nature physics

Article

Nishimori transition across the error threshold for constant-depth quantum circuits

Received: 13 October 2023

Accepted: 7 October 2024

Published online: 16 December 2024

Check for updates

Edward H. Chen[1,2] ✉, Guo-Yi Zhu[3] ✉, Ruben Verresen[4] ✉, Alireza Seif[5], Elisa Bäumer[6], David Layden[1,5], Nathanan Tantivasadakarn[4,7], Guanyu Zhu[5], Sarah Sheldon[5], Ashvin Vishwanath[4], Simon Trebst[3] & Abhinav Kandala[5]

Quantum computing involves the preparation of entangled states across many qubits. This requires efficient preparation protocols that are stable to noise and gate imperfections. Here we demonstrate the generation of the simplest long-range order—Ising order—using a measurement-based protocol on 54 system qubits in the presence of coherent and incoherent errors. We implement a constant-depth preparation protocol that uses classical decoding of measurements to identify long-range order that is otherwise hidden by the randomness of quantum measurements. By experimentally tuning the error rates, we demonstrate the stability of this decoded long-range order in two spatial dimensions, up to a critical phase transition belonging to the unusual Nishimori universality class. Although in classical systems Nishimori physics requires fine-tuning multiple parameters, here it arises as a direct result of the Born rule for measurement probabilities. Our study demonstrates the emergent phenomena that can be explored on quantum processors beyond a hundred qubits.

Traditionally, measurements have been synonymous with extracting information from physical systems. Yet, in the quantum realm, the extraordinary nature of measurements allows them to actively modify and steer quantum states, forging a new route to entanglement generation. Among the more interesting entangled states are those with long-range correlations[1–3]; however, these cannot be prepared by any constant-depth unitary circuits, making them more sensitive to the finite coherence times of current quantum processors[4–6]. In contrast, recent theoretical studies have shown that the use of measurements, which are non-unitary operations, can be used to efficiently create quantum states with long-range order[7–9] and critical quasi-long-range order[10–12]. In essence, measurement-based approaches trade off circuit depth for number of mid-circuit measurements and operations[13] as compared with exclusively unitary approaches.

In this work, we study such measurement-induced long-range order and criticality. In particular, we consider the 'hydrogen atom' of long-range entangled states, the Greenberger–Horne–Zeilinger (GHZ) state $|\mathrm{GHZ}\rangle \propto |00\cdots00\rangle + |11\cdots11\rangle$, which can be thought of as one representative of a more general 'Ising' phase of matter. A necessary condition for realizing a GHZ state is a long-range Ising order that organizes the individual qubits into a macroscopic state. While recent experiments show the practicality of measurement-based protocols to create such Ising-like order in one-dimensional (1D) qubit geometries where stability is not guaranteed[14], theoretical works suggest that this order should be robust against a range of imperfections when using a

[1]IBM Quantum, Almaden Research Center, San Jose, CA, USA. [2]IBM Quantum, Research Triangle Park, NC, USA. [3]Institute for Theoretical Physics, University of Cologne, Cologne, Germany. [4]Department of Physics, Harvard University, Cambridge, MA, USA. [5]IBM Quantum, T. J. Watson Research Center, Yorktown Heights, NY, USA. [6]IBM Quantum, IBM Research, Zurich, Switzerland. [7]Walter Burke Institute for Theoretical Physics and Department of Physics, California Institute of Technology, Pasadena, CA, USA. ✉e-mail:

https://doi.org/10.1038/s41567-024-02696-6

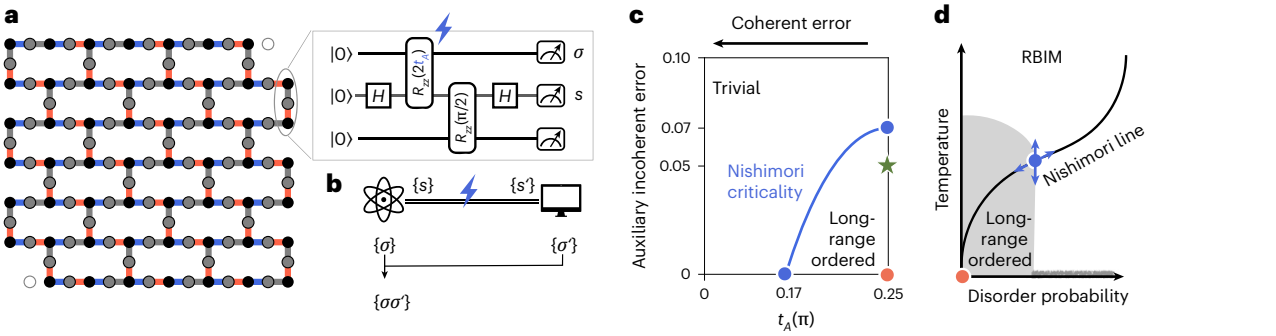


Fig. 1 | Circuit protocol, decoder and phase diagram under coherent and incoherent errors. a, The heavy-hexagonal lattice of 127 qubits. For the 125 active qubits, the inset shows the building block using constant-depth entangling circuits for the three nearest neighbours (grey circles) of each system qubit (black circles) in the presence of noise (blue lightning). The R_{ZZ} gates are executed in order from blue, red then grey bonds within three layers. The auxiliary outcomes, s, on the bonds of the lattice (grey) can be used to inform a decoder for the data outcomes, σ, on the vertices of the lattice (black). **b**, The quantum device outputs a data bit-string $\{\sigma\}$ together with an auxiliary outcome $\{s\}$. In the presence of noise, the auxiliary outcomes become $\{s'\}$ before being passed to a classical decoder to determine a classical replica of the bit-string $\{\sigma'\}$. Their element-wise product, $\{\sigma\sigma'\}$, serves as the decoded bit-string. A measurement error (blue lightning) can corrupt the communication channel between the quantum replica and the classical replica. **c**, The trivial and long-range ordered phases sweep out distinct regions depending on the strength of coherent and incoherent noise. Within a finite threshold, a stable phase (grey), of which the GHZ state is a special case (red circle), exhibits long-range entanglement in the absence of other sources of noise (for example, without dephasing). Even in the presence of dephasing (not shown), classical long-range ordering remains. The boundary separating the trivial and long-range phases is described by the Nishimori criticality. Our experiments have incoherent error rates as low as ~0.05, which is indicated by the green star. **d**, A schematic phase diagram of the classical RBIM. The solid black line is the Nishimori line, which captures the entire phase diagram in **c**.

two-dimensional (2D) protocol[ ]. Here, we implement this 2D protocol on a superconducting qubit processor and, by tuning particular imperfections, we experimentally create a critical ensemble of these long-range ordered states in agreement with theoretical predictions for their stability.

The unavoidable randomness of quantum measurements generates a 'glassy' version of the sought-after long-range Ising order, for example, $|00110\rangle + |11001\rangle$, requiring some form of decoding to tame the structured randomness. This makes it crucial to record the measurement outcomes and then use either post-selection, feed-forward or post-processing to recover the long-range order. In our setup, we implement post-processing to decode the hidden long-range order and determine the decoding threshold beyond which the order is unrecoverable[ ]. This decoding threshold is where our quantum system exhibits a Nishimori transition, or criticality[ ], for both incoherent[ ] and coherent errors[ ]. We argue that the observed Nishimori criticality is, in fact, unavoidable in our protocol and a natural consequence of Born's rule—a striking distinction from materials studies in labs seeking to observe the Nishimori criticality only by fine-tuning disorder within the material against environmental temperatures.

