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

Edward H. Chen,[1, 2, *] Guo-Yi Zhu,[3, †] Ruben Verresen,[4, ‡] Alireza Seif,[5] Elisa Bäumer,[6] David Layden,[1, 5]
Nathanan Tantivasadakarn,[7, 4] Guanyu Zhu,[5] Sarah Sheldon,[5] Ashvin Vishwanath,[4] Simon Trebst,[3] and Abhinav Kandala[5]
[1]*IBM Quantum, Almaden Research Center, San Jose, CA 95120, USA*
[2]*IBM Quantum, Research Triangle Park, NC 27709, USA*
[3]*Institute for Theoretical Physics, University of Cologne, Zülpicher Straße 77, 50937 Cologne, Germany*
[4]*Department of Physics, Harvard University, Cambridge, MA 02138, USA*
[5]*IBM Quantum, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA*
[6]*IBM Quantum, IBM Research, Zurich, Rüschlikon, Switzerland.*
[7]*Walter Burke Institute for Theoretical Physics and Department of Physics,*
California Institute of Technology, Pasadena, CA 91125, USA
(Dated: September 7, 2023)

Preparing quantum states across many qubits is necessary to unlock the full potential of quantum computers. However, a key challenge is to realize efficient preparation protocols which are stable to noise and gate imperfections. Here, using a measurement-based protocol on a 127 superconducting qubit device, we study the generation of the simplest long-range order—Ising order, familiar from Greenberger-Horne-Zeilinger (GHZ) states and the repetition code—on 54 system qubits. Our efficient implementation of the constant-depth protocol and classical decoder shows higher fidelities for GHZ states compared to size-dependent, unitary protocols. By experimentally tuning coherent and incoherent error rates, we demonstrate stability of this decoded long-range order in two spatial dimensions, up to a critical point which corresponds to a 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—locking the effective temperature and disorder driving this transition. Our study exemplifies how measurement-based state preparation can be meaningfully 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 [1–7]; however, these cannot be prepared by any constant-depth unitary circuits, making them more sensitive to the finite coherence times of current quantum processors [8–10]. 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 [11–23] and critical quasi-long-range order [18, 21, 24]. In essence, measurement-based approaches trade off circuit depth for number of mid-circuit measurements and operations [25] as compared to 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 (LRE) 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 GHZ is a long-range Ising order which 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 qubit geometries where stability is not guaranteed [26], theoretical works suggest that this order should be robust against a range of imperfections when using a two-dimensional (2D) protocol [18, 21, 27]. 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, e.g. $|00110\rangle + |11001\rangle$, 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, feedforward, or post-processing to recover the long-range order. In our setup, we implement post-processing to decode [21, 28] 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 [29], for both incoherent [27] and coherent errors [18, 21]. 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. 1a). We will refer to the Pauli matrices on each qubit as X, Y, Z. To turn an initial product state of system qubits in $+1$ X eigenstates into a GHZ-type

---

[*] ehchen@ibm.com
[†] gzhu@uni-koeln.de
[‡] rubenverresen@g.harvard.edu

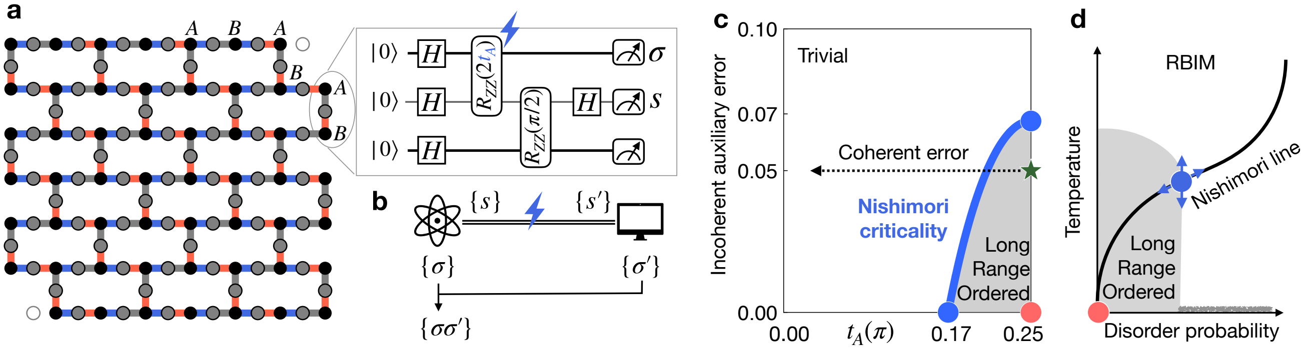

FIG. 1. **Circuit protocol, decoder, and phase diagram under coherent and incoherent errors.** **a.** The heavy-hexagonal lattice of 127 qubits. For the 125 active qubits, the inset shows the building block using constant-depth entangling circuits for the three nearest neighbors (gray circles) of each system qubit (black circles) in the presence of noise (lightning). The R_{ZZ} gates are executed in order from blue, red, then gray bonds within three layers. The auxiliary outcomes, s, on the bonds of the lattice (gray) 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 (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 (gray), of which the GHZ is a special case (red circle), exhibits long-range entanglement in the absence of other sources of noise (e.g. without dephasing). Even in the presence of dephasing (not shown), classical long-range *ordering* remains. The boundary separating the trivial and long-range phase 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.** Schematic phase diagram of the classical RBIM. The solid black line is the Nishimori line, which captures the *entire* phase diagram in **c**.

state, we measure the ZZ parities on all nearest neighbor 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. 1a). By varying t_A away from $\pi/4$ (the Clifford limit), we can perform tunable weak measurements or, equivalently, control the level of coherent errors. 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. The decoder produces an estimate of the quantum sample based on its limited knowledge in the form of $\{s'\}$ [21, 27, 28], where $\{s'\}$ is a copy of $\{s\}$ corrupted by a finite probability p_s of noise that can come from either the quantum circuit or the classical communication, as shown in Fig. 1b. We employ a fast decoder [21] based on minimum-weight-perfect-matching (MWPM), which outputs a bit-string $\{\sigma' = \pm 1\}$ 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. 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 interactions [30–33] had typical error rates of 0.0077 and square root of Pauli-X gates with error rates of 0.0002 ([34]). The typical device measurement error rates of 0.010, which 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 [18], 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 errors leads to a threshold line as shown in Fig. 1c, 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

$$|\psi(s_{ij})\rangle = e^{-\frac{\beta}{2}\sum_{\langle ij \rangle} s_{ij} Z_i Z_j} |+\rangle^{\otimes N}, \tag{1}$$

where $\beta = 2\tanh^{-1}\tan(t_A)$ [18], and N denotes the number of system qubits. The probability of such a measurement outcome follows from Born's rule

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

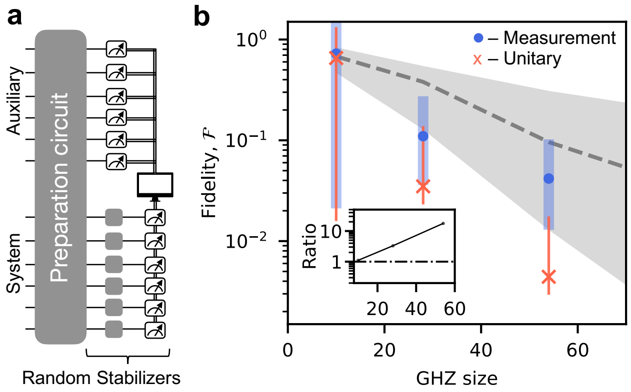

FIG. 2. **Decoded fidelity estimation by randomly sampling GHZ stabilizers. a.** Because our decoder was implemented as Pauli corrections on the system qubits, the characterization of random stabilizers, which is measured in basis rotated by single-qubit rotations (small gray boxes), needed to be done in conjunction with the implemented decoder (symbolized by the monitor). See Methods for details. **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 (See Methods for more details).The theoretically predicted fidelities for measurement-based protocol (dashed gray 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 gray [35]. The inset shows the ratio of the experimentally evaluated measurement- to unitary-based fidelities increasing for system size up to 54 sites.

