

# Quantum critical dynamics in a 5000-qubit programmable spin glass

Andrew D. King,<sup>1,\*</sup> Jack Raymond,<sup>1</sup> Trevor Lanting,<sup>1</sup> Richard Harris,<sup>1</sup> Alex Zucca,<sup>1</sup> Fabio Altomare,<sup>1</sup> Andrew J. Berkley,<sup>1</sup> Kelly Boothby,<sup>1</sup> Sara Ejtemaei,<sup>1</sup> Colin Enderud,<sup>1</sup> Emile Hoskinson,<sup>1</sup> Shuiyuan Huang,<sup>1</sup> Eric Ladizinsky,<sup>1</sup> Allison J.R. MacDonald,<sup>1</sup> Gaelen Marsden,<sup>1</sup> Reza Molavi,<sup>1</sup> Travis Oh,<sup>1</sup> Gabriel Poulin-Lamarre,<sup>1</sup> Mauricio Reis,<sup>1</sup> Chris Rich,<sup>1</sup> Yuki Sato,<sup>1</sup> Nicholas Tsai,<sup>1</sup> Mark Volkmann,<sup>1</sup> Jed D. Whittaker,<sup>1</sup> Jason Yao,<sup>1</sup> Anders W. Sandvik,<sup>2,†</sup> and Mohammad H. Amin<sup>1,3,‡</sup>

<sup>1</sup>*D-Wave Systems, Burnaby, British Columbia, Canada*

<sup>2</sup>*Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, Massachusetts 02215, USA*

<sup>3</sup>*Department of Physics, Simon Fraser University, Burnaby, British Columbia, Canada*

(Dated: April 20, 2023)

Experiments on disordered alloys<sup>1–3</sup> suggest that spin glasses can be brought into low-energy states faster by annealing quantum fluctuations than by conventional thermal annealing. Due to the importance of spin glasses as a paradigmatic computational testbed, reproducing this phenomenon in a programmable system has remained a central challenge in quantum optimization<sup>4–13</sup>. Here we achieve this goal by realizing quantum critical spin-glass dynamics on thousands of qubits with a superconducting quantum annealer. We first demonstrate quantitative agreement between quantum annealing and time-evolution of the Schrödinger equation in small spin glasses. We then measure dynamics in 3D spin glasses on thousands of qubits, where simulation of many-body quantum dynamics is intractable. We extract critical exponents that clearly distinguish quantum annealing from the slower stochastic dynamics of analogous Monte Carlo algorithms, providing both theoretical and experimental support for a scaling advantage in reducing energy as a function of annealing time.

The study of spin glasses initiated an enormously productive exchange between physics and computer science<sup>14,15</sup>. One key byproduct of this exchange was the invention of simulated annealing (SA)<sup>16</sup>, a method of optimization that simulates a gradually cooling system as it settles into a low-energy state. Recent decades have seen annealing brought to bear against countless multivariate optimization applications, seeking low-energy states that translate to low-cost solutions<sup>17</sup>.

Passing through a thermal phase transition—as in SA—is not the only way to evolve a spin system from an “easy” disordered phase into a “hard” glassy phase. One can also pass through a quantum phase transition (QPT), where the ground state undergoes a macroscopic shift in response to changing quantum fluctuations. Both experiments<sup>1</sup> and simulations<sup>4,5</sup> suggested that quantum annealing (QA) can guide a spin glass toward equilibrium faster than thermal annealing. Thus originated QA as a means of both studying quantum critical phenomena and optimizing quadratic objective functions<sup>18,19</sup>. Simulating the Schrödinger dynamics of QA with a classical computer is an unpromising optimization method, since memory requirements grow exponentially with system size. But, as Feynman famously asked<sup>20</sup>, “Can you do it with a new kind of computer—a quantum computer?”

This question motivated not only the development of programmable QA processors<sup>6,21–25</sup>, but a more general effort to probe the capabilities of near-term quantum devices via quantum simulation, including in trapped

ions<sup>26,27</sup>, ultracold atoms<sup>28</sup>, and Rydberg arrays<sup>29</sup> (the latter was recently explored as an annealing optimizer<sup>30</sup>). While D-Wave QA processors have already been used to simulate quantum spin glasses in a decohering thermal bath<sup>31</sup>, it was only recently shown that they can simulate QPTs with negligible interaction with the thermal environment<sup>32</sup>: coherent dynamics of quantum-critical phenomena despite the finite temperature of the apparatus itself. Here we use a QA processor to study the critical dynamics of a spin-glass QPT. The exceedingly slow dynamics of the spin-glass state make this phase transition vitally important in the study of quantum optimization. We compare these dynamics against SA and simulated quantum annealing (SQA), an algorithm based on path-integral Monte Carlo<sup>33</sup> that reproduces thermal equilibrium statistics of QA<sup>7,12,34</sup>.

## QUANTUM ANNEALING AND SCHRÖDINGER DYNAMICS

We use a D-Wave Advantage QA processor (Fig. 1a) whose pairwise-coupled superconducting flux qubits can be programmed to realize a transverse-field Ising model described by the Hamiltonian

$$\mathcal{H}(s) = \Gamma(s)\mathcal{H}_D + \mathcal{J}(s)\mathcal{H}_I \quad (1)$$

$$\mathcal{H}_D = -\sum_i \sigma_i^x \quad (2)$$

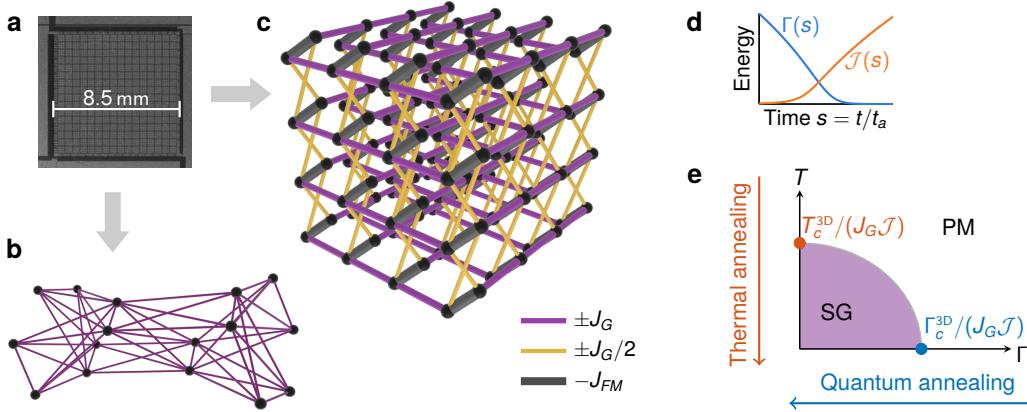
$$\mathcal{H}_I = \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z \quad (3)$$

Here  $\sigma_i^x$ ,  $\sigma_i^z$  are Pauli operators on qubit  $i$ ,  $s$  is a unitless normalized time, the transverse field  $\Gamma(s)$  imparts quantum fluctuations through the driver Hamiltonian  $\mathcal{H}_D$ ,

\* [aking@dwavesys.com](mailto:aking@dwavesys.com)

† [sandvik@buphy.bu.edu](mailto:sandvik@buphy.bu.edu)

‡ [mhsamin@dwavesys.com](mailto:mhsamin@dwavesys.com)



**FIG. 1. Programmable quantum spin glasses.** **a**, QA processor realizing a transverse-field Ising model in pairwise-coupled superconducting flux qubits, into which various lattice geometries can be programmed. **b**, 16-spin graph used for small-scale studies of Schrödinger evolution. Each line represents a coupling, whose energy is set to  $J_G$  or  $-J_G$  at random. **c**, 3D structure, where gray bonds represent ferromagnetically coupled dimers and any two dimers have total coupling  $J_G$ ,  $-J_G$ , or 0 between them. **d**, The QA schedule guides the system from a quantum paramagnet toward a classical state by annealing  $\Gamma(s)/\mathcal{J}(s)$  over time  $t_a$ . **e**, 3D quantum spin-glass phase diagram. Paramagnet and spin-glass phases are separated by a thermal phase transition at temperature  $T = T_c^{3D} > 0$  when  $\Gamma = 0$ , and a quantum phase transition at  $\Gamma_c^{3D}$  when  $T = 0$ .

and  $\mathcal{J}(s)$  is the energy scale of the classical Ising Hamiltonian  $\mathcal{H}_I$ . Over annealing time  $t_a$ ,  $s = t/t_a$  increases from 0 to 1, annealing the system from a quantum paramagnet dominated by  $\mathcal{H}_D$ , to a classical Ising model dominated by  $\mathcal{H}_I$ , following an annealing schedule as in Fig. 1d. The coupling coefficients  $J_{ij}$  can be programmed into a variety of 2D and 3D geometries, among others<sup>12,31,35–38</sup> (the QA processor also provides programmable biases, which we set to zero in this study).

Although our main focus is on large spin glasses, we first seek evidence of coherent quantum dynamics in an ensemble of small spin glasses. Taking the 16-spin graph in Fig. 1b, we generate spin-glass realizations with each coupling set to  $J_{ij} = +1$  or  $-1$  uniformly at random. We select 100 spectrally unique realizations in which  $\mathcal{H}_I$  has two ground states and many first excited states. At this scale we can numerically evolve the time-dependent Schrödinger equation

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \mathcal{H}(t/t_a) |\psi(t)\rangle, \quad (4)$$

for the wavefunction  $|\psi(t)\rangle$ , where  $\mathcal{H}(s)$  is given by Eq. (1). Let  $|\phi_i(s)\rangle$  denote instantaneous eigenstates of  $\mathcal{H}(s)$  with eigenvalues  $E_i(s)$ . Sufficiently slow evolution results in adiabatic quantum optimization (AQO)<sup>39</sup> into the twofold-degenerate final ground states  $|\phi_0(1)\rangle$  and  $|\phi_1(1)\rangle$ . The relevant timescale is proportional to  $\delta_{\min}^{-2}$ , where

$$\delta_{\min} = \min_s |E_2(s) - E_0(s)| \quad (5)$$

is the minimum parity-preserving eigengap. Faster anneals have a higher probability of excitation.

In Fig. 2a, we show spectral gaps  $E_i(s) - E_0(s)$  for three 16-qubit examples: one with a small gap, one with

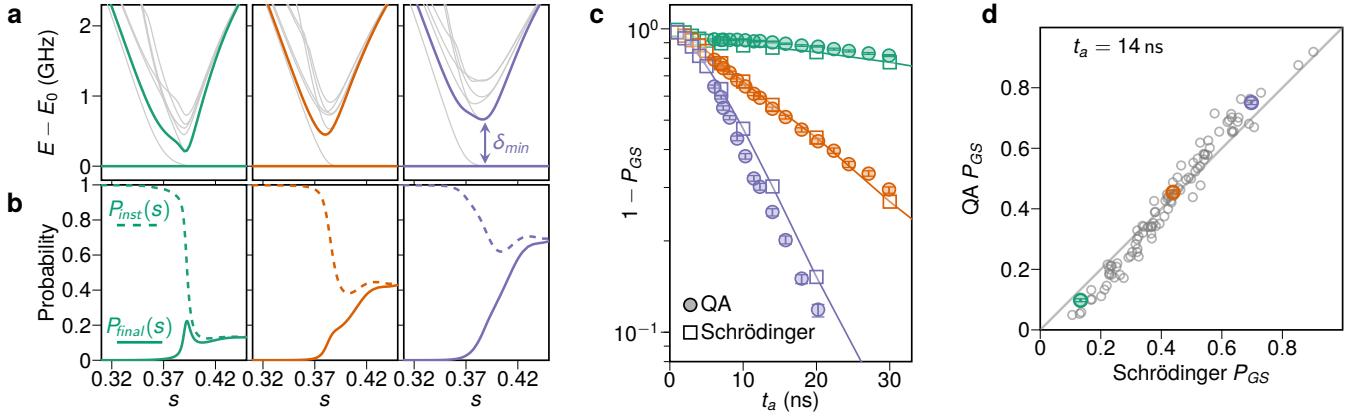
a moderate gap, and one with a large gap. For fixed  $t_a$  we define

$$P_{inst}(s) = \sum_{n \in \{0,1\}} |\langle \phi_n(s) | \psi(s) \rangle|^2, \quad (6)$$

$$P_{final}(s) = \sum_{n \in \{0,1\}} |\langle \phi_n(1) | \psi(s) \rangle|^2, \quad (7)$$

which measure instantaneous probabilities of being in the ground or first excited state of  $\mathcal{H}(s)$  and  $\mathcal{H}(1)$ , respectively. Since the classical ground states are twofold degenerate,  $P_{final}(1)$  gives the success probability of QA under Schrödinger dynamics. Fig. 2b tracks  $P_{inst}(s)$  and  $P_{final}(s)$  through anneals with  $t_a = 14$  ns. The wavefunction begins concentrated on the easily-prepared ground state at  $s = 0$ , and this probability decreases via Landau-Zener excitation in the vicinity of a small gap.

We run QA experiments on each of these 100 instances using 192 disjoint sets of qubits in parallel. The Schrödinger excitation probability  $1 - P_{GS} = 1 - P_{final}(1)$  is compared against experimental QA excitation probability in Fig. 2c; the probabilities are in close agreement with no fitting parameters used, showing an approximately exponential form. Fig. 2d compares  $P_{GS}$  for  $t_a = 14$  ns across the entire 100-instance ensemble. In Extended Data Fig. E1 we compare probability distributions among ground and first excited states for QA, SA, SQA, and Schrödinger dynamics, and find that experimental QA data are better explained by Schrödinger dynamics than by SA and SQA. The quantitative agreement between QA experiment and Schrödinger evolution up to  $t_a = 30$  ns provides strong evidence for coherent quantum dynamics at small scale. We now consider critical dynamics in large 3D spin glasses.



**FIG. 2. Coherent Schrödinger dynamics.** **a**, For three exemplary 16-qubit spin glasses, we show the eight lowest eigengaps of the time-dependent QA Hamiltonian. **b**, We evolve the Schrödinger equation for QA with  $t_a = 14$  ns, tracking the wavefunction’s probability of collapse onto the final (classical) ground-state manifold ( $P_{final}(s)$ ) and the two lowest-energy instantaneous eigenstates ( $P_{inst}(s)$ ).  $P_{final}(1)$  is the final ground-state probability  $P_{GS}$  of the Schrödinger evolution. **c**,  $1 - P_{GS}$  for the same three spin glasses, in experimental QA and Schrödinger evolution for a range of annealing times. **d**, Comparison of  $P_{GS}$  for QA and Schrödinger evolution for an ensemble of 100 16-spin glasses, for  $t_a = 14$  ns. Error bars indicate 95% confidence intervals for the average over parallel QA experiments using 192 different sets of qubits.

## CRITICAL SPIN-GLASS DYNAMICS

When a system is brought slowly (annealed) through a continuous phase transition, its dynamics slow down due to diverging correlations, and its macroscopic properties follow universal behavior described by critical exponents. Here we use extensions of the Kibble-Zurek (KZ) mechanism<sup>40–48</sup>, which describes the generation of excitations as an annealed system falls out of equilibrium. We use a dynamic finite-size scaling (DFSS) ansatz<sup>46</sup> (SM I) to relate time and the growth of correlations as functions of two critical exponents:  $\nu$ , which describes the divergence of correlation length at a phase transition, and  $z$ , which describes divergence of the characteristic timescale of domain fluctuations. We investigate three annealing dynamics—QA, SQA, and SA—and their corresponding phase transitions.

Spin-glass order is quantified via the overlap between two replicas (independently annealed  $N$ -spin states  $S$  and  $S'$ ) of a given realization (set of couplings  $J_{ij}$ ):

$$q = \frac{1}{N} \sum_{i=1}^N S_i S'_i \quad S_i, S'_i \in \{-1, 1\}. \quad (8)$$

The mean-squared Edwards-Anderson order parameter is given by  $\langle q^2 \rangle$ , with  $\langle \cdot \rangle$  denoting an average over both independent replicas and realizations.

Due to restrictions on the available coupling geometry, we program spin glasses in the 3D layout shown in Fig. 1c, which differs from the simple cubic lattice: it has two qubits at every  $(x, y, z)$  coordinate, coupled as a dimer with strong ferromagnetic coupling  $J_{ij} = -J_{FM} = -2$ . Between neighboring cubic coordinates, we program a total coupling of  $\pm J_G$  ( $0 < J_G \leq 1$ )—this coupling uses one coupler in the  $x$  and  $y$  directions, and in the  $z$  direction

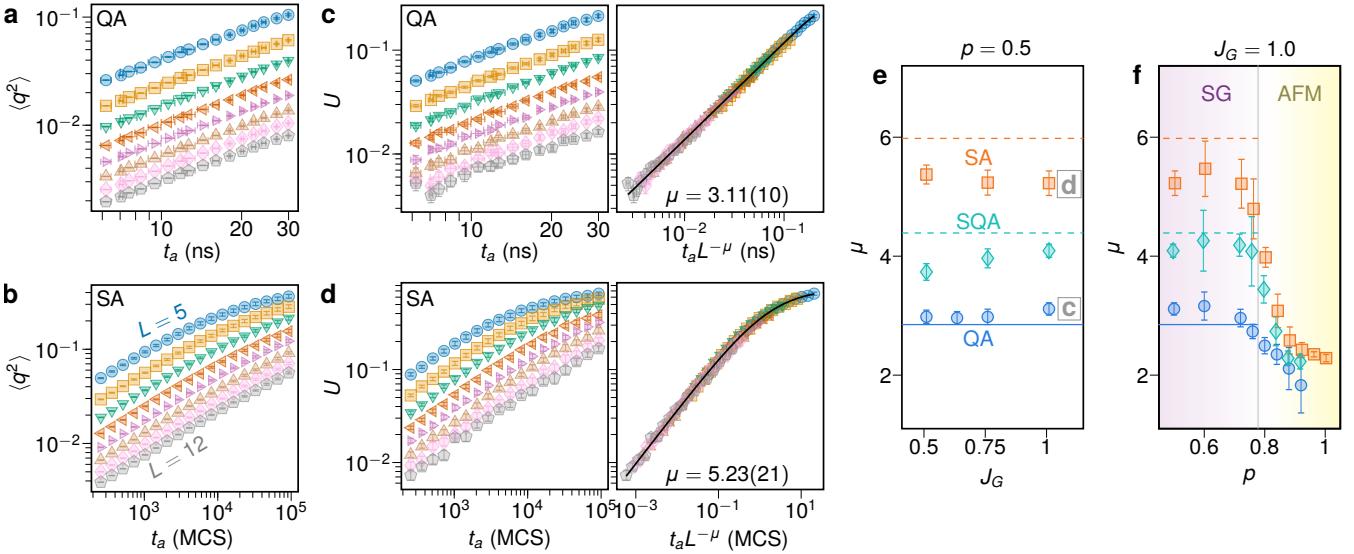
the coupling is equally divided between two couplers. We use open  $x$ - and  $y$ -boundaries and periodic  $z$ -boundaries.

In this model, contracting the strongly-correlated dimers into individual spins yields the  $\pm J_G$  Ising spin glass on a simple cubic lattice, so we expect experimental results to reflect criticality in the same universality class. Fig. 1e shows the model’s phase diagram in the  $\Gamma$ ,  $T$  plane, where a spin-glass (SG) phase is separated from a disordered paramagnetic (PM) phase. At  $T = 0$  there is a quantum critical point  $\Gamma = \Gamma_c > 0$ . At  $\Gamma = 0$  there is a finite-temperature classical transition at  $T = T_c > 0$ . Tuning  $J_G$  while keeping  $J_{FM} = 2$  varies the details of the phase diagram, but not the qualitative picture.

We measure  $\langle q^2 \rangle$  for a range of  $t_a$ , with linear system size  $L$  ranging from 5 to 12, in QA (Fig. 3a), SA (Fig. 3b), and SQA (Extended Data Fig. E6), for  $J_G = 1$ . For Monte Carlo (MC) methods,  $t_a$  is measured in MC sweeps (MCS). Although in all cases we anneal through the critical point rather than stopping at the critical point (see Methods), the system experiences a critical slowing down at the QPT. Due to slower dynamics in the glass phase, measurements in the final state approximately reflect the relevant critical dynamics. The Binder cumulant

$$U = \frac{1}{2} \left( 3 - \frac{\langle q^4 \rangle}{\langle q^2 \rangle^2} \right) \quad (9)$$

provides a statistical signature of phase transitions, and like  $\langle q^2 \rangle$ , also grows with  $t_a$  and  $1/L$ , as seen in Fig. 3c–d. For  $U$ , any post-critical effective scaling dimensions in  $\langle q^2 \rangle^2$  and  $\langle q^4 \rangle$  cancel, but the functional dependence on  $t_a$  remains. Thus, under the DFSS ansatz,  $U(L, t_a)$  is expected to collapse onto a common curve for all system sizes when  $t_a$  is rescaled by  $L^{-z-1/\nu}$  (SM IB), reflecting the fact that the annealing time required for the system



**FIG. 3. Dynamic finite-size scaling in 3D spin glasses.** **a–b**, Squared overlap  $\langle q^2 \rangle$  varies as a function of  $L$  and  $t_a$  in QA (**a**) and SA (**b**) for 3D spin glasses. **c–d**, Binder cumulant  $U$  scales similarly; rescaling  $t_a$  by  $L^\mu$  collapses data onto a single curve for a fit parameter  $\mu$ , providing estimates of the KZ exponent  $\mu = z + 1/\nu$ . **e**, Estimates of  $\mu$  for varying spin-glass coupling strengths  $J_G$ , where dimer coupling  $J_{FM} = 2$ , in QA, SA, and SQA. Solid and dashed horizontal lines indicate literature values<sup>49</sup> and estimates from simple cubic lattices (Extended Data E8) respectively. **f**, Tuning the doping probability  $p$  reveals a finite-size crossover between the SG phase and the AFM phase, separated by a critical doping  $p_c \approx 0.778$  at  $T = 0$ ,  $\Gamma = 0$  (vertical line). Observed dynamics in these phases are characteristic of critical dynamics and coarsening dynamics, respectively. Vertical error bars are 95% statistical confidence intervals and horizontal error bars indicate measurement uncertainty in  $t_a$ .

to remain adiabatic up to a correlation length of  $L$  scales as

$$t_a(L) \sim L^\mu, \quad \mu = z + 1/\nu. \quad (10)$$

Thus we estimate the KZ exponent  $\mu$  via best-fit collapse of  $U$  horizontally along the time axis. Fig. 3c–d show collapses of  $U$ , from which we extract  $\mu$  as a fitting parameter.

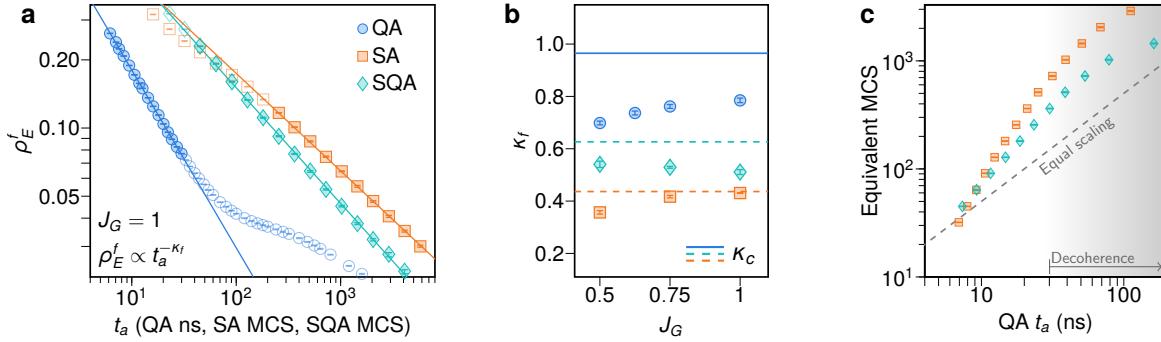
In Fig. 3e we show these estimates for QA, SA and SQA for a range of coupling energies  $J_G$ . In all cases,  $U(L, t_a)$  approximately follows power-law scaling when far from equilibrium<sup>48</sup>. The extracted KZ exponents are clearly distinct, with smaller values indicating faster dynamics. We find  $\mu_{QA}$  between 2.9 and 3.1 depending on  $J_G$ . Adding our MC estimate  $1/\nu_{QA} \approx 1.55$  (SM III) to  $z_{QA} \approx 1.3$ <sup>49</sup> gives  $\mu_{QA} \approx 2.85$ , in close agreement with experimental results.

For SA we find  $\mu_{SA} \approx 5.3$ , much larger than QA but smaller than estimates from simple cubic lattices; for SA we found a value 6.0 (dashed line) compared to previously reported<sup>48</sup> value 6.3. These deviations are largely explained by finite size effects, and dimers delaying the onset of asymptotic behavior (see Figs. E6–E8 and SM Section IV A). No previous estimate of  $\mu_{SQA}$  has been reported; the dashed line in Fig. 3e indicates the extracted value  $\mu_{SQA} = 4.39$  for simple cubic lattices (Fig. E8). Since QA and SQA share an equilibrium exponent  $\nu$ , this corresponds to  $z_{SQA} \approx 2.8$ , which reflects the MC dynamics.

To better understand the role of frustration in the glassy dynamics, we increase the “doping” probability  $p$  of a random inter-dimer coupling being antiferromagnetic ( $+J_G$ ). The SG phase has been shown to persist until  $p_c = 0.778$  (vertical line) for  $T = 0$  and  $\Gamma = 0$ , beyond which the system is a disordered antiferromagnet<sup>50–52</sup>—we might expect that  $\Gamma > 0$  slightly reduces  $p_c$ <sup>53</sup>. Fig. 3f shows the  $p$ -dependence of  $\mu$  extracted by collapsing for even values of  $L$ ; all three dynamics are insensitive to changes in  $p$  until it approaches  $p_c$ . For  $p$  close to 1, QA, SA, and SQA all have  $\mu \approx 2$ , consistent with coarsening dynamics in the AFM phase<sup>47</sup>. In this scenario, the dynamics that occurs after the critical point eliminates small domains, replacing the KZ exponent  $\mu$  with a universal exponent 2. The latter corresponds to correlation length scaling as  $\xi \propto t_a^{-1/2}$ , expected for diffusive dynamics present in the AFM phase. Due to the rough potential landscape in the SG phase, the post-critical dynamics have negligible effect on  $\mu$ , although they affect energy decay as we discuss next.