Protocol and device operation

In our protocol, we divide the qubits on our heavy-hexagonal device into system qubits on the vertex 'sites', and auxiliary qubits on the 'bonds' of a honeycomb lattice (Fig. [ ]). We will refer to the Pauli matrices on each qubit as X, Y, Z. To turn an initial product state of system qubits in $+1X$ eigenstates into a GHZ-type state, we measure the ZZ parities on all nearest-neighbour system (site) qubits, using the auxiliary qubit in between. If the auxiliary outcome is +1, it means the two spins are perfectly anti-ferromagnetic in the −1 eigenstate of ZZ. A crucial element of our protocol is that we implement a coupling to the auxiliary qubit beyond a simple Clifford CNOT gate by an $R_{ZZ}(2t_A) = e^{-it_A ZZ}$ rotation with a control parameter $2t_A$, for the A sublattice (Fig. [ ]). By varying t_A away from π/4 (the Clifford limit) and controling the level of coherent errors, we can perform tunable weak measurements described by Kraus operators

$$K_{ij}(s) = e^{-\frac{\beta}{2} s_{ij} Z_i Z_j} \Big/ \sqrt{2\cosh(\beta)}, \qquad (1)$$

where $\beta = 2\tanh^{-1}\tan(t_A)$ (ref. [ ]). Due to the degree-3 connectivity of the system qubits, we need to repeat the above coupling only three times before simultaneously measuring all the auxiliary qubits—resulting in a constant-depth circuit independent of the number of qubits.

The measurement outcomes of the auxiliary qubits in the X basis, denoted by $s_{ij} = \pm 1$ for each bond $\langle ij \rangle$, are then fed as syndromes to the decoder, operated on a classical computer. As shown in Fig. [ ], the decoder produces an estimate of the quantum sample based on its limited knowledge in the form of $\{s'\}$ (refs. [ ]), where $\{s'\}$ is a copy of $\{s\}$ corrupted by a finite probability p_s of noise that can come from either the quantum device or the classical communication channel with probability distribution: $P(s') = p_s \delta_{s',-s} + (1-p_s)\delta_{s',s}$, where $\delta(\cdot)$ is the Kronecker delta function. We employ a fast decoder[ ] that outputs a bit-string $\{\sigma' = \pm 1\}$ as a classical ground state for each $\{s'\}$. By denoting the bit-string of the system qubits measured in Z basis as $\{\sigma = \pm 1\}$, the element-wise product, $\{\sigma\sigma'\}$, between the quantum sample and the classic replica serves as the decoded bit-string that lower-bounds the nonlinear correlations of the quantum samples[ ]. This is equivalent to correcting the system qubits by one layer of X gates for those sites with $\sigma' = -1$ in a feed-forward manner.

We performed experiments on ibm_sherbrooke, which is one of the IBM Quantum Eagle processors with 127 qubits; entangling gates generated by echoed cross-resonance (ECR) interactions[ ] had typical error rates of 0.0077 and square root of Pauli-X gates with error rates of 0.0002 (Extended Data Fig. 1). The typical device measurement error rates of 0.010 were sufficiently below the decoding threshold needed for the preparation of the long-range ordered state.

Results

Conceptual understanding of protocol

In previous theoretical work by some of the present authors[ ], it was shown that deviations from the Clifford limit by coherent errors induced by $t_A < \pi/4$ are tolerable up to a finite threshold. Here, we expand this perspective by also treating incoherent errors (corrupting the syndromes) in an analytically exact manner and show that the presence of both types of error leads to a threshold line as shown in Fig. [ ], which in its entirety is captured by the Nishimori criticality. To see this, let us consider measuring the auxiliary qubits, which collapses the system's wave function into an unnormalized wave function

https://doi.org/10.1038/s41567-024-02696-6

$$|\psi(s_{ij})\rangle = e^{-\frac{\beta}{2}\sum_{\langle ij\rangle} s_{ij} Z_i Z_j} |+\rangle^{\otimes N} \Big/ (2\cosh(\beta))^{N_b/2}, \quad (2)$$

where $N(N_b)$ denotes the number of system (auxiliary) qubits. The probability of such a measurement outcome follows from Born's rule

$$P(s_{ij}) = \| \psi(s_{ij})\|^2 \propto \sum_{\sigma} e^{-\beta \sum_{\langle ij\rangle} s_{ij}\sigma_i\sigma_j}, \quad (3)$$

which resembles the thermal partition function of the random bond Ising model (RBIM) where the measurement outcomes play the role of disorder. The measurement noises distort the observed measurement records, effectively increasing the disorder. Concretely, by equation ( ) we analytically map our protocol onto a RBIM. The experimental sweep of t_A tracks the Nishimori line along an effective disorder probability $\tilde{p}$ deduced by (Supplementary Section IA)

$$1 - 2\tilde{p} = \sin(2t_A)(1 - 2p_s), \quad (4)$$

as a joint action of both coherent and incoherent errors. Such errors drive the phase transition across the blue line in Fig. , mapped out by equating $\tilde{p}$ with the critical disorder probability 6.75% for RBIM on honeycomb lattice (for square lattice, it is replaced by 10.9% (ref.)). In particular, this implies that every point in the extended transition line shares the same Nishimori criticality, interpolating between the limit $p_s = 0$ (ref.) and the limit $t_A = \pi/4$ (ref.). This scenario for the quantum protocol is quite distinct from the classical RBIM, whose schematic phase diagram is shown in Fig. , where the Nishimori line only occurs at the fine-tuned solid line—demonstrating an unprecedented robustness of Nishimori criticality in the quantum case.

GHZ fidelity in Clifford limit

For a baseline characterization of the measurement-based protocol, we estimated the fidelity of the prepared states in the Clifford limit ($t_A = \pi/4$) relative to the GHZ state. Because the final state in this limit is a stabilizer state, it was sufficient for a desired accuracy to consider only a constant number of randomly sampled measurements of the system qubits . For the specific case of the GHZ state, half the sampled stabilizers contain only Pauli-Z operators, while the other half are combinations of Pauli-X and Pauli-Y operators (see equation ( ) in Methods). To assess the relative performance of our protocol, we also implemented a standard unitary protocol for constructing GHZ states . In Fig. , we see that the fidelities of the measurement-based protocol outperformed the unitary preparation. This can be rationalized by the latter experiencing more errors due to the long idle times of deep circuit with size-dependent depth between $O(N)$ and $O(\log(N))$.

For a system of ten qubits, the measurement-based protocol resulted in a GHZ fidelity above 50%, but with increasing system size the fidelity was found to decrease exponentially (Fig.). We note, however, that this does not imply the absence of long-range order or entanglement for these larger systems. In fact, we expect exponentially decaying GHZ fidelities versus system sizes in the presence of noise for virtually all states in the same phase of matter. We emphasize that no form of error mitigation, for measurement or unitary gates, was used to estimate these fidelities. To explain the experimentally measured fidelities, we compared our results against the predicted fidelities based on a noise model with ~5% incoherent auxiliary errors and ~3% data readout errors—values inferred in the next section. This places us in the long-range ordered phase in Fig. (green star), which in the absence of any additional errors, has long-range GHZ-type entanglement, while its predicted GHZ fidelity shown in grey in Fig. decays exponentially with the number of system qubits. We see that the experimentally obtained values are slightly suppressed with respect to the grey curve, which is probably due to dephasing. This raises the question whether we retain robust long-range order in the presence of such dephasing.