which resembles the partition function of the random bond Ising model (RBIM) [18]. Concretely, by Eq. (2) we analytically map our protocol onto a RBIM precisely tracking the Nishimori line [35] with an effective disorder probability

$$\tilde{p} = \frac{1-(1-2p_s)\sin(2t_A)}{2}\,, \tag{3}$$

as a joint action of *both* coherent and incoherent errors that drives the phase transitions across the blue line in Fig. 1c. In particular, this implies that every point in the extended transition line shares the same Nishimori criticality. This scenario for the quantum protocol is quite distinct from the classical RBIM, whose schematic phase diagram is shown in Fig. 1d, 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 [36, 37]. 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 Methods for more details). To assess the relative performance of our protocol, we also implemented a standard unitary protocol for constructing GHZ states [9]. In Fig. 2, 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 10 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. 2b). 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 estimating these fidelities. To explain the experimentally measured fidelities, we compared our results against the predicted fidelities based on a noise model with $\approx$ 5% incoherent auxiliary errors and $\approx$ 3% data readout errors – values inferred in the next section. This places us in the long-range ordered phase in Fig. 1c (green star), which in the absence of any additional errors, has long-range GHZ-type entanglement, whilst its predicted GHZ fidelity shown in gray in Fig. 2 decays exponentially with the number of system qubits. We see that the experimentally obtained

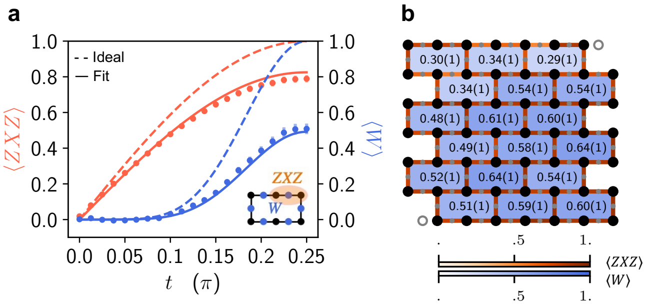

FIG. 3. **Experimentally measured local observables used to generate the state. a.** For two observables, we plot the ideally expected outcomes (dashed lines), the unprocessed experimental data (dots), and a *one parameter* fit (solid line) for each observable for sweeping t_A from 0 (trivial) to $\pi/4$ (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. 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 [35]. The experimental data exhibits 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.** 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. The numbers inside plaquettes show $\langle W \rangle$ with parenthesis showing standard error.

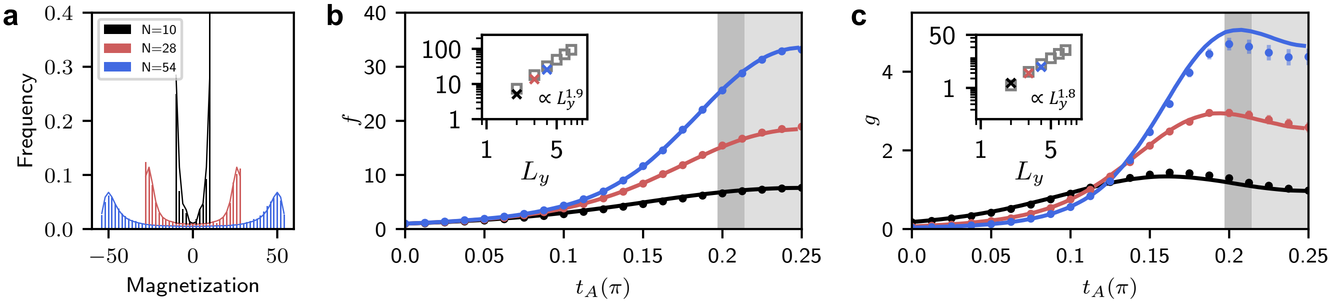

FIG. 4. **Nishimori transition by tuning coherent errors. a.** The distribution of M at $t_A = \pi/4$ for system sizes of 10 (black), 28 (red) and 54 (blue) system qubits using 21, 63, and 125 total qubits, respectively. Solid envelope lines are theoretical estimates using a two-parameter noise model. **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 gray), the state exhibits long-range order (light gray). 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 gray) from short-range correlated phases. The dark gray 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 $\propto L_y^{1.8}$. The noiseless envelopes for all solid curves can be found in the Supplemental Materials.

values are slightly suppressed with respect to the gray curve, which is likely due to dephasing. This raises the question whether we retain robust long-range order in the presence of such dephasing.

Noise analysis. To 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 – we tuned one type of coherent error, via Eq. (1), uniformly across the device; in this sweep, we monitored and fit the experimental observables [18] 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 $\langle W\rangle = (1-2p_s)^6\sin(2t_A)^6$, as shown in Fig. 3a (Sec. IC in Ref. 35). 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 [11] – a precursor state for the GHZ state – with experimental data shown in Fig. 3b. For t_A below $\pi/4$, the implemented circuits become non-Clifford and thus cannot, in general, be efficiently characterized. Nonetheless our modeling 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 $p_\sigma \approx 1.2\% - 2.3\%$ – values consistent with our expectation based on standard calibration benchmarks of the device [35].

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 MWPM decoder [21, 27, 38] 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 [18]. This correction factor approximates the ground state configuration of the RBIM as a classical estimate, $\{\sigma'\}$, of the bitstring from quantum device (Fig. 1b).

The distribution of the decoded bitstrings in the computational basis is shown, for $t_A = \pi/4$, in Fig. 4a, where we sum over the Z expectation values of the individual qubit to obtain a total decoded 'magnetization' $M = \sum_{j=1}^{N} Z_j$. Any bias of this distribution (e.g. 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 $\langle M\rangle \approx 0.02(2)N$ suggests that the global Ising symmetry is largely preserved. Moreover, the decoded, bi-modal experimental distribution (Fig. 4a) 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 $t_A \leq \pi/4$, the decoded system qubit bitstrings and the average two-point, *classical* correlations

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

which is a sum of the correlations $\langle Z_i Z_j\rangle$ for all the system qubits that compose the quantum state. The decoded experimental data is shown in Fig. 4b, where the solid line shows the theoretical benchmark with the noise parameters inferred from Fig. 3a. We observe a hallmark of the long-range ordered phase in the diverging f for increasing system sizes; such divergent behavior 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 divergent behavior for our 2D protocol should be contrasted to results in 1D geometries [35], 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 of the normalized variance of M^2/N:

$$g := \frac{1}{N^3}\left(\langle M^4\rangle - \langle M^2\rangle^2\right) , \tag{5}$$

which quantifies the amount of fluctuations in the squared magnetization [35]. 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 $\tilde{p}_c \approx 6.75\%$ [18, 39] with Eq. (3), which is in very close agreement with the experimental data where the peak locations approach this predicted critical point (Fig. 4c). 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 behavior of the peak height (Fig. 4b inset), where $L_y = 2, 3, 4$ is the number of columns of qubits 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 [39]. While the criticality is exposed in the decoded correlations only, the observable $\langle ZXZ\rangle$ of Fig. 3a 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, before applying the decoder (see Fig. 1b) and thereby chart out a broader phase diagram including varying 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. 5), we observe that the Nishimori critical point t_A^c shifts towards $\pi/4$ and vanishes completely at $p_s \approx 6.75\%$ [39], the decoding threshold [35]. 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 [27]. Our experiments thus not only demonstrate 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 more noisy devices. It also significantly distinguishes a 2D from 1D protocol where the peak quickly converges to $t_A = \pi/4$ without a finite threshold [40]. Thus we claim that our experimentally implemented 2D protocol exhibits long-range order with intrinsic robustness.

