to the device we use the working qubits within selections of $L \times L$ eight-site unit cells from the graph shown in figure 1.

In references [29, 33] it was shown that an optimisation problem on a complete graph with $\sqrt{N}$ vertices can be mapped to an equivalent problem on a chimera graph with N vertices through minor-embedding. The tree width of $\sqrt{N}$ mentioned in the main text arises from this mapping. See Section VI A for additional details about the tree width and tree decomposition of a graph.

III. CLASSICAL ALGORITHMS

A. Simulated annealing

Simulated annealing (SA) is performed by using the Metropolis algorithm to sequentially update one spin after the other. One pass through all spins is called one *sweep*, and the number of sweeps is our measure of the annealing time for SA. Our highly optimised simulated
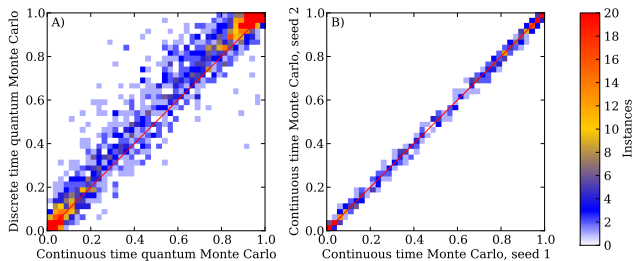
FIG. 2: **Correlation between simulated quantum annealers.** Axes corresponds to success probabilities and pixels are colour-coded according to the number of instances. A) correlations between continuous- and discrete time Monte Carlo simulations. The scatter observed here is a measure for the dependence of success probabilities on details of the simulated quantum annealing implementation, for instances with $N = 108$ spins performing 10,000 sweeps. B) Correlations between two independent sets of 1000 simulations with different initial starting points. Schedule II and 10,000 sweeps are used, see figure 3. Both simulations were performed at $T = 0.1$.

annealing code, based on a variant of the algorithm in Ref. [34, 35], uses multi-spin coding to simultaneously perform 64 simulations in parallel on a single CPU core: each bit of a 64-bit integer represents the state of a spin in one of the 64 simulations and all 64 spins are updated at once. A similar code for GPUs uses 32-bit integers and additionally performs many independent annealing runs and updates many spins in parallel in multiple threads.

The performance of our codes on the classical reference hardware is shown in Table I. We use high-end chips at the time of writing, an 8-core Intel Xeon E5-2670 "Sandy Bridge" CPU and an Nvidia Tesla K20X "Kepler" GPU. To find a ground state of our hardest 108-spin instances with a probability of 99%, this translates to a median annealing 32μs on a single core of the CPU, 4μs on eight cores, and 0.8μs on the GPU, which should be compared to 15μs pure annealing time on the D-Wave device for the same problems.

B. Simulated quantum annealing

For simulated quantum annealing we use both a continuous time algorithm and a discrete time algorithm.

The continuous time algorithm [36] constructs segments of a world line in the (imaginary) time direction

	spin flips per ns	relative speed
Intel Xeon E5-2670, 1 core	5	1
Intel Xeon E5-2670, 8 cores	40	8
Nvidia Tesla K20X GPU	210	42

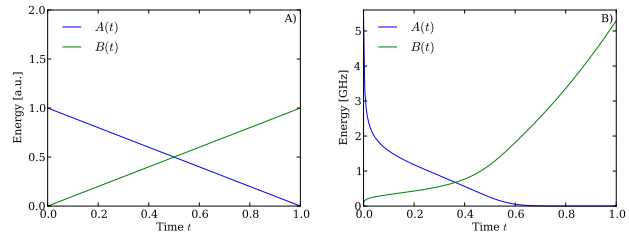TABLE I: Performance of the classical annealer on our reference CPU and GPU.

FIG. 3: **Annealing schedules used for Monte Carlo codes** A) schedule I, the linear schedule. B) schedule II, the schedule of the D-Wave device.

and flips them using the Metropolis algorithm. Specifically, we pick a random site and introduce new cuts in the time direction via a Poisson process. Then we calculate the overlaps with the neighbouring sites and use these overlaps to calculate the Metropolis acceptance ratios P_{Met} and flip a segment with probability $P_{\mathrm{Met}}/2$. We cannot grow the cluster along the space directions as in Ref. 36, which connects segments into larger clusters, since such cluster updates are inefficient in frustrated models like our spin glass.

In order to implement the fastest possible simulated quantum annealing code we also implemented a discrete time algorithm similar to that outlined in Ref. 37. However, unlike Ref. 37. we again used cluster updates along the imaginary time direction, typically with 64 time slices.

To verify that the discrete time code produces results similar to the continuous time algorithm – *i.e.*, that the error in the imaginary time direction is small – we show a correlation plot between the discrete- and continuous time in figure 2A). The scatter observed in this plot is a measure for the dependence of success probabilities on details of the simulated quantum annealing implementation. We also show correlations for a continuous time Monte Carlo using two different sets of random seeds for initial configurations and updates in figure 2B). The scatter of points is within the 3% ($1/\sqrt{1000}$) error expected for the success probabilities when performing 1000 annealing runs per instance.

In Table II, we summarise the performance of these two codes for typical cases using the linear schedule of figure 3A).

We have performed simulated quantum annealing with two different schedules shown in figure 3: A) a linear schedule (referred to as schedule I) where the transverse field is ramped down linearly in time and the Ising couplings are ramped up linearly, and B) the schedule used in the D-Wave device (referred to as schedule II). The performance was similar in both cases. For the scaling plots we used the linear schedule I at an optimised temperature ranging from $T = 0.33$ to $T = 1$ depending on system size, which gives slightly better performance than schedule II.

For the correlation plots we use the continuous time

	spin updates per μs	relative speed
CTQ, 1 core	1.3	1
CTQ, 8 cores	3.8	8
DTQ, 1 core	5.8	4.5
DTQ, 8 cores	46	35

TABLE II: **Performance of simulated quantum anneal-ers.** We show performance figures of both our continuous time (CTQ) and discrete time (DTQ) implementations on an Intel Xeon E5-2670 CPU using schedule II at inverse temperature $\beta = 10$.

code (CTQ) with schedule II – the average over slightly different schedules for the individual qubits on the device – but at up to ten times lower temperature than the device temperature of 20mK (0.4 GHz). The simulated quantum annealer requires a temperature about three times lower than the nominal temperature of the device to exhibit a clear bimodal distribution. This can be explained and motivated as follows. When the transverse field is strong the quantum Monte Carlo updates mimic the quantum tunneling taking place in the device, however when the transverse field becomes smaller these Monte Carlo updates turn into local spin flips of a classical (thermal) annealer. The device in this regime, on the other hand, has high tunneling barriers between the two states of the qubits of the device that suppress thermal tunneling over the barrier. To achieve a similar suppression of thermal excitations we need to lower the temperature in the quantum Monte Carlo simulations by at least a factor two. Lowering by more than a factor of ten does not significantly change histograms or correlations, indicating that at the chosen temperature the simulated quantum annealer is dominated by quantum tunneling and not thermal effects.

C. Exact solvers

We investigated four exact solvers, akmaxsat [38], the biqmac algorithm [39] used in the spin glass server [40], exact belief propagation using bucket sort [27] and a related divide-and-conquer algorithm. The latter is specifically designed for the chimera graph, generalising divide-and-conquer for the square lattice. We consider a chimera graph of $M \times M$ 8-spin unit cells ($8M^2$ spins). For each possible configuration of the $4M$ spins on the left side of the unit cells in the first row (which couple vertically) we find and store the optimal configuration of the remaining $4M$ spins with effort $4M$, giving a total effort of $4M\,2^{4M}$. Finding the 2^{4M} optimal configurations of the next row of unit cells, one builds up the solution row by row, scaling as $\mathcal{O}\left(M^2\,2^{4M}\right)$. Since $M \propto \sqrt{N}$, where N is the number of spins, this scaling demonstrates explicitly the claim made in the main text that exact solutions scale no worse than $\mathcal{O}(\exp(c\sqrt{N}))$, as for the bucket sort algorithm. We present the scaling of these algorithms in
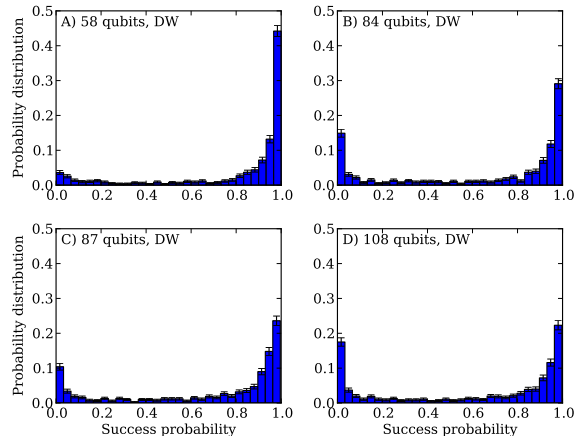

FIG. 4: **Success probability histograms for the D-Wave device for instances without local fields.** A) using 58 qubits, B) 84 qubits, C) 87 qubits and D) 108 qubits. The peak at low success probability grows with the number of qubits, reflecting the increasing hardness of the corresponding problem instances.
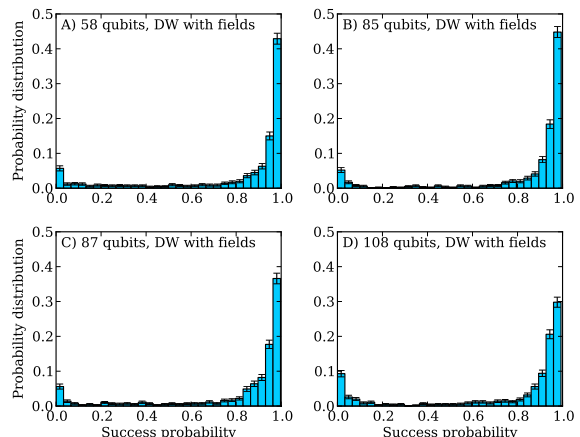

FIG. 5: **Success probability histograms for the D-Wave device for instances with local fields.** A) using 58 qubits, B) 85 qubits, C) 87 qubits and D) 108 qubits.

section VI A.