## ENERGY DECAY

The smaller values of critical exponents  $z$  and  $\mu$ , obtained via data collapse, indicate faster critical dynamics in QA compared to SA and SQA. We now ask whether this leads to a speedup in approximating the ground-state energy in classical Ising models. We first answer



**FIG. 4. Critical scaling of final residual energy.** **a**, Scaling of final residual energy density  $\rho_E^f$  versus  $t_a$ , with lines showing power-law fits to the form  $\rho_E^f \propto t_a^{-\kappa_f}$ . Empty markers are excluded from power-law fit (see text). **b**, Markers show extracted exponents  $\kappa_f$  for QA, SA, and SQA for varying  $J_G$  in 3D spin glasses; horizontal lines show estimates of  $\kappa_c$  (Eq. (12)) using estimates of  $\mu$  shown as horizontal lines in Fig. 3e. **c**, Number of MCS required for SA and SQA to match the residual energy achieved by QA in a given QA annealing time. Dashed line is a guide to the eye showing equal scaling. Shading represents the onset of thermal effects outside the coherent QA regime. Error bars represent 95% bootstrap C.I. over spin-glass realizations.

this question theoretically by considering a hypothetical annealing protocol that is measured at the critical point. Although this is not the real schedule of QA, it makes the energy decay dependent only on the dynamics approaching the critical point, and therefore allows connecting the relevant critical exponents. We define dimensionless residual Ising energy density at the critical point as

$$\rho_E^c = \langle \mathcal{H}_I - E_c \rangle / (N J_G). \quad (11)$$

Here  $N$  is the number of spins, and  $E_c$  is the equilibrium expectation of  $\mathcal{H}_I$  at  $T = T_c$  and  $\Gamma = 0$  for SA, and at  $T = 0$  and  $\Gamma = \Gamma_c$  for QA and SQA. Notice that  $\rho_E^c \rightarrow 0$  as  $t_a \rightarrow \infty$ . It is shown in SM I that  $\rho_E^c$  follows a power-law relation:

$$\rho_E^c \propto t_a^{-\kappa_c}, \quad \kappa_c = (d_s - 1/\nu)/\mu, \quad (12)$$

where  $d_s = d$  for SA and  $d_s = d + z_{\text{QA}}$  for QA and SQA. As expected,  $\kappa_c$  is inversely related to the KZ exponent  $\mu$ , therefore faster critical dynamics (smaller  $\mu$ ) leads to faster decay of energy. Moreover,  $d_s$  is larger for quantum than classical, making a larger contribution to the numerator.

However,  $\rho_E^c$  is not very relevant to optimization because  $E_c$  is far from the ground-state energy  $E_0$  of  $\mathcal{H}_I$ . We therefore consider the corresponding final quantities  $\rho_E^f$  and  $\kappa_f$  obtained by annealing to the low-temperature classical point  $T \ll T_c$ ,  $\Gamma = 0$ :

$$\rho_E^f = \langle \mathcal{H}_I - E_0 \rangle / (N J_G), \quad (13)$$

and fitting  $\rho_E^f \propto t_a^{-\kappa_f}$ . Again the average is over realizations and samples, with annealing according to the full QA schedule. For very long (adiabatic) anneals, we expect  $\rho_E^f \rightarrow 0$ , thus optimal solutions are asymptotically reached.

Fig. 4a shows  $\rho_E^f$  as a function of annealing time for QA, SQA, and SA, for 3D spin glasses on  $N = 5374$  spins

( $15 \times 15 \times 12$  dimers, with some vacancies). The ground-state energy  $E_0$  is estimated by exchange MC (see SM II C). Each dynamics follows a power-law scaling within a window of  $t_a$  (SM V) but deviates outside the window, most notably for QA due to the onset of thermal effects. For experimental QA, deviation from coherent behavior is expected for longer anneals due to the onset of thermal excitations<sup>32</sup>. For SA, the decay of  $\rho_E^f$  settles onto a consistent exponent only after several hundred MCS.

We estimate  $\kappa_f$  for the three dynamics with varying  $J_G$  from power-law fits. Figure 4b shows, as a function of  $J_G$ , the fit values of  $\kappa_f$  (symbols) as well as the critical values  $\kappa_c$  (horizontal lines) obtained using independent MC estimates of  $z$  and  $1/\nu$ , corresponding to the lines in Fig. 3e. Deviations  $\kappa_f < \kappa_c$  are expected beyond the critical point (SM IC 4); we find a modest correction  $\kappa_c - \kappa_f \approx 0.1$  for both SA and SQA (SM IV B), which one might also see in QA if it could be measured at the critical point.

QA shows the fastest energy decay  $\kappa_f$ , followed by SQA and SA. For sufficiently large  $J_G$ ,  $\rho_E^f$  decays roughly quadratically faster in QA than in SA, with SQA in between the two. This experimentally-observed scaling advantage is consistent with the theoretical speedup in critical ordering dynamics (smaller  $\mu$  in Fig. 3e), and faster critical energy decay (larger  $\kappa_c$  in Fig. 4b). In Fig. 4c we show the annealing time (in MCS) required by SA and SQA to match the energy achieved by a given  $t_a$  in QA; within the coherent QA regime, the approximation speedup of QA over SA and SQA increases as a function of annealing time.

## OUTLOOK

We have experimentally demonstrated quantum critical dynamics in programmable spin glasses on thousands of qubits, observing the expected scaling in system size

and annealing time. Simulation accuracy was confirmed via comparison to numerical simulation of Schrödinger dynamics at the 16-qubit scale. For large 3D spin glasses, the simulated many-body quantum dynamics are far beyond the reach of current exhaustive or tensor-based techniques; the former is limited to tens of qubits and the latter can be applied to moderately-sized 2D models<sup>30,54</sup>. We therefore appeal to critical exponents via finite-size scaling analysis, finding good agreement with independent MC estimates. Thus we have presented both microscopic and macroscopic evidence for a coherently annealed programmable quantum spin glass. These exponents indicate, in theory and experiment, that quantum annealing has a dynamical advantage over simulated annealing and simulated quantum annealing in penetrating the spin-glass phase. The predicted speedup was experimentally demonstrated through a scaling advantage in approach to the ground-state energy. These results point to the utility of programmable quantum annealers both as quantum simulators and optimization tools.

For sufficiently large spin systems, the extent of ideal quantum critical scaling is limited in time by qubit decoherence, disorder and noise, and the results shown here indicate that improvements in these areas would pay great dividends. Extending the region of critical scaling would not only facilitate the further study of these dynamics, but also extend their utility in real-world applications, helping QA reach lower-energy solutions. These efforts must be balanced with improvements in qubit connectivity, which allow more flexible problem embeddings, and high coupling energy, which can protect against control error and thermal excitation outside the coherent limit.

Spin glasses represent a paradigmatic hard optimization problem, and provide a robust theoretical framework for understanding and demonstrating quantum critical dynamics. They were instrumental in motivating, via magnetic experiments, the field of quantum annealing itself—here we have answered in the affirmative the foundational question raised over 20 years ago: Is it possible

to engineer a programmable quantum system, in which quantum annealing imparts a dynamical advantage over simulated annealing? Extending this characterization of quantum dynamics to industry-relevant optimization problems, which generally do not allow for analysis via universal critical exponents or finite-size scaling, would mark an important next step in practical quantum computing.

## ACKNOWLEDGMENTS

The authors are grateful to Peter Young, Hideyoshi Nishimori, Sei Suzuki, and Jonas Charfreitag for helpful discussions. A.W.S. was supported by the Simons Foundation under Simons Investigator Award No. 511064. Some of the numerical simulations were carried out using the Shared Computing Cluster managed by Boston University’s Research Computing Services.

## AUTHOR CONTRIBUTIONS

A.D.K., J.R., T.L., R.H., A.Z., A.W.S., and M.H.S. conceived and designed the experiments and analyzed the data. A.D.K., J.R., T.L., and A.W.S. performed the experiments and simulations. T.L., R.H., F.A., A.J.B., K.B., S.E., C.E., E.H., S.H., E.L., A.J.R.M., G.M., R.M., T.O., G.P.-L., M.R., C.R., Y.S., N.T., M.V., J.D.W., J.Y., and M.H.A. contributed to the design, fabrication, deployment, and calibration of the quantum annealing system. A.D.K., J.R., R.H., A.W.S., and M.H.A. wrote the manuscript.

## COMPETING INTERESTS

A.W.S. declares no competing interests. All other authors have received stock options in D-Wave as current or former employees, and declare a competing financial interest on that basis.

- 
- [1] J. Brooke, D. Bitko, T. F., Rosenbaum, and G. Aeppli, Quantum Annealing of a Disordered Magnet, *Science* **284**, 779 (1999).
  - [2] G. Aeppli and T. F. Rosenbaum, Experiments on Quantum Annealing, in *Quantum Annealing and Other Optimization Methods*, edited by A. Das and B. K. Chakrabarti (Springer Berlin Heidelberg, 2005) Chap. 6, pp. 157–169.
  - [3] A. Das and B. K. Chakrabarti, Colloquium: Quantum annealing and analog quantum computation, *Reviews of Modern Physics* **80**, 1061 (2008).
  - [4] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Physical Review E* **58**, 5355 (1998).
  - [5] G. E. Santoro, R. Martonák, E. Tosatti, and R. Car, Theory of Quantum Annealing of an Ising Spin Glass, *Science* **295**, 2427 (2002).
  - [6] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, *et al.*, Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor, *Physical Review B* **82**, 1 (2010).
  - [7] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, *et al.*, Defining and detecting quantum speedup, *Science* **345**, 420 (2014).
  - [8] H. G. Katzgraber, F. Hamze, and R. S. Andrist, Glassy Chimeras Could Be Blind to Quantum Speedup: Designing Better Benchmarks for Quantum Annealing Machines, *Physical Review X* **4**, 021008 (2014).

- [9] I. Hen, J. Job, T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, Probing for quantum speedup in spin-glass problems with planted solutions, *Physical Review A* **92**, 042325 (2015).
- [10] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, Quantum versus classical annealing of Ising spin glasses, *Science* **348**, 215 (2015).
- [11] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, Computational multiqubit tunnelling in programmable quantum annealers, *Nature Communications* **7**, 10327 (2016).
- [12] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. N. Smelyanskiy, J. M. Martinis, and H. Neven, What is the Computational Value of Finite-Range Tunneleing?, *Physical Review X* **6**, 031015 (2016).
- [13] T. Albash and D. A. Lidar, Demonstration of a Scaling Advantage for a Quantum Annealer over Simulated Annealing, *Physical Review X* **8**, 031016 (2018).
- [14] M. Mezard and A. Montanari, *Information, physics, and computation* (Oxford University Press, 2009).
- [15] D. L. Stein and C. M. Newman, *Spin Glasses and Complexity* (Princeton University Press, 2013).
- [16] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, Optimization by Simulated Annealing, *Science* **220**, 671 (1983).
- [17] C. M. Tan, ed., *Simulated Annealing* (InTech, Vienna, 2008).
- [18] T. Albash and D. A. Lidar, Adiabatic quantum computation, *Reviews of Modern Physics* **90**, 015002 (2018).
- [19] A. Das and B. K. Chakrabarti, eds., *Quantum Annealing and Other Optimization Methods*, Lecture Notes in Physics, Vol. 679 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005).
- [20] R. P. Feynman, Simulating physics with computers, *International Journal of Theoretical Physics* **21**, 467 (1982).
- [21] M. W. Johnson, M. H. Amin, S. Gildert, T. Lanting, F. Hamze, et al., Quantum annealing with manufactured spins, *Nature* **473**, 194 (2011).
- [22] W. Lechner, P. Hauke, and P. Zoller, A quantum annealing architecture with all-to-all connectivity from local interactions, *Science Advances* **1**, 1 (2015).
- [23] S. J. Weber, G. O. Samach, D. Hover, S. Gustavsson, D. K. Kim, A. Melville, D. Rosenberg, A. P. Sears, F. Yan, J. L. Yoder, W. D. Oliver, and A. J. Kerman, Coherent Coupled Qubits for Quantum Annealing, *Physical Review Applied* **8**, 014004 (2017).
- [24] S. Novikov, R. Hinkey, S. Disseler, J. I. Basham, T. Albash, A. Risinger, D. Ferguson, D. A. Lidar, and K. M. Zick, Exploring More-Coherent Quantum Annealing, in *2018 IEEE International Conference on Rebooting Computing (ICRC)* (IEEE, 2018) pp. 1–7.
- [25] P. Hauke, H. G. Katzgraber, W. Lechner, H. Nishimori, and W. D. Oliver, Perspectives of quantum annealing: methods and implementations, *Reports on Progress in Physics* **83**, 054401 (2020).
- [26] R. Blatt and C. F. Roos, Quantum simulations with trapped ions, *Nature Physics* **8**, 277 (2012).
- [27] C. Monroe, W. C. Campbell, L.-M. Duan, Z.-X. Gong, A. V. Gorshkov, P. W. Hess, R. Islam, K. Kim, N. M. Linke, G. Pagano, P. Richerme, C. Senko, and N. Y. Yao, Programmable quantum simulations of spin systems with trapped ions, *Reviews of Modern Physics* **93**, 025001 (2021).
- [28] C. Gross and I. Bloch, Quantum simulations with ultracold atoms in optical lattices, *Science* **357**, 995 (2017).
- [29] P. Scholl, M. Schuler, H. J. Williams, A. A. Eberharter, D. Barredo, K.-N. Schymik, V. Lienhard, L.-P. Henry, T. C. Lang, T. Lahaye, A. M. Läuchli, and A. Browaeys, Programmable quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms, *Nature* **595**, 233 (2020).
- [30] S. Ebadi, A. Keesling, M. Cain, T. T. Wang, H. Levine, D. Bluvstein, G. Semeghini, A. Omran, J.-G. Liu, R. Samajdar, X.-Z. Luo, B. Nash, X. Gao, B. Barak, E. Farhi, S. Sachdev, N. Gemelke, L. Zhou, S. Choi, H. Pichler, S.-T. Wang, M. Greiner, V. Vuletic, and M. D. Lukin, Quantum optimization of maximum independent set using Rydberg atom arrays, *Science* **0**, eab06587 (2022).
- [31] R. Harris, Y. Sato, A. J. Berkley, M. Reis, F. Altomare, et al., Phase transitions in a programmable quantum spin glass simulator, *Science* **361**, 162 (2018).
- [32] A. D. King, S. Suzuki, J. Raymond, A. Zucca, T. Lanting, et al., Coherent quantum annealing in a programmable 2,000 qubit Ising chain, *Nature Physics* **18**, 1324 (2022).
- [33] M. Suzuki, Relationship between d-Dimensional Quantal Spin Systems and (d+1)-Dimensional Ising Systems: Equivalence, Critical Exponents and Systematic Approximants of the Partition Function and Spin Correlations, *Progress of Theoretical Physics* **56**, 1454 (1976).
- [34] S. V. Isakov, G. Mazzola, V. N. Smelyanskiy, Z. Jiang, S. Boixo, H. Neven, and M. Troyer, Understanding Quantum Tunneling through Quantum Monte Carlo Simulations, *Physical Review Letters* **117**, 180402 (2016).
- [35] A. D. King, J. Carrasquilla, J. Raymond, I. Ozfidan, E. Andriyash, et al., Observation of topological phenomena in a programmable lattice of 1,800 qubits, *Nature* **560**, 456 (2018).
- [36] K. Nishimura, H. Nishimori, and H. G. Katzgraber, Griffiths-McCoy singularity on the diluted Chimera graph: Monte Carlo simulations and experiments on quantum hardware, *Physical Review A* **102**, 042403 (2020).
- [37] P. Weinberg, M. Tylutki, J. M. Rönnkö, J. Westerholm, J. A. Åström, P. Manninen, P. Törmä, and A. W. Sandvik, Scaling and Diabatic Effects in Quantum Annealing with a D-Wave Device, *Physical Review Letters* **124**, 090502 (2020).
- [38] S. Zhou, D. Green, E. D. Dahl, and C. Chamon, Experimental realization of classical  $\mathbb{Z}_2$  spin liquids in a programmable quantum device, *Physical Review B* **104**, L081107 (2021).
- [39] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem, *Science* **292**, 472 (2001).
- [40] T. W. B. Kibble, Topology of cosmic domains and strings, *Journal of Physics A: Mathematical and General* **9**, 1387 (1976).
- [41] W. H. Zurek, Cosmological experiments in superfluid helium?, *Nature* **317**, 505 (1985).
- [42] A. Polkovnikov, Universal adiabatic dynamics in the vicinity of a quantum critical point, *Physical Review B* **72**, 161201 (2005).
- [43] J. Dziarmaga, Dynamics of a quantum phase transition: Exact solution of the quantum Ising model, *Physical Review Letters* **95**, 1 (2005).

- [44] W. H. Zurek, U. Dorner, and P. Zoller, Dynamics of a quantum phase transition, *Physical Review Letters* **95**, 2 (2005).
- [45] S. Deng, G. Ortiz, and L. Viola, Dynamical non-ergodic scaling in continuous finite-order quantum phase transitions, *Europhysics Letters* **84**, 67008 (2008).
- [46] C. De Grandi, A. Polkovnikov, and A. W. Sandvik, Universal nonequilibrium quantum dynamics in imaginary time, *Physical Review B* **84**, 224303 (2011).
- [47] A. Chandran, A. Erez, S. S. Gubser, and S. L. Sondhi, Kibble-Zurek problem: Universality and the scaling limit, *Physical Review B* **86**, 064304 (2012).
- [48] C.-W. Liu, A. Polkovnikov, A. W. Sandvik, and A. P. Young, Universal dynamic scaling in three-dimensional Ising spin glasses, *Physical Review E* **92**, 022128 (2015).
- [49] M. Guo, R. N. Bhatt, and D. A. Huse, Quantum critical behavior of a three-dimensional Ising spin glass in a transverse magnetic field, *Physical Review Letters* **72**, 4137 (1994).
- [50] A. K. Hartmann, Ground-state behavior of the three-dimensional  $\pm J$  random-bond Ising model, *Physical Review B* **59**, 3617 (1999).
- [51] M. Hasenbusch, F. P. Toldin, A. Pelissetto, and E. Vicari, Critical behavior of the three-dimensional  $\pm J$  Ising model at the paramagnetic-ferromagnetic transition line, *Physical Review B* **76**, 1 (2007).
- [52] M. Hasenbusch, F. P. Toldin, A. Pelissetto, and E. Vicari, Magnetic-glassy multicritical behavior of the three-dimensional  $\pm J$  Ising model, *Physical Review B* **76**, 184202 (2007).
- [53] H. Nishimori, Boundary between the Ferromagnetic and Spin Glass Phases, *Journal of the Physical Society of Japan* **61**, 1011 (1992).
- [54] M. Schmitt, M. M. Rams, J. Dziarmaga, M. Heyl, and W. H. Zurek, Quantum phase transition dynamics in the two-dimensional transverse-field Ising model, *Science Advances* **8**, 1 (2022).
- [55] R. Harris, J. Johansson, A. J. Berkley, M. W. Johnson, T. Lanting, S. Han, P. Bunyk, E. Ladizinsky, T. Oh, I. Perminov, E. Tolkacheva, S. Uchaikin, E. M. Chapple, C. Enderud, C. Rich, M. Thom, J. Wang, B. Wilson, and G. Rose, Experimental demonstration of a robust and scalable flux qubit, *Physical Review B* **81**, 1 (2010).
- [56] M. N. Barber, Finite-Size Scaling, in *Phase Transitions and Critical Phenomena*, Vol. 8, edited by C. Domb and J. L. Lebowitz (Academic Press, 1983).
- [57] V. Privman, Finite-Size Scaling Theory, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, 1990) Chap. 1, pp. 1–98.
- [58] D. Rossini and E. Vicari, Coherent and dissipative dynamics at quantum phase transitions, *Physics Reports* **936**, 1 (2021).
- [59] M. E. Fisher and M. N. Barber, Scaling Theory for Finite-Size Effects in the Critical Region, *Physical Review Letters* **28**, 1516 (1972).
- [60] J. Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, 1996).
- [61] A. Aharony and M. E. Fisher, Nonlinear scaling fields and corrections to scaling near criticality, *Physical Review B* **27**, 4394 (1983).
- [62] S. Jin, A. Sen, W. Guo, and A. W. Sandvik, Phase transitions in the frustrated Ising model on the square lattice, *Physical Review B* **87**, 144406 (2013).
- [63] M. E. Fisher, Renormalization group theory: Its basis and formulation in statistical physics, *Reviews of Modern Physics* **70**, 653 (1998).
- [64] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, 2011).
- [65] H. Rieger and A. P. Young, Zero-temperature quantum phase transition of a two-dimensional Ising spin glass, *Physical Review Letters* **72**, 4141 (1994).
- [66] W. Zurek, Cosmological experiments in condensed matter systems, *Physics Reports* **276**, 177 (1996).
- [67] C. De Grandi, V. Gritsev, and A. Polkovnikov, Quench dynamics near a quantum critical point, *Physical Review B* **81**, 012303 (2010).
- [68] M. Kolodrubetz, B. K. Clark, and D. A. Huse, Nonequilibrium Dynamic Critical Scaling of the Quantum Ising Chain, *Physical Review Letters* **109**, 015701 (2012).
- [69] C.-W. Liu, A. Polkovnikov, and A. W. Sandvik, Dynamic scaling at classical phase transitions approached through nonequilibrium quenching, *Physical Review B* **89**, 054307 (2014).
- [70] C. W. Liu, A. Polkovnikov, and A. W. Sandvik, Quantum versus classical annealing: Insights from scaling theory and results for spin glasses on 3-regular graphs, *Physical Review Letters* **114**, 1 (2015).
- [71] S. J. Rubin, N. Xu, and A. W. Sandvik, Dual time scales in simulated annealing of a two-dimensional Ising spin glass, *Physical Review E* **95**, 052133 (2017), 1609.09024.
- [72] N. Xu, K.-H. Wu, S. J. Rubin, Y.-J. Kao, and A. W. Sandvik, Dynamic scaling in the two-dimensional Ising spin glass with normal-distributed couplings, *Physical Review E* **96**, 052102 (2017).
- [73] H. Shao, W. Guo, and A. W. Sandvik, Monte Carlo Renormalization Flows in the Space of Relevant and Irrelevant Operators: Application to Three-Dimensional Clock Models, *Physical Review Letters* **124**, 80602 (2020).
- [74] T. Blanchard, L. F. Cugliandolo, M. Picco, and A. Tartaglia, Critical percolation in the dynamics of the 2D ferromagnetic Ising model, *Journal of Statistical Mechanics: Theory and Experiment* **2017**, 113201 (2017).
- [75] R. Miyazaki and H. Nishimori, Real-space renormalization-group approach to the random transverse-field Ising model in finite dimensions, *Physical Review E* **87**, 1 (2013).
- [76] D. A. Matoz-Fernandez and F. Romá, Unconventional critical activated scaling of two-dimensional quantum spin glasses, *Physical Review B* **94**, 024201 (2016).
- [77] C. Pich, A. P. Young, H. Rieger, and N. Kawashima, Critical behavior and griffiths-mccoy singularities in the two-dimensional random quantum ising ferromagnet, *Physical Review Letters* **81**, 5916 (1998).
- [78] R. R. Singh and A. P. Young, Critical and Griffiths-McCoy singularities in quantum Ising spin glasses on d-dimensional hypercubic lattices: A series expansion study, *Physical Review E* **96**, 10.1103/PhysRevE.96.022139 (2017).
- [79] E. Marinari, G. Parisi, and J. J. Ruiz-Lorenzo, Numerical simulations of spin glass systems, in *Spin Glasses and Random Fields*, edited by A. P. Young (World Scientific, 1997) pp. 59–98.
- [80] K. Boothby, P. Bunyk, J. Raymond, and A. Roy, Next-Generation Topology of D-Wave Quantum Processors (2020), arXiv:2003.00133.

- [81] J. Charfreitag, M. Jünger, S. Mallach, and P. Mutzel, McSparse: Exact solutions of sparse maximum cut and sparse unconstrained binary quadratic optimization problems, in *2022 Proceedings of the Symposium on Algorithm Engineering and Experiments (ALENEX)*, edited by C. A. Phillips and B. Speckmann (2022) pp. 54–66.
- [82] Z. Zhu, A. J. Ochoa, and H. G. Katzgraber, Efficient cluster algorithm for spin glasses in any space dimension, *Physical Review Letters* **115**, 077201 (2015).
- [83] M. Baity-Jesi, R. A. Baños, A. Cruz, L. A. Fernandez, J. M. Gil-Narvion, *et al.*, Critical parameters of the three-dimensional Ising spin glass, *Physical Review B* **88**, 1 (2013).
- [84] K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, *Reviews of Modern Physics* **58**, 801 (1986).
- [85] N. Kawashima and A. Young, Phase transition in the three-dimensional Ising spin glass, *Physical Review B* **53**, 484 (1996).
- [86] M. Newman and G. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon Press, 1999).
- [87] H. Rieger and N. Kawashima, Application of a continuous time cluster algorithm to the two-dimensional random quantum Ising ferromagnet, *European Physical Journal B* **9**, 233 (1999).
- [88] A. D. King, J. Raymond, T. Lanting, S. V. Isakov, M. Mohseni, *et al.*, Scaling advantage over path-integral Monte Carlo in quantum simulation of geometrically frustrated magnets, *Nature Communications* **12**, 1113 (2021), code: <https://github.com/dwavesystems/dwave-pimc>.
- [89] M. Palassini and S. Caracciolo, Universal Finite-Size Scaling Functions in the 3D Ising Spin Glass, *Physical Review Letters* **82**, 5128 (1999).
- [90] M. Hasenbusch, A. Pelissetto, and E. Vicari, Critical behavior of three-dimensional ising spin glass models, *Physical Review B* **78**, 214205 (2008).
- [91] O. Melchert, autoScale.py - A program for automatic finite-size scaling analyses: A user's guide (2009), [arXiv:0910.5403v1](https://arxiv.org/abs/0910.5403v1).
- [92] P. Kairys, A. D. King, I. Ozfidan, K. Boothby, J. Raymond, A. Banerjee, and T. S. Humble, Simulating the Shastry-Sutherland Ising Model Using Quantum Annealing, *PRX Quantum* **1**, 020320 (2020).
- [93] A. D. King, C. Nisoli, E. D. Dahl, G. Poulin-Lamarre, and A. Lopez-Bezanilla, Qubit spin ice, *Science* **373**, 576 (2021).
- [94] S. V. Isakov, I. Zintchenko, T. Rønnow, and M. Troyer, Optimised simulated annealing for Ising spin glasses, *Computer Physics Communications* **192**, 265 (2015).

## METHODS

### Spin-glass instances

The 3D lattices have open  $x$  and  $y$  dimensions and periodic  $z$  dimension. Instances up to size  $9 \times 9 \times 9 \times 2$  ( $L = 9$ ) are fully yielded, with no site vacancies. Larger instances have over 99.5% site yield. The inputs are heuristically embedded into the qubit connectivity graph of the QA processor, with a structure shown in Extended Data E2. Input construction is discussed in detail in the supplementary material.

### Quantum annealing system and methods

All QA data, except specifically indicated temperature variation experiments discussed in the supplementary material, were taken using one D-Wave Advantage system operating at 12 mK (QPU1). The variable-temperature experiments were performed using a second D-Wave Advantage system of the same design, between 12 mK and 21 mK (QPU2).

Calibration refinement methods were used to balance qubits at degeneracy and to synchronize annealing lines. We follow the same method as described in the supplementary materials for Ref.<sup>32</sup>, but with no tuning of coupling strengths. We describe these methods in detail in SM VII.