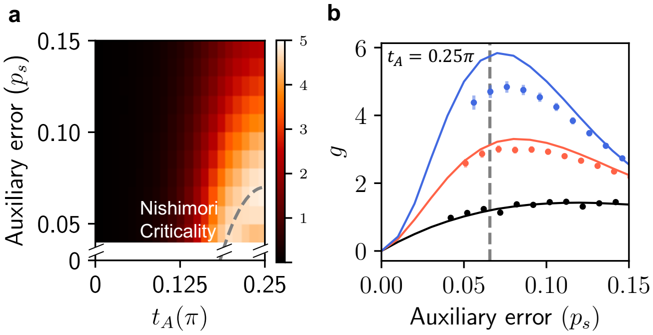

FIG. 5. **Decoding transition out of the long-range ordered phase by increasing auxiliary errors before decoding.** **a.** For the largest system size ($N = 54$), we experimentally mapped the 2D phase diagram for various coherent (t_A) and incoherent (p_s) errors where the color is proportional to the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

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 modeled by a RBIM [41]. 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 [11–24]. This is especially timely since measurements have recently been used to deterministically create exotic long-range entanglement including topological order [42–44]. In such general contexts, stability might inquire additional ingredients, such as using the time-domain [27, 45] or higher—or even fractional—dimensions, opening up a rich territory for exploration.

ACKNOWLEDGMENTS

We thank M. Ware, P. Jurcevic, Y. Kim, A. Eddins, H. Najafi, I. Lauer, G. Jones, J. Summerour, 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 and Prof. Nishimori for thoughtful discussions. 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) 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.). We acknowledge the use of IBM Quantum services for this work.

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 a Echoed Cross-Resonance (ECR) gate with 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 nanoseconds, similar to the method described in [46]. The median square root of 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 ≈ 1244.4 nanoseconds. The qubits under study had a median $T1 \approx 293\ \mu s$ and $T2 \approx 173\ \mu s$. Circuits were executed on the device at a clock rate of 1kHz [33].

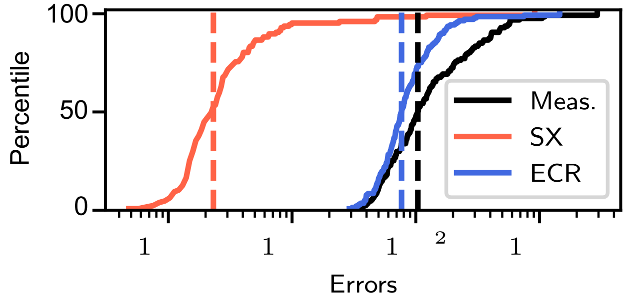

FIG. 6. **Typical error rates,** in cumulative format, for Echoed Cross Resonance (ECR, blue), square-root of Pauli-X (SX, red), and measurement (Meas, black) gates. Dashed lines represent medians of distributions.

Decomposition of ZZ gates

For fidelity comparisons in Fig. 2, 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. 1, 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 [47, 48] 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 benefited 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, and found some schedules to significantly outperform others [35]. 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 sampled up to 100 different non-Z stabilizers (e.g. 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 re-sampling of Z-only and non-Z stabilizers with equal probability to emulate the proposed Monte Carlo sampling method for fidelity state estimation [36, 37]. By sampling k random instances from 2^N

$$\mathcal{F} = \frac{1}{k}\sum_{i=1}^{k} \langle \mathcal{O}_i \rangle \,. \tag{6}$$

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 actually applied [35, 49].

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. I.

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

TABLE I. **Two-parameter noise model.** Fits to experimental data gives p_s which captures errors at the auxiliary qubits, and p_σ at the system qubits.

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

[1] S. Bravyi, M. B. Hastings, and F. Verstraete, Lieb-Robinson Bounds and the Generation of Correlations and Topological Quantum Order, Phys. Rev. Lett. **97**, 050401 (2006).

[2] M. B. Hastings, Locality in quantum systems, in *Quantum Theory from Small to Large Scales*, Les Houches 2010, Session 95 (Oxford University Press, 2012) pp. 171–212.

[3] X. Chen, Z.-C. Gu, and X.-G. Wen, Classification of gapped symmetric phases in one-dimensional spin systems, Phys. Rev. B **83**, 035107 (2011).

[4] X. Chen, Z.-C. Gu, and X.-G. Wen, Complete classification of one-dimensional gapped quantum phases in interacting spin systems, Phys. Rev. B **84**, 235128 (2011).

[5] B. Zeng and X.-G. Wen, Gapped quantum liquids and topological order, stochastic local transformations and emergence of unitarity, Phys. Rev. B **91**, 125121 (2015).

[6] Y. Huang and X. Chen, Quantum circuit complexity of one-dimensional topological phases, Phys. Rev. B **91**, 195143 (2015).

[7] J. Haah, An invariant of topologically ordered states under local unitary transformations, Communications in Mathematical Physics **342**, 771 (2016).

[8] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, Whole-Device Entanglement in a 65-Qubit Superconducting Quantum Computer, Advanced Quantum Technologies **4**, 2100061 (2021).

[9] G. J. Mooney, G. A. L. White, C. D. Hill, and L. C. L. Hollenberg, Generation and verification of 27-qubit Greenberger-Horne-Zeilinger states in a superconducting quantum computer, Journal of Physics Communications **5**, 095004 (2021).

[10] K. X. Wei, I. Lauer, S. Srinivasan, N. Sundaresan, D. T. McClure, D. Toyli, D. C. McKay, J. M. Gambetta, and S. Sheldon, Verifying multipartite entangled Greenberger-Horne-Zeilinger states via multiple quantum coherences, Phys. Rev. A **101**, 032343 (2020).

[11] H. J. Briegel and R. Raussendorf, Persistent Entanglement in Arrays of Interacting Particles, Phys. Rev. Lett. **86**, 910 (2001).

[12] R. Raussendorf, S. Bravyi, and J. Harrington, Long-range quantum entanglement in noisy cluster states, Phys. Rev. A **71**, 062313 (2005).

[13] L. Piroli, G. Styliaris, and J. I. Cirac, Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter, Phys. Rev. Lett. **127**, 220503 (2021).

[14] N. Tantivasadakarn, R. Thorngren, A. Vishwanath, and R. Verresen, Long-range entanglement from measuring symmetry-protected topological phases, (2021), arXiv:2112.01519.

[15] R. Verresen, N. Tantivasadakarn, and A. Vishwanath, Efficiently preparing Schrödinger's cat, fractons and non-Abelian topological order in quantum devices, (2021), arXiv:2112.03061.

[16] S. Bravyi, I. Kim, A. Kliesch, and R. Koenig, Adaptive constant-depth circuits for manipulating non-abelian anyons, (2022), arXiv:2205.01933.

[17] T.-C. Lu, L. A. Lessa, I. H. Kim, and T. H. Hsieh, Measurement as a Shortcut to Long-Range Entangled Quantum Matter, PRX Quantum **3**, 040337 (2022).

[18] G.-Y. Zhu, N. Tantivasadakarn, A. Vishwanath, S. Trebst, and R. Verresen, Nishimori's cat: stable long-range entanglement from finite-depth unitaries and weak measurements, (2022),

arXiv:2208.11136.

[19] N. Tantivasadakarn, A. Vishwanath, and R. Verresen, Hierarchy of Topological Order From Finite-Depth Unitaries, Measurement, and Feedforward, PRX Quantum **4**, 020339 (2023).

[20] N. Tantivasadakarn, R. Verresen, and A. Vishwanath, Shortest route to non-abelian topological order on a quantum processor, Phys. Rev. Lett. **131**, 060405 (2023).

[21] J. Y. Lee, W. Ji, Z. Bi, and M. P. A. Fisher, Decoding Measurement-Prepared Quantum Phases and Transitions: from Ising model to gauge theory, and beyond, (2022), arXiv:2208.11699.

[22] Y. Li, H. Sukeno, A. P. Mana, H. P. Nautrup, and T.-C. Wei, Symmetry-enriched topological order from partially gauging symmetry-protected topologically ordered states assisted by measurements (2023), arXiv:2305.09747 [quant-ph].

[23] H. Buhrman, M. Folkertsma, B. Loff, and N. M. P. Neumann, State preparation by shallow circuits using feed forward (2023), arXiv:2307.14840 [quant-ph].