IV. SUCCESS PROBABILITY HISTOGRAMS

In this section we provide additional experimental and simulation data, complementing the data shown in the main text.
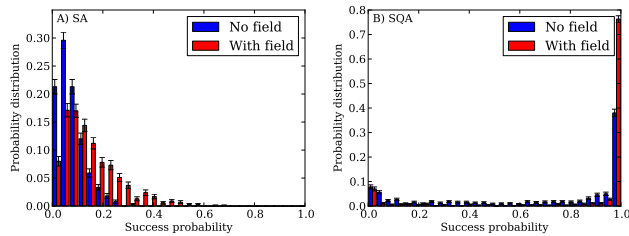
FIG. 6: **Comparison of instances with and without local fields**. A) success probability histograms for the simulated classical annealer with 50 sweeps (updates per spin). B) success probability histograms for the simulated quantum annealer using 10,000 sweeps.

A. Experimental annealing histograms

In addition to the histogram for 108 qubits shown in the main text, we also show histograms for $58, 84, 87$ and 108 qubits without local fields in figure 4 and with local fields in figure 5. By comparing these two figures, one notes that the cases with local fields are in general easier for the D-Wave device. We also verified that this is the case for the simulated annealers - see figure 6.

To check that the bimodality is not due to faulty couplers or qubits we performed the following analysis for the hardest instances of $N = 108$ spin problems where the D-Wave device did not find the ground state with one gauge choice. For each of these instances we looked at the lowest energy configurations reached and determined which spins differ compared to the closest ground state configuration. Closeness is measured in terms of the Hamming distance, which is the number of spins that have to be flipped to reach a ground state configuration. We did not observe a strongly peaked distribution which would have indicated a singly faulty qubit or coupler.

B. Simulated annealing and simulated quantum annealing

Figures 7, 8 and 9 show the success probability histograms for $N = 108$ spin problems for simulated annealing, simulated quantum annealing with different number of annealing sweeps (updates per spin), and simulated quantum annealing at different temperatures. While in the simulated classical annealer the distribution is always unimodal and shifts towards larger success probabilities upon increasing the annealing time, the simulated quantum annealer becomes more strongly bimodal when increasing annealing times.

Figures 10 and 11 show the success probability histograms for simulated quantum annealing for instances without and with local fields when increasing the problem size. As the problem size is increased the weight in the peak at low success probability increases. While for the instances without local fields the bimodality of easy

and hard instances vanishes at about $N = 288$ spins, it remains up to about 512 spins for instances with local fields. This indicates that there are more easy instances here than in the case without fields, and hence that instances with fields can be used to test for quantumness in a device scaled up to more qubits. However, once the "easy" problems disappear beyond 512 spins, and with them the bimodality, alternative analysis methods will be necessary to test for quantumness.

Supplementing the last two figures, we show the evolution of the success probability histograms for different numbers of spins for a simulated classical annealer in figure 12. It is seen that, in contrast to simulated quantum annealing where the weight in the two peaks of the bimodal distribution changes, here at fixed annealing time the unimodal peak gradually shifts to smaller success probabilities.

C. Hardness, gaps and free qubits

Figures 3 and 4 of the main text show that the "hard" instances typically exhibit small gaps during the evolution and often get trapped in excited states with a large Hamming distance to the closest ground state. The correlation between small gaps and large Hamming distance can be understood with a simple perturbative argument in the regime of small transverse fields $-\Gamma \sum \sigma^x$, where most of the small gap avoided level crossings appear. At an avoided level crossing between two states with Hamming distance d, d spins need to be flipped to adiabatically follow the ground state. In perturbation theory, the tunneling matrix element between the two states is of order Γ^d, exponentially small in the Hamming distance. The small matrix element not only poses problems for
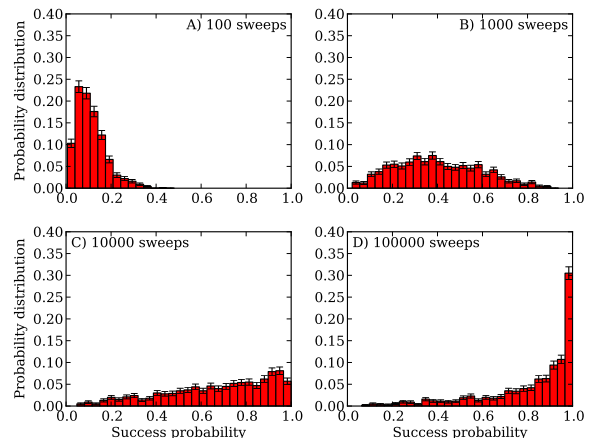

FIG. 7: **Success probability histogram for the simulated classical annealer.** Annealing times are A) 100 sweeps, B) $1,000$ sweeps, D) $10,000$ sweeps and D) $100,000$ sweeps for instances without local fields.
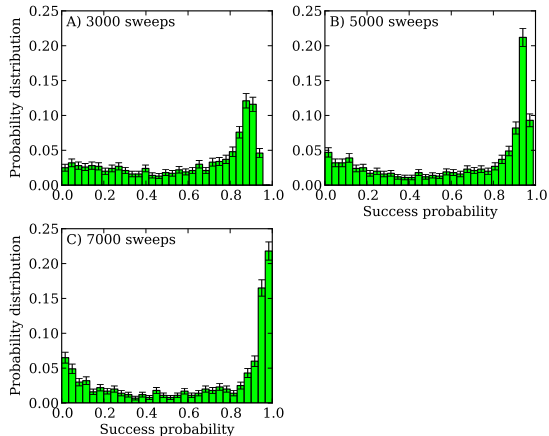
FIG. 8: **Success histograms for simulated quantum annealing.** The bimodal structure becomes more pronounced upon increasing the annealing time from A) $3,000$ sweeps to B) $5,000$ and C) $7,000$ sweeps. All three histograms were obtained for instances without local fields using schedule II and a temperature $T = 0.1$.
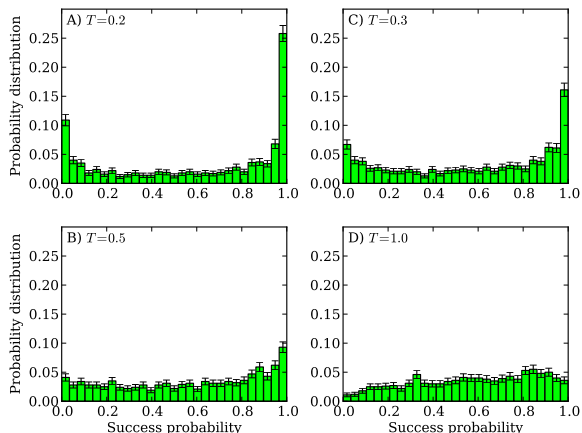

FIG. 9: **Success histograms for simulated quantum annealing as a function of temperature.** The bimodal structure becomes more pronounced at lower temperature and vanishes when increasing the temperature. Shown are results at A) $T = 0.2$ B) $T = 0.3$, C) $T = 0.5$ and D) $T = 1.0$. All histograms were obtained for instances without local fields using schedule II with 7000 sweeps.

adiabatic evolution but also suppresses quantum Monte Carlo updates that connect the two states. This common origin of the hardness in both quantum annealing and simulated quantum annealing explains the observed correlations despite the different underlying dynamics (deterministic *vs* stochastic).

An intuitive understanding of how such small gap avoided level crossings can arise can be obtained by considering degenerate states that are connected by single
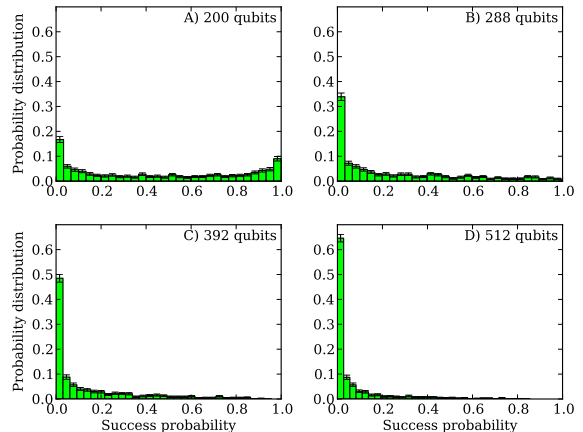

FIG. 10: **Scaling to larger problem sizes for the simulated quantum annealer for instances without local fields.** Histograms of the hardness distribution obtained by simulated quantum annealing with $N_{\text{updates}} = 10000$ Monte Carlo updates per spin using schedule II at $T = 0.1$. The fraction of easy problems rapidly decreases when increasing the problem size beyond 200 spins.

spin flips. The *free* qubits of a state are the qubits that can be flipped without changing the energy. From perturbation theory, a small transverse field Γ breaks the degeneracy of each free qubit [12, 41, 42]. In the simplest case, degenerate states form a hypercube and have the same free qubits. If the unperturbed energy was E_0, the lowest energy state of a hypercube with F free qubits will then have energy $E_0 - \Gamma F$ to first order in perturbation theory. If the low energy excited states have more
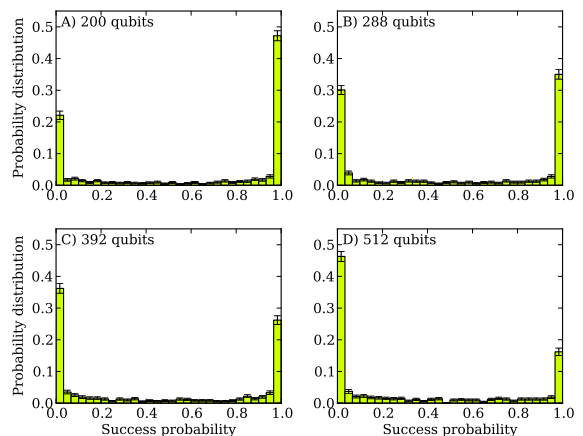