Each call to the QA system resulted in 200 anneals. QA data on 3D spin glasses are generated from 900 calls, cycling through 300 disorder realizations. For MC methods, between 100 and 300 realizations were used. Thus the data points for  $\langle q^2 \rangle$ ,  $U$ , and  $\rho_E$  represent the average over between 100 and 300 spin-glass realizations, with error bars capturing variation between instances. QA Binder cumulants were computed by comparing overlap between annealing samples generated in different QA calls but the same seed and embedding, thus suppressing the effect of correlated biases. The experiments were sufficiently extensive that the Binder cumulants and overlaps computed from samples within the same QA programming give similar results.

### Measuring annealing time

The anneal of the Hamiltonian from  $s = 0$  to  $s = 1$  over annealing time  $t_a$  is achieved through a rapid change in qubit control current. Of the factors limiting the minimum value of  $t_a$ , two are most important: First is the ability to reliably quench  $s$  at a known rate and with tolerable nonlinearity and distortion from filtering; second is the ability to synchronize the qubits to within a reasonable deviation in terms of  $s$ .

Annealing times slower than 20 ns are reliably realized. Annealing times faster than 20 ns deviate significantly

from their requested values due to lowpass filter bandwidth and resolution of digital control electronics, and must be measured independently. We do so in two ways. The recent demonstration of the KZ mechanism in a 1D Ising chain<sup>32</sup> showed very good agreement with theory, in particular a scaling of residual energy  $\rho_E^f \propto t_a^{-1/2}$ . Therefore extrapolating 1D chain data from  $t_a = 20$  ns to faster anneals provides a reliable measurement of effective annealing time. The second measurement we perform is a direct measurement in which one qubit is quenched and measured by a witness qubit, allowing us to estimate the effective annealing time. Data for these two measurements are compared in Extended Data E3. We take the average of the two measurements as our value of  $t_a$ . Error bars are obtained by adding, in quadrature, the difference between the two measurements and the deviation in annealing slope before and after software filtering (< 0.05 relative error).

### Quantum annealing schedule

Simulating the time-dependent Schrödinger equation requires an accurate annealing schedule  $(\Gamma(s), \mathcal{J}(s))$ . To achieve this, we diagonalize a time-dependent many-body flux-qubit Hamiltonian for a small representative system of coupled qubits, using a flux-qubit model<sup>55</sup> whose parameters are given in Extended Data E4. Then, for each  $s$  in our range of interest, we compute  $\Gamma(s)$  and  $\mathcal{J}(s)$  as best-fit parameters that give an Ising model whose eigengaps closely match those of the diagonalized superconducting quantum interference device (SQUID) Hamiltonian.

Diagonalizing the many-body flux-qubit Hamiltonian to high accuracy is computationally challenging even at small scales because the flux qubits have multiple energy levels, unlike the model two-level Ising spins. Thus as a representative system we take eight qubits, and divide them up into four dimers such that the two qubits in each dimer have similar interactions with the other qubits. We then treat each dimer as a six-level object, and diagonalize the system of four dimers.

Reflecting the frustration in spin glasses, our eight-spin system, shown in Extended Data E5, has frustration in its twofold-degenerate classical ground state. Performing a two-parameter fit with  $\Gamma(s)$  and  $\mathcal{J}(s)$  becomes numerically unstable for  $s > 0.40$  due to the small first gap. For  $s > 0.40$  we fit  $\Gamma(s)$  to the expected exponential decay form and extract only  $\mathcal{J}(s)$  as a fitting parameter. Extended Data E5 shows the nominal and extracted schedules.

## Classical Monte Carlo dynamics

### *Simulated annealing*

SA uses a geometric annealing schedule from  $\beta = 0.001J_G$  to  $\beta = 10J_G$ . Each MCS processes  $N$  spins, sampled randomly with replacement. Spin updates are accepted or rejected in a Metropolis-Hastings algorithm.

### *Simulated quantum annealing*

SQA uses an annealing schedule based on the QA schedule; to accelerate our experiments we begin when  $\Gamma/\mathcal{J} \approx 6$ —a fast-equilibrating regime where we assume the system is quasistatic with respect to the changing Hamiltonian—and perform 10 MCS before proceeding through the QA schedule until  $\Gamma/\mathcal{J} \approx 1/25$ , far into the ordered phase for the models studied. Inverse temperature  $\beta$ , which is 64 except where specified, is relative to  $\Gamma(s) = \mathcal{J}(s)$  at the crossing point of  $\Gamma$  and  $\mathcal{J}$ . Swendsen-Wang cluster updates are used in imaginary time.

## Time-dependent Schrödinger evolution

For the data in Fig. 2 we used an iterative method that follows the annealing schedule from  $s = 0.1$ , where the QA wavefunction  $\psi$  is concentrated on the instantaneous ground state, and  $s = 0.7$ , where dynamics are negligible. We track populations for only the lowest 80 of  $2^{16}$  eigenstates; this exceeds the number of ground and first excited states for the classical models studied. Step sizes in  $s$  are determined adaptively based on the minimum eigengap, and range between 0.00008 and 0.01.

## Statistical methods and error analysis

Error bars on the order parameter  $\langle q^2 \rangle$ , Binder cumulant  $U$ , and residual energy  $\rho_E^f$  were generated by treating each random seed as an independent experiment and performing a resampling bootstrap on the set of statistics. This bootstrapping method gives a population for each statistic estimate. We use this population in two ways. First, we take the middle 95% of the population as our confidence interval for the statistic. Second, from the population we compute a variance on the logarithm, which we use to determine error weights for our data collapses.

## Data collapse

To collapse measurements of the Binder cumulant  $U$  for varying system sizes onto a common target curve, we need to find a best-fit value for  $\mu$ . This fit minimizes

a weighted sum of squared distances (in the logarithm) from the target curve. Weights in the sum are inversely proportional to the variance of the logarithm of the estimator. The form of the target curve must capture a crossover between the power-law form of the KZ regime and the equilibrium ( $t_a \rightarrow \infty$ ) limit. To achieve this, for each putative value of  $\mu$  we find a best-fit target curve (nested within the  $\mu$ -optimizing method) whose power-law slope varies as a logistic function. Our target curve has the form

$$f(x) = a_0 + a_1 \log(1 + \exp(a_2(x - a_3))), \quad (14)$$

which fits  $\log(U)$  as a function of  $\log(t_a)$ .

To generate error estimates for  $\mu$  we perform a jackknife on the measurements in  $L$  and  $t_a$  and add the resulting standard errors in quadrature. To approximate the 95% C.I. in the data points we use error bars that span  $\pm 2\sigma$ .

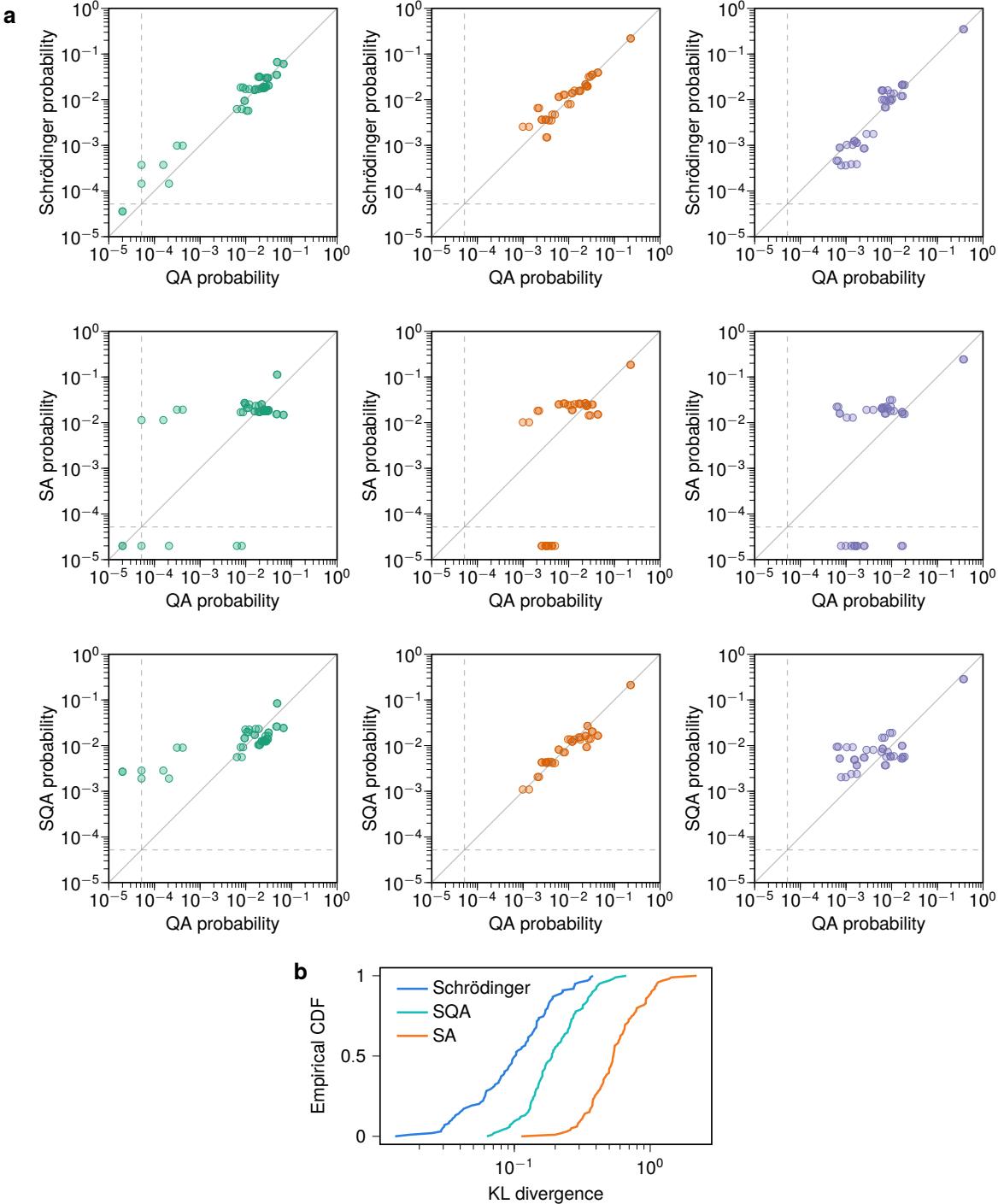
Collapse in  $\langle q^2 \rangle$  is achieved with the same approach. Selection of ranges of  $t_a$  over which we collapse data is described in the supplementary material.

## DATA AVAILABILITY

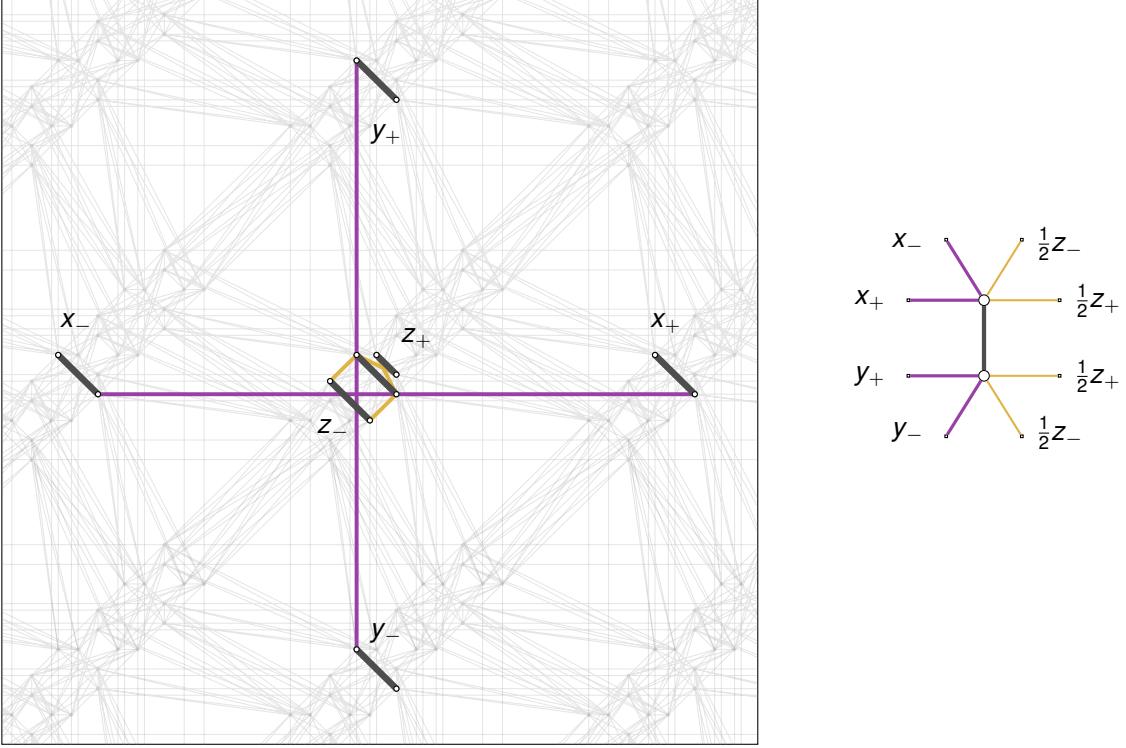
The data supporting the findings are being prepared for deposition in an online repository, and are currently available upon request.

## CODE AVAILABILITY

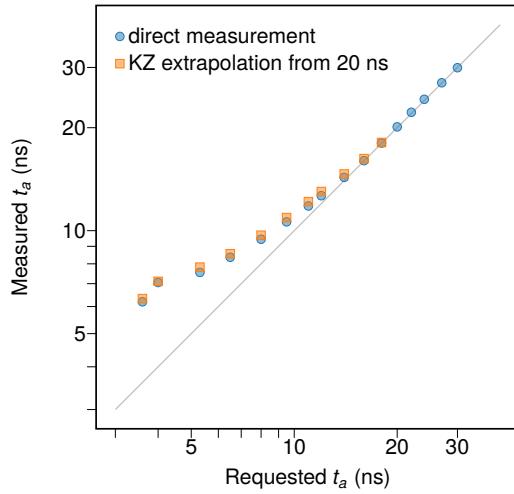
An open-source version of the SQA code used in this work is available at <https://github.com/dwavesystems/dwave-pimc>.



Extended Data E1. **State probabilities for 16-spin glasses.** **a**, We show observed probabilities for ground states and first excited states in Schrödinger evolution ( $t_a = 14$  ns), SA ( $t_a = 200$  MCS), and SQA ( $t_a = 100$  MCS) compared with experimental measurements from QA ( $t_a = 14$  ns). The three columns contain data for the three exemplary instances, with colors corresponding to those in Fig. 2. Annealing times for SA and SQA were chosen to have good agreement with Schrödinger evolution in average ground state probability. Each dynamics was run 19,200 times; dashed lines indicate the statistical floor, i.e., the probability if a state is seen exactly once. Unobserved states are represented by points below the statistical floor. **b**, Kullback-Leibler (KL) divergence in the probability distribution among first excited states, with QA used as a reference distribution, measures deviation between two dynamics for a given realization. Empirical CDF (proportion of 100 realizations below a given KL divergence) is shown. Data indicate that coherent quantum (Schrödinger) dynamics agrees more closely with experimental data better than does SA or SQA.



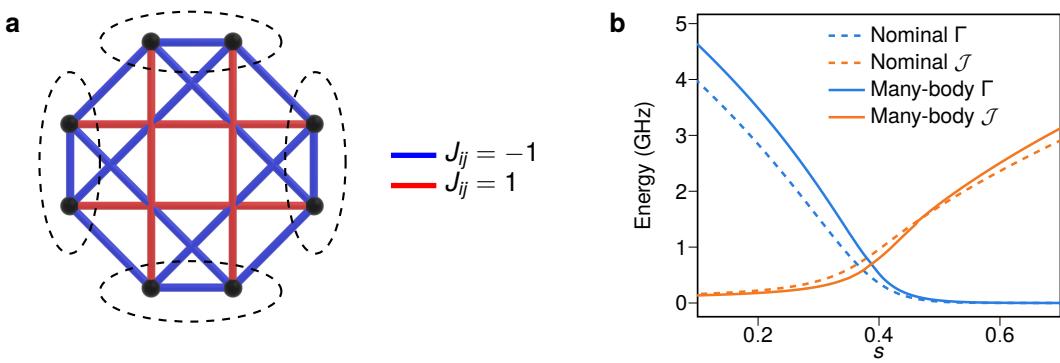
**Extended Data E2. 3D lattice structure in qubit connectivity graph.**  $L \times L \times (\max(L, 12)) \times 2$  lattices are found by heuristic search given a basic structure in which horizontal and vertical (long) couplings form two dimensions, and the interior of unit cells forms the third dimension. One dimer (thick gray line) and its six neighboring dimers are shown as thick gray lines. As in Fig. 1, purple and yellow lines represent glass couplings of  $\pm J_G$  and  $\pm J_G/2$  respectively.

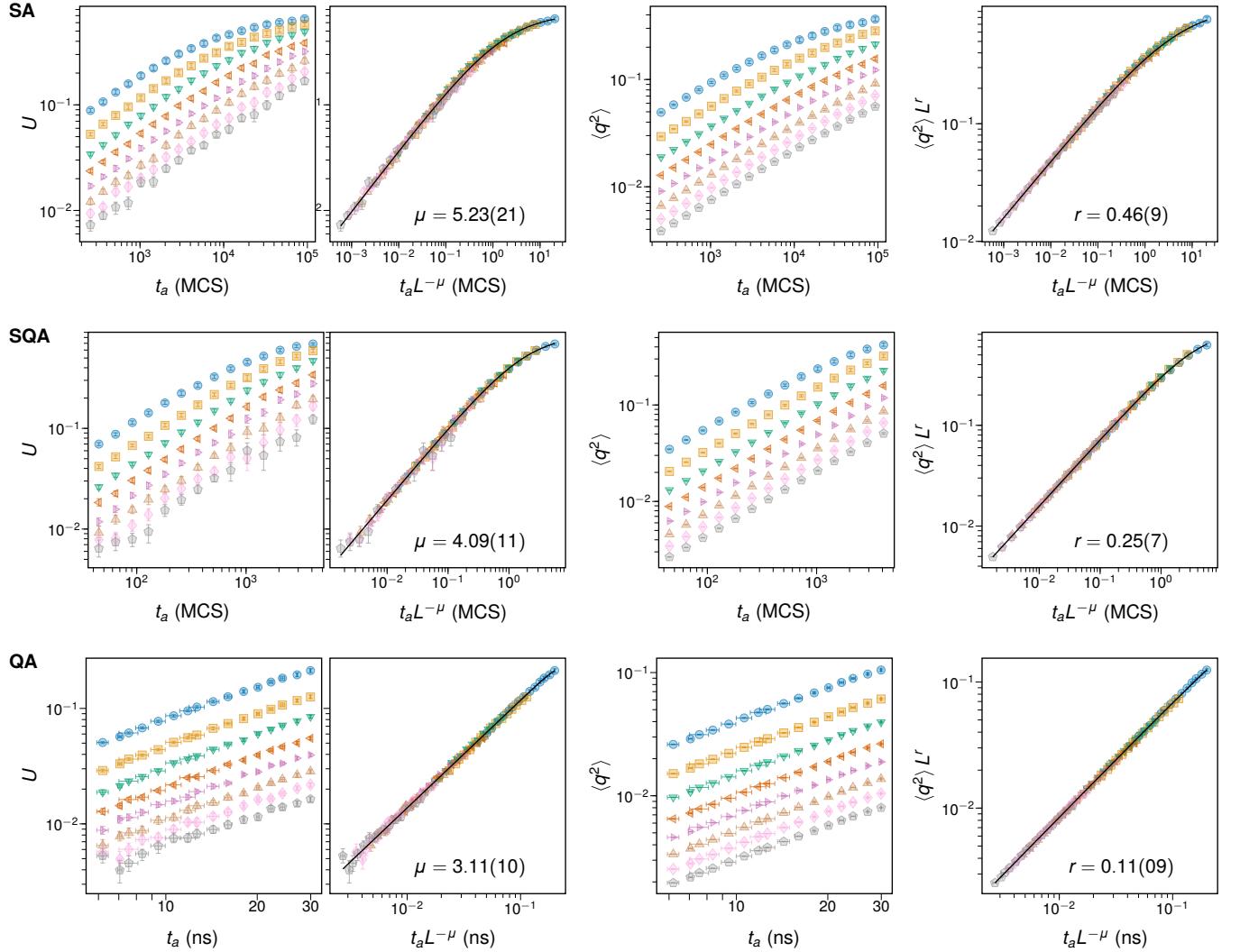


**Extended Data E3. Measurement of effective QA annealing time.** Two independent measurement methods are used to estimate  $t_a$  for fast anneals ( $< 20$  ns). First is a direct measurement using a witness qubit. Second is an extrapolative measurement that assumes a quantum KZ scaling in a 1D chain and assumes a kink density  $n \propto t_a^{-1/2}$  for  $t_a < 20$  ns. For the fastest anneals,  $t_a$  deviates significantly from the values requested from the control electronics. However, the two independent measurement approaches give consistent results.

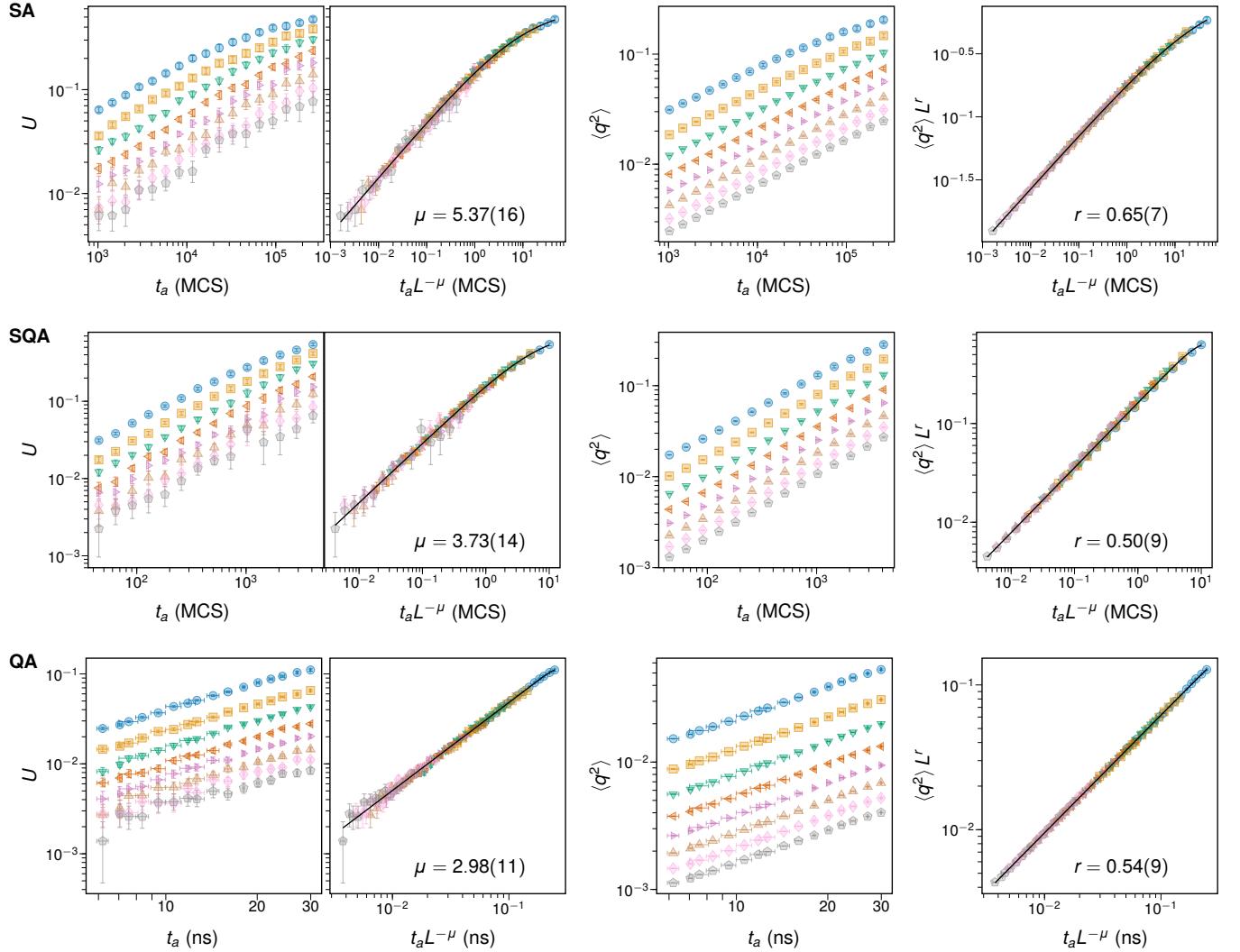
Parameter	Symbol	Value
Mutual inductance for $J = 1$	$M_{\text{AFM}}$	1.65 pH
Initial external applied flux	$\Phi_{\text{CCJJ}}^i$	$-0.621 \Phi_0$
Final external applied flux	$\Phi_{\text{CCJJ}}^f$	$-0.717 \Phi_0$
Qubit inductance	$L_q$	371 pH
Qubit capacitance	$C_q$	118 fF
CJJ loop capacitance	$C_l$	5 fF
CJJ junction capacitance	$C_{\text{CJJ}}$	25 fF
Qubit critical current	$I_c$	2.10 $\mu\text{A}$

Extended Data E4. Physical properties for the QA processor.