[24] T.-C. Lu, Z. Zhang, S. Vijay, and T. H. Hsieh, Mixed-state long-range order and criticality from measurement and feedback (2023), arXiv:2303.15507.

[25] A. J. Friedman, C. Yin, Y. Hong, and A. Lucas, Locality and error correction in quantum dynamics with measurement, (2023), arXiv:2206.09929.

[26] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, J. G. Bohnet, N. C. Brown, N. Q. Burdick, W. C. Burton, S. L. Campbell, J. P. Campora III, C. Carron, J. Chambers, J. W. Chen, Y. H. Chen, A. Chernoguzov, E. Chertkov, J. Colina, M. DeCross, J. M. Dreiling, C. T. Ertsgaard, J. Esposito, B. Estey, M. Fabrikant, C. Figgatt, C. Foltz, M. Foss-Feig, D. Francois, J. P. Gaebler, T. M. Gatterman, C. N. Gilbreth, J. Giles, E. Glynn, A. Hall, A. M. Hankin, A. Hansen, D. Hayes, B. Higashi, I. M. Hoffman, B. Horning, J. J. Hout, R. Jacobs, J. Johansen, T. Klein, P. Lauria, P. Lee, D. Liefer, S. T. Lu, D. Lucchetti, A. Malm, M. Matheny, B. Mathewson, K. Mayer, D. B. Miller, M. Mills, B. Neyenhuis, L. Nugent, S. Olson, J. Parks, G. N. Price, Z. Price, M. Pugh, A. Ransford, A. P. Reed, C. Roman, M. Rowe, C. Ryan-Anderson, S. Sanders, J. Sedlacek, P. Shevchuk, P. Siegfried, T. Skripka, B. Spaun, R. T. Sprenkle, R. P. Stutz, M. Swallows, R. I. Tobey, A. Tran, T. Tran, E. Vogt, C. Volin, J. Walker, A. M. Zolot, and J. M. Pino, A Race Track Trapped-Ion Quantum Processor, (2023), arXiv:2305.03828.

[27] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, Journal of Mathematical Physics **43**, 4452 (2002).

[28] S. J. Garratt, Z. Weinstein, and E. Altman, Measurements conspire nonlocally to restructure critical quantum states, (2022), arXiv:2207.09476.

[29] H. Nishimori, Internal Energy, Specific Heat and Correlation Function of the Bond-Random Ising Model, Progress of Theoretical Physics **66**, 1169 (1981).

[30] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Simple All-Microwave Entangling Gate for Fixed-Frequency Superconducting Qubits, Phys. Rev. Lett. **107**, 080502 (2011).

[31] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, Procedure for systematically tuning up cross-talk in the cross-resonance gate, Phys. Rev. A **93**, 060302 (2016).

[32] M. Malekakhlagh, E. Magesan, and D. C. McKay, First-principles analysis of cross-resonance gate operation, Phys. Rev. A **102**, 042605 (2020).

[33] N. Sundaresan, I. Lauer, E. Pritchett, E. Magesan, P. Jurcevic, and J. M. Gambetta, Reducing unitary and spectator errors in cross resonance with optimized rotary echoes, PRX Quantum **1**, 020318 (2020).

[34] IBM Quantum. https://quantum-computing.ibm.com/, 2021. (Downloaded August 6, 2023).

[35] See Supplemental Material.

[36] S. T. Flammia and Y.-K. Liu, Direct fidelity estimation from few pauli measurements, Phys. Rev. Lett. **106**, 230501 (2011).

[37] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Practical characterization of quantum devices without tomography, Phys. Rev. Lett. **107**, 210404 (2011).

[38] O. Higgott, PyMatching: A Python package for decoding quantum codes with minimum-weight perfect matching, arXiv:2105.13082.

[39] S. L. A. de Queiroz, Multicritical point of Ising spin glasses on triangular and honeycomb lattices, Phys. Rev. B **73**, 064410 (2006).

[40] See Extended Data Fig. 7.

[41] K. Binder and A. P. Young, Spin glasses: Experimental facts, theoretical concepts, and open questions, Rev. Mod. Phys. **58**, 801 (1986).

[42] M. Iqbal, N. Tantivasadakarn, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, A. Hankin, N. Hewitt, C. V. Horst, M. Matheny, T. Mengle, B. Neyenhuis, A. Vishwanath, M. Foss-Feig, R. Verresen, and H. Dreyer, Topological Order from Measurements and Feed-Forward on a Trapped Ion Quantum Computer (2023), arXiv:2302.01917.

[43] M. Foss-Feig, A. Tikku, T.-C. Lu, K. Mayer, M. Iqbal, T. M. Gatterman, J. A. Gerber, K. Gilmore, D. Gresh, A. Hankin, N. Hewitt, C. V. Horst, M. Matheny, T. Mengle, B. Neyenhuis, H. Dreyer, D. Hayes, T. H. Hsieh, and I. H. Kim, Experimental demonstration of the advantage of adaptive quantum circuits (2023), arXiv:2302.03029.

[44] M. Iqbal, N. Tantivasadakarn, R. Verresen, S. L. Campbell, J. M. Dreiling, C. Figgatt, J. P. Gaebler, J. Johansen, M. Mills, S. A. Moses, J. M. Pino, A. Ransford, M. Rowe, P. Siegfried, R. P. Stutz, M. Foss-Feig, A. Vishwanath, and H. Dreyer, Creation of Non-Abelian Topological Order and Anyons on a Trapped-Ion Processor (2023), arXiv:2305.03766.

[45] M. B. Hastings and J. Haah, Dynamically Generated Logical Qubits, Quantum **5**, 564 (2021).

[46] Y. Kim, A. Eddins, S. Anand, K. X. Wei, E. Van Den Berg, S. Rosenblatt, H. Nayfeh, Y. Wu, M. Zaletel, K. Temme, *et al.*, Evidence for the utility of quantum computing before fault tolerance, Nature **618**, 500 (2023).

[47] N. Earnest, C. Tornow, and D. J. Egger, Pulse-efficient circuit transpilation for quantum applications on cross-resonance-based hardware, Phys. Rev. Res. **3**, 043088 (2021).

[48] Y. Kim, C. J. Wood, T. J. Yoder, S. T. Merkel, J. M. Gambetta, K. Temme, and A. Kandala, Scalable error mitigation for noisy quantum circuits produces competitive expectation values, Nature Physics **19**, 752 (2023).

[49] E. van den Berg, Z. K. Minev, and K. Temme, Model-free readout-error mitigation for quantum expectation values, Phys. Rev. A **105**, 032620 (2022).

EXTENDED DATA

As discussed in the main text, the 2D protocol exhibited robustness over the 1D protocol; the key signature being based on the scaling of the average of two-point correlations, f, as a function of system size (Fig. 7). Whereas in 2D (Fig. 9), we not only observed the expected $f \propto L_y^{1.9}$ behavior in the long-range ordered state, but also that the criticality occurs below the GHZ point ($t_A^c < \pi/4$). The 1D behavior (Fig. 8), in contrast, exhibited no growth in f with system size from 28 to 54, and had peak variances at the GHZ value of $t_A = \pi/4$.

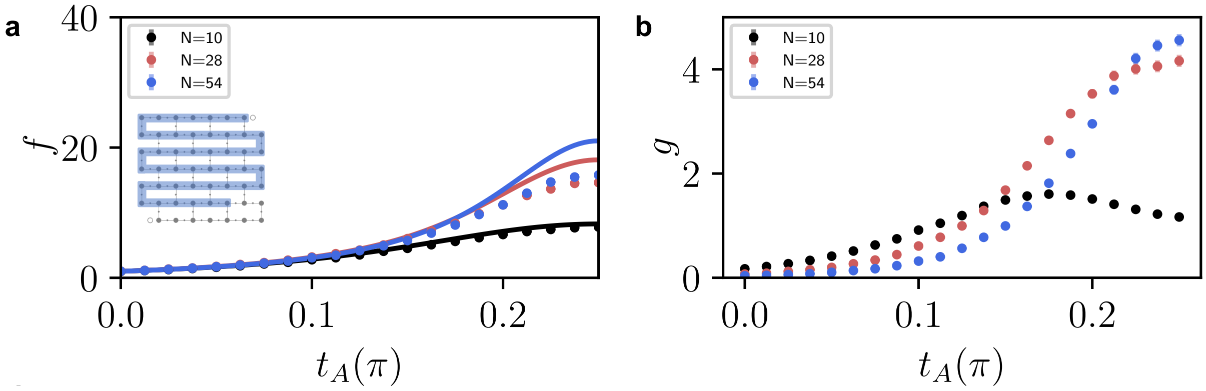

FIG. 7. **Absence of finite threshold in one-dimensional protocol**, for comparison with Fig. 4. (a) f grows with increasing system size but converges to finite value that depends on t_A. (b) The peak of g converges to $t_A = \pi/4$ indicative of absence of finite threshold for coherent error.