FIG. 11: **Scaling to larger problem sizes for the simulated quantum annealer for instances with local fields.** Histograms of the hardness distribution obtained by simulated quantum annealing with $N_{\text{updates}} = 10000$ Monte Carlo updates per spin using schedule II with $T = 0.1$.
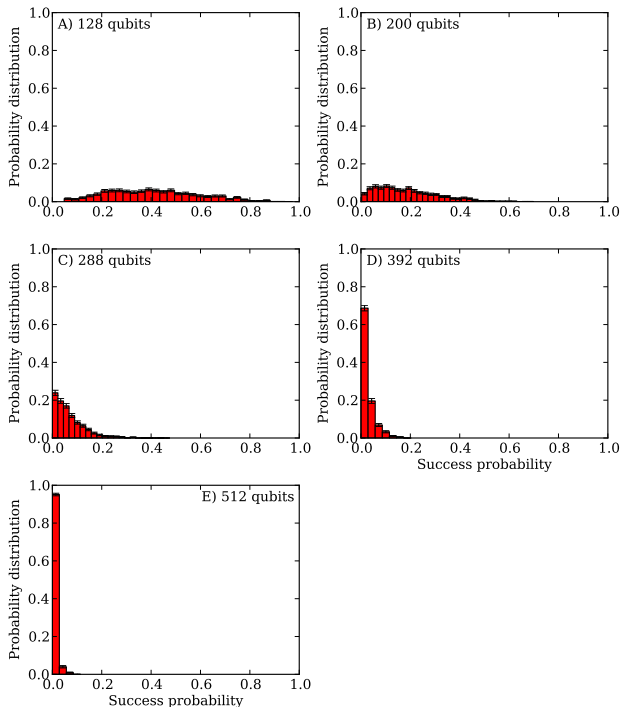
FIG. 12: **Scaling to larger problem sizes for the simulated classical annealer.** Histograms of the hardness distribution obtained by simulated annealing with $N_{\mathrm{updates}} = 10000$ Monte Carlo updates per spin for instances without local fields.
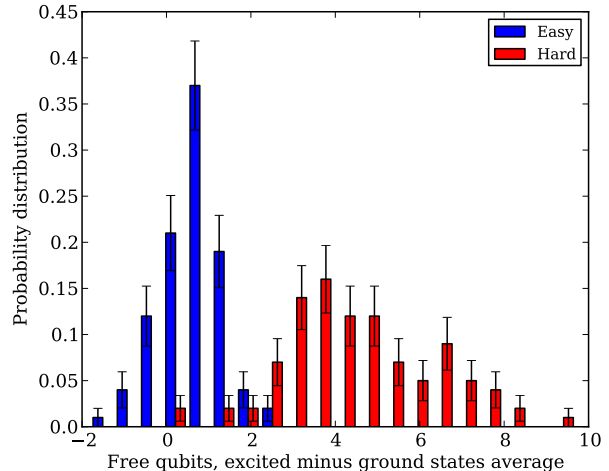

FIG. 13: **Degeneracies and hardness.** To illustrate correlations between hardness and degeneracies we calculate the average number of free qubits in excited states and ground states found by the D-Wave device. The histogram shows ⟨free qubits in excited states⟩ - ⟨free qubits in ground states⟩. Hard instances (here the hardest 10% from 1000 instances) are typically trapped in excited states with many free qubits, which lead to avoided level crossings. The easiest decile of instances, in contrast, typically does not have more free qubits in excited states than in ground states. Results from the simulated quantum annealer are very similar.

D. Improving the final state by fixing single spin errors

Single spin flip errors can be fixed with a linearly scaling effort by checking once for each spin whether the total energy can be lowered by flipping it, and then flipping the spin that lowers the energy the most. To illustrate the effect of this procedure we show in figure 14A), the success probability histogram for $N = 108$ spin instances with local fields, with and without error reduction. Without error reduction only ground states (Hamming distance $d = 0$) count as a successful annealing run, while with error reduction also the runs giving states a Hamming distance $d = 1$ away from the ground state will, after error reduction, end up in the ground state. It is clear from figure 14A) that while "hard" instances do not benefit from this error reduction scheme, such single spin flip errors are a dominant error source for the "easy" instances and their failure rate is substantially reduced using error reduction. The most likely source for such errors are thermal excitations or readout errors.

To understand the dependence on problem size N and hardness, we consider the percentiles of the success probability distribution as a function of N in figure 14B) to E). We see that the success probability decreases upon increasing N, i.e., the probability of such errors increases for larger problems. Independently of N, single-bit error reduction is ineffective in increasing the success proba-

free qubits than the ground states, their energy will be lowered more, resulting in avoided level crossings and small gaps. As seen in figure 13, hard instances tend to have more free qubits in low energy excited states. This phenomenon has been previously observed in the random subcubes model, a specially constructed toy model [43]. It can also be seen that on the device the experimental average of free qubits for first excited states is higher than the unweighted average over all excited states. This is not the case in simulated annealing.

In the spin glasses with ±1 couplings chosen here, we can expect to find a significant number of free qubits per state. This is because many qubits have six couplings, that can cancel each other. We also expect more low energy excited states than ground states, and consequently some low energy excited states will have more free qubits than most ground states. As argued above, both physical and simulated quantum annealing are sensitive to this problem: many spins updates are needed to move between the states at both sides of the avoided crossing. On the other hand, classical simulated annealers, not having a transverse field, do not suffer from these avoided crossings. This might explain why simulated quantum annealing scales slightly worse than classical annealing for the ±1 spin glasses considered in this work.
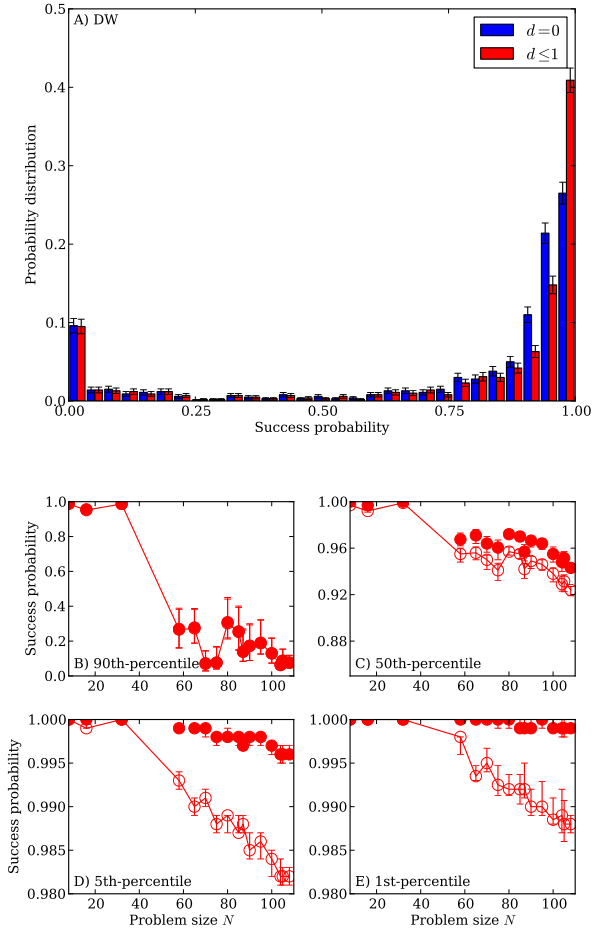
FIG. 14: **Effect of fixing single spin flip errors.** Shown are results for $N = 108$ spin instances with local fields: A) Histogram of the success probability s with $(d \leq 1)$ and without $(d = 0)$ single spin flips. The dependence of the effect of such error reduction on problem size N can be seen by following the percentiles of success probability as a function of system size for B) the 90% percentile, C) 50% percentile D) 5% percentile and E) 1% percentile. Empty (filled) symbols are the success probabilities without (with) single spin flips.

bility for the "hard" instances (high percentiles). It becomes more and more effective as problems become easier (low percentiles), improving the success probability to very nearly 1, independently of N, at the 1% percentile.

Variants of this simple error reduction scheme can be derived that, also with linear scaling, fix not only one single spin errors, but also "disconnected" single spin errors (*i.e.*, flips of spins that are not connected by one of the couplings). One such example is zero temperature Glauber dynamics, flipping randomly selected spins as long as they lower the energy. However, we saw only minimal improvements using such alternative schemes and sometimes worse results when the "wrong" spins are flipped and one reaches a different local minimum. The reason may be that disconnected single spin errors are
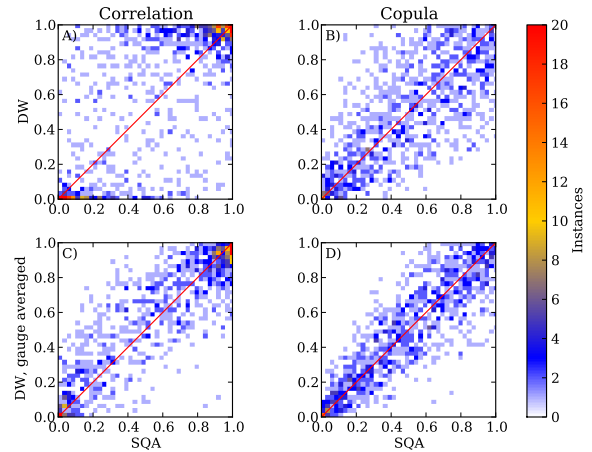

FIG. 15: **Correlations and copulae** between the simulated quantum annealer and the D-Wave device. Axes corresponds to success probabilities and pixels are colour-coded according to the number of instances. The simulation was carried out using schedule II at $T = 0.3$ using 7000 sweeps.

not common in the problem sizes we studied, but that may change in larger problems.

V. CORRELATIONS

A. Copulae

To better understand the correlations we here show copulae in addition to the correlations shown in the main text. Copulae

$$c(x_1, x_2) = \frac{f(x_1, x_2)}{f_1(x_1) f_2(x_2)}. \tag{1}$$

factor out the marginal distributions $f_1(x_1)$ and $f_2(x_2)$ (*e.g.*, the strong bimodal distribution) from the joint probability function $f(x_1, x_2)$ describing the correlation density. For independent sets the copula density is $c(x_1, x_2) = 1$ while for perfect correlations it is a delta function $c(x_1, x_2) = \delta(x_1 - x_2)/f_1(x_1)$.

To plot copula densities of success probabilities we replace each success probability by its rank after sorting the success probabilities divided by the number of values, and then plot the correlation densities of these normalised ranks. In figure 15 we plot copulae corresponding to the correlations shown in the main text, enhancing the visibility of the correlations especially in the corners.