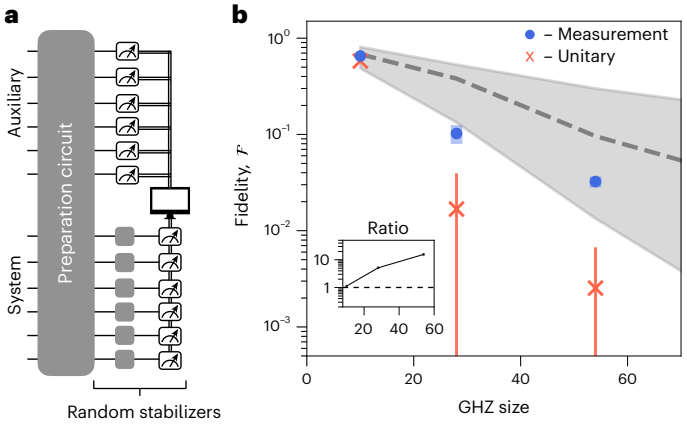


Fig. 2 | Decoded fidelity estimation by randomly sampling GHZ stabilizers. **a**, Because our decoder was implemented as Pauli corrections on the system qubits, the characterization of random stabilizers, which is measured in bases rotated by single-qubit rotations (small grey boxes), needed to be done in conjunction with the implemented decoder (symbolized by the monitor). **b**, Estimated fidelities relative to GHZ states for measurement-based (filled blue circles) and unitary-based (red X marks) preparation of long-range Ising-ordered states on two dimensions. The error bars represent the standard deviation of the fidelities estimated from bootstrap resampling random sets of stabilizers. For 10, 28 and 54 qubits, we sampled up to 46 (105), 111 (111) and 147 (148) stabilizers for measurement (unitary)-based estimation (Monte Carlo sampling discussed in Methods and Supplementary Section IIB). The error bars for the unitary results (red) are comparable to the fidelity itself, thereby extending far below what is visible on a logarithmic plot; for exact values, see data. The theoretically predicted fidelities for measurement-based protocol (dashed grey line) were based on an inferred noise model with auxiliary and site readout errors with a range of parameters giving rise to a 25th–75th percentile confidence interval in shaded grey (Supplementary Section IIB). The inset shows the ratio of the experimentally evaluated measurement- to unitary-based fidelities increasing for system size up to 54 sites.

Noise analysis

We now determine where in the phase diagram our experimental protocol accessed the GHZ state relative to the criticality threshold—implicitly bounding the amount of other sources of errors that were present in our experiments. To do so, we tune one type of coherent error via equation ( ) uniformly across the device. In this sweep, we monitored and fit the experimental observables associated with every bond to $\langle ZXZ\rangle = -(1-2p_\sigma)^2(1-2p_s)\sin(2t_A)$, and experimental observable of every plaquette to $\langle W\rangle = (1-2p_s)^6 \sin(2t_A)^6$, with two fit parameters p_s and p_σ, as shown in Fig. (Supplementary Section IC). Here, p_σ accounts for the readout error of system qubits while p_s captures both readout error on the auxiliary qubits and some of the noise during the entangling process. For $t_A = \pi/4$, the bond and plaquette observables should ideally approach unity (dashed lines) because they capture, partially, the quality of the constituent cluster states —a precursor state for the GHZ state—with experimental data shown in Fig. . For t_A below $\pi/4$, the implemented circuits become non-Clifford and thus cannot, in general, be efficiently characterized. Nonetheless our modelling of coherent and incoherent noise sources turns out to be sufficiently comprehensive to quantitatively explain the observed experimental data, even for experiments involving up to 125 qubits. This allows us to infer the amount of noise afflicting the auxiliary (p_s) and system (p_σ) qubits when sweeping t_A. This led to an estimate for the amount of incoherent errors present in the experiment to be in the range of $p_s \approx 4.2$–5.6% and $p_\sigma \approx 1.2$–2.3% (Table)—values consistent with our expectation based on standard calibration benchmarks of the device (Supplementary Section IC).

https://doi.org/10.1038/s41567-024-02696-6

Table 1 | Two-parameter noise model

System size (N)	L_y	p_s (auxiliary)	p_σ (system)
10	2	0.042	0.012
28	3	0.051	0.018
54	4	0.056	0.023

Fits to experimental data give p_s, which captures errors at the auxiliary qubits, and p_σ, which captures errors at the system qubits.

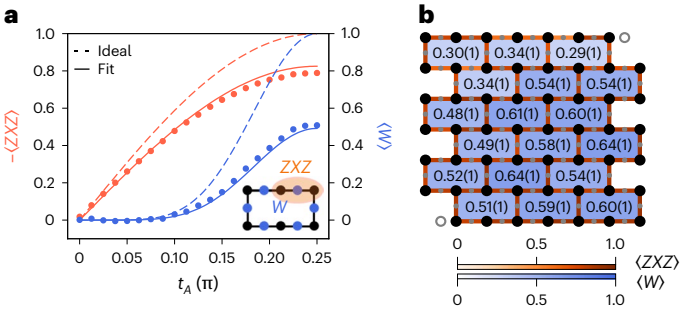


Fig. 3 | Experimentally measured local observables used to generate the state. a, For two observables, we plot the ideally expected outcomes (dashed lines), the unprocessed experimental data from 20,000 experimental samples (dots), and a one-parameter fit (solid line) for each observable from sweeping t_A from 0 (trivial) to π/4 (long-range ordered). The average 3-qubit bond (red) observable reached as high as 0.8 across the 72 total bonds, while the average 6-qubit plaquette (blue) observable reached 0.5 across the 18 plaquettes. The plotted error bars for the 3-qubit bond and 6-qubit plaquette are the standard deviations across the 72 and 18 samples, respectively. Although in a noiseless setting both were expected to reach unity, the measured values agree well with the fit by $p_s = 5.6\%$, and $p_\sigma = 2.3\%$, which are approximately consistent with the known errors on the device and a classical simulation with circuit-level noise (Supplementary Section IIB). The experimental data exhibit an absence of a singularity in these observables, consistent with expectations for both local shallow quantum circuit, and the internal energy of Nishimori line. **b**, In total, 125 of the 127 qubits used on ibm_sherbrooke where each bond ($\langle ZXZ\rangle$) and plaquette ($\langle W\rangle$) observable values are shaded according to the measured value at $t_A = \pi/4$. The numbers inside the plaquettes (**b**) show $\langle W\rangle$, with the standard error in parentheses.

Nishimori transition for tunable coherent errors

Having established the incoherent noise level of our device, we can now proceed to validate the existence of a stable, long-range, Ising-ordered phase when experimentally sweeping the level of coherent errors in our protocol. To reveal the hidden order, we applied a decoder to process every classical snapshot for the auxiliary qubits in the X basis and the system qubits in the Z basis. The basic idea is to perform a correction based solely on the auxiliary readout. This correction factor approximates the ground state configuration of the RBIM as a classical estimate, $\{\sigma'\}$, of the bit-string from the quantum device (Fig.).

The distribution of the decoded bit-strings in the computational basis is shown, for $t_A = \pi/4$, in Fig. , where we sum over the decoded Z expectation values of the individual qubit to obtain a total decoded 'magnetization' $M = \sum_{j=1}^{N} \sigma' Z_j$. Any bias of this distribution (for example, towards positive values) may be explained by an Ising asymmetric error originating from physical mechanisms such as amplitude damping or relaxation. Such errors would reduce the amount of classical correlations, and the small value $\langle M\rangle \approx 0.02(2)N$ (for $N = 54$) suggests that the global Ising symmetry is largely preserved. Moreover, the decoded, bimodal experimental distribution (Fig.) agrees well with the theoretical prediction (solid lines), lending confidence to the two-parameter noise model we used.