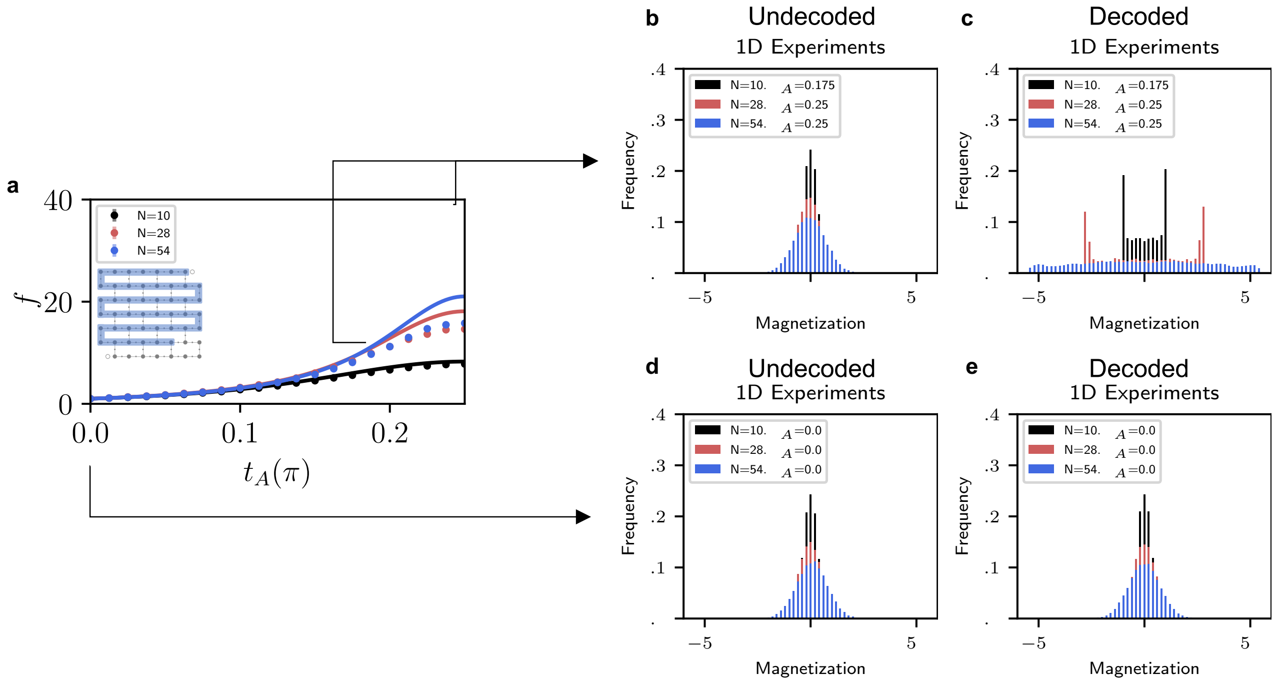

FIG. 8. **Magnetization of 1D experiments with and without decoding at different t_A values** (**a**) Two-point correlations in 1D experiments for sweeps of t_A. The histograms at values of t_A where variances were maximized for undecoded (**b**) and decoded (**c**). Although the bimodal distribution 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 ($t_A = 0$).

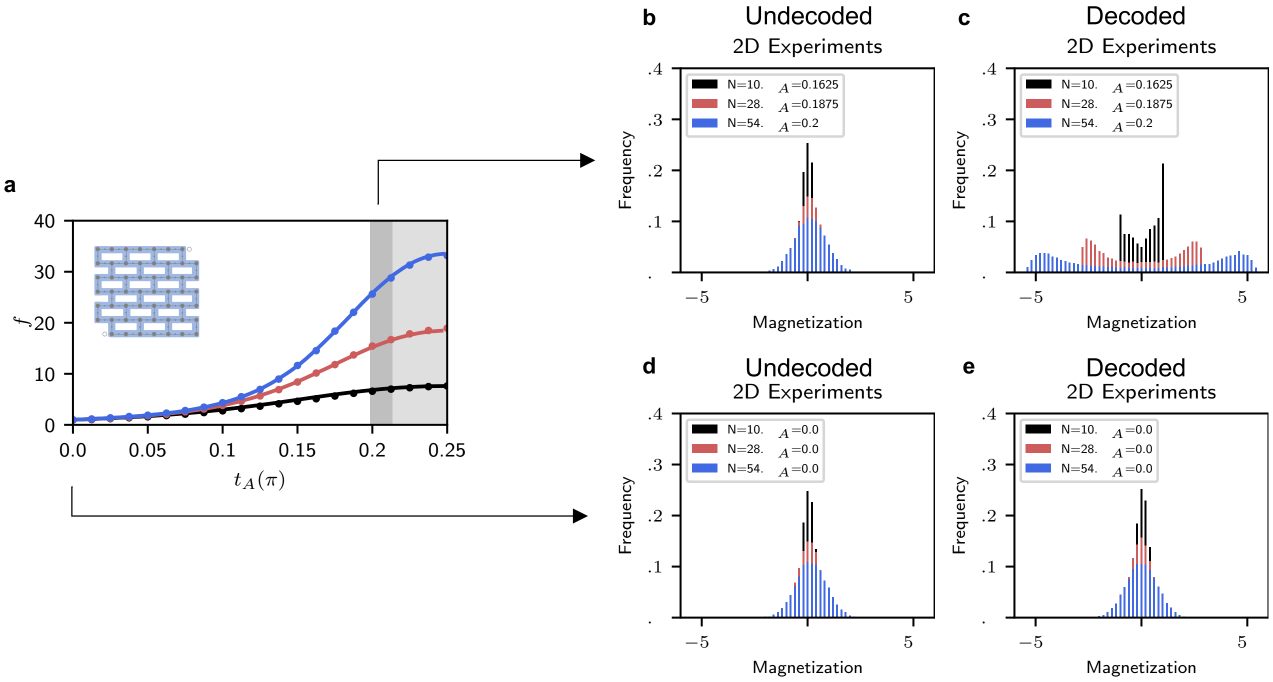

FIG. 9. **Magnetization of 2D experiments with and without decoding at different t_A values** (**a**) Two-point correlations in 2D experiments for sweeps of t_A. The histograms at values of t_A where variances were maximized for undecoded (**b**) and decoded (**c**). In contrast to the 1D cases (Fig. 8), 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 ($t_A = 0$).