In figure 16 we show the copulae between the hardness for akmaxsat and the D-Wave device and the simulated classical and quantum annealers. We do not observe any correlations.
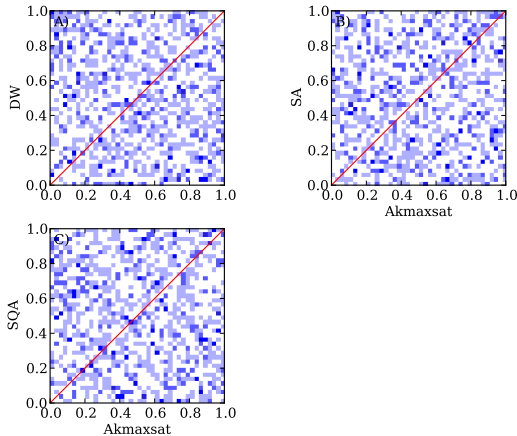
FIG. 16: **Copulae of the hardness between exact optimisers akmaxsat and the annealers.** A) for the D-Wave device, B) the simulated annealer and C) the simulated quantum annealer, showing the absence of correlations.
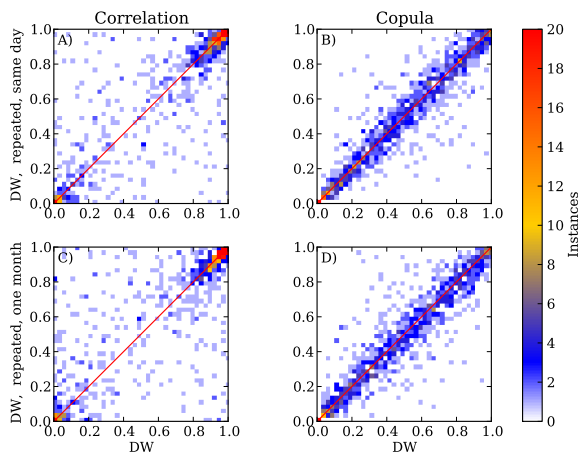

FIG. 17: **Reproducibility of the experiments**. We show correlations and copulae for repetitions of experiments: A) and B) on the same day, and, C) and D) one month apart. A) and C) show correlations between success probabilities, while B) and D) show the corresponding copulae.

B. Reproducibility

To verify reproducibility of the QA data we performed experiments on $N = 108$ spin instances three times, with a month between the first and the second two repetitions. The correlations (see figure 17) show that the device is stable over the time of a month.

Strong deviations are seen for a small fraction of the instances. These are most likely due to 1/f noise or "programming errors" when flux quanta are loaded into the programmable magnetic memories to program a specific set of couplings into the device. Since these program-
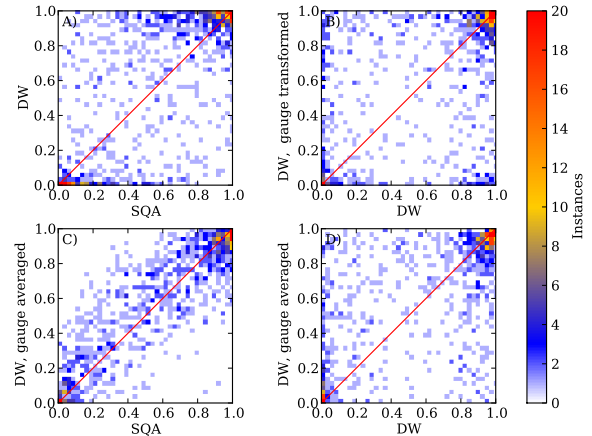

FIG. 18: **Correlations.** Shown are scatter plots of correlations of the success probabilities of 1000 instances of $N = 108$ spins. The colour scale indicates how many of the instances are found in a pixel of the plots. A) between a simulated quantum annealer (SQA) and the D-Wave device (DW) B) the quantum device and a gauge-transformed encoding of the same instance on the device where the sign of all couplings is changed. C) the SQA and an average over 16 random gauges on the device. D) a single gauge choice on the device against the average over 16 random gauges. The simulations were done at $T = 0.3$ with schedule II using 7000 sweeps.

ming errors will limit the correlations between any model and the device, no better correlations than shown in this figure can be expected between our simulations and the device.

C. Gauge averaging

The spectrum of an Ising spin glass is invariant under a gauge transformation that changes spins $\sigma_i^z \rightarrow a_i \sigma_i^z$, with $a_i = \pm 1$, when at the same time changing the couplings $J_{ij} \rightarrow a_i a_j J_{ij}$ and the local fields as $h_i \rightarrow a_i h_i$. While the simulated annealers are invariant under such a gauge transformation, the calibration of the D-Wave device is not perfect and breaks the gauge symmetry. Different gauges hence realise slightly different physical systems with different success probabilities. We average over gauges to reduce calibration uncertainties.

Panel A of figure 18 shows a scatter plot of the hardness of instances in the simulated quantum annealer (SQA) and the D-Wave device (DW). The 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 quantum device and a simulated quantum annealer. However, a small percentage of instances appears in the other corners (hard for one method but easy for the other). We attribute these to calibration errors of the couplers in the device which vary by about 10% [31]. To test for calibration errors we performed a
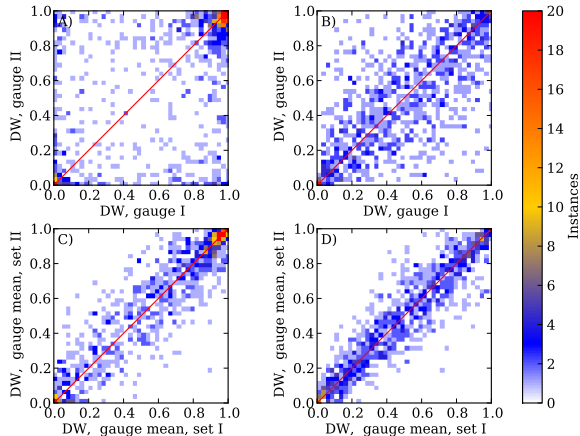
FIG. 19: **Effect of gauge averaging on correlations**. Shown are in A) and B) are the correlations and copulae between two different gauge choices of 1000 instances with $N = 108$ spins. In C) and D) we show correlations and copulae of results averaged over eight gauge choices against averages over eight different gauge choices. Gauge averaging significantly increases correlations.

gauge transformation on each instance to realise a different encoding of the same spin glass instance on the device. The correlations between two different encodings (gauge choices), shown in panel B, turn out to be comparable to those between the SQA and the device, demonstrating that the observed deviations can be attributed to calibration errors.

To minimise calibration errors, we then performed annealing with multiple encodings of the same instance related by gauge transformations and averaged the success probabilities. The resulting correlations between the simulated quantum annealer and the gauge-averaged results from the D-Wave device, shown in panel C (identical to the plot shown in the main text), are significantly improved compared to a single gauge choice, with a substantial reduction of the number of extreme outliers. These correlations are even better than those between a single embedding and the gauge averaged results on the device (see panel D).

In figure 19 we show correlations between the arithmetic average over eight gauges compared against the average over eight other gauges. Gauge averaging significantly increases correlations compared to single gauge choices. For the marginal distributions (see figure 20), the number of hard instances (weight of the peak close to zero) is reduced by gauge averaging, but the distribution remains bimodal even after averaging over many gauge choices. While calibration errors enhance the bimodality, the convergence to a bimodal distribution after averaging many gauges shows that the bimodality is intrinsic and not solely due to calibration errors.

For scaling plots of the total effort we use geometric means of failure rates instead of arithmetic means of suc-
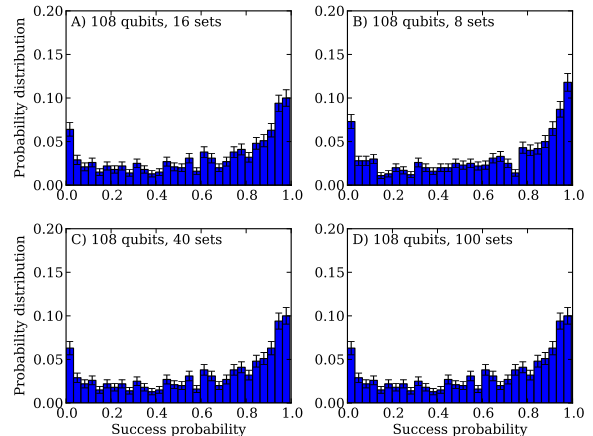

FIG. 20: **Gauge averaged histograms.** Arithmetic averaging over success probabilities obtained using different gauge transformations, using A) 2 gauges, B) 8 gauges, C) 40 gauges and D) 100 gauges. One sees that the marginal distribution converges after roughly 40 gauges.

cess probabilities. If the probability for finding a ground state for a specific percentile and gauge choice g is denoted by s then the probability of achieving a ground state at least once in R repetitions of the annealing is

$$P = 1 - (1 - s)^R. \tag{2}$$

Splitting the R repetitions into R/G repetitions for each of G gauge choices and denoting the success probabilities for a specific gauge choice by s_g the total success probability becomes

$$P^{(G)} = 1 - \prod_{g=1}^{G} (1 - s_g)^{R/G}, \tag{3}$$

which can be written in a form similar to equation (2)

$$P = 1 - (1 - \overline{s})^R, \tag{4}$$

by using the *geometric mean* of the failure rates to define $\overline{s}$ as

$$\overline{s} = 1 - \prod_{g=1}^{G} (1 - s_g)^{1/G}. \tag{5}$$

The higher percentiles grow much more rapidly in the absence of gauge averaging, as can be seen by comparing figure 21 to figure 4A) in the main text. Without gauge averaging the device actually fails to find the ground states for the 5% hardest instances in 1000 repetitions and those percentiles are thus not shown.