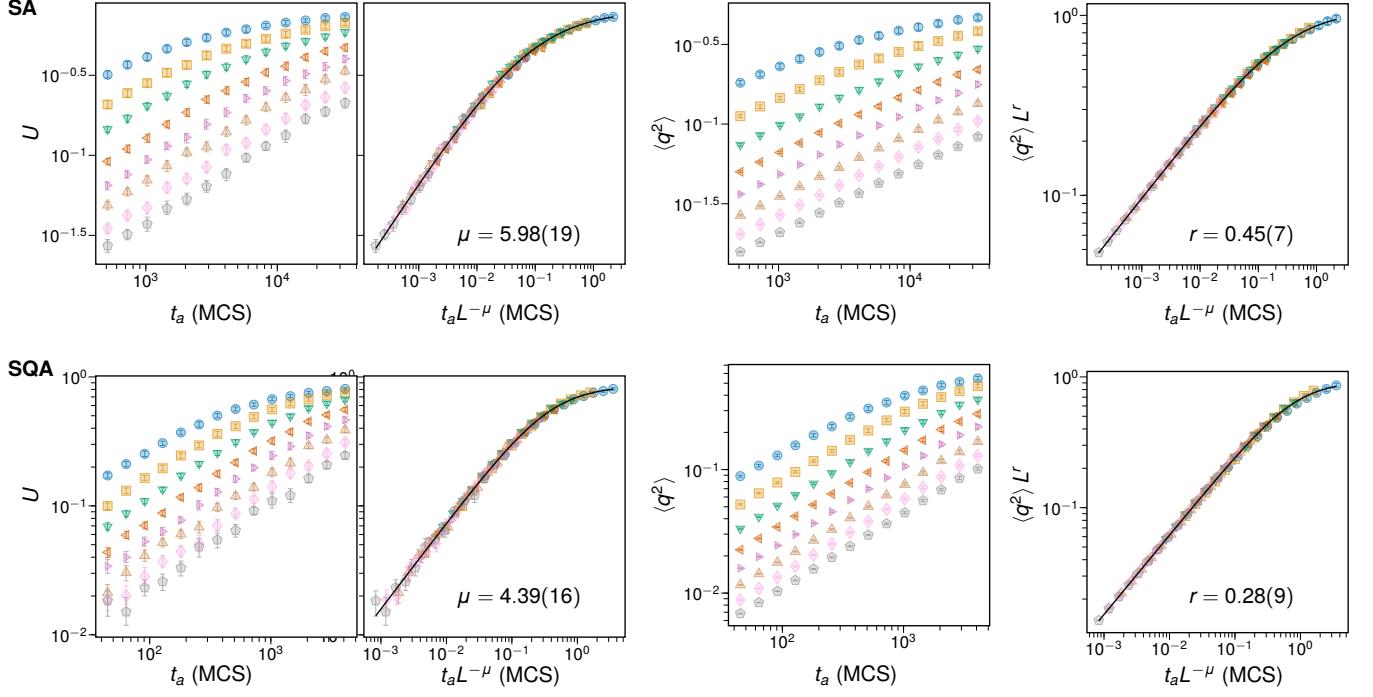
Extended Data E5. **Extracting Ising model from flux-qubit model.** **a**, Eight-qubit gadget used to extract an effective annealing schedule in the transverse-field Ising model based on a many-body flux-qubit Hamiltonian. Dimers indicated by dashed ellipses are treated as six-level objects and combined to diagonalize a many-body Hamiltonian. **b**, General-purpose (nominal) annealing schedule based on single-qubit measurements, and extracted many-body effective schedule, used for Schrödinger evolution.



Extended Data E6. **Data collapse for 3D spin glasses with  $J_G = 1$ .** Best-fit exponent  $\mu$  collapses  $U$  by rescaling data horizontally based on  $L$ , and  $r$  collapses  $\langle q^2 \rangle$  by scaling vertically based on  $L$ , given a horizontal rescaling by  $L^{-\mu}$ .



Extended Data E7. **Data collapse for 3D spin glasses with  $J_G = 1/2$ .** Best-fit exponent  $\mu$  collapses  $U$  by rescaling data horizontally based on  $L$ , and  $r$  collapses  $\langle q^2 \rangle$  by scaling vertically based on  $L$ , given a horizontal rescaling by  $L^{-\mu}$ .



Extended Data E8. **Data collapse for 3D spin glasses on simple cubic lattices.** Best-fit exponent  $\mu$  collapses  $U$  by rescaling data horizontally based on  $L$ , and  $r$  collapses  $\langle q^2 \rangle$  by scaling vertically based on  $L$ , given a horizontal rescaling by  $L^{-\mu}$ .

**Supplementary Materials:  
Quantum critical dynamics in a 5000-qubit programmable spin glass**

**CONTENTS**

I.	Finite-size scaling	19
A.	Equilibrium finite-size scaling	19
1.	Classical phase transitions	19
2.	Quantum phase transitions	23
B.	Dynamic finite-size scaling	25
C.	Application to annealing experiments and simulations	27
1.	Spin glass order parameter	28
2.	Scaling of the spin-glass order parameter	29
3.	The Binder ratio	30
4.	Residual Ising energy	31
II.	3D spin glasses	32
A.	Construction	32
B.	Broken FM dimers in the 3D ground state	33
C.	Estimating ground state energies	33
D.	Quantum and classical phase transitions	34
III.	Equilibrium analysis of critical phenomena by Monte Carlo methods	34
A.	Finite-size scaling in classical and path-integral models	35
B.	Monte Carlo sampling	35
C.	Classical critical scaling	36
D.	Quantum critical scaling	36
IV.	Nonequilibrium dynamics of SA and SQA	39
A.	Freezing of dynamics in SA	39
B.	Energy decay at the critical point and inside the glass phase	39
V.	Data collapse domains	40
VI.	Effect of temperature on QA	40
VII.	Calibration refinement	40
VIII.	Residual energy decay	42

## I. FINITE-SIZE SCALING

Here we provide the required background on the dynamic finite-size scaling (FSS) ansatz used in this work. The main goal is to justify the extraction of the KZ exponent  $\mu$  by collapsing, as in Fig. 2, the Binder cumulant  $U$  for multiple annealing times and system sizes, as well as the exponent  $\kappa_f$  of the power law quantifying the reduction in excess energy with increasing annealing time in Fig. 4. Since the annealing process stops inside the glassy phase not only for QA, but also in this work for SA and SQA, our analysis here must go beyond the simplest case of dynamic scaling at a critical point. We will show that many aspects of KZ scaling remain valid under the circumstances prevailing in the annealing experiments conducted here, including the power-law forms governed by the exponent  $\mu$ , even though the order parameter evolves beyond its critical form.

There are many excellent reviews of conventional critical phenomena and finite-size scaling, e.g., Refs.<sup>56,57</sup>, but for ease of reference in the later sections we begin in Sec. IA by a concise summary. In particular, we explain the relationships between the conventional critical exponents and the scaling dimensions appearing within the renormalization group (RG) framework. The classical case of transitions driven by thermal fluctuations at critical temperature  $T_c > 0$  is considered first, followed by the formally simple generalization to quantum phase transitions, where the ground state ( $T_c = 0$ ) of a system changes as a function of a model parameter regulating the quantum fluctuations. In Sec. IB we review the extensions of the finite-size scaling formalism to the case of a system brought through a classical or quantum phase transition by an annealing process, where the annealing rate regulates the correlation length by the KZ mechanism and there is a generalized FSS ansatz involving both the system size and the annealing rate. There are also extensive reviews of this topic, e.g., the very recent Ref.<sup>58</sup>; here we outline the formalism underlying our analysis of simulations and experimental data. In Sec. IC, we focus on the particular conditions pertaining to the quantum annealing device and how some aspects of critical scaling can persist also when driving the system past the quantum phase transition into an ordered (here glassy) state.

### A. Equilibrium finite-size scaling

#### 1. Classical phase transitions

Consider a system described by some Hamiltonian  $H$  at a short distance  $\delta = T - T_c$  away from a classical continuous phase transition with critical temperature  $T_c$ . There is a characteristic length scale in the infinite system, the correlation length  $\xi$  governed by the exponent  $\nu > 0$ ,

$$\xi \propto \delta^{-\nu}, \quad (\text{S1})$$

where the divergence takes place for  $\delta \rightarrow 0$  both from above and below (and for  $\delta < 0$  the absolute value  $|\delta|$  is implied here and henceforth in similar power laws). Also consider some other quantity  $n$  (an expectation value of, typically, a volume-normalized density) that in the thermodynamic limit is governed by another critical exponent  $\sigma$ ,

$$n \propto \delta^\sigma, \quad (\text{S2})$$

where  $\sigma$  depends on the quantity  $n$  and the universality class of the transition. For example, if  $n$  is the order parameter,  $n = \langle m \rangle$ , then in the conventional nomenclature  $\sigma \equiv \beta$ . The expectation value  $\langle \cdot \rangle$  involves also averaging over realizations in a system with some type of intrinsic disorder, like the spin glasses of interest here.

On a finite lattice with  $d$  spatial dimensions and volume  $N = L^d$ , the singular behavior breaks down when the correlation length becomes of the order of  $L$ . According to FSS theory<sup>56,57,59</sup>, for finite  $L$  the critical form (S2) should be replaced by

$$n = L^{-\sigma/\nu} g(\delta L^{1/\nu}), \quad (\text{S3})$$

where the scaling function  $g(x)$  has the property  $g(x) \rightarrow \text{constant}$  when  $x \rightarrow 0$ . In order for the form (S2) in the thermodynamic limit to be recovered when  $L \rightarrow \infty$  (at fixed small  $\delta$ ), the large- $x$  limit of  $g(x)$  must be of the form

$$g(x) \rightarrow x^\sigma h(x), \quad (\text{S4})$$

where  $h(x) \rightarrow \text{constant}$  when  $x \rightarrow \infty$ . Then there is no longer any  $L$  dependence of Eq. (S3) and the correct exponent on  $\delta$  in Eq. (S2) is obtained.

In the case  $n = \langle m \rangle$ , it should be noted that a finite system does not break any symmetry spontaneously; thus  $\langle m \rangle = 0$  for any finite  $L$  at any temperature. This apparent problem can be easily circumvented by considering

the squared order parameter  $\langle m^2 \rangle$ , which vanishes with increasing  $L$  for  $T \geq T_c$  and approaches the square of the infinite-size order parameter for  $T < T_c$ . The FSS form Eq. (S3) holds for  $\langle m^2 \rangle$  in the neighborhood of  $T_c$ , with  $\sigma = 2\beta$ .

The critical exponents can be related to the RG scaling dimensions appearing in field theory<sup>60</sup>. Consider an extensive operator  $P$  (a function of the system's degrees of freedom) defined as a sum over the entire system of local density operators  $p(\mathbf{r})$ :

$$P = \sum_{\mathbf{r}} p(\mathbf{r}). \quad (\text{S5})$$

If this operator is added with some factor  $\delta_p$  to the Hamiltonian at  $T = T_c$ ,  $H \rightarrow H + \delta_p P$ , the system would typically be driven away from the critical point and critical scaling forms such as those discussed above for  $\xi$  and  $n$  will apply with  $\delta \rightarrow \delta_p$ . The factor  $\delta_p$  is referred to as the field conjugate to  $P$ .

Classical statistical mechanics relies on the Boltzmann weight  $e^{-H/T}$ . At  $T = T_c$ , if the perturbing operator  $P$  is just the Hamiltonian itself,  $H \rightarrow H + \delta H$ , the offset  $\delta H$  is equivalent to just shifting the inverse temperature,  $1/T = (1/T_c)(1 + \delta)$  with  $H$  kept unchanged. Therefore, instead of shifting  $H$  we consider deviations from the critical temperature  $T_c$  and refer to  $\delta = T - T_c$  as the thermal field. Strictly speaking  $\delta$  is the conjugate field to the entropy  $S$ , which appears in the free energy  $F = E - TS$  along with the internal energy  $E = \langle H \rangle$ . By the reverse of the above arguments, we can also regard  $1/T - 1/T_c \approx -\delta/T_c^2$  as the conjugate field to  $H$ . Not worrying about factors, we can say that  $\delta$  is the field conjugate to both the entropy and the internal energy.

A field of strength  $\delta_p$  conjugate to an arbitrary operator  $P$  is associated with a scaling dimension  $y_p = 1/\nu_p$  governing finite-size forms such as Eq. (S3),

$$n = L^{-\sigma/\nu} g(\delta_p L^{y_p}), \quad (\text{S6})$$

when perturbing the critical system by an arbitrary interaction  $\delta_p P$ . In the above case of just changing the temperature, the thermal exponent  $\nu_t$  is conventionally just called  $\nu$  as in Eqs. (S1) and (S3), and for simplicity we also do not attach any subscript on  $\delta = t \equiv T - T_c$ . It is important to note that the symmetry of the system does not change when the temperature is changed, and  $\delta$  is therefore also referred to as a symmetric scaling field.

For generic perturbations  $P$ , the symmetry of  $H(\delta_p \neq 0)$  can be lower than that of  $H(0)$ , e.g., in the common case of  $P = M$ , where  $\langle m \rangle = \langle M \rangle/N$  is the order parameter density of a system hosting a symmetry-breaking transition. For instance, in the ferromagnetic Ising model,  $M$  is the total magnetization and  $\delta_m$  is the external magnetic field. Thus, the spin-inversion symmetry is violated when  $\delta_m \neq 0$ , and when  $\delta_m = 0$  the system spontaneously breaks the symmetry,  $\langle m \rangle \neq 0$ , for  $T < T_c$  in the thermodynamic limit.

It should be pointed out that a microscopic field strength  $\delta_p$  is never strictly speaking equivalent to the scaling field defined in a continuum theory. The term  $\delta_p P$  that we add (or imagine adding) to the Hamiltonian can be thought of as consisting of an infinite sum of operators  $O_j$  multiplied by their conjugate scaling fields  $\lambda_j$ ,

$$\delta_p P = \sum_j \lambda_j O_j, \quad (\text{S7})$$

and these fields enter scaling functions in the form  $\lambda_j L^{y_j}$  in the same way as in Eq. (S6). More precisely, the equality above should be interpreted as a correspondence between the lattice operator  $P$  and operators  $O_j$  in a continuum field theory. If  $y_j > 0$  the field  $\lambda_j$  (and operator  $O_j$ ) is said to be *relevant*, as its effect increases with the system size ( $\lambda_j L^{y_j} \rightarrow \infty$  when  $L \rightarrow \infty$ ), and, conversely,  $y_j < 0$  for an irrelevant field ( $\lambda_j L^{y_j} \rightarrow \infty$ , with the marginal case  $y_j = 0$  is associated with logarithmic scaling corrections, which we will not discuss here).

Naturally the fields  $\lambda_j$  in Eq. (S7) are proportional to  $\delta_p$  when the latter is small, but they also depend on other parameters of a given model and the full form is nonlinear<sup>61</sup>, which causes corrections to scaling when taken into account in FSS forms such as (S6). Here we neglect the scaling corrections that can (depending on the kind of analysis performed) be generated by the nonlinear scaling field, and, since constants of proportionality are not important, we can use  $\delta_p$  as the scaling parameter. The scaling dimension  $y_p$  in Eq. (S6) corresponds to the largest among the  $y_j$  values of the fields  $\lambda_j$  in (S7), and normally we just refer to this  $y_p$  as the scaling dimension of  $\delta_p$  even though, strictly speaking, this is only the leading scaling dimension.

We typically consider a transition tuned by the thermal field  $\delta$ , but scaling theory in principle relies on the possibility of arbitrary perturbations of the critical system. A key fact is that, in the entire space of possible infinitesimal deformations of  $H$ , there is only a small number of relevant fields (just  $\delta \equiv \delta_t$  and  $\delta_m$  in the most common case, with multi-critical points having additional relevant fields), whose infinitesimal presence destabilize the critical point. Adding an irrelevant field  $\delta_i$  just extends the critical point at  $(\delta = 0, \delta_m = 0)$  to a line of critical points in the parameter space, with the universality class not affected. In practice, most lattice operators would contain all the

scaling fields (conjugate to continuum field operators) allowed by symmetry, and adding them at weak strength to  $H$  only changes the location of the critical point, again without changing the universality class. A simple example is the Ising model with next-nearest-neighbor interactions  $J'$  added to the conventional model with nearest-neighbor interactions  $J$ , where  $T_c$  depends on  $J$  and  $J'$  but the universality class of the  $J$ -only model persists (unless  $J$  and  $J'$  are strongly frustrated, whence other phases and transitions are realized<sup>62</sup>).

A related important point is that it is normally only by symmetry that a given lattice operator  $P$  can be guaranteed to contain a single relevant field. In the Ising model, a change in temperature (conjugate to  $P = H$ ) at zero external field only corresponds to “touching” the thermal field. However, at the critical point of water (which is also of Ising type), changing the temperature (or the pressure) touches both the thermal field and the order-parameter field, because there is no microscopic symmetry like the spin-inversion symmetry of the Ising model (though such symmetry effectively emerges close to the critical point). Spin models in general have the advantage that the thermal field and the order parameter field can be separated thanks to symmetry.

In addition to the scaling fields  $\delta_p$ , the local operators  $p(\mathbf{r})$  in Eq. (S5) are also associated with their own scaling dimensions  $\Delta_p$ , which can be concretely defined in terms of the asymptotic distance dependence of the corresponding real-space correlation function at the critical point,

$$C_p(r) = \langle p(\mathbf{r})p(0) \rangle - \langle p \rangle^2 \propto \frac{1}{r^{2\Delta_p}}, \quad (\text{S8})$$

where, depending on the operator,  $\langle p \rangle$  (same as  $\langle p(\mathbf{r}) \rangle$  in a uniform system or in a system with randomness when averaged over realizations) may or may not vanish. More precisely, each operator in the sum (S7) has its own scaling dimension and  $\Delta_p$  is the smallest of these, i.e., (S8) represents the leading form of the correlations.

In the case of the order parameter,  $P = M$ ,  $\langle m \rangle = 0$  at  $T_c$  and the exponent  $2\Delta_m$  is related to the conventional critical exponents by

$$2\Delta_m = 2\beta/\nu = d - 2 + \eta, \quad (\text{S9})$$

where  $\eta$  is called the anomalous dimension. Normally this kind of relationship is not used for any other operator besides the order parameter, and no subscript is therefore attached to  $\beta$  and  $\eta$ .

The first equality in Eq. (S9) follows from writing the squared order parameter  $\langle m^2 \rangle = \langle M^2 \rangle/N^2$  as a sum (converted to an integral) of the order-parameter correlation function  $\langle m(\mathbf{r})m(0) \rangle$  over  $\mathbf{r}$  up to a cut-off length  $L$ , which gives  $\langle m^2 \rangle \propto L^{-2\Delta_m}$  from the form analogous to Eq. (S8). Then, comparing with Eqs. (S2) and (S3), where by definition  $\sigma = 2\beta$ , the equality follows. The second equality in Eq. (S9) can be regarded as the definition of  $\eta$ , but this exponent is also fundamentally related to the fractal dimensionality of critical domains in the system.

One can show, as we will below, that the scaling dimensions  $y_p$  and  $\Delta_p$  are related to each other by

$$y_p \equiv \frac{1}{\nu_p} = d - \Delta_p. \quad (\text{S10})$$

If  $\delta_p$  drives the transition (typically but not necessarily by changing the temperature, whence  $\delta_p = \delta$  and  $y_p = 1/\nu$ ), the FSS ansatz (S6) of an observable  $n$  in the neighborhood of the critical point can be expressed using the scaling dimension  $y_p$  of the driving field and the scaling dimension  $\Delta_n$  of the observable  $n$ :

$$n(L, \delta_p) = L^{-\Delta_n} g(\delta_p L^{y_p}). \quad (\text{S11})$$

Again, more precisely,  $\Delta_n$  is the smallest of the scaling dimensions of all operators in field theory that may be contained in the lattice operator  $n$ .

It should be noted that the correlation function (S8) from which a scaling dimension can be extracted may also include a phase depending on  $\mathbf{r}$ . As a simple example, in a square-lattice antiferromagnet the sign of the spin correlation function  $\langle S(\mathbf{r})S(0) \rangle$  is  $(-1)^{r_x+r_y}$ , where  $r_x$  and  $r_y$  are the (integer) lattice coordinates. For simplicity we do not consider any such phases here, but they can be easily handled by, e.g., considering absolute values. For the order parameter  $M$  defined using a sum of local operators as in Eq. (S5), a corresponding phase should then also be included, i.e., the Fourier transform should be taken at the relevant ordering wave-vector, which is  $(\pi, \pi)$  in the above case of the antiferromagnet with local operators  $S(\mathbf{r})$ . These phases do not in any other way affect the scaling formalism discussed here, and we will assume that they are taken into account if needed.

In the common case of a simple critical point with only two relevant perturbations, we can extend the FSS form Eq. (S11) of an observable to

$$n = L^{-\Delta_n} g(\delta L^{1/\nu}, \delta_m L^{y_m}, \delta_i L^{-|y_i|}, \dots), \quad (\text{S12})$$

where  $\delta_i$  is just the first of an infinite number of irrelevant scaling fields that are normally present,  $\delta_i \neq 0$ , and additional relevant fields are present at multi-critical points. An irrelevant field has scaling dimension  $y_i < 0$  and only produces finite-size scaling corrections in Eq. (S11), as seen when Taylor expanding in the argument  $\delta_i L^{y_i} = \delta_i L^{-|y_i|}$  when  $L$  is sufficiently large. Here we will henceforth neglect irrelevant fields and scaling corrections.

The most fundamental among the finite-size scaling forms in classical statistical mechanics is that of singular part of the free-energy density  $f$ , which has scaling dimension equal to the system's dimensionality,  $\Delta_f = d$  (for reasons that we do not further motivate here but refer to the literature<sup>59,60,63</sup>);

$$f = L^{-d} g(\delta L^{1/\nu}, \delta_m L^{y_m}, \dots). \quad (\text{S13})$$

Scaling forms for arbitrary physical observables can be derived from this free-energy density by taking appropriate derivatives with respect to added perturbations. As an example, the magnetization (the order parameter of a ferromagnet for simplicity) is the derivative of  $f$  with respect to the magnetic field  $h = \delta_m$ ,

$$\langle m \rangle = \frac{\partial f}{\partial \delta_m} = L^{-d+y_m} g'(\delta L^{1/\nu}, \delta_m L^{y_m}), \quad g'(x, y) = \frac{\partial g(x, y)}{\partial y}, \quad (\text{S14})$$

where we see that, as implied by Eq. (S10), indeed  $\Delta_m = d - y_m$ . Setting  $\delta = 0$  and taking  $L \rightarrow \infty$  we can argue, in analogy with Eq. (S4), that the function  $g'$  must become a power law;  $g' \rightarrow (\delta_m L^{y_m})^x$ , where the exponent must take the value  $x = \Delta_m/y_m = \Delta_m/(d - \Delta_m)$  so that the size dependence is eliminated. Thus,  $\langle m \rangle = \delta_m^x$ . In the conventional nomenclature  $x$  is called  $1/\delta$ , which conflicts with our notation for the thermal field. Therefore, continuing to use  $x$  for this exponent here, we have the well known relationship  $1/x = (d\nu - \beta)/\beta$ .

Another important case is when  $n$  in Eq. (S12) is the internal energy density  $E/N = \langle H \rangle/N$ . In this case, the infinite-size critical value (which by definition is its regular part) also has to be subtracted, and we define the singular part of the energy density as (now for simplicity considering the tuning only of the thermal field  $\delta$ )

$$\rho_E(L, \delta) = E(L, \delta)/L^d - \lim_{\ell \rightarrow \infty} E(\ell, 0)/\ell^d. \quad (\text{S15})$$

This residual energy density has the finite-size critical form  $\rho_E = L^{-\Delta_e} g(\delta L^{1/\nu})$ . Since, as discussed above, the energy is conjugate to the thermal field, we must have  $\Delta_e = d - 1/\nu$  by Eq. (S10). Therefore, in the thermodynamic limit  $\rho_E \propto \delta^{d\nu-1}$ , from which we can easily obtain the more often considered critical form of the specific heat:  $c = d\rho_E/d\delta \propto \delta^{d\nu-2}$ . The exponent governing the specific heat is customarily called  $\alpha$ , and we see that the well known exponent relation  $\alpha = d\nu - 2$  is obtained.

An energy-energy correlation function  $C_e(r)$  defined with local energy density operators  $e(\mathbf{r})$  can in principle be used to determine the correlation-length exponent  $\nu$  in numerical model simulations. At the critical point

$$C_e(r) \sim \frac{1}{r^{2\Delta_e}} = \frac{1}{r^{2(d-1/\nu)}}. \quad (\text{S16})$$

Since the expectation value of the energy does not vanish, the correlation function is defined with a subtraction of  $\langle E(\infty, 0)/N \rangle^2$ , or, equivalently, the operators  $e(\mathbf{r})$  can be defined with  $E(\infty, 0)/N$  subtracted.

In practice, it is often easier to determine  $\nu$  using the finite-size scaling form (S11), preferably with some dimensionless quantity  $n$  ( $\Delta_n = 0$ ), e.g., one of the Binder ratios  $R_k$  defined by

$$R_k = \frac{\langle m^{2k} \rangle}{\langle m^k \rangle^2}, \quad (\text{S17})$$

for which the scaling dimensions cancel. Data for  $R_k$  (normally with  $k = 2$ ) versus  $x = (T - T_c)L^{1/\nu}$  for different  $T$  (sufficiently close to  $T_c$ ) and  $L$  (sufficiently large) then collapse onto the common scaling function  $g(x)$ , provided that  $T_c$  and  $\nu$  have their correct values. These values can be determined by optimizing the quality of the data collapse. Alternatively,  $T_c$  can be obtained by investigating the flow of crossing points  $T_{12} = T(L_1, L_2)$  for which  $R_k(L_1, T_{12}) = R_k(L_2, T_{12})$  for different system size pairs  $(L_1, L_2)$ , e.g.,  $L_1 = L$ ,  $L_2 = 2L$ . The slope  $dR_k/d\delta$  at  $T_c$  scales as  $L^{1/\nu}$  and can be used once  $T_c$  has been determined, or the maximum slope for each  $L$  can be used in the same way even without prior knowledge of  $T_c$ . The location (temperature) of the maximum, or any feature in a quantity governed by a scaling form such as Eq. (S3), shifts with  $L$  as  $L^{1/\nu}$ . Details of the practicalities of finite-size scaling have been discussed extensively in the literature, and in Sec. III we illustrate procedures in our application to the classical 3D spin glass.

The above formalism and scaling methodology also apply to a quantum system undergoing a phase transition at a critical temperature  $T_c > 0$ , because the fluctuations at the longest (diverging) length scale are always thermal unless  $T = 0$ . Quantum fluctuations can drive continuous transitions with  $T_c = 0$ , as we discuss next.

## 2. Quantum phase transitions

A quantum system in  $d$  spatial dimensions can be mapped through the path integral approach to an effective statistical-mechanics system with an additional imaginary-time dimension<sup>60,64</sup>. In some cases, this effective system is space-time isotropic beyond a simple scale factor (a velocity) that relates space and time, but in general the time dimension is very different. The difference is quantified by the dynamic exponent  $z$ , which relates the spatial correlation length  $\xi$  to a relaxation time  $\xi_\tau$  (a correlation length in imaginary time) according to  $\xi_\tau \propto \xi^z$ . A critical real-space correlation function  $C(r) \propto r^{-a}$  translates by  $r \rightarrow \tau^{1/z}$  into a corresponding critical time correlation function  $C(\tau) \propto \tau^{-a/z}$ . In scaling dimensions and relationships where the dimensionality appears, the classical spatial dimensionality  $d$  should be replaced by  $d + z$  (where we still use  $d$  for the spatial dimensionality of the quantum system), following from the scaling dimension of the Hamiltonian density at  $T = 0$  (which replaces the classical free-energy density at  $T > 0$ ). With these generalizations, the formalism discussed in the preceding section remains valid, with important extensions described below.

A classical phase transition takes place at  $T_c > 0$  as a consequence of competition between the internal energy and the entropy. As discussed in the preceding section, the energy and the entropy can both be regarded as conjugate quantities to the temperature, which we call the thermal field when given relative to the critical temperature;  $\delta = T - T_c$ . In a quantum system at  $T = 0$  it is instead the competition between different terms in the Hamiltonian that cause the ground state transition. The fluctuations that allow for a continuous transition can in this case be traced to non-commuting terms of the Hamiltonian, leading to quantum fluctuations when the ground state is expressed in some local basis, e.g., with the spin- $z$  eigenvalues  $S_i^z$  ( $i = 1, \dots, N$ ) in a spin system..