To more rigorously characterize the long-range order, we examined, for $t_A \le \pi/4$, the decoded system qubit bit-strings and the average two-point classical correlations

$$f := \frac{1}{N}\left(\langle M^2\rangle - \langle M\rangle^2\right), \tag{5}$$

which is a sum of the correlations $\langle \sigma'\sigma' Z_i Z_j\rangle$ for all the system qubits that compose the quantum state. The decoded experimental data is shown in Fig. , where the solid line shows the theoretical benchmark with the noise parameters inferred from Fig. . We observe a hallmark of the long-range ordered phase in the linearly scaling f for increasing system sizes; such divergent tendency for f is expected throughout the ensemble of long-range ordered states, or phase, even away from $t_A = \pi/4$ up to a finite threshold, t_A^c. In fact, we have confirmed that, in our two-parameter theory model, f indeed grows unbounded above $t_A^c \approx 0.20\pi \sim 0.21\pi$. In contrast, for small t_A far below the threshold, f is apparently bounded and does not grow with increasing size. This scaling behaviour for our 2D protocol should be contrasted to results in 1D geometries (Supplementary Section IE), where we found f to stop growing for larger system sizes in line with theoretical expectations that f is bounded by a finite correlation length in the presence of infinitesimal weak errors.

To determine the threshold, or critical point, a practical way is to use the normalized variance of M^2/N

$$g := \frac{1}{N^3}\left(\langle M^4\rangle - \langle M^2\rangle^2\right), \tag{6}$$

which quantifies the amount of fluctuations in the squared magnetization (Supplementary Section IE). In the presence of 5% incoherent auxiliary errors, the peak location is expected to converge to a critical value of $t_A^c \approx 0.205\pi$, by translating the Nishimori critical point $\tilde{p}_c \approx 6.75\%$ (refs. ,) with equation (), which is in very close agreement with the experimental data where the peak locations approach this predicted critical point (Fig.).

Moreover, at this transition, we also observe that f exhibits steep increases as one would expect for a critical system. The three experimental values for f for increasing system sizes agree well with noisy classic simulations exhibiting a $\propto L_y^{1.9}$ scaling behaviour of the peak height (Fig. , inset), where $L_y = 2$, 3 and 4 is the number of columns of qubits ($N = 10$, 28 and 54) in a brickwall lattice; this experimentally observed scaling is in close agreement with the scaling exponent calculated value of 1.8(1) for the RBIM at the Nishimori point. While the criticality is exposed in the decoded correlations only, the observable $\langle ZXZ\rangle$ of Fig. is another, direct probe of Nishimori physics; it corresponds to the internal energy of the classical RBIM along the Nishimori line, which we experimentally confirm to be free of any singularity at the phase transition and in agreement with theoretical predictions.

Decoding transition by tuning incoherent errors

As we have shown, the long-range ordered phase created by our 2D protocol is unveiled only after using a decoder, whose performance critically depends on the quality of the auxiliary measurements. While the auxiliary error is lower-bounded by the quantum device, we can inject additional errors, in post-processing without any additional experiments, before applying the decoder (Fig.). Using the procedure outlined in Supplementary Section IIC, we charted out a broader phase diagram at various rates of incoherent errors. By again monitoring the degree of fluctuations, g, now as a function of an increasing level of incoherent errors and system size (Fig.), we observe that the Nishimori critical point t_A^c shifts towards π/4 and vanishes completely at $p_s \approx 6.75\%$ (ref.), the decoding threshold (Supplementary Section IA). The origin of this limit can be readily understood as being equivalent to the decoding transition of a repetition code on a honeycomb lattice with bit-flip errors. Our experiments thus not only demonstrate

https://doi.org/10.1038/s41567-024-02696-6

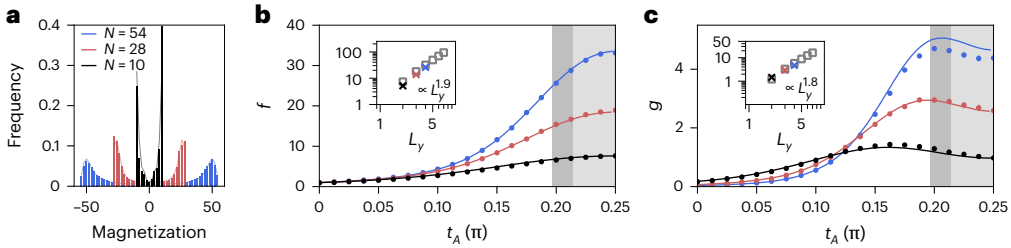


Fig. 4 | Nishimori transition by tuning coherent errors. a, The distribution of M at $t_A = \pi/4$ for system sizes of 10 (black), 28 (red) and 54 (blue) system qubits using 21, 63 and 125 total qubits, respectively. The solid envelope lines are theoretical estimates using a two-parameter noise model. **b**, The sum of two-point correlation function f signals the growth of long-range correlation when increasing t_A and system sizes. Beyond a critical threshold for $t_A \geq t_A^c \approx 0.20\pi \sim 0.21\pi$ (dark grey), the state exhibits long-range order (light grey). The estimated t_A^c varied depending on the system size studied. The inset shows the size scaling of experimentally measured f (X marks) at the peak location of g agreeing well with the theoretically predicted noiseless values (square markers) scaling with $\propto L_y^{1.9}$. **c**, The peak locations of g converge to the finite threshold that separates the long-range ordered phase (light grey) from short-range correlated phases. The dark-grey shading illustrates the theoretically predicted critical point (t_A^c), according to the previously inferred two noise parameters, that spans a finite width because of the variation of noise probabilities. In the inset, the experimental (X marks) values agree well with the theoretically predicted values scaling as $\propto L_y^{1.8}$. For both **b** and **c**, the error bars represent standard deviations from taking 20,000 experimental samples. The noiseless envelopes for all solid curves can be found in Supplementary Information.

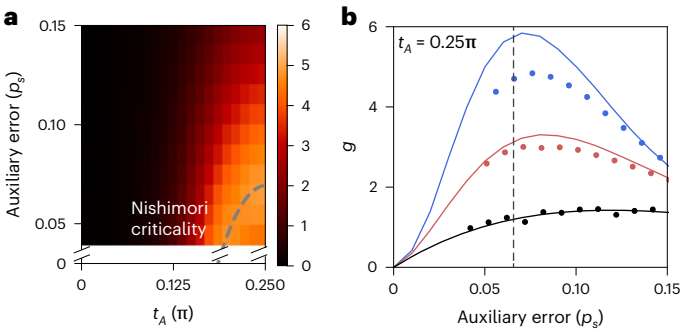


Fig. 5 | Decoding transition out of the long-range ordered phase by increasing auxiliary errors before decoding. a, For the largest system size ($N = 54$), we experimentally mapped the 2D phase diagram for various coherent (t_A) and incoherent (p_s) errors where the colour is proportional to the amount of variance in the magnetization squared, g. The analytically derived contour (dashed grey) shows close agreement for incoherent, auxiliary errors starting from approximately 0.05. **b**, For the lowest amount of injected coherent error ($t_A = \pi/4$, dashed grey line in **a**), the experimentally estimated variance g (circles) is maximized at the theoretically expected (solid lines) decoder transition of approximately 6.75% (vertical dashed grey) for all three system sizes (10: black, 28: red, 54: blue). The plotted error bars represent standard deviations for all three system sizes, each taken with 20,000 experimental samples.

the stability of the long-range ordered phase separated from a trivial one via a Nishimori transition, but also quantify when it would fail for noisier devices. It also distinguishes a 2D from 1D protocol where the peak quickly converges to $t_A = \pi/4$ without a finite threshold (Extended Data Figs. 2–4). Thus we claim that our experimentally implemented 2D protocol exhibits long-range order with intrinsic robustness.