D. Correlations for the simulated annealer

To complement figures 2 and 3 in the main text we show correlations between the D-Wave device and a sim-
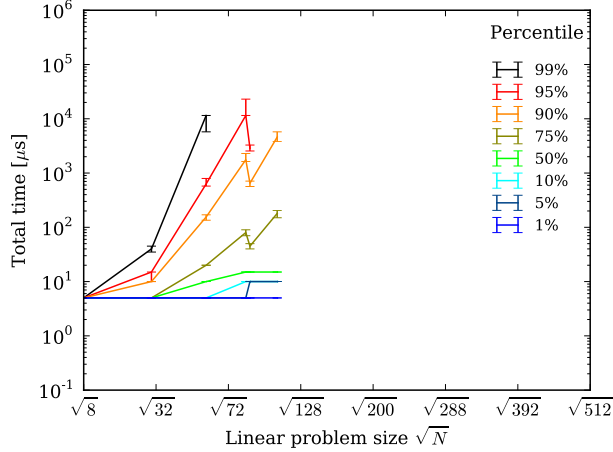
FIG. 21: **Scaling without gauge averaging**. Shown is the total effort on the D-Wave device for a single gauge choice. The higher percentiles are not shown for large systems since 1000 repetitions of the annealing failed to find the ground states for these hardest instances.

ulated classical annealer in figure 22 and the energy-success distribution $p(s, \Delta)$ of the simulated classical annealer in figure 23. As can already be expected from the different success distributions (figure 1 of the main text), the classical annealer does not correlate well with the D-Wave device and $p(s, \Delta)$ is significantly different.
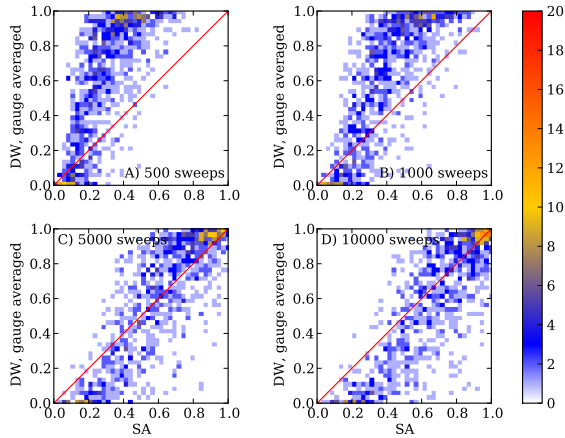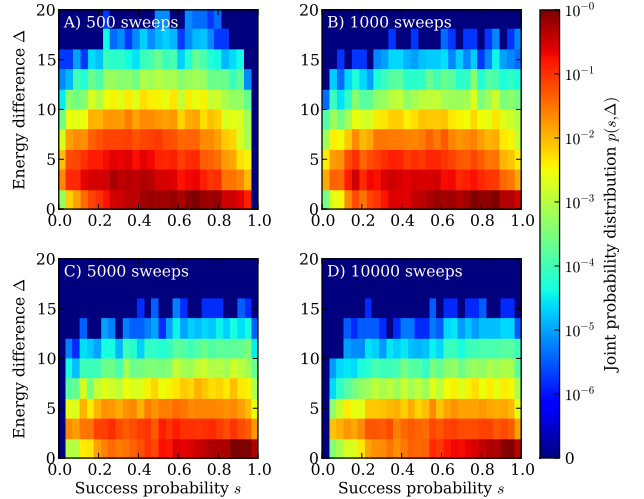

FIG. 23: **Energy-success distributions for the simulated classical annealer.** Shown is the joint probability distribution $p(s, \Delta)$ (colour scale) of success probability s and the final state energy Δ measured relative to the ground state. Shown are results for the simulated annealer at A) 500 sweeps, B) 1000 sweeps, C) 5000 sweeps and D) 10000 sweeps. Increasing the annealing time increases the success probability and decreases the energy of states found. The distribution is always significantly different from that of the D-Wave device.

VI. SCALING

A. Scaling of the exact solvers

In this section we present the scaling of the time to solution for the various exact solvers used. Timings are from our reference CPU, except for the biqmac algorithm

FIG. 22: **Correlations between the D-Wave device and a simulated classical annealer**. Shown are correlations at A) 500 sweeps, B) 1000 sweeps, C) 5000 sweeps and D) 10000 sweeps. The correlations are worse than for the simulated quantum annealer.
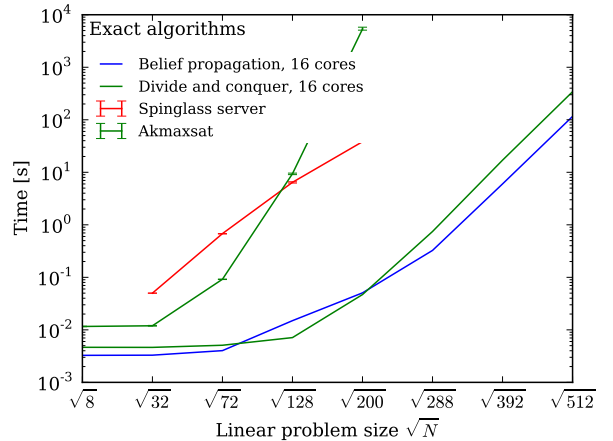

FIG. 24: **Scaling with problem size for the exact solvers.**

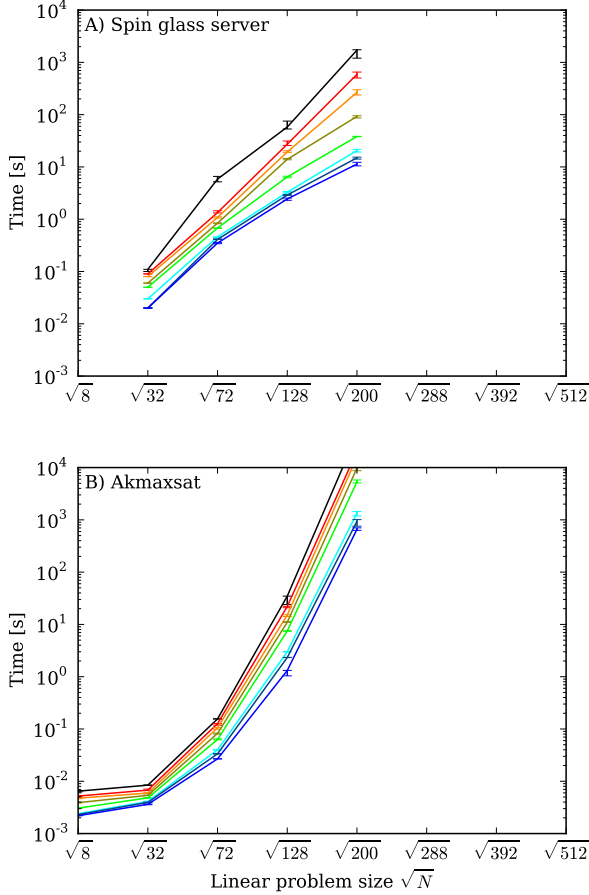FIG. 25: **Scaling plot for exact solvers and instances without fields** for A) the spin glass server and B) akmaxsat.

FIG. 26: **Scaling plot for exact solvers and instances with fields** for A) the spin glass server and B) akmaxsat.

where the timings data from the spin glass server was used.

Figure 24 shows the time scaling of the exact solvers for instances without fields. Two of the algorithms, an exact belief propagation algorithm using bucket sort [27] and the divide and conquer algorithm presented in section III C do not depend on the specific instance. The time to find a solution is the same for all instances with and without field. These algorithms make use of the structure of the chimera graph with a tree width of $\sqrt{N}$ [29, 33] and thus show the expected scaling $\mathcal{O}(\exp(c\sqrt{N}))$.

The tree width of a tree decomposition of a graph is the size of the largest vertex set of the tree (minus one according to some definitions [29, 33]). A tree decomposition of a graph G is a tree whose nodes are sets of vertices of G, and that satisfies the following properties:

1. Every vertex of G is in at least one node of the tree.

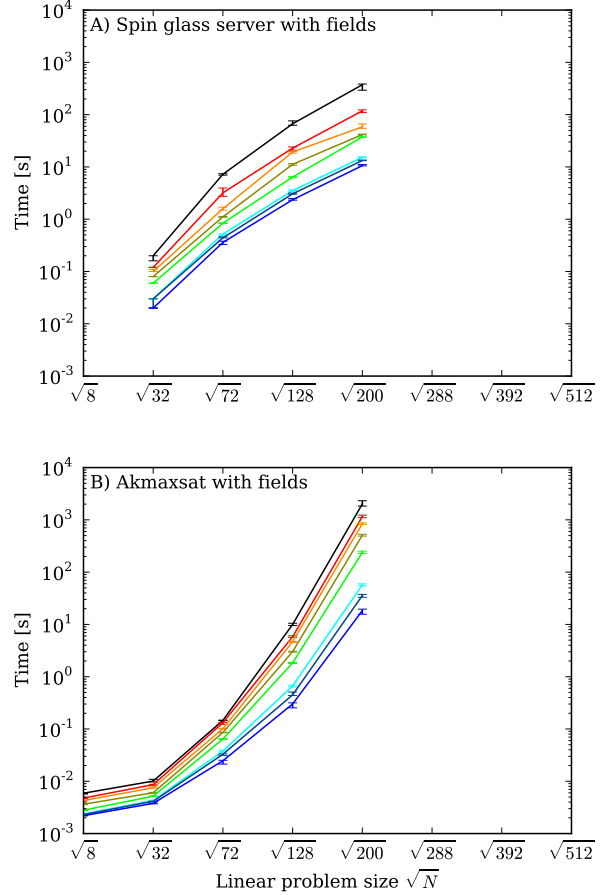2. For every edge (v, w) in G there is at least one node of the tree which includes both v and w.

3. The nodes of the tree that contain any given vertex v of G form a connected subtree.

The cost of doing exact belief propagation (or dynamical programming) on a tree decomposition of a graph G is exponential in the tree width. Belief propagation proceeds from the leaves down. A tree node t with a set of vertices $\{v_1, \ldots, v_l\}$ will calculate the minimum of the subgraph corresponding to the tree above t, conditional on all possible assignments to $\{v_1, \ldots, v_l\}$. The cost of this calculation is exponential in l, and the tree width is the largest l. A tree decomposition for chimera graphs with tree width $\mathcal{O}(\sqrt{N})$ can be seen in Ref. 44. Because a complete graph with $\mathcal{O}(\sqrt{N})$ vertices can be minor-embedded in the same chimera graph [29], this scaling is optimal for exact belief propagation.

The time to solution of the other two solvers, akmaxsat [38], and the biqmac algorithm [39] used in the spin glass server [40] depends on the specific instance. The scaling for the various percentiles from the easiest (1% percentile) to the hardest (99% percentile) is shown in figures 25 and 26 for instances with and without local fields.
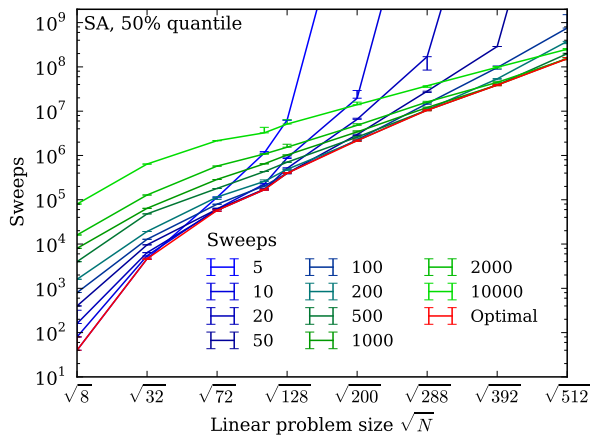
FIG. 27: **Scaling with problem size at fixed annealing time.** Shown is the scaling of the median total effort of the simulated annealer at constant annealing times measured in sweeps (one sweep is one attempted update per spin). This graph demonstrates that extrapolating results for fixed annealing times does not give the true asymptotic behaviour.