The quantum fluctuations are regulated by changing some model parameter. A non-trivial Hamiltonian can always be formally decomposed into two non-commuting terms,  $H = aH_A + bH_B$ , so that the transition of the ground state is driven by relative changes in the prefactors  $a$  and  $b$ . Since an overall factor does not matter for the ground state wave function, we can fix  $b = 1$ . Then we can define the critical Hamiltonian  $H_c$  and consider deviations from it by tuning the parameter  $a$  only (for simplicity):

$$H(\delta_a) = H_c + \delta_a H_A. \quad (\text{S18})$$

Assuming that this tuning does not change the symmetry of  $H$  (i.e.,  $H_c$  and  $H_A$  have exactly the same symmetries, which normally is apparent),  $\delta_a$  is the symmetric scaling field whose inverse scaling dimension defines the conventional correlation length exponent  $\nu$ . Thus, if  $H_A$  consists of a sum of local terms  $h_A(\mathbf{r})$  in the same way as the generic perturbing operator  $P$  in Eq. (S5), we can define the correlation function, evaluated at  $\delta_a = 0$ , and its asymptotic power-law form will be dictated by the scaling dimension of the operator as in the classical case:

$$C_A(\mathbf{r}) = \langle h_A(\mathbf{r})h_A(0) \rangle - \langle h_A \rangle^2 \propto \frac{1}{r^{2\Delta_a}}. \quad (\text{S19})$$

Then, in a generalization of the classical relationship Eq. (S10) between scaling dimensions, the correlation-length exponent  $\nu = 1/y_a$  is given by

$$\frac{1}{\nu} = d + z - \Delta_a. \quad (\text{S20})$$

Here it should be noted that, we can also go away from the critical point by letting  $H = H_c + \delta_b H_B$ , and, therefore, the scaling dimensions of the operators  $H_A$  and  $H_B$  must be the same;  $\Delta_a = \Delta_b$ . Since it does not matter what term is tuned, we will also just use  $\delta$  without subscript for the deviation from the critical point. Similarly, we do not attach any index on  $\nu$  when there is a single relevant symmetric field, i.e., at a regular critical point that can be reached by tuning a single parameter in  $H$ .

The dynamic exponent  $z$  governs the low-energy excitations of the system, with the dispersion (energy–momentum) relation  $\epsilon_k \propto k^z$  and a finite-size gap

$$\Delta_L \propto L^{-z}. \quad (\text{S21})$$

This gap scaling corresponds directly to the scaling dimension of the total Hamiltonian  $H$  being  $z$ , i.e., the ground-state energy density has scaling dimension  $d + z$ . The corresponding quantity in the classical case is the free-energy density at  $T > 0$ , which has scaling dimension  $d$ , as stated explicitly in the finite-size form (S13). Thus, the replacement  $d \rightarrow d + z$  when converting classical scaling forms (obtained as derivatives of the free energy) to the quantum case (corresponding derivatives of the ground state energy).

It is important to recognize that the scaling dimension of the total Hamiltonian  $H = aH_a + bH_b$  is different from that of  $H_a$  and  $H_b$ . The latter two terms can drive a transition, as discussed above, while just changing the overall

quantum-mechanical Hamiltonian by  $H \rightarrow H + \delta_h H$  does not change the ground state. It is interesting to note that the relationship Eq. (S20) with  $\Delta_a$  replaced by the scaling dimension of the Hamiltonian density,  $\Delta_h = d + z$ , gives  $1/\nu_h = 0$ . Thus, a scaling form  $L^{-\Delta_n} g(\delta_h L^{1/\nu_h})$  of some operator  $n$  under the change of  $\delta_h$  reduces to just  $n = L^{-\Delta_n} g(\delta_h)$  without size dependence in  $g$ . In many cases the scaling function  $g(\delta_h)$  must be a constant, since the ground state does not depend on  $\delta_h$ . An exception is the finite-size gap, which depends on the overall scale of  $H$ , whence Eq. (S21) can be written as  $\Delta_L \propto (1 + \delta_h)L^{-z}$ , i.e.,  $g(\delta_h) \propto 1 + \delta_h$ .

A quantum phase transition takes place in the ground state, but the path integral representation is valid for any temperature, and the length  $L_\tau$  of the  $d + 1$  dimensional system in the imaginary-time dimension is proportional to the inverse temperature  $1/T$ . Finite-size scaling can in general be formulated with both the spatial size  $L$  and the temporal size  $L_\tau$ . In order for expectation values to converge to their ground state limits, the temperature has to be much smaller than the gap;  $T \ll L^{-z}$ . This convergence and cross-over from finite-temperature behavior can be expressed by an extended scaling function with an argument  $L_\tau/L^z$ :

$$n = L^{-\sigma/\nu} g(\delta L^{1/\nu}, L_\tau/L^z). \quad (\text{S22})$$

If  $z$  of a system is not known, it can be obtained by analyzing the dependence of some suitable quantity  $n$  on  $L$  as well as the inverse temperature scaled with some exponent  $z'$ , i.e.,  $1/T \propto L_\tau \propto L^{z'}$ . If  $z' < z$ , the argument  $L_\tau/L^z \rightarrow 0$  when  $L$  increases, while if  $z' > z$  we have  $L_\tau/L^z \rightarrow \infty$ . Depending on the quantity  $n$ , the scaling function behaves very different in these two limits, and the correct value  $z' = z$  can be deduced as a separatrix<sup>65</sup>. An example of this type of analysis for the 3D spin glass is presented in Sec. III.

If  $z$  is known, the inverse temperature can just be scaled as  $1/T \propto L^z$  and the scaling function (S22) then effectively has a single argument  $\delta L^{1/\nu}$ . In numerical simulations, an alternative approach if  $z$  is not *a priori* known is to study low enough temperatures so that all finite-size observables have converged to their ground state values. Then, effectively the limit  $L_\tau/L^z \rightarrow \infty$  is taken for each individual  $L$ , and again the scaling function in Eq. (S22) becomes one with a single argument:  $g(x, y) \rightarrow \tilde{g}(x)$ . As we will see below, the dynamic exponent can still be accessed because it appears in scaling dimensions when the classical scaling forms are modified by setting  $d \rightarrow d + z$ .

An important aspect of quantum-critical scaling is that some physical observables are defined at equal time, while others are integrals over imaginary time. Consider first the equal time expectation value  $\langle m^2 \rangle$  of the squared order parameter, which, following the same procedures as in the classical case, we can express as a correlation function integrated over  $\mathbf{r}$ , resulting in  $\langle m^2 \rangle \propto L^{-2\Delta_m}$ . Often, this quantity is multiplied by  $N = L^d$  for what is called the static structure factor, and it scales as

$$S_m = \langle M^2 \rangle / N = L^d \langle m^2 \rangle \propto L^{d-2\Delta_m} = L^{2-z-\eta}, \quad (\text{S23})$$

where in the last step we have used the quantum mechanical analogue of Eq. (S9) for the scaling dimension of  $m^2$ ;

$$2\Delta_m = d + z - 2 + \eta. \quad (\text{S24})$$

The scaling form of  $S_m$  also formally applies to a classical system by setting  $z = 0$ .

The perhaps most well known example of a quantity involving integration over imaginary time  $\tau$  (a Kubo integral) is the susceptibility corresponding to the order parameter field:  $\chi = d\langle m \rangle / d\delta_m$ . In the common case where  $m = M/N$  and  $H$  do not commute (e.g., in the case of the transverse-field Ising model), we have

$$\chi_m \propto \frac{1}{N} \int_0^{1/T} \langle M(\tau)M(0) \rangle d\tau, \quad (\text{S25})$$

which can be readily demonstrated by writing  $e^{-\beta H}$  (with  $\delta_m M$  included in  $H$ ) in the thermal expectation value  $\langle m \rangle$  as  $(e^{-\Delta_\tau H})^k$ , with small  $\Delta_\tau = \beta/k$ , so that the  $\delta_m$  derivative can be taken properly in each of the factors (“time slices”) and summed up in the form of an integral when  $k \rightarrow \infty$ . In order for the susceptibility to produce scaling reflecting the properties of the ground state (which is the only case we consider here, though finite-temperature scaling can also be considered and is an important aspect of quantum criticality<sup>64</sup>),  $T$  has to be below the lowest excitation gap of the system, which has the scaling form (S21). Thus, the upper integration bound  $1/T$  in Eq. (S25) has to be of order  $L^z$  or larger. Formally the cut-off should be set at  $\tau \propto L^z$  even when the limit  $T \rightarrow 0$  is taken, because the time correlations decay exponentially for  $\tau \gg L^z$ .

Again writing  $M$  as a sum of local operators, now  $M(\tau) = \sum_{\mathbf{r}} m(\mathbf{r}, \tau)$ , and integrating  $C(\mathbf{r}, \tau) = \langle m(\mathbf{r}, \tau)m(0, 0) \rangle$  over both space and (imaginary) time, we have  $d$  spatial coordinates  $\mathbf{r} = (r_1, \dots, r_d)$  and a time coordinate  $\tau$ . The critical correlation function has the asymptotic form

$$C(r, \tau) = \frac{1}{(r^2 + \tau^{2/z})^{\Delta_m}}. \quad (\text{S26})$$

To integrate this over the anisotropic space-time volume, we can define  $r_{d+1} \equiv \tau^{1/z}$  and use  $R$  for the length of the vector  $\mathbf{R} = (r_1, \dots, r_{d+1})$ . Then all components are integrated up to  $L$ , and the scaling form of  $\chi$  becomes (neglecting angular integrals that only produce unimportant factors)

$$\chi_m \propto \int^L dR \frac{R^d R^{z-1}}{R^{2\Delta_m}} \propto L^{d+z-2\Delta_m} = L^{2-\eta}. \quad (\text{S27})$$

Thus, compared to the structure factor (S23), there is an additional factor  $L^z$  arising from the time integral, and the dynamic exponent can be singled out, e.g.,  $\chi_m/S_m \propto L^z$ .

There is no obvious classical counterpart to the above scaling form of  $\chi_m$ , which is reflected in the fact that there is no  $z$  left that can be set to zero, unlike Eq. (S23). The definition (S25) of the susceptibility can of course still be used classically, in which case  $M(\tau) = M(0) = M$  (because  $M$  and  $H$  commute) and the integral only produces a factor  $1/T$ . Thus, classically  $\chi_m = S_m/T$ , which holds at any temperature.

From Eqs. (S23) and (S27) we see that both the exponents  $z$  and  $\eta$  can, in principle, be determined by investigating the size dependence of the structure factor and the susceptibility. In current quantum annealing devices, only equal-time quantities are accessible, in the form of the probability distribution of the spin- $z$  configurations at the end of the annealing process. This state can contain “memories” of the wave function close to the quantum-critical point, through the critical slowing-down mechanism first discussed by Zurek<sup>41,66</sup> in classical statistical mechanics (following related work on cosmological topological defects by Kibble<sup>40</sup>) and later generalized to quantum phase transitions as well<sup>42–44</sup>. The dynamics of quantum annealing depends on  $z$ , thus offering a path to extracting this exponent from experiments. We next discuss the generalization of FSS to a system undergoing an annealing process, first in general terms and then focusing specifically on the spin-glass annealing experiments and simulations.

## B. Dynamic finite-size scaling

In a classical or quantum annealing process in which a parameter like  $\delta$  varies linearly as a function of time  $t$ , the finite-size scaling form (S11) attains a second scaling argument;

$$n = L^{-\Delta_n} g(\delta L^{1/\nu}, vL^\mu), \quad (\text{S28})$$

where  $v$  is the velocity by which  $\delta$  changes,  $v = d\delta(t)/dt$  taken at  $\delta = 0$  (and if  $v = 0$  according to this definition, the approach discussed here can be generalized to the second derivative or any higher derivative<sup>46,67</sup>), and we consider  $v > 0$  for simplicity. Since constants of proportionality do not matter in scaling forms, we can also simply define the scaling velocity as the inverse of the total annealing time;  $v = 1/t_a$ . The exponent  $\mu$  controlling the velocity scaling is defined in the same way as in the main paper;

$$\mu = z + 1/\nu. \quad (\text{S29})$$

The FSS form (S28) has its origin in the KZ mechanism<sup>42–44,66</sup> and its use for analyzing numerical simulation data and experiments under various dynamical protocols was further elaborated and tested in numerous later works<sup>45–47,68,69</sup>.

We here assume that the process starts with the system in a disordered phase and approaches an ordered phase, though the opposite process can also be treated in a similar way. The basic starting point leading to (S28) is that a finite system approaching a continuous classical or quantum phase transition at rate  $v$  “freezes out” when the remaining time to reach the critical point is less than the required time for the system to equilibrate. In the thermodynamic limit, the correlation length reached before the system falls out of equilibrium is similarly

$$\xi_v \propto v^{-1/\mu}. \quad (\text{S30})$$

The argument  $vL^\mu$  in Eq. (S28) is simply a power of the dimensionless ratio  $L/\xi_v$ .

The KZ mechanism is often discussed in terms of defects generated by the out-of-equilibrium dynamics, but in general, except for some often studied 1D systems, it is not easy to directly relate a defect density to observed physical quantities. However, the scaling approach relies only on the existence of some dynamic process associated with a relaxation time  $\xi_t$ , which should be related to a spatial correlation length  $\xi$  by  $\xi_t \propto \xi^z$ . No other information on the dynamic process or the nature of the defects is required to derive universal scaling forms<sup>47,69</sup>.

In a quantum system undergoing an annealing process of the type considered here,  $z$  is exactly the dynamic exponent discussed in the previous section. Though there  $z$  was introduced in the context of imaginary (Euclidean) time, the relevant time scale is the same in real and imaginary time:  $\xi_t \propto \xi_\tau \propto \xi^z$ . At a classical equilibrium phase transition, this quantum dynamic exponent, which appears in the scaling dimensions, e.g., in Eq. (S20), is formally

zero. However, when a classical system is subjected to some stochastic dynamics, e.g., in a Monte Carlo simulation, there is a dynamic exponent corresponding to the particular process used (Metropolis or cluster updates). Then, in the KZ exponent Eq. (S29),  $z$  is the stochastic dynamic exponent. A third possibility is a quantum system with intrinsic dynamic exponent  $z$  that is simulated using a Monte Carlo process, such as the SQA simulations used as benchmarks here. Then the intrinsic (quantum) dynamic exponent should still be used in Eq. (S20), but  $z$  in the KZ exponent should be the relevant exponent corresponding to the imposed Monte Carlo dynamics (which also depends on the system studied)<sup>70</sup>. Here, for simplicity of the notation, we will first assume an actual quantum system undergoing Hamiltonian dynamics, where there is only a single dynamic exponent, but we will later return to the other cases as well.

The easiest way to analyze data within the dynamic finite-size scaling ansatz is to consider the quantity  $n$  exactly when the critical point has been reached:  $\delta = 0^{46,69}$ . Then we can simply write Eq. (S28) as

$$n = L^{-\Delta_n} g(vL^\mu), \quad (\text{S31})$$

which can be tested with data for different values of  $v$  and system sizes  $L$ . For  $v \rightarrow 0$ , we should of course recover the standard FSS form  $n \propto L^{-\Delta_n}$ , which means  $g(x) \rightarrow \text{constant}$  when  $x \rightarrow 0$ . For  $v > 0$ , the scaling function can be extracted by graphing  $y = nL^{\Delta_n}$  versus  $x = vL^\mu$  for sufficiently large  $L$  and a range of  $v$  values, whence all the data points should collapse onto the common function  $g(x)$ . If the exponents are not known, they can be found by adjusting their values for optimal collapse.

Quantities  $n$  with known scaling dimension  $\Delta_n = 0$  are particularly useful, since any uncertainties in (or lack of knowledge of) the leading finite-size dependence can then be avoided. A well known example is the Binder ratio defined with two different powers of the order parameter, Eq. (S17), for which the scaling dimensions of the numerator and denominator cancel. Normally the  $k = 2$  case,  $R_2 = \langle m^4 \rangle / \langle m^2 \rangle^2$ , is considered, and a corresponding cumulant is defined as  $U \equiv U_2 = a(1 - bR_2)$ , where the coefficients  $a$  and  $b$  are determined using the symmetry of the order parameter such that  $U \rightarrow 0$  in the disordered phase and  $U \rightarrow 1$  in the ordered phase. The definition of this quantity in a system with random couplings is further discussed below in Sec. IC3.

The asymptotic form of  $g(x)$  in Eq. (S31) can be argued as follows: For  $v > 0$  the correlation length  $\xi_v$  is finite. In limit  $L \gg \xi_v$  the system is disordered and the size dependence becomes trivial, though its form depends on the quantity  $n$  considered. In the case of a squared order parameter, we have  $\langle m^2 \rangle = \langle M^2 \rangle / N^2 \propto 1/N = L^{-d}$ , which can be easily seen by expressing  $\langle M^2 \rangle$  using the exponentially decaying (asymptotically) correlation function  $C_m(\mathbf{r})$  of corresponding local operators [defined with a generic operator  $P$  in Eq. (S5)]:

$$\langle M^2 \rangle = N \sum_{\mathbf{r}} C_m(\mathbf{r}) \propto N \quad (\text{for short-ranged } C_m). \quad (\text{S32})$$

In order to obtain the  $L^{-d}$  dependence of  $\langle m^2 \rangle$  from Eq. (S31), the scaling function must take an asymptotic power-law form,  $g(x) \rightarrow x^{-c} \tilde{g}(x)$  with the function  $\tilde{g}(x)$  approaching a constant when  $x \rightarrow \infty$  and the exponent  $c$  satisfying  $c\mu + \Delta_{m^2} = d$ , with  $\Delta_{m^2} = 2\Delta_m$ , i.e.,

$$\langle m^2 \rangle \propto L^{-2\Delta_m} (vL^\mu)^{-c} \tilde{g}(vL^\mu) \rightarrow L^{-d} v^{-c}, \quad c = \frac{d - 2\Delta_m}{\mu} = \frac{d - 2\beta/\nu}{\mu}. \quad (\text{S33})$$

This is the KZ FSS form far from equilibrium, which must eventually break down when  $v$  is very large, i.e., when the correlation length  $\xi_v$  is of order unity (the lattice spacing in a lattice model). The ultimate high-velocity form is just  $\langle m^2 \rangle = L^{-d} g(v)$ , where the function  $g$  is analytic in  $1/v$ , and this form crosses smoothly over into the algebraic KZ form (S33) as  $v$  is reduced. When  $v$  is further reduced, for given  $L$  this form turns smoothly into the equilibrium form of Eq. (S31),  $\langle m^2 \rangle \propto L^{-2\beta/\nu}$ , when  $v \rightarrow 0$ .

For the excess energy density defined in Eq. (S15), the above arguments can be repeated with the modification that now  $\rho_E$  must be size independent at high velocities since it is a simple density with only exponentially small finite-size corrections in the disordered initial state. Thus, the algebraic KZ form for large  $vL^\mu$  (but still  $v \ll 1$ ) is

$$\rho_E \propto L^{-\Delta_{\rho_E}} (vL^\mu)^\kappa \tilde{g}(vL^\mu) \rightarrow v^\kappa, \quad \kappa = \frac{\Delta_{\rho_E}}{\mu} = \frac{d + z - 1/\nu}{\mu}. \quad (\text{S34})$$

This form evolves gradually into a constant when  $v \rightarrow \infty$ , and in the other extreme it crosses over to the equilibrium form  $\rho_E \propto L^{-\Delta_e} = L^{-(d+z-1/\nu)}$  when  $v$  is of order  $L^{-\mu}$  (i.e., when  $\xi_v$  grows to order  $L$ ). Note that the exponents  $c$  and  $\kappa$  in Eqs. (S34) and (S33) are both positive but they appear with different signs when the powers of  $\delta$  are taken, reflecting the decrease in the residual energy and increase in the order parameter as the annealing velocity is reduced.

For the quantities  $n$  considered above, and many others, the algebraic KZ scaling form can be investigated by graphing  $y = nL^{\Delta_n}$  versus  $x = vL^\mu$  after a process stopping at the critical point. The power-law forms discussed

above will then cross over into constants for sufficiently low  $v$ , which is possible because of the function  $\tilde{g}(vL^\mu)$  in Eqs. (S33) and (S34). Upon increasing  $v$ , data for small systems will successively peel off from the common scaling function when the correlation length  $\xi_v$  is of the order of the lattice spacing (loosely speaking, when  $v$  is of order one). In an alternative method,  $y = n$  (in the case of the residual energy density) or  $y = nL^d$  (in the case of the squared order parameter) can be graphed versus  $v$  or  $v^{-1}$ , in which case the power-law behaviors in Eqs. (S33) and (S34) are observed in the algebraic KZ regime but data collapse does not apply when equilibrium is approached (where curves for different  $L$  gradually peel off from the common power-law form). Instead, data collapse now applies also in the high- $v$  limit, where constant behaviors pertain. The two ways of analyzing SA data have been demonstrated for uniform 2D and 3d classical Ising models (using both local and cluster Monte Carlo updates)<sup>69</sup> as well as 2D<sup>71,72</sup> and 3D<sup>48</sup> classical spin glasses. Similar methods were also applied to imaginary-time QA (implemented in quantum Monte Carlo simulations with imaginary-time dependent interactions) of the transverse-field Ising model with uniform<sup>46</sup> and random (3-regular graphs)<sup>70</sup> interactions.

In the case of the Binder cumulant (or ratio), the scaling dimension is zero but the asymptotic form of the equilibrium scaling function  $U = u(\delta L^{1/\nu})$  in the disordered phase is non-trivial and still not well understood<sup>57,73</sup>. The size dependence of  $U$  in a disordered state is not just a constant [as in Eq. (S34)] or  $\propto L^{-d}$  [as in Eq. (S33)] but a different power-law  $L^k$  that is not known generically. The exponent  $b$  governing the KZ power law  $U \propto v^{-b}$  when  $L \gg \xi_v$  can therefore also not be obtained in a simple manner using just the arguments leading to Eqs. (S33) and (S34). The scaling form  $U = u(vL^\mu)$  still of course applies (including an asymptotic power-law form with some exponent  $b$ ) and can be extracted empirically as in other cases by graphing  $y = U$  versus  $x = vL^\mu$  with  $\mu$  optimized.

When annealing slightly into an ordered phase (taken as  $\delta > 0$ ) the extended KZ form (S28) can be Taylor expanded as long as  $\delta L^{1/\nu}$  remains small, i.e., asymptotically for large  $L$  the system tends to the critical point. When annealing more significantly into the ordered phase, but still within the critical window  $\delta \propto L^{-1/\nu}$ , power laws appear similar to those discussed above on the disordered side of the transition. When annealing far into an ordered state the situation is much more complicated and non-generic, requiring additional knowledge of the dynamical processes inside the phase. While some progress on this front has been reported recently<sup>54</sup>, in general the issue remains largely open. We next discuss annealing into a spin-glass phase and present insights and hypotheses of direct relevance to the quantum annealing experiments and simulations reported in this work.

### C. Application to annealing experiments and simulations

For a system annealed through a phase transition and slightly into an ordered phase, we can expect the extended KZ finite-size scaling form (S28) to apply. In principle, once the system is deep enough in the ordered phase, some other dynamical process beyond the critical fluctuations underlying the KZ mechanism will take over and eventually dominate. In some cases, a dynamical process in the ordered phase is faster than the KZ dynamics, e.g., in the case of coarsening dynamics of a system with no randomness—see Ref.<sup>74</sup> for discussions of how the KZ scaling is violated in the ordered phase in a simple Ising model. Then KZ dynamics should in general not be expected, though recent work suggests that an extension of the scheme, with modified power laws, is valid slightly inside in ordered phase in some cases like the standard uniform 2D Ising model in a transverse field<sup>54</sup>.

Inside a classical or quantum glass phase, all dynamical processes should be extremely slow. Therefore, KZ scaling can be anticipated even if the annealing process stops (and the outcome is observed only) far beyond the transition point. In the QA device used in our experiments, there are also technical reasons related to the shape of the annealing protocol (see Fig. 1b) why the system should “freeze out” and not evolve significantly at the latter stages of the process ending at the classical point  $s = 1$  (because the transverse field becomes very small already at  $s \approx 0.7$ ). In MC simulations, KZ scaling can be measured by observing the system exactly at the critical point (as was done in an SA study of the simple-cubic 3D Ising spin-glass<sup>48</sup>). However, in this work we anneal deep into the glassy phase even in SA and SQA for two reasons. First, we wish to qualitatively reproduce the QA protocol. Second, annealing toward the  $T = 0$ ,  $\Gamma = 0$  point guarantees a final equilibrium near the classical ground state.

Fig. 3f shows quite clearly the expected crossover when the density of AFM bonds is increased from  $p = 1/2$ , between glassy KZ dynamics with an exponent  $\mu$  that is close to its known KZ value ( $\mu = z + 1/\nu$ ) and a smaller value  $\mu \approx 2$ . The smaller value may indicate coarsening dynamics inside the AFM phase for  $p \gtrsim 0.8$ . However, coarsening in quantum systems has not been extensively investigated using reliable unbiased simulations, and no direct evidence of such dynamics was found at the early stages of entering the ferromagnetic phase of the transverse-field Ising model in Ref.<sup>54</sup>.

Focusing here on the glass phase, from the analysis of the Edwards-Anderson (EA) order parameter  $\langle q^2 \rangle$  (where again  $\langle \cdot \rangle$  also involves an average over disorder realizations), while the exponent  $\mu$  is close to its KZ value, the scaling dimension of the quantity itself (the exponent  $r$  in Figs. E6 and E7) is far from the anticipated critical scaling dimension. We will show here that such behavior over a range of relatively fast annealing rates and small to

moderate system sizes can be explained by an extended KZ scaling form supplemented by mild, physically reasonable assumptions pertaining to the glassy state.

### 1. Spin glass order parameter

Before considering scaling behavior, we briefly discuss the EA order parameter in light of the definitions of exponents and scaling dimensions explained in the previous sections. We use the conventional overlap parameter of the spins  $S^{(a)}$  and  $S^{(b)}$  in two replicas,  $a$  and  $b$ ,

$$q_{ab} = \frac{1}{N} \sum_i S_i^{(a)} S_i^{(b)}, \quad S_i^{(a,b)} \in \{-1, +1\}, \quad (\text{S35})$$

obtained from the same realization of the Ising couplings but in different runs with the QA device or in independent MC simulations. The expectation value (quantum or thermal, depending on the case considered) of the squared order parameter for a given coupling realization is

$$\langle q_{ab}^2 \rangle = \frac{1}{N^2} \sum_{i,j} \langle S_i^{(a)} S_j^{(a)} \rangle \langle S_i^{(b)} S_j^{(b)} \rangle, \quad (\text{S36})$$

where the expectation value factors because the two replicas are statistically independent. We can now take averages also over the replicas. Because the coupling realization is the same in all replicas, the replica-averaged expectation values in Eq. (S36) will be the same. In fact, over a sufficiently long time a finite system is ergodic, and no average over  $a$  and  $b$  replicas need to be formally taken. However, in reality simulations of glasses are extremely slow, and replicas have to be considered in order to sufficiently sample the configuration space. In annealing experiments and simulations out of equilibrium, the systems are by definition not ergodic, and averaging over replicas is an integral aspect of such studies. Once replica and disorder averages have been taken, we drop the indices  $a$  and  $b$  and have

$$\langle q^2 \rangle = \frac{1}{N^2} \sum_{i,j} \langle S_i S_j \rangle^2. \quad (\text{S37})$$

When averaging over coupling realizations, translational symmetry is also restored, and we can introduce a squared correlation function  $C^2(\mathbf{r})$  as the averaged  $\langle S_i S_j \rangle^2$ , with  $\mathbf{r} = \mathbf{r}_i - \mathbf{r}_j$ , to write

$$\langle q^2 \rangle = \frac{1}{N} \sum_{\mathbf{r}} C^2(\mathbf{r}). \quad (\text{S38})$$

Such a sum without squaring  $C(\mathbf{r})$  would vanish because of coupling averaging when the fraction of AFM couplings is 1/2, and the use of the overlap  $q$  defined with replicas is a convenient way to solve this problem. The fact that the squared EA order parameter formally is a sum over squared correlation functions in Eq. (S38) then has to be taken into account when using exponent relationships such as (S9), where  $d$  and  $z$  cannot depend on what correlation function is considered and, therefore, the factor in front of the scaling dimension must be 4 instead of 2.