Discussion

The Nishimori multicritical point arises from a delicate balance between disorder and temperature—a condition that is largely inaccessible in experiments on real, physical materials modelled by a RBIM. This should be contrasted to our experiments using a shallow circuit protocol on a quantum system, where the Nishimori transition shows remarkable robustness even in a noisy device of significant size. We argue that this can in fact be traced back to Born's rule, which naturally enforces the delicate balance of Nishimori physics: the auxiliary qubits play the role of quenched disorder by being measured, whose probability is exactly the wave function squared amplitude of the system qubits.

Our systematic study and generation of long-range ordered states using measurements shows that such protocols can be robust against certain errors and even outperform unitary approaches on existing quantum hardware. Improvements in coherence and measurement fidelity should further improve the performance of our measurement-based protocol. Our work emphasizes the importance of spatial geometry in measurement-based protocols; by tuning errors across an error threshold, we observed a stable phase that persists in 2D but is absent in 1D. While the experimentally accessible order parameters, f and g, were observed to be below the theoretically predicted noiseless values due to the presence of noise, we expect to still be able to determine the universal critical exponents using equivalently noisy but larger devices, up to system sizes of 180, where finite-sized effects play less of a role.

It would be interesting to similarly explore the (in)stability of measurement-induced long-range entanglement upon tuning coherent and incoherent errors for other proposals in the literature. This is especially timely since measurements have recently been used to deterministically create exotic long-range entanglement including topological order and related states with dynamic quantum circuits utilizing feed-forward operations. In such general contexts, stability might require additional ingredients, such as using the time-domain or higher—or even fractal—dimensions, opening up a rich territory for exploration.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at .

References

1. Bravyi, S., Hastings, M. B. & Verstraete, F. Lieb–Robinson bounds and the generation of correlations and topological quantum order. *Phys. Rev. Lett.* **97**, 050401 (2006).
2. Hastings, M. B. in *Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School: Volume 95, August 2010* (eds Frohlich, J., Salmhofer, M., Mastropietro, V., De Roeck, W. & Cugliandolo, L. F.) 171–212 (Oxford Univ. Press, 2012).
3. Chen, X., Gu, Z.-C. & Wen, X.-G. Classification of gapped symmetric phases in one-dimensional spin systems. *Phys. Rev. B* **83**, 035107 (2011).

4. Chen, X., Gu, Z.-C. & Wen, X.-G. Complete classification of one-dimensional gapped quantum phases in interacting spin systems. *Phys. Rev. B* **84**, 235128 (2011).
5. Zeng, B. & Wen, X.-G. Gapped quantum liquids and topological order, stochastic local transformations and emergence of unitarity. *Phys. Rev. B* **91**, 125121 (2015).
6. Huang, Y. & Chen, X. Quantum circuit complexity of one-dimensional topological phases. *Phys. Rev. B* **91**, 195143 (2015).
7. Haah, J. An invariant of topologically ordered states under local unitary transformations. *Commun. Math. Phys.* **342**, 771–801 (2016).
8. Mooney, G. J., White, G. A. L., Hill, C. D. & Hollenberg, L. C. L. Whole-device entanglement in a 65-qubit superconducting quantum computer. *Adv. Quantum Technol.* **4**, 2100061 (2021).
9. Mooney, G. J., White, G. A. L., Hill, C. D. & Hollenberg, L. C. L. Generation and verification of 27-qubit Greenberger–Horne–Zeilinger states in a superconducting quantum computer. *J. Phys. Commun.* **5**, 095004 (2021).
10. Wei, K. X. et al. Verifying multipartite entangled Greenberger–Horne–Zeilinger states via multiple quantum coherences. *Phys. Rev. A* **101**, 032343 (2020).
11. Briegel, H. J. & Raussendorf, R. Persistent entanglement in arrays of interacting particles. *Phys. Rev. Lett.* **86**, 910–913 (2001).
12. Raussendorf, R., Bravyi, S. & Harrington, J. Long-range quantum entanglement in noisy cluster states. *Phys. Rev. A* **71**, 062313 (2005).
13. Piroli, L., Styliaris, G. & Cirac, J. I. Quantum circuits assisted by local operations and classical communication: transformations and phases of matter. *Phys. Rev. Lett.* **127**, 220503 (2021).
14. Tantivasadakarn, N., Thorngren, R., Vishwanath, A. & Verresen, R. Long-range entanglement from measuring symmetry-protected topological phases. *Phys. Rev. X* **14**, 021040 (2024).
15. Verresen, R., Tantivasadakarn, N. & Vishwanath, A. Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices. Preprint at [redacted] (2021).
16. Bravyi, S., Kim, I., Kliesch, A. & Koenig, R., Adaptive constant-depth circuits for manipulating non-abelian anyons. Preprint at [redacted] (2022).
17. Lu, T.-C., Lessa, L. A., Kim, I. H. & Hsieh, T. H. Measurement as a shortcut to long-range entangled quantum matter. *PRX Quantum* **3**, 040337 (2022).
18. Zhu, G.-Y., Tantivasadakarn, N., Vishwanath, A., Trebst, S. & Verresen, R. Nishimori's cat: stable long-range entanglement from finite-depth unitaries and weak measurements. *Phys. Rev. Lett.* **131**, 200201 (2023).
19. Tantivasadakarn, N., Vishwanath, A. & Verresen, R. Hierarchy of topological order from finite-depth unitaries, measurement, and feedforward. *PRX Quantum* **4**, 020339 (2023).
20. Tantivasadakarn, N., Verresen, R. & Vishwanath, A. Shortest route to non-abelian topological order on a quantum processor. *Phys. Rev. Lett.* **131**, 060405 (2023).
21. Lee, J. Y., Ji, W., Bi, Z. & Fisher, M. P. A. Decoding measurement-prepared quantum phases and transitions: from ising model to gauge theory, and beyond. Preprint at [redacted] (2022).
22. Li, Y., Sukeno, H., Mana, A. P., Nautrup, H. P. & Wei, T.-C. Symmetry-enriched topological order from partially gauging symmetry-protected topologically ordered states assisted by measurements. *Phys. Rev. B* **108**, 115144 (2023).
23. Buhrman, H., Folkertsma, M., Loff, B. & Neumann, N. M. P. State preparation by shallow circuits using feed forward. Preprint at [redacted] (2023).
24. Lu, T.-C., Zhang, Z., Vijay, S. & Hsieh, T. H. Mixed-state long-range order and criticality from measurement and feedback. *PRX Quantum* **4**, 030318 (2023).
25. Friedman, A. J., Yin, C., Hong, Y. & Lucas, A. Locality and error correction in quantum dynamics with measurement. Preprint at [redacted] (2023).
26. Moses, S. A. et al. A race track trapped-ion quantum processor. *Phys. Rev. X* **13**, 041052 (2023).
27. Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *J. Math. Phys.* **43**, 4452–4505 (2002).
28. Nishimori, H. Internal energy, specific heat and correlation function of the bond-random Ising model. *Progress Theor. Phys.* **66**, 1169–1181 (1981).
29. Nishimori, H. Optimum decoding temperature for error-correcting codes. *J. Phys. Soc. Jpn* **62**, 2973–2975 (1993).
30. Garratt, S. J., Weinstein, Z. & Altman, E. Measurements conspire nonlocally to restructure critical quantum states. *Phys. Rev. X* **13**, 021026 (2023).
31. Lee, J. Y., You, Y.-Z. & Xu, C. Symmetry protected topological phases under decoherence. Preprint at [redacted] (2023).
32. Garratt, S. J. & Altman, E. Probing postmeasurement entanglement without post-selection. *PRX Quantum* **5**, 030311 (2024).
33. Chow, J. M. et al. Simple all-microwave entangling gate for fixed-frequency superconducting qubits. *Phys. Rev. Lett.* **107**, 080502 (2011).
34. Sheldon, S., Magesan, E., Chow, J. M. & Gambetta, J. M. Procedure for systematically tuning up cross-talk in the cross-resonance gate. *Phys. Rev. A* **93**, 060302 (2016).
35. Malekakhlagh, M., Magesan, E. & McKay, D. C. First-principles analysis of cross-resonance gate operation. *Phys. Rev. A* **102**, 042605 (2020).
36. Sundaresan, N. et al. Reducing unitary and spectator errors in cross resonance with optimized rotary echoes. *PRX Quantum* **1**, 020318 (2020).
37. de Queiroz, S. L. A. Multicritical point of Ising spin glasses on triangular and honeycomb lattices. *Phys. Rev. B* **73**, 064410 (2006).
38. Flammia, S. T. & Liu, Y.-K. Direct fidelity estimation from few Pauli measurements. *Phys. Rev. Lett.* **106**, 230501 (2011).
39. da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. *Phys. Rev. Lett.* **107**, 210404 (2011).
40. Cao, S. et al. Generation of genuine entanglement up to 51 superconducting qubits. *Nature* **619**, 738–742 (2023).
41. Higgott, O. Pymatching: A python package for decoding quantum codes with minimum-weight perfect matching. *ACM Trans. Quantum Comput.* **3**, 1–16 (2022).
42. Morvan, A. et al. Phase transitions in random circuit sampling. *Nature* **634**, 328–333 (2024).
43. Binder, K. & Young, A. P. Spin glasses: experimental facts, theoretical concepts, and open questions. *Rev. Mod. Phys.* **58**, 801 (1986).
44. Iqbal, M. et al. Topological order from measurements and feed-forward on a trapped ion quantum computer. *Commun. Phys.* **7**, 205 (2024).
45. Foss-Feig, M. et al. Experimental demonstration of the advantage of adaptive quantum circuits. Preprint at [redacted] (2023).
46. Iqbal, M. et al. Non-abelian topological order and anyons on a trapped-ion processor. *Nature* **626**, 505–511 (2024).
47. Bäumer, E. et al. Efficient long-range entanglement using dynamic circuits. *PRX Quantum* **5**, 030339 (2024).
48. Hastings, M. B. & Haah, J. Dynamically generated logical qubits. *Quantum* **5**, 564 (2021).