While the biqmac algorithm scales as $\mathcal{O}(\exp(c\sqrt{N}))$, akmaxsat scales worse and – unlike the other solvers – does not benefit from the limited tree width of the chimera graph. Unfortunately, the spin glass server constrained us to instances of up to 200 spins. It would be interesting to see if a crossover is present, i.e., to check whether the best exact code scales better than $\mathcal{O}(\exp(c\sqrt{N}))$.

B. Optimising the total annealing time

For the purpose of scaling comparisons we consider the total annealing time needed to find a ground state with a probability of 99%. Using Eq. (2) we find that the number of repetitions to find the ground state at least once with probability $p = 0.99$ is

$$R = \left\lceil \frac{\log(1-p)}{\log(1-s)} \right\rceil, \qquad (6)$$

where s is the probability for a given percentile. The total annealing time is Rt_f. Note that this expression assumes uncorrelated repetitions. On the D-Wave device we have seen statistically significant correlations between repetitions in 15% of the instances. In those instances the observed positive autocorrelations increase the length of "runs" of consecutive failures or successes on average by about a factor of two. This will result in an increase of the number of repetitions R required on the device, as given by equation (6) above. The simulated annealers, on the other hand, show no detectable correlations.

In the main text we argued that extrapolating the total annealing time on the D-Wave device at a fixed annealing t_f is tempting but misleading. It can only be used

to estimate the performance for slightly larger problem sizes, but will not give the true asymptotic scaling. To see this, consider figure 27 where we show the scaling of the median time for a simulated annealer for various fixed annealing times (similar behaviour is observed for a simulated quantum annealer). The scaling at *fixed* annealing time t_f increases only modestly for small system sizes, and then suddenly shoots upwards once t_f becomes too short for the problem size.

What needs to be done instead is to optimise the total annealing time for each problem size N. To do this we vary t_f, measure the success probability s, calculate the required number of repetitions R, and plot the total annealing time Rt_f as a function of t_f. The required number of repetitions R diverges when the annealing time t_f is too short, causing diabatic transitions. When the annealing time is too long it becomes disadvantageous to perform repetitions since a single run already optimises the success probability; however, Eq. (6) always yields $R \geq 1$. We thus pick the optimal t_f as the time that minimises Rt_f and use it in the scaling plots.

In figure 28 we show typical data for simulated classical and quantum annealing (where time is measured in number of spin flips for SA and spin updates for SQA) and for the D-Wave device (measured in μs). For the simulated classical and quantum annealer we observe that both the optimal time and the required number of repetitions R increase exponentially. Note that a minimum is not found for the D-Wave device but instead a monotonic increase, indicating that even the fastest annealing time of the hardware of $t_f = 5\mu$s is slower than the optimal time. All timings reported for the device are thus only *upper bounds* on the optimal time.

While the total annealing time for the D-Wave device can simply be given in microseconds, the annealing times of the simulated classical and quantum annealers depend on compiler options and the specific machine used. We thus give the total effort in terms of the number of attempted spin flips for the simulated classical annealer. For the simulated quantum annealer we specify the total effort as the number of spins updated multiplied by the inverse temperature β, to account for the complexity of updating the imaginary time world lines of the spins of length β.

Scaling plots of the total effort for instances with fields, complementing the results for instances without fields shown in the main text, are shown in figure 29. In panel A, we show the scaling of the D-Wave device. Again simulated classical annealing scales slightly better than simulated quantum annealing. Comparing the scaling for instances without fields (figure 4 in the main text) and instances with local random fields (figure 29), we see that problems with fields are not only easier for small problem sizes but the total annealing time also scales better when going to larger problem sizes.
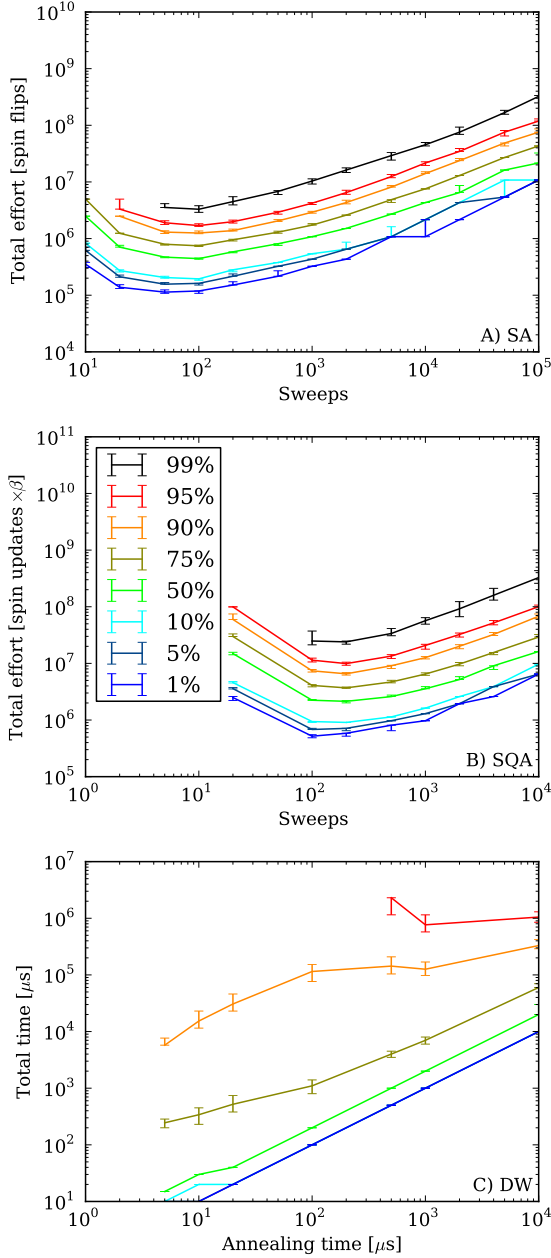
FIG. 28: **Optimisation of the total annealing time for $N = 108$ spin instances without local fields**. Shown is the total annealing time run time Rt_f time as function of the annealing time t_f for a single annealing run.

FIG. 29: **Scaling plot with fields** for A) QA, B) SA and C) SQA. Comparing with figure 6, B), D) and E) in the main text, which shows the same plots for problem instances without fields, we see that problems with fields are easier for SA and SQA.

C.   Detecting quantum speed up

Quantum speedup of a hardware quantum annealer can be detected by comparing the scaling of the total annealing times to that of the simulated classical or quantum annealer. To draw valid conclusions about a speedup one must ensure that the experimental annealing times are optimal. As we pointed out in the previous subsec-
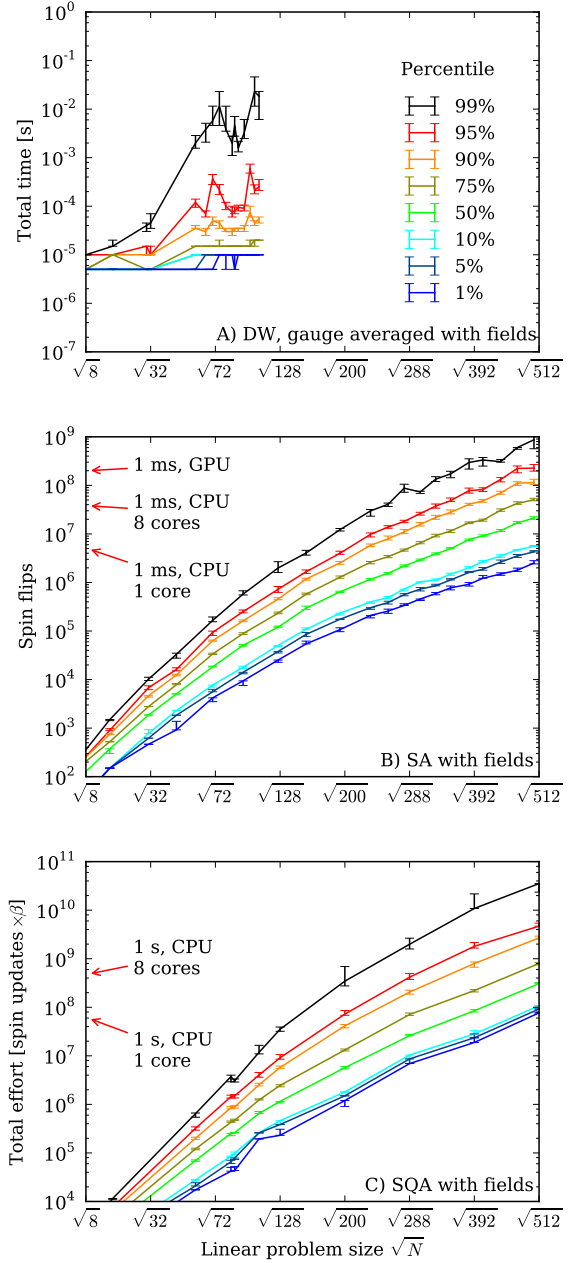
tion, this is not the case for the current device: the annealing time of 5μs is suboptimal, as demonstrated in figure 28. It follows that the inferred total annealing times are only upper bounds, and those bounds are worse for smaller problem sizes, which leads to scaling plots with an underestimated scaling. To see the pitfall it suffices to consider extremely long fixed annealing times where a single repetition $R = 1$ might be enough. We would then
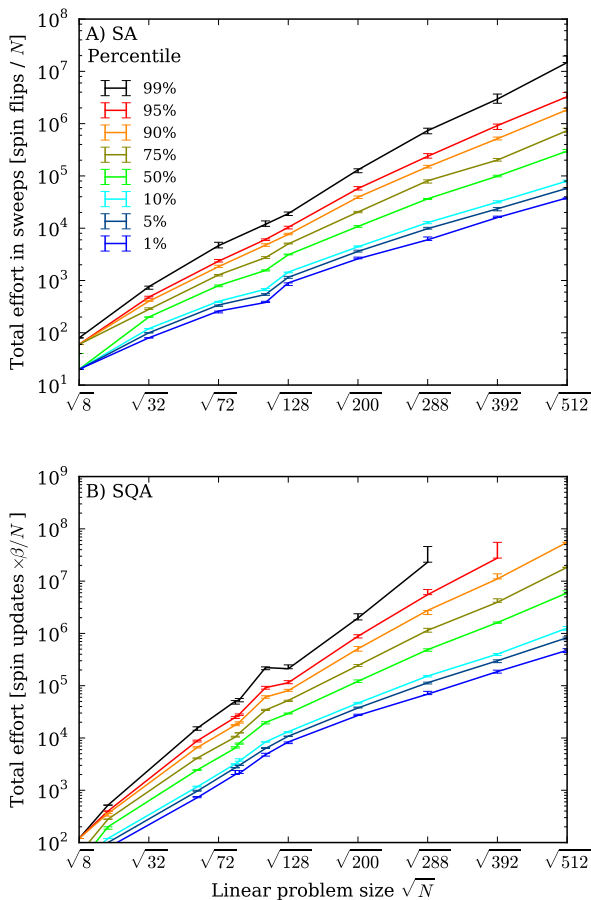
FIG. 30: **Scaling of total annealing effort assuming parallel updates of spins for instances without fields**. Since experimental annealers (quantum or classical) have intrinsic parallelism of updating all spins simultaneously, we need to divide the total effort by the number of spins N to obtain scaling plots against which quantum speedup can be detected.

see a *constant* time needed to find the ground state.