Supplementary Information for:
"Realizing the Nishimori transition across the error threshold for constant-depth quantum circuits"

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] and Abhinav Kandala[5]

[1]*IBM Quantum, Almaden Research Center, San Jose, CA 95120, USA*
[2]*IBM Quantum, Research Triangle Park, NC 27709, USA*
[3]*Institute for Theoretical Physics, University of Cologne, Zülpicher Straße 77, 50937 Cologne, Germany*
[4]*Department of Physics, Harvard University, Cambridge, MA 02138, USA*
[5]*IBM Quantum, T. J. Watson Research Center, Yorktown Heights, NY 10598, USA*
[6]*IBM Quantum, IBM Research, Zurich, Rüschlikon, Switzerland.*
[7]*Walter Burke Institute for Theoretical Physics and Department of Physics,*
California Institute of Technology, Pasadena, CA 91125, USA

arXiv:2309.02863v1 [quant-ph] 6 Sep 2023

CONTENTS

I. THEORY

A. Threshold for incoherent error

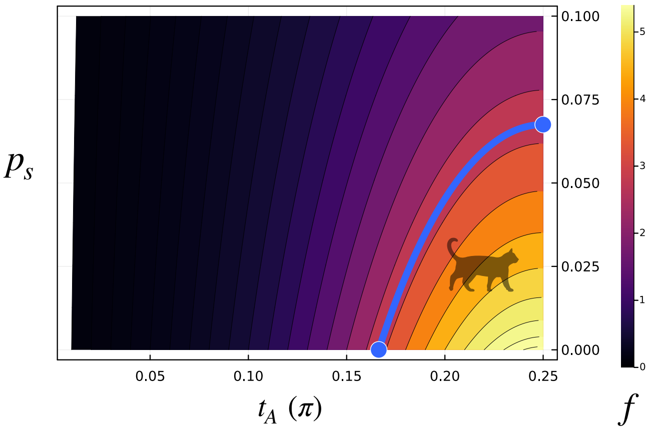

FIG. 1. **Theoretical phase diagram with contours for f.** The blue line illustrates the Nishimori critical line. Eq. (1) indicates that all parameter points on the same contour shares the same decoded bit-string ensemble. Decoded f is computed from noiseless classical calculations at $p_s = 0$ for 4×7 brickwall lattice with 54 sites.

Here, we prove that the protocol under both coherent error t_A and incoherent bit flip error upon auxiliary qubits is still mapped to the random bond Ising model and has its threshold determined by the Nishimori critical point.

Let us denote the event of an error flip for the auxiliary qubit by $s' = +1$, and absence of flip as $s' = -1$, under probability $e^{-\beta' s'}/(2\cosh\beta')$. The decoded LRE order parameter is the element-wise product between the quantum replica and the classical decoder replica. The classical decoder serves as an estimate of the observable $\langle\sigma\rangle_{s;\beta}$ in the quantum replica. However, based on the distorted and thus *imprecise* knowledge of the auxiliary read out i.e. $s \cdot s'$, it provides $\langle\sigma\rangle_{ss';\beta_d}$ instead, where β_d is the temperature parameter that can be freely tuned in the classical decoder, to optimize its performance. Under incoherent error, the decoded order parameter as quantum-classic-replica-correlation can be represented and

---

[*] ehchen@ibm.com
[†] gzhu@uni-koeln.de
[‡] rubenverresen@gmail.com

simplified as

$$\begin{aligned} q &= \sum_{s,s'} P(s)\cdot\langle\sigma\rangle_{s;\beta}\cdot P(s')\cdot\langle\sigma'\rangle_{ss';\beta_d} \\ &\propto \sum_{s,s',\sigma} e^{-\beta s\sigma\sigma}\sigma\cdot e^{-\beta' s'}\cdot\langle\sigma'\rangle_{ss';\beta_d} \\ &\propto \sum_{s,s'} e^{-\beta s}\cdot e^{-\beta' s'}\cdot\langle\sigma\sigma'\rangle_{ss';\beta_d} \\ &\propto \sum_{s} e^{-\tilde{\beta} s}\langle\sigma\rangle_{s;\beta_d}\,. \end{aligned} \tag{1}$$

Here, for simplicity we abbreviate the summation notation over bonds; from the first to the second line we make use of the property that $P(s)$ is the wave function amplitude, which cancels out the denominator for $\langle\sigma\rangle_{s;\beta}$; from the second to the third line we apply a gauge transformation $s\to s\sigma\sigma$ which is crucial to yielding the analytic Nishimori properties; from the third to the fourth line we combine s and s' as the effective disorder, with probability

$$\tilde{p} = \frac{1}{1+e^{2\tilde{\beta}}} = \frac{1-\tanh\beta\tanh\beta'}{2} = \frac{1-\sin(2t_A)(1-2p_s)}{2}\,. \tag{2}$$

This effective disorder, $\tilde{p}$, is a joint action of both the coherent gate error that results in β and the incoherent auxiliary error β' (Fig. 1). Note that their contributions are symmetric: $\beta\leftrightarrow\beta'$ leaves it invariant.

Now let us discuss the optimal decoder. As is rigorously proved by Nishimori [1], for a fixed disorder probability $\tilde{p}$, the optimal magnetization is reached only at the Nishimori temperature i.e. $\beta_d\to\tilde{\beta}$. Thus we prove that for arbitrary coherent gate and incoherent bit flip error rates (β,β'), the optimal threshold is uniquely pinned by the Nishimori critical point at $\tilde{p}=p_c\approx 6.75(5)\%$ on honeycomb lattices [2] (smaller than $\approx 10(1)\%$ on square lattices [3]).

For simplicity the derivation above only shows the mapping for the correlation of a single qubit between two replicas being mapped to RBIM Nishimori line. It *mutatis mutandis* carries over to arbitrary-point correlations, and thus any polynomial functions $f(\{\sigma\})$. This includes the fidelity with the GHZ state, when dephasing is absent i.e. $\prod_j X_j=+1$:

$$F = |\langle GHZ|\psi\rangle|^2 = \langle\prod_{\langle ij\rangle}\frac{1+\sigma_i\sigma_j}{2}\rangle\,. \tag{3}$$

B. Using MWPM decoder

An estimate [4] of $\langle\sigma\rangle_{\{s\}}$ based on the random bond configuration $\{s\}$ is to use the minimum-weight-perfect-matching (MWPM) [5], which finds the shortest path to pair up the syndromes. This is equivalent to finding a classical ground state of the Ising spins, where the shortest path costs minimal domain-wall energy (Fig. 2). The decoded observables from classical calculations is shown in Fig. 3. By using Eq. 1 to map this noiseless data to the noisy scenario, we obtain the benchmark data for our experimental data in Fig. 4 and Fig. 5 of the main text.

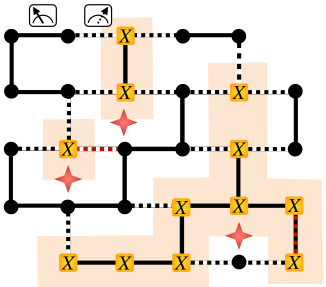

FIG. 2. **Using MWPM decoder**. The black dashed (solid) bond denotes positive (negative) measurement outcome. The red stars show the syndromes, a plaquette surrounded by an odd number of black dashed bonds. The shortest path that pairs up the syndromes (found by MWPM) goes across the red dashed bond. By cutting the dashed bonds, the sites are grouped into domains, where the qubits inside the domain shall align but the adjacent domains shall have opposite $\langle\sigma\rangle$. We generate such configuration by fixing the top-left site, and applying X flip operations to every other domains as shown by the yellow squares and the shaded regions.

C. Benchmark data for two-dimensional experiments

The inferred auxiliary (incoherent) error rates, p_s, allow us to estimate the expected critical transition, t_A^C. In this manner, we were able to generate a range of critical threshold values in Table. I.

System size (N)	p_s (auxiliary)	t_A^C (asymptotic)
10	0.042	0.20
28	0.051	0.21
54	0.056	0.21

TABLE I. **Inferred asymptotical critical transitions, t_A^C, from auxiliary errors (p_s) for different system sizes**. The critical point is deduced from Eqn.1 in the main text i.e. $\sin(2t_A^C)=(1-2\tilde{p})/(1-2p_s)$, by inputting the disorder probability of Nishimori critical point $\tilde{p}\approx 6.75\%$ [2].

And finally, the experimentally measured values of f and g, extracted at the values of t_A where g was maximized, were compared against the theoretically predicted, noiseless values (Fig. 3) in Table. II. These exact values were used in the inset for Figures 4b and 4c in the main text.