Note that  $\langle q^2 \rangle$  is a perfectly valid order parameter also for a system without disorder, e.g., an Ising ferromagnet. Then it is clear that the decay of the squared correlation function should be associated with four times the scaling dimension of the order parameter,  $C_m^2(r) \propto 1/r^{4\Delta_m}$ , because of the exponent  $2\Delta_m$  in the conventional (not squared) correlation function (S8). Thus, if we wish to use conventional definitions of the critical exponents in the case of the EA order parameter, obeying forms analogous to Eq. (S20), we would have to write the FSS form of  $\langle q^2 \rangle$  implied by Eq. (S38) as

$$\langle q^2 \rangle \propto L^{-2(d+z-2+\eta)}. \quad (\text{S39})$$

However, normally the exponents are defined differently in spin glasses, by treating  $q$  as a regular order parameter and defining  $\eta$  based on the scaling form for susceptibilities such as (S27), where  $\chi_q$  would be defined with  $M = Nq$  in Eq. (S25)<sup>65</sup>. Such a definition of  $\eta$ —let us call it  $\eta'$ —does not satisfy the common relationship to a scaling dimension, i.e.,  $2\Delta_q \neq d + z - 2 + \eta'$ .

To conform with the standard notation in the spin glass literature, we will here use the common definition of the order parameter exponent  $\beta$  for spin glasses, i.e., we treat the EA order parameter as a conventional order parameter with the form  $\langle q^2 \rangle \propto \delta^{2\beta}$  in the thermodynamic limit, so that the critical finite-size form is  $\langle q^2 \rangle \propto L^{-2\beta/\nu}$ . Note that the above issue with exponent relationships does not affect the exponent  $\nu$  when extracted using data-collapse

methods, or the intrinsic (quantum) dynamic exponent  $z$  when it is extracted from the dependence on the aspect ratio  $L_\tau/L^z$  in Eq. (S22). The KZ exponent  $\mu = z + 1/\nu$  is also not affected by the specific definitions of  $\beta$  and  $\eta$ . We will use a conventional scaling dimension further below when discussing the residual energy, whose definition is independent of the use of replicas.

## 2. Scaling of the spin-glass order parameter

Applying the extended KZ form (S28) to the EA order parameter, we have

$$\langle q^2 \rangle = L^{-2\beta/\nu} g(vL^\mu, \delta L^{1/\nu}), \quad (\text{S40})$$

where  $\delta = s - s_c$  and  $s(t) \in [0, 1]$ . In theoretical work and simulations, the simple linear protocol  $J \propto s$ ,  $\Gamma \propto 1 - s$  is often used for transverse-field Ising models. In the QA experiments, the protocol is nonlinear (Fig. 1d) but linearity still holds for  $s$  in the close neighborhood of the critical point and Eq. (S40) is valid with  $v$  defined as the actual coupling derivative, or, for scaling purposes, the inverse of the total annealing time.

Though there are some claims in the literature that  $z \rightarrow \infty$  at the glass transition in the 2D and 3D Ising spin glasses considered here<sup>75,76</sup> (similar to random-field models<sup>36,77</sup>), in our opinion the available numerical evidence leans in the favor of finite, relatively small values of  $z$ <sup>65,78</sup>. In our own simulations and QA experiments (SI Sec. II and Figs. E6-E8), we also do not see any evidence of large or significantly flowing (with increasing system size) values of  $z$ . Though the possibility of an infinite-disorder fixed point (where formally  $z = \infty$ ) cannot be completely ruled out, we here consider only scaling with finite  $z$ .

As discussed in the preceding sections, the thermodynamic-limit form  $\langle q^2 \rangle \propto \delta^{2\beta}$  demands that  $g(0, x)$  with  $x = \delta L^{1/\nu}$  reduces to  $x^{2\beta}$  when  $x \rightarrow \infty$  inside the spin-glass phase close to the transition point. Moving further inside the glass phase, this power-law form can no longer hold. To enable the cross-over to a different (unspecified) form in equilibrium, we can consider a more flexible version of Eq. (S40),

$$\langle q^2 \rangle = L^{-2\beta/\nu} f(vL^\mu, \delta, L), \quad (\text{S41})$$

where  $\delta$  and  $L$  appear as two different arguments in the function. For small  $\delta$  we know that the function must simplify to the standard scaling form (S40), where the two arguments  $\delta$  and  $L$  are replaced by just the single argument  $\delta L^{1/\nu}$ .

We can express the cross-over from the two arguments  $(L, \delta)$  to effectively a single argument by introducing another function as an argument to the scaling function,

$$\langle q^2 \rangle = L^{-2\beta/\nu} g(vL^\mu, h(\delta, L)\delta L^{1/\nu}), \quad (\text{S42})$$

where  $h \rightarrow 1$  in the regime of  $\delta$  and  $L$  where the simpler form (S40) applies. This functional form is still as general as Eq. (S41) and can in principle describe the behavior for any  $\delta$  and  $L$ . To recover the form  $\langle q^2 \rangle \propto \delta^{2\beta}$  from (S42) in the thermodynamic limit when  $v$  is small, a power-law of the second argument must again form

$$\langle q^2 \rangle \rightarrow L^{-2\beta/\nu} h^{2\beta}(\delta, L)(\delta L^{1/\nu})^{2\beta} g(vL^\mu) = \delta^{2\beta} h^{2\beta}(\delta, L)g(vL^\mu), \quad (\text{S43})$$

where we have also assumed that, to a good approximation, we do not need to keep a second argument of the function  $g$  when the power law form has set in, which is equivalent to saying that the function  $g$  now is of the form  $g(vL^\mu, k(\delta, L))$ , where the function  $k$  is almost constant in the relevant regime of  $L$  and  $\delta$ .

To simplify further, we can re-define the function  $h$  in Eq. (S43) without the exponent  $2\beta$ , and we can also absorb the factor  $\delta$ . Then

$$\langle q^2 \rangle = h(\delta, L)g(vL^\mu). \quad (\text{S44})$$

This form is clearly valid for  $v \rightarrow 0$ , since it can capture any dependence on  $\delta$  and  $L$ , and we know that  $h(\delta, L) \rightarrow h(\delta L^{1/\nu})$  close to the critical point. We have of course made assumptions leading to  $g(vL^\mu)$  appearing as a separate factor, for which we have no proof but which can be tested in experiments and simulations.

Inside the glass phase, the squared order parameter converges with  $L$  to a non-zero value when  $v \rightarrow 0$ . Thus, for sufficiently large  $L$  inside the ordered phase, in Eq. (S44)  $h(\delta, L) \rightarrow k(\delta)$  and

$$\langle q^2 \rangle = k(\delta)g(vL^\mu), \quad (\text{S45})$$

where the function  $k$  represents whatever the  $\delta$  dependence is in equilibrium,

$$k(\delta) \equiv \langle q^2(\delta, v = 0) \rangle, \quad (\text{S46})$$

and  $g(x \rightarrow 0) \rightarrow 1$ .

The final forms (S44) and (S45) cannot be strictly correct, as KZ dynamics should at least not be valid at the final stage (the most extreme case) when the order saturates. However, this just means that there is some other (assumed to be slower) dynamics that takes over, but we have assumed only one dynamical mechanism. The assumptions and resulting forms may be valid as long as the annealing time is not too long, so that the dynamics associated with the glass phase can be considered as a weak perturbation to the critical dynamics with exponent  $z$ . Then also  $\langle q^2 \rangle$  will remain small though its scaling may reflect (as we will see) a cross-over from critical to ordered behavior.

For  $\delta$  not small and  $L$  not large enough to eliminate the  $L$  dependence of  $h$  in (S44), we must still keep the function with two arguments. It is not clear what the functional form of this  $h(\delta, L)$  is, but it is at least plausible that a power law applies for some range of system sizes, i.e.,

$$h(\delta, L) \rightarrow \tilde{k}(\delta)L^{-r}. \quad (\text{S47})$$

Here, for  $\delta \rightarrow 0$  (i.e., in the limit of conventional FSS and KZ scaling) we must have  $r = 2\beta/\nu$  and  $k(\delta)$  approaching a non-zero constant. For  $L \rightarrow \infty$  and  $\delta \neq 0$ , the exponent  $r$  must tend to zero and  $\tilde{k}(\delta)$  takes the equilibrium form as in Eq. (S46). For fixed  $\delta$  inside the glass phase, the exponent  $r$  must then be flowing but may still evolve only slowly with  $L$ , thus essentially constant when observed only over a small range of system sizes. Indeed, in our experiments and simulations, we observe an effective exponent  $r$  between 0 and  $2\beta/\nu$ .

### 3. The Binder ratio

According to our arguments above, an effective scaling dimension different from that at the critical point can be manifested when annealing a system into a spin glass phase;

$$\langle q^2 \rangle = \tilde{k}(\delta)L^{-r_L}g(vL^\mu), \quad (\text{S48})$$

with an exponent  $r_L$  that changes only slowly with  $L$  from  $r_L = 0$  to  $r_L = 2\beta/\nu$ . As in the case of conventional critical scaling, the effective scaling dimension of  $\langle m^4 \rangle$  should be twice that of  $\langle m^2 \rangle$ , and the Binder ratio (and cumulant) then remains dimensionless. Thus, it is the preferred quantity to analyze when the goal is to extract the exponent  $\mu$  in annealing experiments where the state exactly at the transition point  $s_c$  is inaccessible.

The Binder ratio for a system with intrinsic randomness (e.g., a spin glass) can be defined with the average over realizations taken either before or after evaluating the ratio in Eq. (S17). Now using  $[.]$  to denote a disorder average and  $\langle . \rangle$  for only thermal and quantum averages (including replicas), we can define the following two ratios:

$$R_a = \left[ \frac{\langle q^4 \rangle}{\langle q^2 \rangle^2} \right], \quad (\text{S49a})$$

$$R_b = \frac{[\langle q^4 \rangle]}{[\langle q^2 \rangle]^2}. \quad (\text{S49b})$$

While Eq. (S49a) has been frequently used (e.g., Ref.<sup>65</sup>), Eq. (S49b) has become more common (see also Ref.<sup>79</sup>). They are both dimensionless quantities and in principle equally valid for use in scaling analysis.

Equation (S49a) has a drawback of being biased when evaluated with a finite number of measurements to estimate the expectation values. Consider statistical deviations  $\epsilon_2$  and  $\epsilon_4$  of  $\langle q^2 \rangle$  and  $\langle q^4 \rangle$  from their exact values  $\langle . \rangle_{\text{ex}}$  for a given coupling realization, based on  $\Lambda$  measurements. Then the ratio Eq. (S49a) is

$$R = \frac{\langle q^4 \rangle}{\langle q^2 \rangle^2} = \frac{\langle q^4 \rangle_{\text{ex}} + \epsilon_4}{(\langle q^2 \rangle_{\text{ex}} + \epsilon_2)^2} = \frac{\langle q^4 \rangle_{\text{ex}}}{\langle q^2 \rangle_{\text{ex}}^2} + O(\epsilon_2^2, \epsilon_2 \epsilon_4), \quad (\text{S50})$$

where odd powers of the error vanish on average. Since the magnitudes of  $\epsilon_4$  and  $\epsilon_2$  scale as  $\Lambda^{-1/2}$ , the expected bias in the realization-averaged ratio  $R_a$ , Eq. (S49a), scales as  $\Lambda^{-1}$ . While this bias decays faster than the expected overall  $\propto \Lambda^{-1/2}$  statistical error in  $R$ , if the number of measurements  $\Lambda$  is small there can still be a non-negligible effect of bias left. This bias is smaller by a factor  $N_r^{-1/2}$  in  $R_b$ , Eq. (S49b), when both the numerator and denominator are averaged over  $N_a$  independent disorder realizations.

The leading bias can in principle be removed by combining results for different  $\Lambda$ . In the simulations and experiments for the cumulant  $U = (3 - R)/2$  presented here,  $\Lambda$  is sufficiently large for the bias to be insignificant. We nevertheless use Eq. (S49b), though (S49a) gives compatible results.

#### 4. Residual Ising energy

For the residual Ising energy  $E$ , we proceed in a slightly different way. Since the Ising part of the Hamiltonian drives the transition, its scaling dimension  $\Delta_e = d + z - 1/\nu$  is given by Eq. (S20) with  $\nu$  being the conventional correlation-length exponent. The KZ-FSS scaling form is

$$\rho_E = L^{-\Delta_{\rho_E}} f(vL^\mu, \delta L^{1/\nu}), \quad (\text{S51})$$

where it is assumed that the regular part of the energy, i.e., the ‘‘background’’ that is not part of the singular behavior, has been subtracted off. The need for a subtraction here is an important aspect of the residual energy that is absent in the case of the order parameter. The subtracted part should normally be the equilibrium critical Ising energy in the thermodynamic limit, as in Eq. (S15). However, nothing prohibits us from instead subtracting the equilibrium finite-size Ising energy  $E(\delta, L)$ , which has the same scaling dimension as  $\rho_E$ . Then we have the following expected FSS form [keeping the same symbol  $\rho_E$  as before even though its meaning has changed slightly]:

$$\rho_E(L, \delta, v) = L^{-\Delta_{\rho_E}} [g(vL^\mu, \delta L^{1/\nu}) - g(0, \delta L^{1/\nu})]. \quad (\text{S52})$$

Here the difference between the two scaling functions is just another scaling function, and we can write

$$\rho_E(L, \delta, v) = L^{-\Delta_{\rho_E}} k(vL^\mu, \delta L^{1/\nu}). \quad (\text{S53})$$

For any  $\delta$ , by definition (because of what we have decided to subtract), this  $\rho_E$  must vanish when  $v \rightarrow 0$ . That means that a positive power of  $v$  must appear as a factor, and  $\delta L^{1/\nu}$  can only appear in a function multiplying this overall  $v$  dependence. The most natural way to express this behavior is for  $k$  to develop a power law of its first argument, i.e.,

$$k(vL^\mu, \delta L^{1/\nu}) \rightarrow (vL^\mu)^\kappa h(vL^\mu, \delta L^{1/\nu}). \quad (\text{S54})$$

Note that the new scaling function  $h$  can still also depend on  $vL^\mu$ , and mathematically there is no approximation here because the right-hand side still depends only on the two arguments  $x = vL^\mu$  and  $y = \delta L^{1/\nu}$  (i.e., it is some function of these two arguments, as on the left-hand side).

The exponent  $\kappa$  can again be extracted as in Eq. (S34) if we demand no  $L$  dependence apart from that in the function  $h$ , and we now call this critical exponent  $\kappa_c$ . We know that this scaling must hold when  $\delta = 0$  and the KZ correlation length is smaller than the system size. At least for small  $\delta$ , as long as the function  $h(x, y)$  is well behaved, the exponent on  $v$  cannot change. Then, the final residual Ising energy in a QA process stopping inside the glassy phase takes the form

$$\rho_E(L, \delta, v) = v^{\kappa_c} h(vL^\mu, \delta L^{1/\nu}), \quad (\text{S55})$$

with the same exponent  $\kappa_c$  as when stopping at the critical point,

$$\kappa_c = \frac{\Delta_{\rho_E}}{z + 1/\nu} = \frac{d + z - 1/\nu}{\mu}. \quad (\text{S56})$$

It should be emphasized that Eqs. (S55) and Eq. (S53) are two completely equivalent forms of  $\rho_E$ , and whichever form to use just depends on the type of scaling analysis one wishes to perform. They are both scaling functions of the single variable  $x = vL^\mu$  at the critical point  $\delta = 0$ .

For a process not stopping exactly at the critical point, a scaling function of a single variable can still be obtained when considering the thermodynamic limit of Eq. (S55). For  $\delta \neq 0$ , when  $L \gg \xi$  the system size can no longer appear.  $L$  can only be eliminated from the scaling function  $h(x, y)$  with  $x = vL^\mu$  and  $y = \delta L^{1/\nu}$  if the function of two arguments evolves into one of a single argument,  $h[z(x, y)]$ , where  $z$  does not depend on  $L$ . To cancel  $L$ , we can take  $z = x^{1/\mu\nu} y^{-1}$  and obtain a scaling function  $e(z)$  of a single variable, whence Eq. (S55) becomes

$$\rho_E(\delta, v) = v^{\kappa_c} e(v^{1/\mu\nu} \delta^{-1}). \quad (\text{S57})$$

Recall that here we must have  $\delta \neq 0$  so that the limit where  $L$  exceeds  $\xi$  exists. A physical argument for the new scaling variable is that  $\delta$  in this limit should not be compared with  $L^{-1/\nu}$  as in equilibrium finite size scaling (with the single scaling argument  $\delta L^{1/\nu}$ ). Instead,  $L$  must be replaced by the finite velocity-dependent KZ length scale  $\xi_v$ , which is smaller than  $L$ . Thus,  $L^{1/\nu}$  is replaced by  $\xi_v^{1/\nu} = v^{1/\mu\nu}$  and the proper scaling argument is  $\delta v^{-1/\mu\nu} = z^{-1}$ , or, alternatively, the inverse thereof as in (S57), where the low-velocity limit corresponds to  $e(z \rightarrow 0)$ .

The way in which we have defined the excess Ising energy  $\rho_E$ , it vanishes when  $v \rightarrow 0$  for any  $\delta$ . However, pure critical KZ behavior,  $\rho_E \propto v^{\kappa_c}$ , should only apply very close to  $\delta = 0$  for some range of velocities before a cross-over

into a form governed by the dynamics of the glassy state when  $v \rightarrow 0$ . For a process stopping significantly inside the glassy state, there may be no range of velocities for which the exponent  $\kappa_c$  strictly applies. However, if  $v$  is not too low, i.e., the argument  $x = v^{1/\nu\mu}\delta^{-1}$  is relatively large in  $e(x)$  in Eq. (S57), it is plausible that the scaling function takes a power-law form,  $e(x) \sim x^k$  with some unknown exponent  $k$  (and also note that the scaling function cannot be analytic in the glass phase, because  $\kappa_c$  does not asymptotically describe the dynamics). Since  $\delta$  is just a fixed number, this behavior would imply a change in the exponent on  $v$  in Eq. (S57),  $\rho_E \sim v^{\kappa_f}$  with  $\kappa_f = \kappa_c + k/\mu\nu$ . Given that the dynamics of the quantum spin glass should be slower than the KZ dynamics, it is natural to expect  $\kappa_f < \kappa_c$ , as we indeed also found in our QA experiments (Fig. 4b).

The above analysis presumes that  $\delta$  drives a quantum phase transition at which the Hamiltonian is the sole provider of dynamics. At a classical phase transition of a model with no intrinsic dynamics,  $d+z$  in the expression for  $\kappa_c$ , Eq. (S56), should be just  $d$ , and the dynamic exponent in the KZ exponent  $\mu = z+1/\nu$  should be that of the stochastic process used to evolve the system with SA,

$$\kappa_{SA} = \frac{d - 1/\nu}{z_{SA} + 1/\nu}, \quad (S58)$$

where  $z_{SA}$  also depends on the system studied (i.e., it is a universal exponent given a critical point and a type of stochastic process).

The case of SQA is a mix of QA and SA in the sense that it implements the imaginary-time dimension through a path integral, hence  $d+z$ , with the same  $z$  as in QA, appears in the numerator of  $\kappa_{SQA}$ . However, the  $d+1$  dimensional effective statistical-mechanics system is simulated by an MC procedure as in SA, and it is the exponent corresponding to that process that allows the simulation to relax toward its equilibrium. Thus, in this case, the exponent contains two dynamic exponents

$$\kappa_{SQA} = \frac{d + z - 1/\nu}{z_{SQA} + 1/\nu}, \quad (S59)$$

where  $z_{SQA}$  is not *a priori* known, and, like the classical exponent  $z_{SA}$ , depends on both the system studied and the MC procedures applied. To our knowledge, this exponent had not been computed previously for the 3D Ising spin glass in a transverse field. In our SQA work, we extracted  $\mu_{SQA} = z_{SQA} + 1/\nu$  from the data collapse of the Binder cumulant (Fig. E6).

## II. 3D SPIN GLASSES

### A. Construction

The instances we study for size  $L$  are on  $L \times L \times \min(L, 12) \times 2$  spins, indexed by  $(x, y, z, w)$  with  $x, y \in \{0, \dots, L-1\}$ ,  $z \in \{0, \dots, \min(L-1, 11)\}$ , and  $w \in 0, 1$ . For a given 3D coordinate  $x, y, z$ , sites  $(x, y, z, 0)$  and  $(x, y, z, 1)$  are coupled with a strong ferromagnetic coupling,  $J = -J_{FM}$ ; in this work  $J_{FM} = 2$ . For  $L > 9$ , instances contain site vacancies due to inoperable qubits.

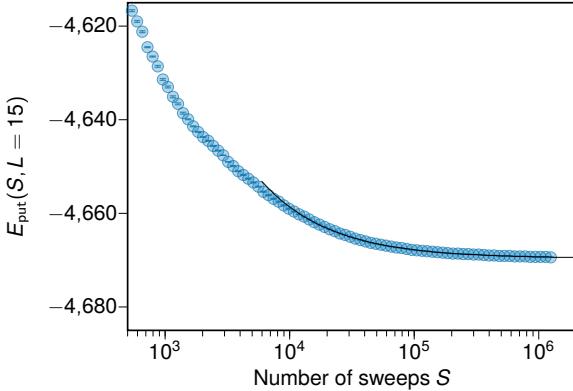
In the  $x$  dimension, couplings use only the  $w = 0$  sites; the coupling between  $(x, y, z, 0)$  and  $(x+1, y, z, 0)$  for  $x < L$  is  $\pm J_G$ :  $J_G$  with probability  $p$ , and  $-J_G$  otherwise. Random choices of spin-glass couplings are independent. Except where stated,  $p = 0.5$ . The  $x$  dimension has open boundaries.

In the  $y$  dimension, couplings use only the  $w = 1$  sites; the coupling between  $(x, y, z, 1)$  and  $(x, y+1, z, 1)$  for  $y < L$  is  $\pm J_G$ . The  $y$  dimension has open boundaries.

The  $z$  dimension is periodic. For  $z' \equiv z + 1 \pmod{L}$ , the coupling between  $(x, y, z, 1)$  and  $(x, y, z', 0)$  is equal to the coupling between  $(x, y, z, 0)$  and  $(x, y, z', 1)$ ; each is  $\pm J_G/2$ , meaning that their total is  $\pm J_G$ . Said differently, the four spins  $(x, y, z, 0)$ ,  $(x, y, z, 1)$ ,  $(x, y, z', 0)$ , and  $(x, y, z', 1)$  induce an unfrustrated loop.

In the limit  $J_{FM}/J_G \rightarrow \infty$ , FM-coupled dimers can be identified into single “logical” spins, so the model is equivalent to a bimodal spin glass (random bond model) on a standard 3D cubic lattice with open  $x$ - and  $y$ -boundaries and periodic  $z$ -boundaries.

The  $x$ - and  $y$ -dimensions are isotropic with respect to one another; the  $z$ -dimension is anisotropic in the following three ways. First, the pair-to-pair couplings are split between two couplers. Second, the dimension is periodic, while the  $x$ - and  $y$ -dimensions are open. Third, for  $L > 12$ , the  $z$ -dimension has length bounded by 12. This is because the “Pegasus”<sup>80</sup> layout of qubits is roughly described as a  $16 \times 16$  grid of 24-qubit unit cells, with some loss around the boundaries; the 24 qubits are used for 12 2-qubit dimers, forming the  $z$  and  $w$  dimensions. For this reason when studying finite-size scaling, we restrict our attention to  $L \leq 12$ . When studying energy scaling, we use the largest size:  $15 \times 15 \times 12 \times 2$ , with inoperable qubits in the particular processor used leaving only 5374 of 5400 qubits in the inputs.



**FIG. S1. Convergence of mean ground-state energy under the PT-ICM heuristic.** Shown are average energies over 300 spin-glass realizations with  $L = 15$  for a parallel tempering method used to estimate ground-state energies, running for up to  $S \geq 10^6$  sweeps. The asymptote (power-law fit to last 50 data points, shown in black, with limit  $-4669.6$ ) appears to be well established with bias and uncertainty negligible in comparison to the experimentally relevant residual energies.

### B. Broken FM dimers in the 3D ground state

There are differences between the 3D spin glasses studied in this work and the corresponding glasses on the simple cubic lattice that one obtains by contracting each two-qubit dimer into a single spin. In Section III we will discuss the effect on critical exponents, which explains some of the deviation seen between QA values of the KZ exponent  $\mu = z + 1/\nu$  from previously reported values for 3D quantum Ising spin glasses<sup>49</sup>. Here we consider the separate question of how the classical ground states relate to one another between the “embedded” lattice that we study, versus the “logical” counterpart in a simple cubic lattice.

Let us consider a single spin-glass realization, which has an embedded version and a logical version. When  $J_{FM} > 3J_G$ , it is easy to show that no ground state in the embedded problem can have a broken dimer, since flipping a spin to repair a broken dimer (one spin up, one spin down) in such a state would reduce the energy. Likewise, when  $J_{FM} \geq 3J_G$  it is easy to see that the problem must have at least one ground state with no broken dimers. However, when  $J_{FM} < 3J_G$  it is possible that all ground states could have at least one broken dimer.

As an example, we begin with a completely ferromagnetic instance, where all couplings have negative sign. Now, for some fixed  $x_0$ ,  $y_0$ , and  $w_0$ , the spins with coordinates  $(x_0, y_0, z, w_0)$  form an independent set. We flip the sign of all glass couplings incident to this set. It is a simple exercise to confirm that this instance is valid in the sense that it has a well-defined logical counterpart. Furthermore the ground state is twofold degenerate, with broken dimers (with  $x = x_0$  and  $y = y_0$ ) in each ground state. We confirmed this with a computer search for the small case  $L = 4$ .