In other words, since the fastest possible annealing time $t_f = 5\mu s$ on the D-Wave device is longer than the optimal time for the considered problem sizes, our experimental data is in the initial transient regime of modest increase, and thus cannot be used for reliable extrapolation or determination of quantum speedup.

The optimal annealing time defined in the previous section (see figure 27) increases exponentially with problem size for SA. It is expected to increase exponentially with problem size also for QA. Therefore, assuming that the minimum programmable annealing time on future D-Wave devices does not increase too rapidly, we expect that perhaps already for a device with $N = 512$ spins, or possibly $N = 2048$ spins, the optimal annealing time t_f for a single run will be within reach. This will allow us to determine the optimal total annealing times Rt_f for QA.
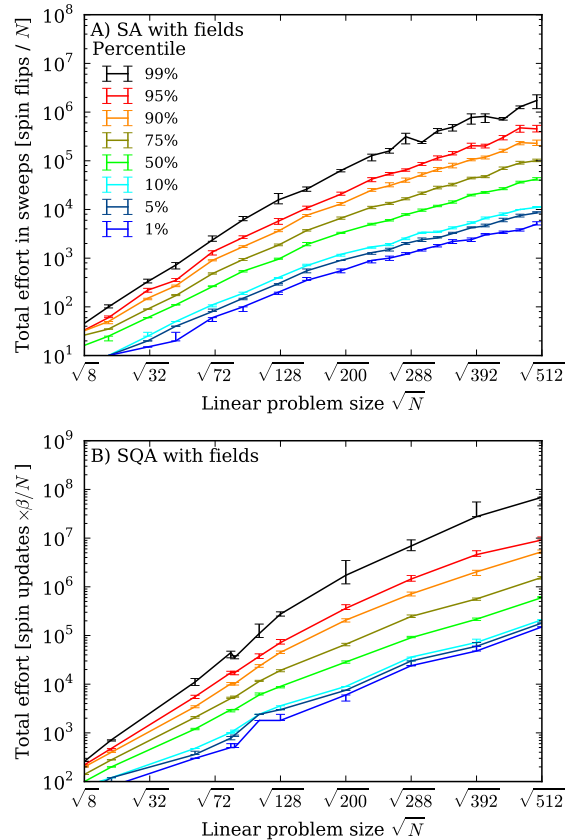

FIG. 31: **Scaling of total annealing effort assuming parallel updates of spins for instances with fields**. As in figure 30 the total effort is divided by the system size due to the intrinsic parallelism of an experimental quantum annealer.

To detect quantum speedup one should then compare to the total effort of the simulated annealers *divided by the number of spins* N. This division is necessary to compensate for the trivial parallelism of the hardware annealer, which updates N spins in parallel, since an analog classical annealing device would have the same parallelism. For completeness we provide the scaling plots of total effort in units of sweeps (spin flips divided by N) in figure 30.

VII. GAP CALCULATION

We finally describe how the excitation gaps are obtained using a method similar to that of Refs. [24, 25]. For simplicity we consider the transverse field Ising spin glass Hamiltonian without fields,

$$H = -\sum_{i<j} J_{ij}\sigma_i^z\sigma_j^z - \Gamma\sum_i \sigma_i^x, \qquad (7)$$

where Γ is the strength of a transverse magnetic field and for our instances $J_{ij} = \pm 1$.
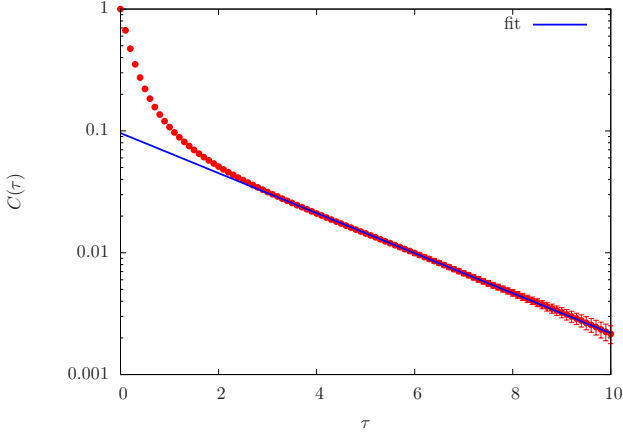
FIG. 32: **Correlation function** $C(\tau)$. The line shows an exponential fit. The gap can be obtained from the slope of the line.

The excitation gap can be obtained from the connected correlation function in imaginary time

$$C(\tau) = \langle \hat{O}(\tau)\hat{O}(0)\rangle - \langle \hat{O}\rangle^2, \qquad (8)$$

where $\hat{O}$ is an observable with non-vanishing matrix element between the ground state and first excited state. The correlation function is given by a sum:

$$C(\tau) = \sum_i c_i \exp(-\Delta_i \tau), \qquad (9)$$

where Δ_i is a gap to the ith eigenvalue above the ground state; $i = 0$ corresponds to the lowest gap. Only one term survives when τ is large enough:

$$C(\tau) = c_0 \exp(-\Delta_0 \tau). \qquad (10)$$

Thus $\Delta \equiv \Delta_0$ can be obtained by fitting $C(\tau)$ at large values of τ. We use periodic boundary conditions in the imaginary time direction. In this case, $C(\tau)$ can be efficiently calculated at discrete points using the Fast Fourier Transformation.

The observable $\hat{O}$ must be chosen carefully because for a poorly chosen observable, c_0 can be much smaller than $c_1, c_2, \ldots$ leading to very small values of $C(\tau)$ (comparable to statistical noise) at large τ and making the gap extraction very difficult. This issue is discussed in Ref. 45. We use a simple observable, the local magnetisation $\sigma_j^z(\tau)$ and its correlation function, given by

$$C(\tau) = \frac{1}{N}\sum_{j=0}^{N} \langle \sigma_j^z(\tau)\sigma_j^z(0)\rangle - \langle \sigma_j^z(0)\rangle^2, \qquad (11)$$

where the sum runs over all the spins N.

We use the continuous time algorithm described in section III B and "anneal" the system from $\Gamma = 3$ to small values of Γ in steps of 0.1. The simulation at a transverse

field Γ_i is started from the final configuration obtained in the previous simulation $\Gamma_{i-1} = \Gamma_i - 0.1$, except for the initial easy simulation at a strong transverse field $\Gamma = 3$, where the simulation is started from a random configuration. $200000/\Gamma$ Monte Carlo sweeps (one Monte Carlo sweep consists of $2N$ site updates) are performed for equilibration before measurements. Simulations are done at inverse temperatures $\beta = 100$ and $\beta = 200$.

Extraction of the gap for many instances and many values of Γ can be a cumbersome task. To automate the
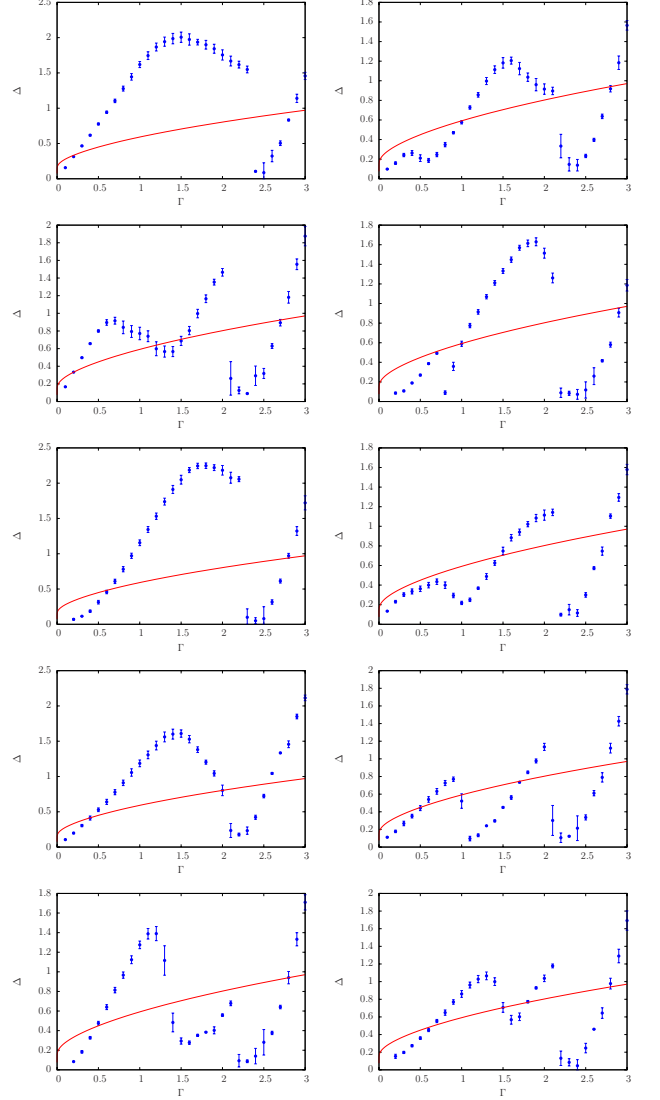
FIG. 33: **Evolution of the excitation gap**. Shown is the lowest excitation gap for the two instances presented in the main text, along with eight additional, complementary instances. The gap is given in units of $B(t)$ so that it is a property of the model and not of the annealing schedule. "Easy" instances are shown on the left (success probabilities from top to bottom: 0.98, 0.986, 0.995, 0.99, 0.932) and 'hard' instances are shown on the right (success probabilities from top to bottom: 0.08, 0.001, 0.063, 0.000625, 0.0000625). The red line indicates the temperature in units of $B(t)$ for schedule II.

process, we use the following approach. We obtain the gap by fitting the correlation function to the exponential function given by Eq. (10) in the range from τ_0 to $\tau_0 + 5/J$, where τ_0 is chosen empirically in such a way that $C(\tau) - C(\beta/2) > s_0$ for all $\tau < \tau_0$, $s_0 = 0.008 f(\Gamma)$ and $f(\Gamma) = 0.2 + 0.3\Gamma$. Let us denote this gap as Δ_0. To obtain the error bars, we calculate another two gaps Δ_1 and Δ_2 by performing two extra fits with $s_1 = 0.006 f(\Gamma)$ and with $s_2 = 0.011 f(\Gamma)$. Then the error bar is just $\max(\Delta_2 - \Delta_0, \Delta_0 - \Delta_1)$. This procedure can be fully automated. We show an example of such a fit in figure 32.