Methods

127 superconducting qubit device

We performed all experiments on ibm_sherbrooke, a 127-qubit Eagle r3 processor. The entangling gate has a native ZX interactions and is known as an ECR gate. Using Clifford randomized benchmarking[49,50], we observed a median error of 0.0077, with a 50% confidence interval of 0.006–0.008. The two-qubit gate times across the device were uniformly set to 533.3 ns, similar to the method described in ref. [6]. The median square root of the Pauli-X error rate was 0.0002 (0.0002–0.0004). The readout error was 0.010 (0.007–0.021) with typical measurement times of ~1,244.4 ns. The qubits under study had a median $T_1 \approx 293$ µs and $T2 \approx 173$ µs. The distribution of these error rates is plotted in Extended Data Fig. 1. Circuits were executed on the device at a clock rate of 1 kHz (ref. [6]). For all data found in Figs. [3]–[5], experimental error bars were calculated from the standard error on 20,000 shots at the 1 kHz clock rate.

Decomposition of ZZ gates

For fidelity comparisons in Fig. [3], we compiled the $ZZ(t = \pi/4)$ gate into a single ECR gate, which is the native basis gate on the device, and single-qubit rotations. For Fig. [4], the $ZZ(t)$ gates were all decomposed into two ECR gates with a virtual-$R_Z(t)$ gate in between, resulting in a depth-6 unitary circuit followed by a layer of measurements. We note that further improvements could be accessed by shortening the $ZZ(t)$ gate time with fractional ZX rotations [49] that are accessible on the device. Hadamard gates, decomposed into two square root of Pauli-X gates and virtual-$R_Z(t)$ gates, were also used for the preparation and readout of qubits.

Quantum circuit transpilation

For both unitary and measurement-based experiments, dynamical decoupling (DD) was used in the same fashion. All single- and two-qubit gates were scheduled as late as possible after initialization in the ground state, and all idle periods after the first operations were replaced with a $X_{+\pi} - X_{-\pi}$ sequence in which the total idling period was divided proportionally according to a 1:2:1 ratio surrounding the X gates. Consequently, the unitary-based protocols benefitted more from applying DD than for measurement-based circuits.

Furthermore, we identified at least 12 different ways to schedule entangling gates for the measurement-based circuits and found some schedules to significantly outperform others (Supplementary Section IID). This is consistent with our expectation that certain gates, when executed in parallel, can induce frequency collisions on the device that reduce the fidelity of the entangling gates.

Monte Carlo sampling of GHZ stabilizer observables

For size-N GHZ states generated at the fixed point of the Nishimori line, we randomly measured up to S different non-Z stabilizers (for example, weight-N observables containing only X and Y Paulis). We also included measurements of the system qubits in the all-Z basis, which allows us to reconstruct any of the 2^{N-1} possible Z-only stabilizers of the GHZ state. In practice, however, we perform a binomial resampling of Z-only and non-Z stabilizers with equal probability to emulate the proposed Monte Carlo sampling method for fidelity state estimation[49,50]. By sampling $2S'$ random instances from 2^N

$$\mathcal{F} = \frac{1}{2S'}\sum_{i=1}^{2S'}\langle \mathcal{O}_i \rangle, \tag{7}$$

an average fidelity was estimated. The exact samples given N sites, S samples and k resampling instances are, in N, S, k notation: (10, 30, 500), (28, 75, 500) and (54, 100, 500). For each of the k resampling instance, a total of $2S'$ observables were sampled but drawn randomly without replacement from the S non-Z stabilizers or from the 1,000 Z-only stabilizers. Thus, $S' \approx 3S/4$ of the non-Z stabilizers were sampled without replacement from S, and another $S' \approx 3S/4$ Z-only stabilizers were sampled also without replacement from 1,000 unique Z-only stabilizers. Both sets S' combined to give ~$2S'$ observables as part of this binomial resampling procedure.

In all expectation values above, we randomly applied X gates before readout of the system qubits and, after applying the correcting spin flip to sites on which Z and Y Paulis were supported, calculated the expectation values of the random stabilizers. Although readout was 'twirled', the model-free readout correction was not applied[6].