D. Analytic solution for one-dimensional protocol

First, the decoded two-point correlation in 1D is equivalent to the gauge invariant string correlation [6], which in the presence of errors can be derived as

$$\langle\sigma_i s\cdots s\sigma_j\rangle = (1-2p_\sigma)^2((1-2p_s)\sin(2t_A))^{|j-i|}\,. \tag{4}$$

Here i,j denotes only the sites of system qubits σ, and s denotes the measurement outcomes on the bonds between them.

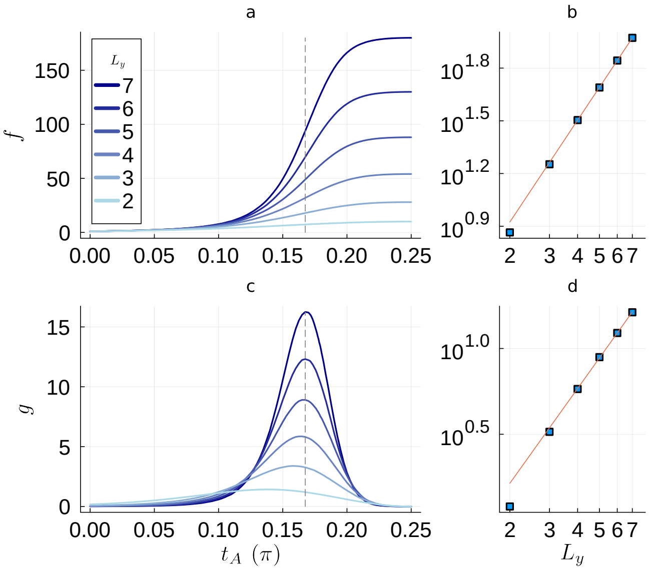

FIG. 3. **Noiseless classical calculations as benchmark for quantum experiment**, where the brickwall lattice is chosen to fit the IBM processors from Falcon to Osprey generations (with L_y columns and $2L_y - 1$ rows of system qubits). (a)(c) Decoded observables by sweeping t_A for increasing system sizes. Dashed line illustrates the peak location of g that indicates the critical point $t_A^c \approx 0.1675\pi$ (which is translated to critical disorder probability 6.568%, slightly smaller than the critical probability 6.75(5)% in Ref. [2] due to our use of MWPM decoder). (b)(d) Scaling of f and g at critical point. The red lines indicate the power-law fitting, where the first two system sizes are discarded. We find $f \sim L^{1.9}$; $g \sim L^{1.8}$.

System size (N)	L_y	f_C (data)	f_C (theory)	g_C (data)	g_C (theory)	$t_A^{\max g}(\pi)$ (data)
10	2	5.1	7.3	1.4	1.2	0.1625
28	3	13.7	17.9	3.0	3.3	0.1875
54	4	25.6	31.9	4.7	5.8	0.2000
88	5	-	48.9	-	8.9	-
130	6	-	69.7	-	12.3	-
180	7	-	93.8	-	16.3	-

TABLE II. **Comparison of f and g values from experimentally measured data at the peak locations of g, against noiseless theoretical values at critical point.** System size (N) describes the total number of particles in the RBIM, and L_y the linear dimension of the 2D system. The subscripts C denote the values of f and g at the *critical* values. For the experimental data, the values for f and g were taken at discrete values indicated by $t_A^{\max g}$. The theoretical values were taken at $t_A = 0.1675$ for all system sizes - in other words, these are the points shown in Supplemental Fig. 3b and 3d.

Note that the 1D protocol yields the same result as the 2D protocol if and only if $t_A = \pi/4$ and $p_s = 0$. The presence of infinitesimal coherent ($t_A < \pi/4$) or incoherent error ($p_s > 0$) would contribute to an exponential decaying factor. Thus the correlation is bounded by a finite correlation length $\xi \sim -1/\ln((1-2p_s)\sin(2t_A))$. This makes it qualitatively different from the 2D protocol. Using Eq. (4) one can fit p_σ and p_s from the experimental data of every 5 qubits e.g. $\langle Z_1 X_2 \mathbb{I}_3 X_4 Z_5\rangle = (1-2p_\sigma)^2(1-2p_s)^2\sin(2t_A)^2$ and $\langle Z_1X_2Z_3\rangle\langle Z_3X_4Z_5\rangle = (1-2p_\sigma)^4(1-2p_s)^2\sin(2t_A)^2$.

Secondly, by summing over the string correlation, we obtain the decoded squared magnetization:

$$\begin{aligned} f &= \langle M^2\rangle/N = 1 + \frac{1}{N}\sum_{i\neq j}\langle \sigma_i s \cdots s \sigma_j\rangle \\ &= 1 + 2(1-2p_\sigma)^2 \sum_{d=1}^{N-1} \frac{N-d}{N}\left((1-2p_s)\sin(2t_A)\right)^d . \end{aligned} \tag{5}$$

Using the fit noise terms in Table III, we plotted Eqn. 5 in Fig. 7a of the main text.

System size (N)	p_s (auxiliary)	p_σ (data)
10	0.018	0.020
28	0.020	0.023
54	0.028	0.034

TABLE III. **Two-parameter noise model for the one-dimensional results in Fig. 7 of the main text.** Fits to experimental data gives p_s which captures errors at the auxiliary qubits, and p_σ at the data qubits.

II. EXPERIMENT

A. Comparing circuit depths

To further emphasize the improvement in circuit efficiency of measurement-based schemes over unitary-based schemes, we plot the number of layers needed to prepare the state as a function of system size in Fig. 4. As an illustration, the exact depth-10 circuit used for the largest system size characterized in Fig. 2 of the main text is shown in Fig. 5.

B. Circuit scheduling for 2D measurement protocol

C. Sampling of Stabilizers in quantum device

$$\begin{aligned} F &= \mathrm{Tr}\left(\rho\,|GHZ\rangle\langle GHZ|\right) \\ &= \langle \frac{1+\prod_j X_j}{2}\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2}\rangle \\ &= \frac{1}{2}\left(\langle\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2}\rangle + \langle\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2}\prod_j X_j\rangle\right) \\ &= \sum_{n=1}^{2^N}\langle P_n\rangle, \end{aligned} \tag{6}$$

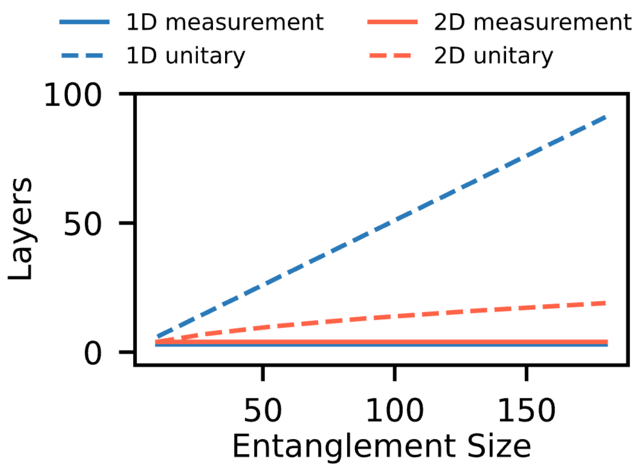

FIG. 4. Number of layers needed for measurement-based (solid) versus unitary-based (dashed) protocols for preparing entanglement of sizes up to 180 qubits. The number of layers for the unitary protocol decreases from 1D (blue) to 2D (red) protocols, but remains much less efficient than measurement-based protocols which only require constant-depth circuits irrespective of the system size.

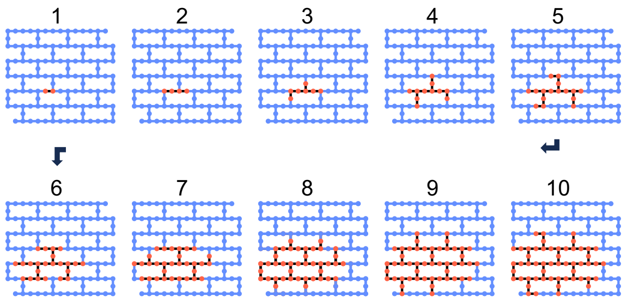

FIG. 5. Unitary protocol used to prepare entangle a system size of 54 with a circuit of depth 10.