In figure 4 of the main text we showed the excitation gap in units of temperature. In figure 33 we instead show the gap in units of $B(t)$, so that the gap shown is a property of the model only and not of the annealing schedule. While the physical temperature is constant ($17\,\mathrm{mK}$), it increases as a function of Γ when shown in units of $B(t)$. The gap and temperature are both shown in figure 33 for the two instances presented in figure 4 of the main text, as well as for four "easy" (left column) and four "hard" (right column) additional instances. Note that the gap closes trivially around $\Gamma = 2.2$, related to a global Z_2 symmetry breaking. Once the gap becomes very small it can no longer be detected by our procedure since the

decay of $C(\tau)$ becomes too slow and is indistinguishable from a constant. Our procedure then picks up the gap to the next excited state, which results in an apparent jump of the gap to a bigger value. Generally, all gaps shown here are upper bounds for the gap to the lowest excited state.

The gap always closes also in the limit of a weak transverse field $\Gamma \longrightarrow 0$, where multiple ground states become degenerate. Some of the instances, however, have an additional small gap (relative to the temperature) that can be associated with an avoided level crossing. These instances are "hard" for both the D-Wave device and SQA. Usually, the instances that do not have such a small gap are "easy" for both the D-Wave device and SQA. However, note that the last "easy" instance shown in figure 33 does have a small gap. This shows that a small gap is not a sufficient condition for making a problem "hard". The fact that this last instance is "easy" can have multiple reasons, for example a second avoided level crossing with a small gap that takes the annealing back to a ground state or an avoided level crossing with a small Hamming distance from the ground state, from where thermal relaxation is not difficult.

[1] Muhly, J. The beginnings of metallurgy in the old world. In Maddin, R. (ed.) *The beginning of the use of metals and alloys*, 2–20 (Cambridge, MA: 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. *The Journal of Chemical Physics* **21**, 1087–1092 (1953). URL `http://link.aip.org/link/?JCP/21/1087/1`.

[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). URL `http://link.aps.org/doi/10.1103/PhysRevB.39.11828`.

[5] Finnila, A., Gomez, M., Sebenik, C., Stenson, C. & Doll, J. Quantum annealing: A new method for minimizing multidimensional functions. *Chemical Physics Letters* **219**, 343 – 348 (1994). URL `http://www.sciencedirect.com/science/article/pii/0009261494001170`.

[6] Kadowaki, T. & Nishimori, H. Quantum annealing in the transverse Ising model. *Phys. Rev. E* **58**, 5355–5363 (1998). URL `http://link.aps.org/doi/10.1103/PhysRevE.58.5355`.

[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). URL `http://link.aps.org/doi/10.1103/PhysRevB.66.094203`.

[8] Santoro, G. E., Martoňák, R., Tosatti, E. & Car, R. Theory of quantum annealing of an Ising spin glass. *Science* **295**, 2427–2430 (2002). URL `http://dx.doi.org/10.1126/science.1068774`.

[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). URL `http://www.sciencemag.org/content/284/5415/779.abstract`. http://www.sciencemag.org/content/284/5415/779.full.pdf.

[11] Johnson, M. W. *et al.* Quantum annealing with manufactured spins. *Nature* **473**, 194–198 (2011). URL `http://dx.doi.org/10.1038/nature10012`.

[12] Boixo, S., Albash, T., Spedalieri, F. M., Chancellor, N. & Lidar, D. A. Experimental signature of programmable quantum annealing (2012). URL `http://arxiv.org/abs/1212.1739`. 1212.1739.

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

[14] Barahona, F. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General* **15**, 3241 (1982). URL `http://stacks.iop.org/0305-4470/15/i=10/a=028`.

[15] Lucas, A. Ising formulations of many NP problems (2013). URL `http://arxiv.org/abs/1302.5843`. 1302.5843.

[16] Farhi, E. *et al.* A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem. *Science* **292**, 472–475 (2001). URL `http://dx.doi.org/10.1126/science.1057726`.

[17] 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). URL http://link.aps.org/doi/10.1103/PhysRevLett.104.207206.

[18] Hen, I. & Young, A. P. Exponential complexity of the quantum adiabatic algorithm for certain satisfiability problems. *Phys. Rev. E* **84**, 061152 (2011). URL http://link.aps.org/doi/10.1103/PhysRevE.84.061152.

[19] 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). URL http://link.aps.org/doi/10.1103/PhysRevA.86.052334.

[20] Bian, Z., Chudak, F., Macready, W. G., Clark, L. & Gaitan, F. Experimental determination of Ramsey numbers with quantum annealing (2012). URL http://arxiv.org/abs/1201.1842. 1201.1842.

[21] Perdomo-Ortiz, A., Dickson, N., Drew-Brook, M., Rose, G. & Aspuru-Guzik, A. Finding low-energy conformations of lattice protein models by quantum annealing. *Scientific Reports* **2** (2012).

[22] Wang, L. *et al.* Comment on: 'Classical signature of quantum annealing' (2013). URL http://arxiv.org/abs/1305.5837. arXiv:1305.5837.

[23] Smolin, J. A. & Smith, G. Classical signature of quantum annealing (2013). URL http://arxiv.org/abs/1305.4904. arXiv:1305.4904.

[24] 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). URL http://link.aps.org/doi/10.1103/PhysRevB.59.1162.

[25] 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). URL http://link.aps.org/doi/10.1103/PhysRevLett.101.170503.

[26] Altshuler, B., Krovi, H. & Roland, J. Anderson localization makes adiabatic quantum optimization fail. *Proceedings of the National Academy of Sciences* **107**, 12446–12450 (2010). URL http://www.pnas.org/content/107/28/12446.abstract. http://www.pnas.org/content/107/28/12446.full.pdf+html.

[27] Dechter, R. Bucket elimination: A unifying framework for reasoning. *Artificial Intelligence* **113**, 41–85 (1999).

[28] McGeoch, C. C. & Wang, C. Experimental evaluation of an adiabatic quantum system for combinatorial optimization. In *Proceedings of the 2013 ACM Conference on Computing Frontiers* (2013).

[29] Choi, V. Minor-embedding in adiabatic quantum computation: II. Minor-universal graph design. *Quantum Information Processing* **10**, 343–353 (2011). URL http://dx.doi.org/10.1007/s11128-010-0200-3.

[30] Harris, R. *et al.* Experimental demonstration of a robust and scalable flux qubit. *Physical Review B* **81** (2010). URL http://link.aps.org/doi/10.1103/PhysRevB.81.134510.

[31] Harris, R. *et al.* Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor. *Phys. Rev. B* **82**, 024511 (2010). URL http://link.aps.org/doi/10.1103/PhysRevB.82.024511.

[32] Berkley, A. J. *et al.* A scalable readout system for a superconducting adiabatic quantum optimization system. *Superconductor Science and Technology* **23**, 105014 (2010). URL http://stacks.iop.org/0953-2048/23/i=10/a=105014.

[33] Choi, V. Minor-embedding in adiabatic quantum computation: I. The parameter setting problem. *Quantum Information Processing* **7**, 193–209 (2008). URL http://dx.doi.org/10.1007/s11128-008-0082-9.

[34] Bhanot, G., Duke, D. & Salvador, R. A fast algorithm for the Cyber 205 to simulate the 3d Ising model. *J. Stat. Phys.* **44**, 985–1002 (1986).

[35] Heuer, H.-O. A fast vectorized Fortran 77 program for the Monte Carlo simulation of the three-dimensional Ising system. *Comput. Phys. Commun.* **59**, 387 – 398 (1990). URL http://www.sciencedirect.com/science/article/pii/0010465590901865.

[36] Rieger, H. & Kawashima, N. Application of a continuous time cluster algorithm to the two-dimensional random quantum Ising ferromagnet. *Eur. Phys. J. B* **9**, 233–236 (1999).

[37] 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). URL http://link.aps.org/doi/10.1103/PhysRevB.66.094203.

[38] Kuegel, A. Improved exact solver for the weighted max-sat problem. In Berre, D. L. (ed.) *POS-10*, vol. 8 of *EPiC Series*, 15–27 (EasyChair, 2012).

[39] Rendl, F., Rinaldi, G. & Wiegele, A. Solving Max-Cut to optimality by intersecting semidefinite and polyhedral relaxations. *Math. Programming* **121**, 307 (2010).

[40] Spin glass server. URL http://www.informatik.uni-koeln.de/spinglass/.

[41] Moessner, R. & Sondhi, S. L. Ising models of quantum frustration. *Phys. Rev. B* **63**, 224401 (2001). URL http://link.aps.org/doi/10.1103/PhysRevB.63.224401.

[42] Amin, M. H. S. & Choi, V. First order quantum phase transition in adiabatic quantum computation. *Phys. Rev. A* **90**, 062326 (2009).

[43] Bapst, V., Foini, L., Krzakala, F., Semerjian, G. & Zamponi, F. The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective. *Physics Reports* **523**, 127 – 205 (2013). URL http://www.sciencedirect.com/science/article/pii/S037015731200347X. ¡ce:title¿The Quantum Adiabatic Algorithm Applied to Random Optimization Problems: The Quantum Spin Glass Perspective¡/ce:title¿.

[44] Klymko, C., Sullivan, B. D. & Humble, T. S. Adiabatic quantum programming: Minor embedding with hard faults. *arXiv:1210.8395* (2012). URL http://arxiv.org/abs/1210.8395.

[45] Hen, I. Excitation gap from optimized correlation functions in quantum Monte Carlo simulations. *Phys. Rev. E* **85**, 036705 (2012). URL http://link.aps.org/doi/10.1103/PhysRevE.85.036705.