Fits to noise model

Two most basic linear observables are analytically known in the noiseless limit: $\langle W \rangle = \sin(2t_A)^6$, $\langle ZXZ \rangle = \sin(2t_A)$. Let us consider two phenomenological errors: readout errors on the the auxiliary qubit measured in X basis, with probability p_s; and that on the system qubit measured in Z basis, with probability p_σ. Note that they also include the consequence of some of the mid-circuit bit-flip or phase-flip errors that propagate to yield the same effect in the end, such as the bit-flip (phase-flip) for system (auxiliary) qubits at the moment after the R_{ZZ} gates. These two error rates turn the expectation values of the above observables into $\langle W \rangle = (1-2p_s)^6 \sin(2t_A)^6$, $\langle ZXZ \rangle = (1-2p_s)(1-2p_\sigma)^2 \sin(2t_A)$. We can perform a linear fit to extract such phenomenological error rates per bond and plaquette, which are then averaged over the lattice for mean values and standard deviations. The averaged effective errors per qubit weakly grows with the total number of qubits in our three experimental implementations, as seen in Table [1].

For the 1D protocol where we do not have the Wilson loop $\langle W \rangle$, we can use two Wilson lines of different lengths, for example, $\langle ZXZ \rangle$ and $\langle ZX\mathbb{1}XZ \rangle$, to extract the two parameters for auxiliary and system qubits, respectively.

From measurement disorder to uncorrelated disorder

The measurement-induced disorder obeys Born's rule, which leads to a highly correlated arrangement. Nevertheless, equation ([1]) has emergent gauge symmetry [5], which allows us to reinterpret the disorder as gauge symmetrizing an independent and identically distributed (i.i.d.) random bond disorder. Namely, we first generate the i.i.d. disorder s with probability $e^{-\beta s}/(2\cosh(\beta))$. Second, we perform a Z_2 gauge transformation, denoted by $\tau_j = \pm 1$ as a set of random binary number on every site. It transforms $s_{ij} \to s_{ij}\tau_i\tau_j$, $\sigma_j \to \sigma_j\tau_j$. Finally, we sum over all possible gauge transformations τ_j with equal probability. By noting that $\frac{1}{2^N}\sum_\tau \frac{e^{-\beta\sum_{ij}s_{ij}\tau_i\tau_j}}{(2\cosh(\beta))^{N_b}} = P(s)$, we obtain

$$\sum_{s=\pm 1}\frac{e^{-\beta\sum_{ij}s_{ij}}}{(2\cosh(\beta))^{N_b}}\langle\sigma_i\sigma_j\rangle_s = \sum_{s=\pm 1}P(s)\langle\sigma_i\sigma_j\rangle_s^2, \tag{8}$$

which relates the ferromagnetic correlation of the i.i.d. disorder model to the spin glass correlation of the gauge symmetric correlated disorder model.

Data availability

The data supporting the findings of this study can be found via figshare at [redacted] (ref. [6]).

Code availability

Simulation and data analysis code may be made available upon reasonable request.

References

49. Magesan, E., Gambetta, J. M. & Emerson, J. Characterizing quantum gates via randomized benchmarking. *Phys. Rev. A* **85**, 042311 (2012).
50. Helsen, J., Roth, I., Onorati, E., Werner, A. H. & Eisert, J. General framework for randomized benchmarking. *PRX Quantum* **3**, 020357 (2022).

51. Kim, Y. et al. Evidence for the utility of quantum computing before fault tolerance. *Nature* **618**, 500–505 (2023).
52. Earnest, N., Tornow, C. & Egger, D. J. Pulse-efficient circuit transpilation for quantum applications on cross-resonance-based hardware. *Phys. Rev. Res.* **3**, 043088 (2021).
53. Kim, Y. et al. Scalable error mitigation for noisy quantum circuits produces competitive expectation values. *Nat. Phys.* **19**, 752–759 (2023).
54. van den Berg, E., Minev, Z. K. & Temme, K. Model-free readout-error mitigation for quantum expectation values. *Phys. Rev. A* **105**, 032620 (2022).
55. Chen, E. et al. Realizing the Nishimori transition across the error threshold for constant-depth quantum circuits. *figshare* [redacted] (2024).

Acknowledgements

We thank M. Ware, P. Jurcevic, Y. Kim, A. Eddins, H. Nayfeh, I. Lauer, D. McKay, G. Jones and J. Summerour for assistance with performing experiments and B. Mitchell, D. Zajac, J. Wootton, L. Govia, X. Wei, R. Gupta, T. Yoder, T. Soejima, K. Siva, M. Motta, Z. Minev, S. Pappalardi, S. Garratt, E. Altman, F. Valenti and H. Nishimori for thoughtful discussions. We thank H. Nishimori for careful reading of the paper. The Cologne group was partially funded by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769 and within the CRC network TR 183 (project grant no. 277101999: G.-Y.Z. and S.T.) as part of projects A04 and B01. The classical simulations were performed on the JUWELS cluster at the Forschungszentrum Juelich. R.V. is supported by the Harvard Quantum Initiative Postdoctoral Fellowship in Science and Engineering. A.V. is supported by a Simons Investigator grant and by NSF DMR-2220703. A.V. and R.V. are supported by the Simons Collaboration on Ultra-Quantum Matter, which is a grant from the Simons Foundation (618615, A.V.). G.Z. is supported by the US Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704. We acknowledge the use of IBM Quantum services for this work.

Author contributions

E.H.C. led the execution and analysis of the experimental data. G.-Y.Z. and R.V. led the theoretical developments; G.-Y.Z. developed all numerical simulations and additional experimental analysis code. A.S., E.B. and D.L. contributed technical insights and code related to characterizing the states. N.T., G.Z., A.V. and S.T. contributed theoretical insights related to the critical transition; S.T. provided additional insights related to entanglement generation and verification. S.S. and A.K. provided experimental support and access, and contributed to the design of the experiments. E.H.C., G.-Y.Z. and R.V. drafted the paper and supplementary material; all authors contributed to revising both.

Competing interests

The authors declare no competing interests.

Additional information

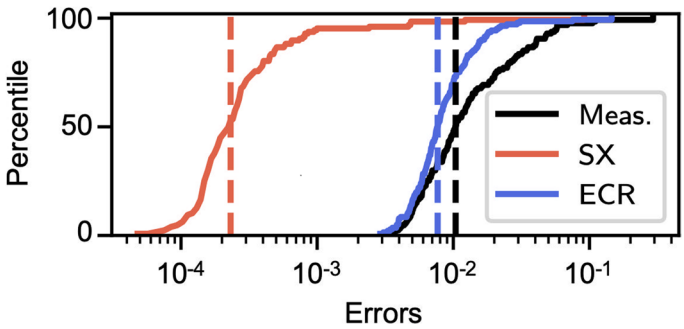
Extended data is available for this paper at [redacted].

Supplementary information The online version contains supplementary material available at [redacted].

Correspondence and requests for materials should be addressed to Edward H. Chen, Guo-Yi Zhu or Ruben Verresen.