Therefore when  $J_G > J_{FM}/3$  we cannot guarantee the existence of a ground state of the embedded lattice that maps to a ground state of the logical lattice. However, this construction is pathological, and appears to rely on a specific configuration spanning the periodic  $z$ -dimension. Thus it seems to be highly unlikely in random inputs, and we have not observed such an input in this study.

### C. Estimating ground state energies

To understand mean residual energy in the context of approximate optimization, we must estimate ground state energy—this task is NP-hard for cubic lattices and practically challenging in the typical case for the largest system sizes we consider. Bearing in mind the preceding discussion, we simplify the task by searching for ground states in the logical model, i.e., spin glasses on a simple cubic lattice.

The spin glasses studied here are too large to solve with exhaustive approaches such as branch-and-cut<sup>81</sup>. We instead use a parallel tempering algorithm with isoenergetic cluster moves (PT-ICM)<sup>82</sup> to find putative ground states for each instance, and show by a self-consistent method that any bias introduced by failure to uncover the exact ground state in some subset of instances is negligible with respect to the average residual energies studied in this work.

We employ an adaptive form of the algorithm. The temperature range is initialized so that the lowest and highest chains sample at  $T = \infty$  (spin flip proposals accepted with 50% probability) and  $T = 0$  (no upward energy proposals accepted). Intermediate temperatures are then inserted iteratively, so as to approach replica exchange rates limited

to the range  $[0.2, 0.5]$ . The temperature set stabilizes after several thousand sweeps of all replicas. One sweep entails updating all spins in all replicas in some fixed order. Isoenergetic cluster moves are performed for all temperatures every 10 sweeps. After  $S$  sweeps of the algorithm we can estimate the ground state as the best state observed across the entire algorithm run.

This random estimator produces an  $S$ -dependent upper bound on the ground state  $E_{put}(S)$  in the lowest temperature chain that tightens with  $S$ . For our analysis we require a model average per system size,  $E_{put}(L, S) = [E_{put}(S)]_L$  where square brackets denote our expectation with respect to the set of instances used experimentally. This estimator is subject to a variance and bias, which increase with system size. We can determine confidence intervals by considering many independent instantiations of the algorithm at each  $S$ . The bias is expected to decrease with  $S$  as shown in Fig. S1;  $E_{put}(L, S)$  appears to converge and we can assume the limit  $E_{put}(L) = \lim_{S \rightarrow \infty} E_{put}(L, S)$  with reasonable confidence. The black line in Fig. S1 indicates a power-law fit to the last 50 data points, which has a limit  $E_{put}(15) = -4669.6$ , which would imply that our estimate of the mean ground state energy is correct to within 0.00005 relative error, or 0.00004 energy per site—negligible in comparison to experimental values.

We needn't (and cannot) claim to have found the ground state for every instance, at every size by this method for any finite  $S_{max}$ . We only need convincing evidence that the error in the estimate of ground state energy is negligible compared to the residual energies that we use to compare SA, QA, and SQA.

We can assume  $E_{put}(L, S)$  converges smoothly in  $S$  to obtain a sufficient estimator for our mean residual energy analysis. Assuming a power-law form of residual energy gives one estimate of  $E_{put}(L, S)$ , which describes the data closely (see Fig. S1). In practice, we find it sufficient to approximate

$$E_{put}(L) \approx \min_{\text{independent runs}} E_{put}(L, S_{max}), \quad (\text{S60})$$

the best energy observed amongst 5 independent replicas at some large (but experimentally reasonable) number of sweeps. The convergence  $E_{put}(L)$  is shown in Figure S1 for the most challenging problem scale ( $L = 15$ ), and it is clear the error in the ground state energy across 300 instances is a negligible factor in our analysis.

We emphasize that an error in ground state energy affects the computed residual energy for all dynamics equally.

#### D. Quantum and classical phase transitions

Here we briefly describe the phase diagram for the 3D spin glass shown in Fig. 1e. We first focus our attention on the  $\pm J$  spin glass on a simple cubic lattice. The classical case has been studied in detail using the Janus special-purpose computer<sup>83</sup>. Although it was not initially clear<sup>84,85</sup>, it has been established that there is a finite-temperature transition at  $T_c^{3D} \approx 1.1019^{83}$ . The critical correlation exponent and anomalous dimension have been estimated, respectively, as  $\nu = 2.562$ ,  $\eta = -0.39^{83}$ . In the main text we referenced an estimate of the dynamic exponent  $z = 5.93$  for the  $\pm J$  case<sup>48</sup>, but in that paper<sup>48</sup> (Table I) it is pointed out that a range of estimates has been reported over the last 20 years.

Our picture of the quantum phase transition is mostly informed by a study of Guo, Bhatt and Huse from 1994<sup>49</sup>, which gives estimates of  $z \approx 1.3$ ,  $\nu \approx 1/1.3$ . Making use of relatively recent algorithmic advances—primarily, continuous-time PIMC—we perform similar MC simulations on both the simple cubic lattice and the embedded 3D lattices studied in this work. We discuss these simulations in the next section.

### III. EQUILIBRIUM ANALYSIS OF CRITICAL PHENOMENA BY MONTE CARLO METHODS

As highlighted in the main text, the universal and non-universal critical behavior of 2D and 3D transverse field Ising spin glass models has been studied in previous works<sup>49,83</sup>. In this section we use path-integral Monte Carlo—in the limit of continuous imaginary time—to establish the equilibrium critical behavior of 3D spin glasses both on simple cubic lattices and on the embedded models studied in the main text, as well as 2D square ferromagnets. We generate data for ferromagnetic and spin-glass models at comparable scales to those employed in the QA studies. Our aims in this section are: to verify previously reported critical exponents with an independent numerical analysis; to test the viability of finite-size scaling analysis at the experimented scales; and to verify universality of the embedded 3D spin glasses via the universal correlation exponent  $\nu$ , regardless of microscopic model details. We use a methodology inspired by previous studies of universal behavior, but adapted to allow for a continuous imaginary-time integral limit of the Trotterized model<sup>49</sup>.

### A. Finite-size scaling in classical and path-integral models

As discussed in Section I, the properties of a classical or quantum spin-glass phase transition may be determined by a finite-size scaling analysis of  $\langle q^2 \rangle$ <sup>49,65,86</sup>.

In the classical case we have a model over  $N$ -spin states  $\vec{s}$ :

$$P(\vec{s}) = \exp \left( -\beta \sum_{i < j} J_{ij} s_i s_j \right), \quad (\text{S61})$$

with

$$\langle q^2 \rangle = \sum_{\vec{s}^{(1)}} \sum_{\vec{s}^{(2)}} P(\vec{s}^{(1)}) P(\vec{s}^{(2)}) \left( \frac{1}{N} \sum_i s_i^{(1)} s_i^{(2)} \right)^2. \quad (\text{S62})$$

The anticipated scaling form is

$$\langle q^2 \rangle = L^b B(\delta L^{1/\nu}) \quad (\text{S63})$$

with  $\delta = (T - T_c)/T_c$ .

Critical behavior in the quantum model may be established by analysis of the density matrix  $\exp(-\beta \mathcal{H}(s))$ , where  $\mathcal{H}(s)$  is the Hamiltonian of our experimental Ising system (1) and  $\beta$  is the inverse temperature. By a process of Trotterization this density matrix can be transformed into a classical model, sufficient to establish the distribution of projected states and other statistics. The Trotterized model is defined by a modified Hamiltonian over a space of worldlines, in which each qubit  $i$  is replaced by  $L_z$  time-indexed classical spin-states  $s_{i,t}$ . For each time slice, variables interact in accordance with the problem Hamiltonian, scaled to  $L_z$ . Spins are coupled ferromagnetically in imaginary time, subject to periodic boundary conditions  $s_{i,L_z} = s_{i,0}$ . The probability distribution is given by

$$P(\vec{s}, L_z) \propto \exp \left( \left( -\frac{1}{2} \log \tanh \left( \frac{\beta \Gamma(s)}{L_z} \right) \right) \sum_{t=0}^{L_z-1} \sum_i s_{i,t} s_{i,t+1} - \frac{\beta J(s)}{L_z} \sum_{i < j} J_{ij} s_{i,t} s_{j,t} \right). \quad (\text{S64})$$

An order parameter  $q$  can be defined as the overlap of two states projected into the computational basis, the projected state distribution is determined by the statistics of a single time slice (say  $t=0$ ) in our model<sup>49,65</sup>. The spin-glass susceptibility is defined as the expected square value of this quantity, averaged over spin-glass realizations, i.e.,  $\langle q^2 \rangle$ :

$$\langle q^2 \rangle = \sum_{\vec{s}^{(1)}} \sum_{\vec{s}^{(2)}} P(\vec{s}^{(1)}, L_z) P(\vec{s}^{(2)}, L_z) \left( \frac{1}{N} \sum_i s_{i,0}^{(1)} s_{j,0}^{(2)} \right)^2. \quad (\text{S65})$$

Statistics of the quantum model are correctly established by taking the continuous limit in imaginary time<sup>1</sup>,  $L_z \rightarrow \infty$ . In this limit the scaling of susceptibility is expected to take the form

$$\langle q^2 \rangle = L^{b_q} B(\delta L^{1/\nu}, \beta/L^z), \quad (\text{S66})$$

for a two-parameter collapse function  $B$ . Our analysis of the phase transition involves variation of  $\Gamma = \Gamma(s)/J(s)$  at fixed  $J(s)$ , so that the reduced transverse field is defined as  $\delta = (\Gamma - \Gamma_c)/\Gamma_c$  for a critical field  $\Gamma_c$ .

### B. Monte Carlo sampling

Monte Carlo methods can be used to sample both the classical model (S61) and the Trotterized model (S64). For the latter, we take the numerically convenient value of  $L_z = 2^{16}$ , which is indistinguishable from the continuous-time limit ( $L_z \rightarrow \infty$ ) for the relevant parameter ranges. Since in this limit the coupling strength in the imaginary time direction is strong, it is necessary to employ Swendsen-Wang cluster updates<sup>87</sup>. Our algorithm thus proceeds by

<sup>1</sup> Numerically we work at  $L_z = 2^{16}$  with varying  $\beta$ . Because  $\beta \ll L_z$  in all experiments conducted, this captures the continuous-

time limit.

iterating over sites  $i$ , and applying a cluster update (single-site, many-times) to  $s_{i,\cdot}$  for each  $T$  or  $\Gamma$ . For our study of embedded problems where disjoint pairs of qubits (dimers) are strongly ferromagnetically coupled (with relative strength  $J_{\text{FM}}/J_G$ —see Fig. 1) we can apply a multi-site multi-time update per set; clustering occurs both in space and time with respect to each dimer. The cluster updates are described, and code published, in previous studies<sup>35,88</sup>. Since we are interested only in equilibrium behavior it is sufficient to consider a fixed sequence of updates covering all sites; one such sequence is called a sweep.

For the classical systems, we employ a standard parallel tempering (exchange Monte Carlo) method<sup>83</sup>. In the quantum model we must sample across a larger parameter space, and with a slower worldline sampling update. For efficiency of data collection we replace the parallel tempering routine with an annealing procedure. For a fixed model (disorder realization, at some size  $L$ ) and fixed  $\Gamma$  we vary inverse temperature at each sweep according to a geometric schedule from a low value (where worldlines are fast-mixing) through the critical region to a low-enough temperature for purposes of establishing critical scaling. In this way we collect single-sample (equilibrated, but correlated) data at a wide range of temperatures. Statistics are averaged based upon 400 independent anneals drawing one realization of the bond disorder per anneal. Uncertainty is estimated with bootstraps with respect to the disorder realizations. In both the classical and quantum cases, we can maintain two independent Monte Carlo processes (replica) in parallel, from which the overlaps can be calculated.

In the parallel tempering routine temperatures are spaced inclusive of high temperature fast mixing models and critical region models to ensure sufficient replica exchange rates. Statistics are collected sequentially with exponentially increasing burn-in and sampling times until convergence through the critical region is apparent. In the annealing procedure fair sampling through the critical region requires a sweep rate (sweeps per unit change in  $\log(\beta)$ ) decreasing with system size. In our methodology  $\langle q^2 \rangle$  converges from below, and this process becomes slow for large  $L$ , small  $\Gamma$  and low temperature. We truncated our data collection at small size in part for this reason, and ensured that statistics were for practical purposes insensitive in the data presented to variation by a factor of 4 in the sweep rate (estimators agreeing within confidence intervals at each system size and parameter set presented). The rate of progress,  $d \log(\beta)$  per sweep, was chosen between between  $2 \times 10^{-6}$  and  $7 \times 10^{-4}$ , for the data presented.

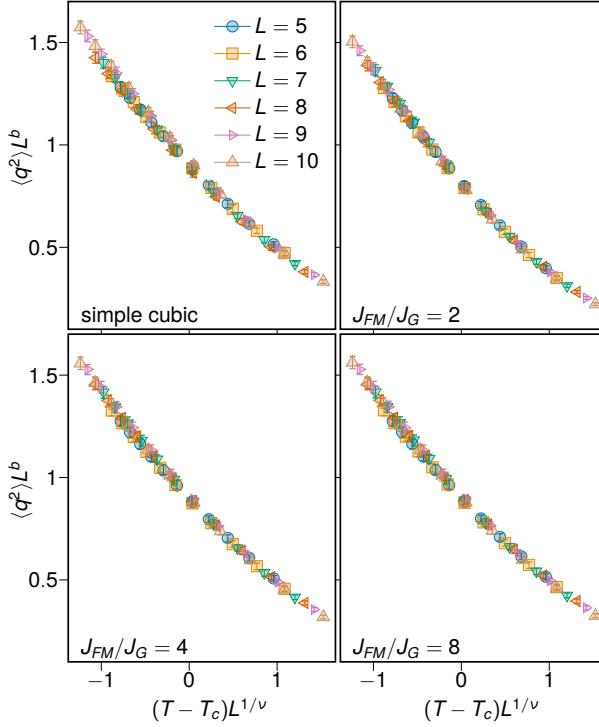
### C. Classical critical scaling

Before showing collapse results in the quantum model, we digress to consider the classical model. In this case we can collect data by a similar methodology but set  $\Gamma = 0$  and employ in independent code a Metropolis method supplemented with cluster flip moves for the dimers. Estimates of  $T_c$ ,  $\nu$ , and  $\eta$  satisfying the classical scaling form (S63) have evolved over the years<sup>83,85,89,90</sup>, with the most recent of these papers<sup>83</sup> estimating  $T_c \approx 1.1$ ,  $\nu \approx 2.56$ , and  $\eta \approx -0.39$ . Earlier estimates of  $\nu$ <sup>85,89</sup> were affected by unaccounted corrections to finite-size scaling, leading to significantly lower estimates of  $\nu$ . In Fig. S2 we show collapses of  $\langle q^2 \rangle L^b$  as functions of  $tL^{1/\nu}$  for  $T_c = 1.1$ ,  $b = 0.61$ , and  $\nu = 1.5$ , for the logical (simple cubic) model and the embedded model with  $J_{\text{FM}} = 2, 4$ , and  $8$ . The value of  $\nu$  is far from recent estimates, but it provides a good collapse for the modest system sizes shown. For larger systems, the appropriate collapse is achieved for  $\nu \approx 2.5$ . Here, the collapses simply illustrate the consistency of critical scaling with respect to the dimer embedding and  $J_{\text{FM}}$ . The evidence is consistent with little or no change in  $T_c$ ,  $\eta$ , and  $\nu$  between the logical and embedded classical 3D spin glass models.

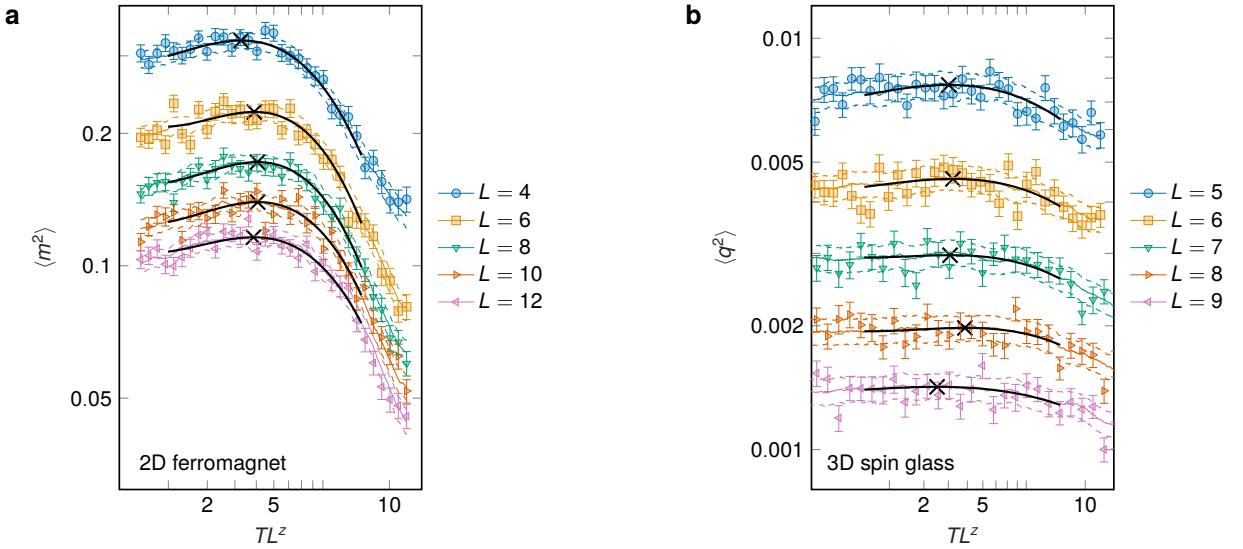
### D. Quantum critical scaling

We can now move on to collapses for the quantum model. An additional difficulty in performing a collapse of data for the quantum model arises from the need to collapse a two-dimensional form (S66). In order to reduce this to a one-dimensional problem we adapt the method of Guo, Bhattacharyya, and Huse<sup>49</sup> to the continuous-time setting. The effect of increasing inverse temperature in our model (S64) is a nonlinear decrease in coupling strength in imaginary time, with a linear increase in the coupling strength between sites. Therefore, at zero temperature the correlation length  $\xi_z$  is much smaller than  $L_z$  in imaginary time, whereas spatial correlation length  $\xi$  approaches zero at high temperature. Between these two extremes there exists—near the critical point—a value for which  $\xi_z \approx L_z$  and  $\xi \approx L$  can be determined. This defines the point of maximum susceptibility which we presume (and show) to be approached smoothly from high temperature; consequently  $\beta \sim L^z$  for this point.

We verify this intuition in a ferromagnetic model, then use it in a spin-glass model, as shown in Fig. S3. Since we can determine by this method a  $\beta$  sufficient for the correlation length in imaginary time to match  $L_z$ , we are able to reduce (S66) to a one-dimensional fit  $\langle q^2 \rangle = L^{b_q} \max_{\beta'} B(\gamma L^{1/\nu}, \beta')$ . As in the classical case we can extract parameters  $\Gamma_c$ ,  $\nu$ , and  $b_q$  from the fit. We used an open-source library, autoScale.py<sup>91</sup>, to fit this data, with the  $J_{\text{FM}} = 2J_G$  example shown in Figure S4. For a goodness of fit  $S$  (with small  $S$  indicating a good fit; see<sup>91</sup>), error



**FIG. S2. Collapse of  $\langle q^2 \rangle$  in the critical region for classical spin glasses.** Finite-size collapses are shown for classical 3D spin glasses on the simple cubic lattice, along with the dimer-embedded lattices with varying  $J_{FM}/J_G$ , using identical parameters  $T_c = 1.1$ ,  $\nu = 1.5$ ,  $b = 0.61$ . The value of  $\nu$  is far from recent estimates but provides a good collapse for these small systems, regardless of embedding, suggesting that dimer-embedding does not cause a large change to  $T_c$ . The largest deviations from the simple cubic model are seen for small  $J_{FM}/J_G$ , as one would expect.



**FIG. S3. Susceptibility as a function of temperature for quantum models.** **a**, 2D ferromagnet ( $\Gamma = 3.06$ ) and **b**, embedded 3D spin glass with  $J_{FM}/J_G = 2$  ( $\Gamma = 2.9$ ). In the case of the ferromagnet, we use the standard linear susceptibility  $\langle m^2 \rangle$  in place of  $\langle q^2 \rangle$ . Near the critical  $\Gamma$ , susceptibility is small at high temperature and reaches a peak value as temperature is decreased. Each model and size has data for 1024 temperatures. Solid colored lines indicate a 64-temperature moving average, dashed lines indicate moving averages of bootstrap 95% confidence intervals, and data with error bars indicate a subset of individual temperatures and bootstrap confidence intervals, selected for visual clarity. Black lines indicate polynomial fits to log-log data, with maxima marked with  $\times$ . Bootstrapping is done over 400 independently annealed realizations (all realizations being identical for the ferromagnet). Rescaling the temperature by  $L^z$  with  $z = 1$  (left) and  $z = 1.3$  (right) leads to reasonably well horizontally-aligned maxima, consistent with previous studies.

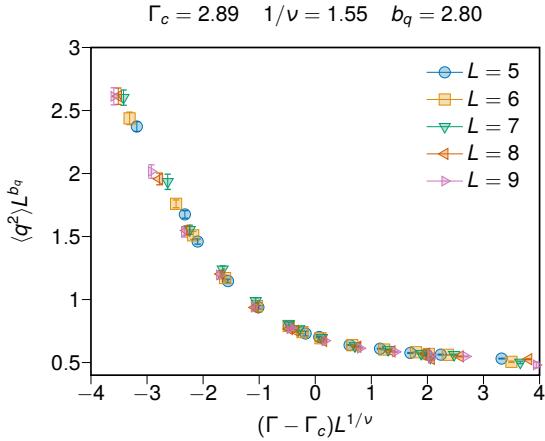


FIG. S4. **Collapse of peak susceptibility for embedded 3D quantum spin glass.** Scaled susceptibility  $\langle q^2 \rangle L^{b_q}$  collapses as a function of reduced transverse field  $(\Gamma - \Gamma_c)L^{1/\nu}$  in the vicinity of the QPT, for the embedded 3D system with  $J_{\text{FM}}/J_G = 2$ . From fitting the collapse we can determine  $\nu$ ,  $\Gamma_c$ , and  $b_q$ . Error bars indicate  $S + 1$  goodness-of-fit intervals (see text).

Model	$\Gamma_c$	$1/\nu$	$b_q$	Goodness-of-fit $S$
2D Spin Glass	2.11(1)	1.02(16)	1.76(3)	2.05
3D Spin Glass	3.02(2)	1.52(15)	2.79(6)	1.01
3D Spin Glass ( $J_{\text{FM}}/J_G = 8$ )	5.07(2)	1.56(12)	2.77(3)	0.84
3D Spin Glass ( $J_{\text{FM}}/J_G = 4$ )	3.75(1)	1.55(6)	2.79(3)	0.87
3D Spin Glass ( $J_{\text{FM}}/J_G = 2$ )	2.89(1)	1.55(11)	2.80(3)	0.99

TABLE I. Optimized collapse parameters for quantum spin glasses, with restriction of data to a rescaled interval  $|\Gamma - \Gamma_c|/\Gamma_c < 1$ . Collapse values are determined by a  $\chi^2$  method<sup>91</sup>. Parameters  $b$  and  $\nu$  deviate slightly downward from expectations based on the literature<sup>49,65</sup>, both for embedded and unembedded models. Aside from a large shift in the non-universal critical point, foreseeable based on modification of local tunneling rates, fit quality is similar across all 3D models.

bars indicate values that lead to goodness of fit at most  $S + 1$ . Outcomes for other models are shown in Table I. We note that all models show  $1/\nu$  near 1.55, significantly different from the previously reported  $1/\nu \approx 1.3$ <sup>49</sup>, and we use this new estimate in the main text.

As shown in Figure S3 the peak susceptibility in  $T$  for a given  $L$  and  $\Gamma$  can be established with reasonable confidence. We expect  $b_q$  to describe the scaling of the peak height in the critical region, i.e.  $\langle q^2 \rangle_{\text{max}} \sim L^{-b_q}$ ; indeed we find very good agreement with this power-law form and find very consistent estimates  $b_q \approx 2.8$  for both logical and embedded 3D quantum spin glasses (Table I). In these simulations we were restricted in the largest value of  $L$  by convergence requirements at low  $T$  and small  $\Gamma$ .

Similarly to the peak height, the peak location (in  $T$ ) is expected to scale as  $L^{-z}$ , so in principle we could use this information to extract an estimate of the dynamic exponent  $z$ . Although the data in S3b are consistent with the reported value of  $z \approx 1.3$ <sup>49</sup>, the data are not reliable enough to confidently estimate  $z$ , owing to the flat and noisy nature of the peak for large systems. We note that in a random ferromagnetic transverse field model study based on a similar methodology, an alternative collapse form in temperature was successful<sup>77</sup>. If a similar *infinite randomness fixed point* applied in the spin glass model, an exponential collapse form might prove valuable, but we were unable based upon our data to establish  $z$  with confidence based on such a hypothesis. (Our collapses use a peak susceptibility value, and hence are robust with regards the scaling form in temperature, a strength of the method.) We maintain the hypothesis that  $z \approx 1.3$ , which despite the challenges of classical simulation, agrees well with our quantum simulation data from the QA processor. The observed value of  $b_q \approx 2.8$  differs from Guo, Bhatt, and Huse's estimate  $b_q \approx 3.4$ <sup>49</sup>.

In general, the FSS analysis appears to provide a good qualitative description of the quantum models at these small scales. As in the classical case, collapses indicate that the scaling of the embedded models is very close to that of the unembedded model. For these unembedded models we find values that are—with the exception of  $\eta$ —not far from those reported in the literature, but with important deviations in  $\nu$ , where we use our estimate  $1/\nu \approx 1.55$  rather than the previous estimate  $1/\nu \approx 1.3$ . Differences in methodology may account for these deviations.

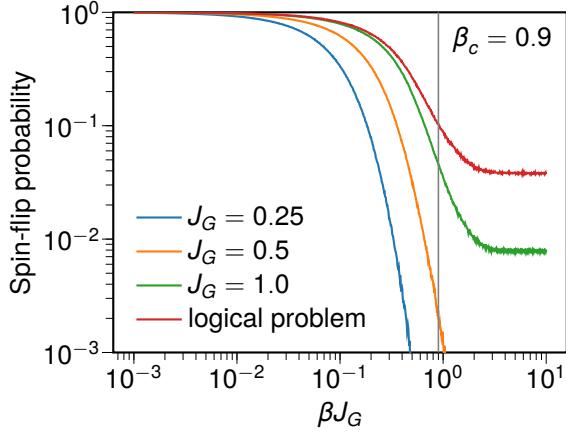


FIG. S5. **Freezing in SA.** Spin-flip acceptance probability varies as a function of  $J_G$ , resulting in a freezing of dynamics at high temperature for small  $J_G$ . Three values of  $J_G$  are shown, along with data for the logical model on a simple cubic lattice.