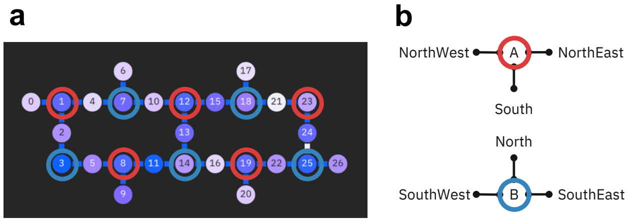

FIG. 6. a. The site qubits on the heavy-hexagonal lattice can be categorized into ‘A‘ and ‘B‘ sub-lattices. b. For each site type, there are three directions in which entangling gates can be performed. We label each direction uniquely for each site to identify all possible schedules for the depth-3 circuit such that there no overlapping entangling gates.

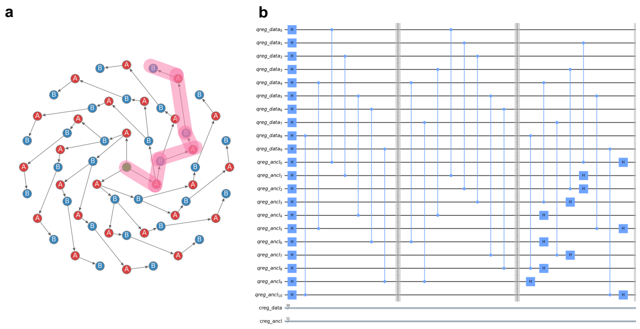

FIG. 7. a. The three concentric circles represent all the possible circuit schedules. Each layer represents a circuit layer, each of which consists of A-type and B-type entangling gates. The outer-most ring contains 12 total leaves, indicating that there are exactly 12 ways to schedule the finite depth circuit such that no overlapping gates occur. b. An example circuit schedule based on one path of traced out in (a).

where P_n specifies a Pauli string in the stabilizer group, as arbitrary product of the generators Z_iZ_j and $\prod_j X_j$. The exponential sum can be approximated by a Monte Carlo direct sampling with guaranteed precision given the sample size [7]. We corroborated their results with a simulation, using a depolarizing model, of increasing GHZ system sizes with a fixed number of randomly sampled stabilizers and observed fidelities agree to within 1% (See Fig. 8).

D. Fidelity and parity oscillation

The stabilizers in Eq. (6) can be divided into two sets, the diagonal ones generated by only Z_iZ_j bilinears, and the off-diagonal ones that involve $\prod_j X_j$. The average of the former set yields the GHZ population i.e. the probability of the two-dimensional logical subspace spanned by $|000\cdots\rangle$ and $|111\cdots\rangle$:

$$\langle\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2}\rangle = \|\langle 000\cdots|\psi\rangle\|^2 + \|\langle 111\cdots|\psi\rangle\|^2 . \quad (7)$$

The average of the latter set yields the many-body coherence which is the expectation value of $\prod_j X_j$ supported only in the logical subspace. By noting that the diagonal projector can be Fourier expanded in the phase space:

$$\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2} = \frac{1}{N}\sum_{\phi=2\pi/N}^{2\pi} e^{i\phi\sum_j Z_j} , \quad (8)$$

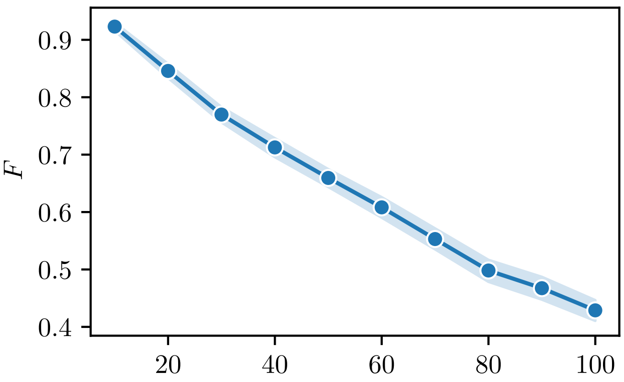

FIG. 8. Estimated fidelity (F) using 100 random stabilizers with 100 shots for all GHZ system sizes (n). The points (mean) indicate the mean of the fidelity, and error bars (shading) the 95% bootstrap confidence interval. As expected the fidelity decays with system size for a fixed depolarizing rate of 1% per two-qubit gates.

we can expand the coherence term as

$$
\begin{aligned}
&\langle\prod_{\langle ij\rangle}\frac{1+Z_iZ_j}{2}\prod_j X_j\rangle\\
=&\frac{1}{N}\sum_{\phi=2\pi/N}^{2\pi}\langle e^{i\frac{\phi}{2}\sum_j Z_j}\prod_j X_j e^{-i\frac{\phi}{2}\sum_j Z_j}\rangle \qquad (9)\\
=&\frac{1}{N}\sum_{\phi=2\pi/N}^{2\pi}\langle\prod_j\sigma_j^{\phi}\rangle\,,
\end{aligned}
$$

which is the average of the expectation value of product of Pauli operators lying in the XY plane in an angle ϕ with respect to X axis. In other words, the off-diagonal quantity characterizes the sensitiveness of the state in response to a phase rotation, manifesting in the parity oscillation period. The phase rotation can be implemented by a single-body gate rotation before the final measurement of the data qubits. Experimental result is shown in Fig. 9 for a 1-D chain of qubits.

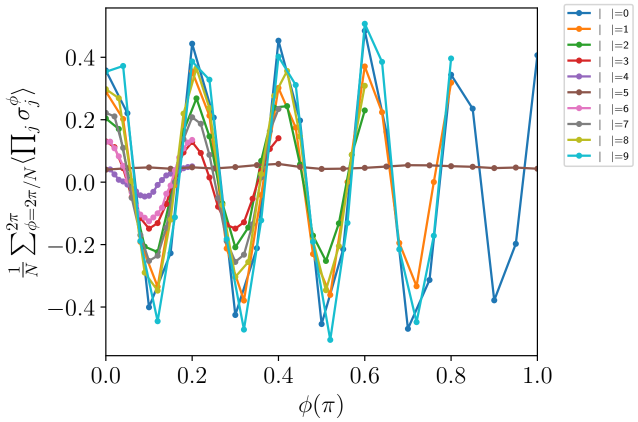

FIG. 9. Example of observed parity oscillations measured on a 1-D chain of 19 qubits for a system size of $N = 10$ (9 auxiliary qubits). While the system qubits were measured in a basis determined by ϕ (See Eqn. 9), the Pauli-X corrections were done in post-processing based on the Hamming weight of the auxiliary outcomes ($|R|$). The protocol results in a binomial distribution of $|R|$, and thus becomes exponentially unlikely to result in the extremal values of $|R| = 0, N$. Nonetheless, the fidelity of the entangled state is approximately the average of the amplitudes – in this case around 30%.

[1] H. Nishimori, Internal Energy, Specific Heat and Correlation Function of the Bond-Random Ising Model, Progress of Theoretical Physics **66**, 1169 (1981).

[2] S. L. A. de Queiroz, Multicritical point of Ising spin glasses on triangular and honeycomb lattices, Phys. Rev. B **73**, 064410 (2006).

[3] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, Journal of Mathematical Physics **43**, 4452 (2002).

[4] J. Y. Lee, W. Ji, Z. Bi, and M. P. A. Fisher, Decoding Measurement-Prepared Quantum Phases and Transitions: from Ising model to gauge theory, and beyond, (2022), arXiv:2208.11699.

[5] O. Higgott, PyMatching: A Python package for decoding quantum codes with minimum-weight perfect matching, arXiv:2105.13082.

[6] G.-Y. Zhu, N. Tantivasadakarn, A. Vishwanath, S. Trebst, and R. Verresen, Nishimori's cat: stable long-range entanglement from finite-depth unitaries and weak measurements, (2022), arXiv:2208.11136.

[7] S. T. Flammia and Y.-K. Liu, Direct fidelity estimation from few pauli measurements, Phys. Rev. Lett. **106**, 230501 (2011).