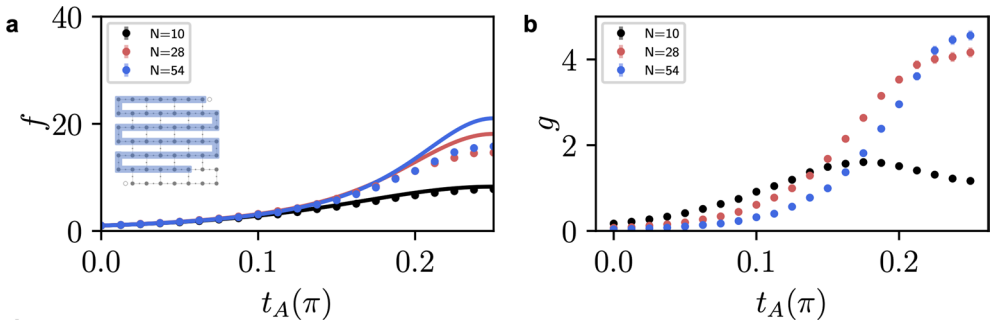
Peer review information *Nature Physics* thanks Matteo Ippoliti, Morten Kjaergaard and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

https://doi.org/10.1038/s41567-024-02696-6



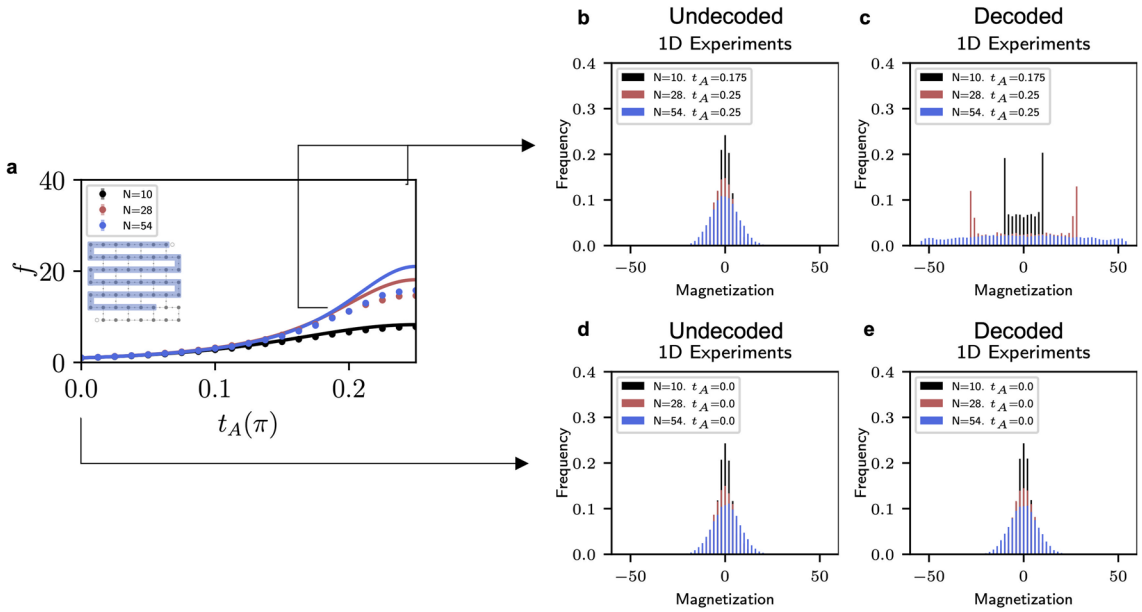
Extended Data Fig. 1 | Typical device error rates. Plotted in cumulative format, for Echoed Cross Resonance (ECR, blue), square-root of Pauli-X (SX, red), and measurement (Meas, black) gates. Dashed lines represent medians of distributions.

https://doi.org/10.1038/s41567-024-02696-6



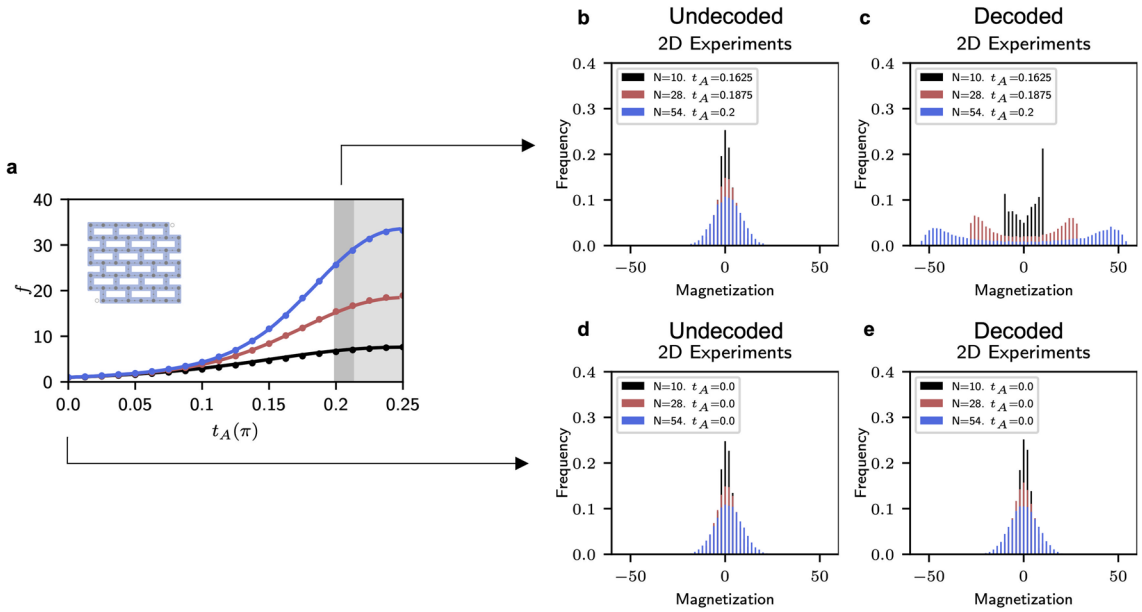
Extended Data Fig. 2 | Absence of finite threshold in one-dimensional protocol. As discussed in the main text, the 2D protocol exhibited robustness over the 1D protocol (seen here); the key signature being based on the scaling of the average of two-point correlations, f, as a function of system size. For comparison with Fig. [redacted] in the main text. (**a**) f grows with increasing system size but converges to finite value that depends on t_A. (**b**) The peak of g converges to $t_A = \pi/4$ indicative of absence of finite threshold for coherent error. For both (**a**) and (**b**), the three system sizes were measured with 20,000 experimental samples giving rise to the small standard deviations (bars).

https://doi.org/10.1038/s41567-024-02696-6



Extended Data Fig. 3 | Magnetization of 1D experiments with and without decoding at different t_A values. Plotting the same data set from Extended Data Fig. 2, we observe that the 1D behavior exhibited no growth in f with system size from 28 to 54 and had peak variances at the GHZ value of $t_A = \pi/4$. (**a**) Two-point correlations in 1D experiments for sweeps of t_A. The histograms at values of t_A where variances were maximized for undecoded (**b**) and decoded (**c**). Although the bimodal distribution persisted up to a system size of 28, at 54 the distribution became uniform. And as expected, both the undecoded (**d**) and decoded (**e**) exhibited a binomial distribution in the trivial state ($t_A = 0$). The error bars in (**a**) are standard deviations based on 20,000 experimental samples, while the histograms in the other subplots are based on the same data sets.

https://doi.org/10.1038/s41567-024-02696-6



Extended Data Fig. 4 | Magnetization of 2D experiments with and without decoding at different t_A values. (**a**) Two-point correlations in 2D experiments for sweeps of t_A. The histograms at values of t_A where variances were maximized for undecoded (**b**) and decoded (**c**). In contrast to the 1D cases (Extended Data Fig. 3), the bimodal distribution persisted up to a system size of 54. And similarly to the 1D case, both the undecoded (**d**) and decoded (**e**) exhibited a binomial distribution in the trivial state ($t_A = 0$). The error bars in (**a**) are standard deviations based on 20,000 experimental samples, while the histograms in the other subplots are based on the same data sets.