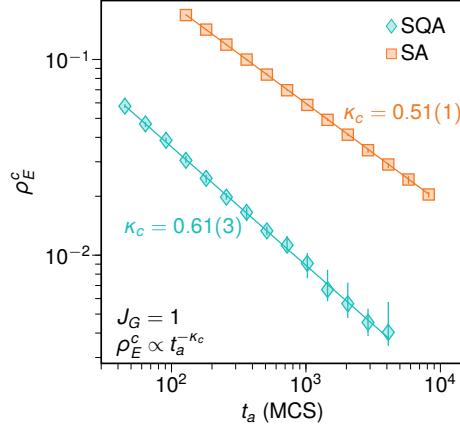


FIG. S6. **Scaling of residual Ising energy density at the critical point.** Results shown are analogous to Fig. 4a, with SQA and SA anneals stopped at their respective critical points. SQA has  $\kappa_c = 0.61(3)$  with equilibrium Ising energy per site  $\approx -0.63$ , whereas SA has  $\kappa_c = 0.51(1)$  with equilibrium Ising energy per site  $\approx -1.79$ . Error bars indicate 95% bootstrap confidence intervals.

#### IV. NONEQUILIBRIUM DYNAMICS OF SA AND SQA

##### A. Freezing of dynamics in SA

The SA dynamics studied here proposes and accepts or rejects one spin update at a time, and proceeds along a geometric schedule in  $\beta/J_G$ . When  $J_{FM} \gg J_G$ , dynamics “freeze”—spin flips become very unlikely to be accepted—at a high temperature. To probe this phenomenon, we annealed  $L = 9$  embedded 3D spin glasses with  $J_G \in \{1/4, 1/2, 1\}$ , as well as the corresponding logical (simple cubic) spin glasses, and tracked the probability of a spin flip being accepted at each sweep. Data are shown in Fig. S5.

For sufficiently long anneals and large systems, this effect should not change extracted KZ exponents. However, there is a significant effect for the system sizes and anneal lengths probed in the main text, as evidenced by Extended Data Figs. E6—E8.

##### B. Energy decay at the critical point and inside the glass phase

As discussed in Section I, the form of energy decay at the critical point is subject to corrections within the glass phase, and it is not obvious that Eq. (12) should hold even approximately. Here we investigate this question.

In Fig. 4, mean Ising energy at the end of the anneal for QA, SA, and SQA is compared to the putative ground-state energy, giving a final residual Ising energy density  $\rho_E^f$ . To study energy decay at the (quantum or classical) critical point, we stop the anneal at the critical point—here this is  $T_c = 1.1$  and  $\Gamma_c/J = 2.89$  for SA and SQA respectively, with  $J_G = 1$  (see Section III). We take the mean Ising energy for a given anneal time  $t_a$ , and subtract off the equilibrium Ising energy to obtain a critical Ising residual energy density  $\rho_E^c$ .

Here, the equilibrium Ising energy is estimated as a fitting parameter to a best-fit power-law form  $\rho_E^c \propto t_a^{-\kappa_c}$ . This fit is shown in Fig. S6. Based on the correction to Eq. (12), we may expect both SA and SQA to show  $\kappa_f < \kappa_c$ . Indeed, for SA we see  $\kappa_c \approx 0.51$ ,  $\kappa_f \approx 0.42$ . For SQA we see  $\kappa_c \approx 0.61$ ,  $\kappa_f \approx 0.51$ . We note that the equilibrium Ising energy per site is estimated at  $-0.63$  at the quantum critical point for SQA, much higher than the value  $-1.79$  at the thermal critical point for SA. Eq. (12) predicts  $\kappa_c$  to be very close to the measured  $\kappa_f$  rather than the measured  $\kappa_c$ ; we attribute this mainly to finite size effects.

Although we are not able to projectively read out QA states at the critical point, this evidence suggests that one might expect QA to show  $\kappa_c > \kappa_f$ , more in line with Eq. (12).

## V. DATA COLLAPSE DOMAINS

All dynamics in this work are studied on the same geometries with the same boundary conditions (one periodic dimension, 2 open dimensions). For this reason collapses are ideally performed over a range of  $t_a$  such that the observable, e.g.,  $U$ , follows a consistent power-law scaling. Thus we discard fast annealing times that deviate from this scaling. We also discard long annealing times where correlation length approaches system size, since this can lead to anomalous boundary effects. This window of  $t_a$  varies from one dynamics to another. In particular, we restrict our collapse for QA to the region  $t_a \leq 30$  ns to minimize complications arising from decoherence and noise, which causes a smooth increase in the observed KZ exponent  $\mu$ .

In Fig. S7 we plot  $\mu$  for QA, SA, and SQA across sliding windows of  $t_a$ . These windows are chosen to span a given dynamic range in  $t_a$ :  $t_{\min} \leq t_a \leq 3t_{\min}$  for QA and  $t_{\min} \leq t_a \leq 6t_{\min}$  for SA and SQA; using different dynamic ranges is justified due to the polynomial speedup in QA compared to the software solvers. Only windows containing at least 6 measured annealing times are considered. As explained in the Methods section, confidence intervals are generated from combined jackknife standard errors across system sizes and annealing times. The confidence intervals for the individual windows overlap the confidence intervals for the overall region of data collapse; annealing times that are deemed too fast or too slow are shown in gray.

There are several potential causes for the observed increase in  $\mu$  for increasing  $t_a$  in QA. Decoherence and noise, as mentioned above, are leading causes. Deviation between the rf-SQUID flux qubit model and the transverse-field Ising model may also play a role.

## VI. EFFECT OF TEMPERATURE ON QA

Measurements in the main text were collected on a first quantum processing unit (QPU1) at a cryostat set-point of 12 mK. To probe the effect of temperature, we performed measurements on a second QPU (QPU2) at temperatures varying from 12 mK to 21 mK. KZ exponents  $\mu$  were extracted using annealing times ranging from 8.1 ns to 30 ns, with results shown in Fig. S8. Measurements are insensitive to temperature over this parametric range, supporting a hypothesis of a coherent regime with negligible interaction with the environment.

## VII. CALIBRATION REFINEMENT

Symmetries in the Ising Hamiltonian provide an opportunity to suppress calibration imperfections. This has been shown to be very effective for geometrically-frustrated low-dimensional systems<sup>35,36,88,92,93</sup>. In this work we study thousands of spin-glass realizations, and it is impractical to extensively refine the calibration for each one. Instead, we tune only two aspects of the calibration: First, we balance qubits at degeneracy—with average magnitude zero—using flux offsets. Second, we synchronize the eight annealing lines that control the annealing schedule of eight sets of qubits, using anneal offsets. The latter is most relevant for the fastest anneals, since desynchronization between annealing lines is on the order of 1 ns or less. For both of these refinements we use the same approach as was taken in Ref.<sup>32</sup>, without tuning individual couplings; we refer the interested reader to the supplementary material of Ref.<sup>32</sup> for more detail. For each selection of parameters ( $L$ ,  $J_G$ ,  $t_a$ , and  $p_{\text{AFM}}$ ) we perform an independent iterative shim for both flux offsets and anneal offsets; each of these offsets is programmable on a per-qubit basis.

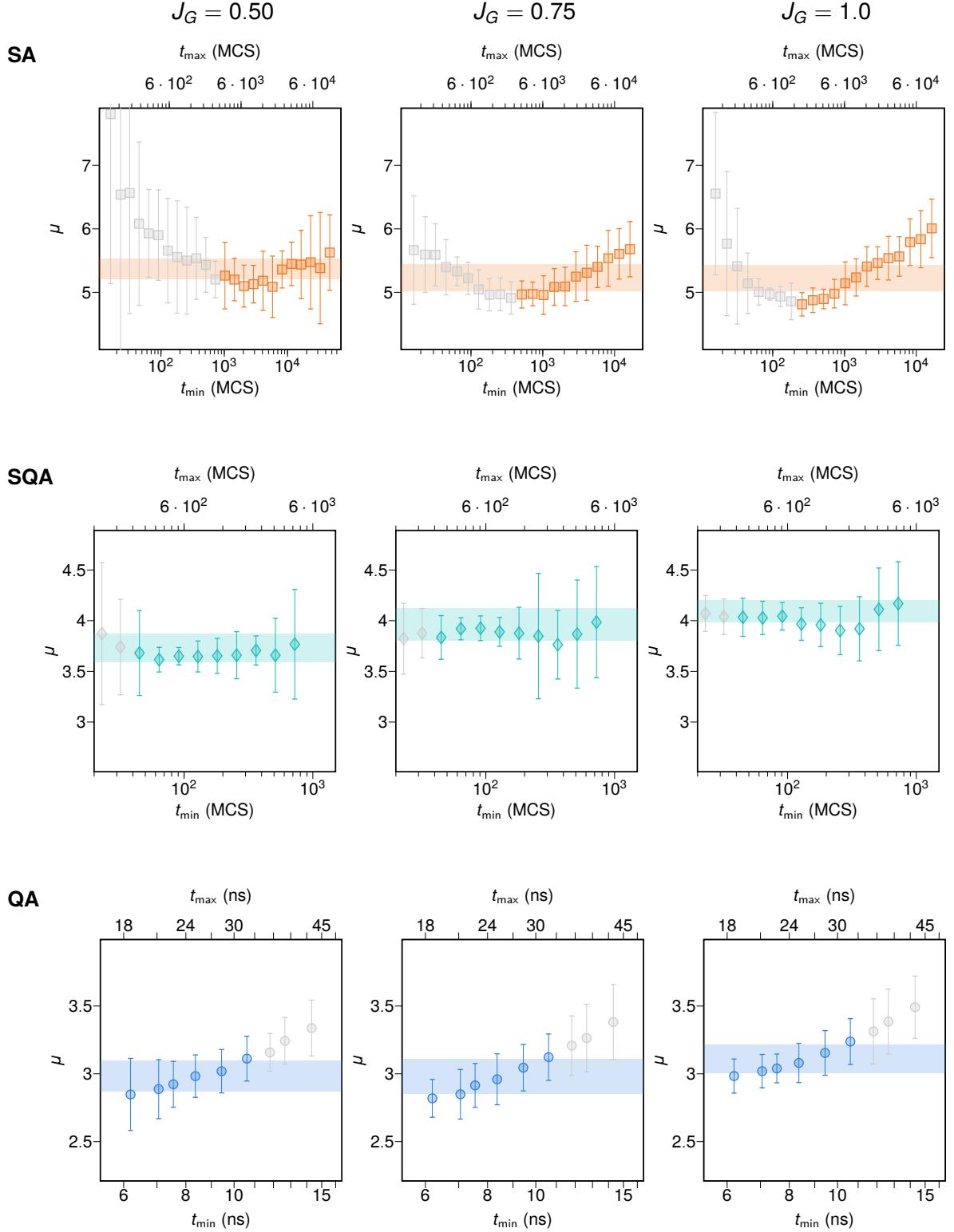
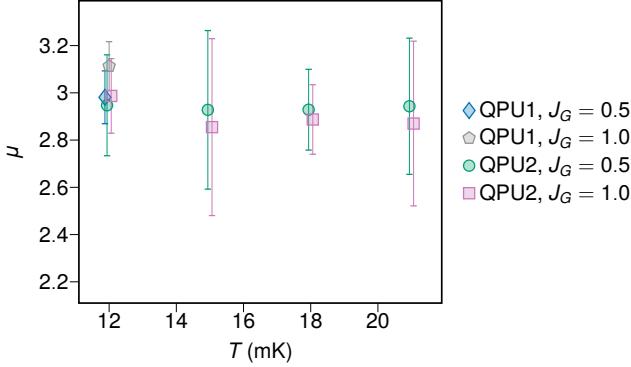


FIG. S7. **Data collapse windows.** Each dynamics (SA, SQA, and QA) is run for varying annealing times; cumulants and order parameters are collapsed over a subset of these times. For each dynamics and each  $J_G \in \{0.50, 0.75, 1.0\}$ , the 95% confidence interval for the extracted exponent  $\mu$  (Fig. 3e) is shown as a shaded region. To test self-consistency, we extract  $\mu$  for annealing time windows  $[t_{\min}, kt_{\min}]$  where  $k = 3$  for QA,  $k = 6$  for SA and SQA (see text). Windows within the collapse range used to determine  $\mu$  are shown in color; windows intersecting unused annealing times are shown in gray.



**FIG. S8. Kibble-Zurek exponents for varying temperature.** Main-text measurements on QPU1 at 12 mK are compared against a second QA processor, QPU2, operating at a range of temperatures. Error bars indicate jackknife 95% confidence intervals over system sizes and anneal times.

For  $i \in \{1, \dots, 8\}$  let  $V_i$  denote the set of qubits on annealing line  $i$ , and for  $i, j \in \{1, \dots, 8\}$  let  $E_{ij}$  denote the set of couplings coupling a qubit in  $V_i$  to a qubit in  $V_j$ .

We perform the shim based on two assumptions:

- All qubits should have average magnetization zero.
- The average frustration probability of a coupler in  $E_{ij}$ —that is, the average observed probability of a nonzero coupler between  $V_i$  and  $V_j$  having a positive contribution to the energy, where the average is taken over both samples and realizations—should be effectively independent of the choice of  $i$  and  $j$ .

The first assumption is trivially justified because there are no longitudinal fields used in the Ising Hamiltonian in this work. For the second, we assume that the sets  $E_{ij}$  are large and sufficiently spatially uncorrelated from the position in the 3D lattice position. This assumption is reasonable because the annealing line assignments follow a regular geometric pattern and the 3D lattice embeddings are determined with a heuristic random approach.

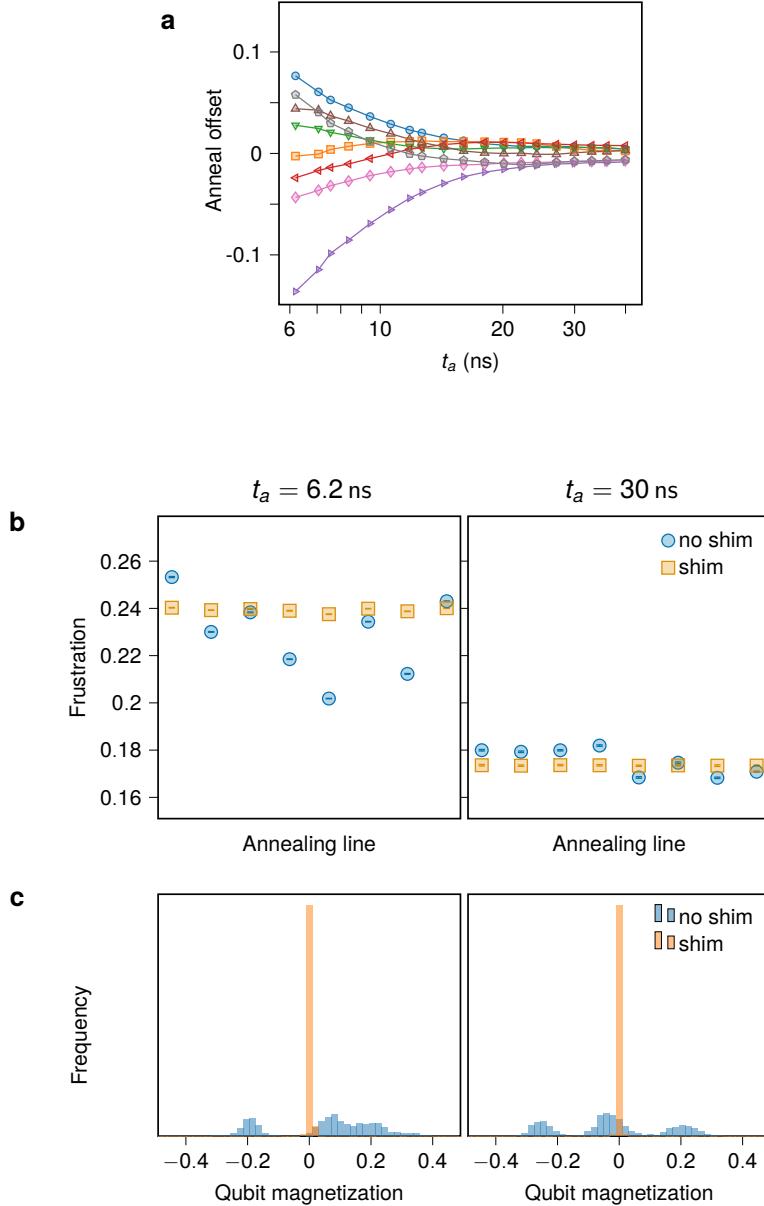
In Fig. S9a we show the final anneal offsets for the eight annealing lines after 1200 iterations on  $L = 12$  3D spin glasses. In Fig. S9b we show the distribution of average frustration in a glass coupling incident to each annealing line. Data are shown with and without the anneal offset and flux offset shim, for  $t_a = 6.2$  ns and 30 ns. The calibration refinement shim has a clear homogenizing effect on the per-line frustration. In Fig. S9c we show the average magnetization of each qubit over the final 300 of 1200 iterations with and without the shim, which has a clearly beneficial effect in balancing the qubits at zero magnetization.

In Fig. S10 we show the effect of calibration refinement on the dynamic scaling of the order parameter and Binder cumulant. These data make it clear that the calibration refinement is necessary to obtain reasonable estimates of critical exponents. However, we are focusing on a regime far outside the specifications of the calibration being refined, which is intended for  $t_a \geq 500$  ns.

### VIII. RESIDUAL ENERGY DECAY

In Fig. S11 we show decay of residual energy for three dynamics: QA, SA, and SQA. For the MC solvers (i.e., SA and SQA), we show this both in terms of MCS and in terms of computation time. We measured times per MCS on a CPU (Intel® Core™ i7-7700HQ CPU @ 2.80GHz): 0.4 ms for SA and 8.5 ms for SQA. We call out two caveats:

- SQA time per sweep is approximately linearly dependent on  $\beta$ , and we have used a high value  $\beta = 64$  throughout (relative to crossing point  $\Gamma(s) = J(s)$ ; see Methods), to minimize thermal effects. However, even with a lower  $\beta$ , SQA is not competitive with SA on the systems studied here.
- The codes used are reasonably fast but are written to be general, without optimizations such as lattice-specific memory structure, function lookup tables, static spin ordering for sweeps (which deviates from standard interpretations of quasi-physical dynamics), and random number reuse. SA in particular can be sped up significantly, so we show SA annealing time using the quoted time of 2.42 spin flips per nanosecond (0.4 ns per spin flip) for a highly optimized version of SA measured on an Intel® Xeon™ E5-2670 CPU @ 2.60GHz (94, Table 4, `an_ss_ge_fi`).



**FIG. S9. Calibration refinement.** **a**, Anneal offsets (unitless, in  $s$ ) are tuned to synchronize eight annealing lines for fast anneals. **b**, Anneal offsets are learned through a loss function related to inhomogeneity of frustration of couplings incident to the qubits on each annealing line. Flux offsets are learned through a loss function that minimizes nonzero average magnetization of individual qubits. Tuning these improves the homogeneity of frustration with respect to annealing line. **c**, Systematic nonzero magnetization for the qubits is also reduced.

As previously noted in Section II C, the ground state energies of these “embedded” spin glasses, which use two-qubit ferromagnetic dimers, can generally be found by instead solving a reduced “logical” problem on a simple cubic lattice. Therefore we also compare QA performance on the embedded spin glasses against MC dynamics solving the logical spin glasses. Although the primary aim of this work is to study quantum critical spin-glass dynamics, it is notable that QA outperforms SA—both in scaling (in the coherent QA regime) and in absolute terms—even when QA has an embedding overhead that does not affect SA.

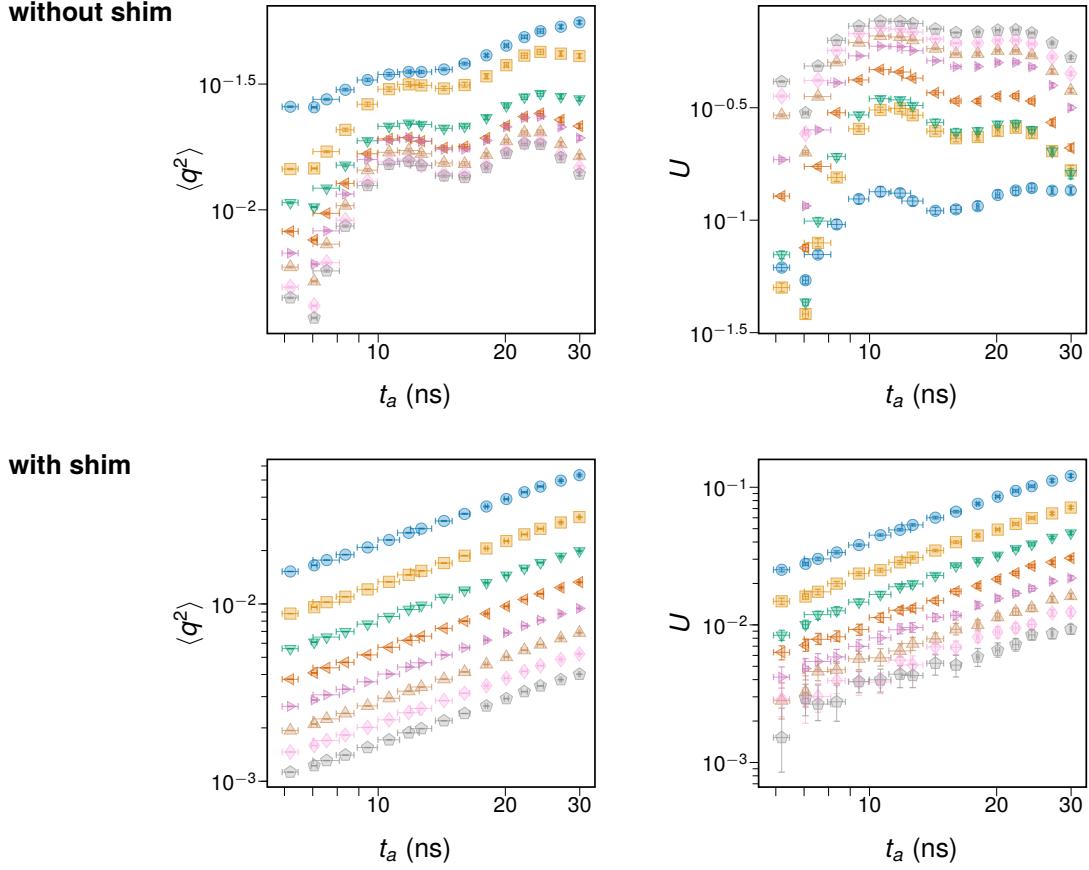
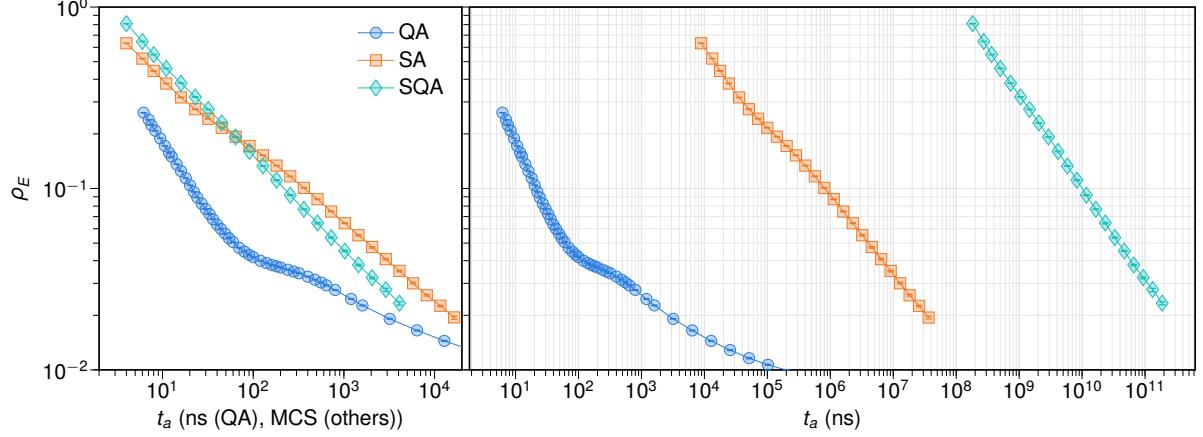
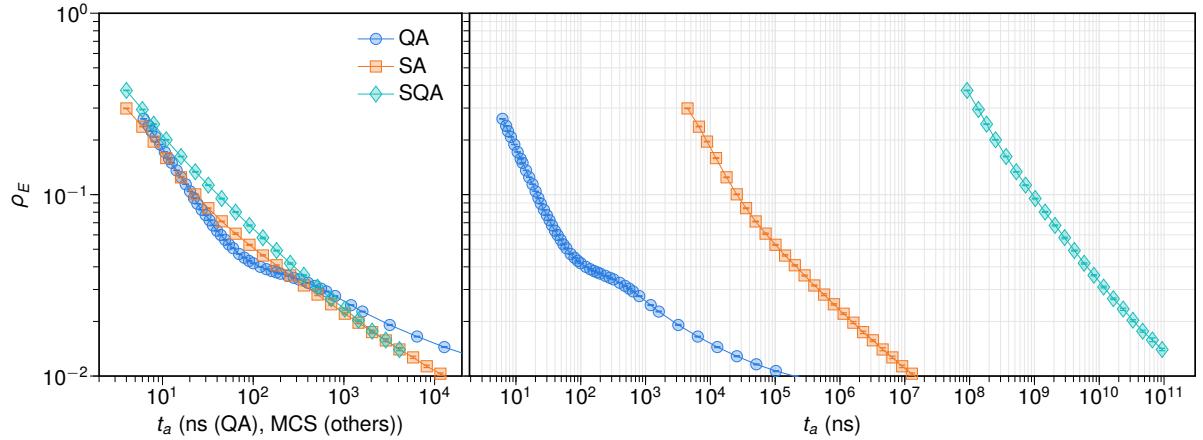


FIG. S10. **Effect of shim on dynamic scaling.** Shown in the top and bottom rows, respectively, are Binder cumulant  $U$  and order parameter  $\langle q^2 \rangle$  with and without the anneal offset (synchronization) and flux offset (balancing) shim. Data are for 3D lattices,  $J_G = 0.5$ .

**Monte Carlo dynamics solve same embedded 3D spin glass as QA**

**Monte Carlo dynamics solve reduced logical problem on simple cubic lattice**


**FIG. S11. Residual energy decay.** We compare  $\rho_E$  as a function of  $t_a$  for QA, SA, and SQA. The top row shows data for MC dynamics running on the same Ising model as QA, with two-qubit FM dimers. The bottom row shows the same data for QA, but with MC dynamics running on a reduced Ising model on a simple cubic lattice, in which two-qubit FM dimers have been contracted into single logical variables, reducing both time per sweep and the number of sweeps required to reach a given  $\rho_E$ . The right-hand plots show annealing time in nanoseconds. For SQA we use measured sweep times; for SA we use a time of 0.4 ns per spin update as reported in Ref.<sup>94</sup> for a highly optimized SA code. Our SQA timescales rely on worldline updates, for which open-source code is provided<sup>88</sup>.