

# Exact Decoding of Repetition Code under Circuit Level Noise

Hanyan Cao,<sup>1,2,\*</sup> Shoukuan Zhao,<sup>3,\*</sup> Dongyang Feng,<sup>1,2,4,\*</sup> Zisong Shen,<sup>1,2</sup> Haisheng Yan,<sup>3</sup>  
Tang Su,<sup>3</sup> Weijie Sun,<sup>3</sup> Huikai Xu,<sup>3,†</sup> Feng Pan,<sup>1,5,6,‡</sup> Haifeng Yu,<sup>3,7</sup> and Pan Zhang<sup>1,4,§</sup>

<sup>1</sup>CAS Key Laboratory for Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup>School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>Beijing Academy of Quantum Information Sciences, Beijing 100193, China

<sup>4</sup>School of Fundamental Physics and Mathematical Sciences,

Hangzhou Institute for Advanced Study, UCAS, Hangzhou 310024, China

<sup>5</sup>Centre for Quantum Technologies, National University of Singapore, 117543, Singapore

<sup>6</sup>Science, Mathematics and Technology Cluster, Singapore University of Technology and Design, 8 Somapah Road, 487372 Singapore

<sup>7</sup>Hefei National Laboratory, Hefei 230088, China

Repetition code forms a fundamental basis for quantum error correction experiments. To date, it stands as the sole code that has achieved large distances and extremely low error rates. Its applications span the spectrum of evaluating hardware limitations, pinpointing hardware defects, and detecting rare events. However, current methods for decoding repetition codes under circuit level noise are suboptimal, leading to inaccurate error correction thresholds and introducing additional errors in event detection. In this work, we establish that repetition code under circuit level noise has an exact solution, and we propose an optimal maximum likelihood decoding algorithm called PLANAR. The algorithm is based on the exact solution of the spin glass partition function on planar graphs and has polynomial computational complexity. Through extensive numerical experiments, we demonstrate that our algorithm uncovers the exact threshold for depolarizing noise and realistic superconductor SI1000 noise. Furthermore, we apply our method to analyze data from recent quantum memory experiments conducted by Google Quantum AI, revealing that part of the error floor was attributed to the decoding algorithm used by Google. Finally, we implemented the repetition code quantum memory on superconducting systems with a 72-qubit quantum chip lacking reset gates, demonstrating that even with an unknown error model, the proposed algorithm achieves a significantly lower logical error rate than the matching-based algorithm.

Quantum computing has the potential ability to surpass classical computers in tackling challenging practical problems such as crytography [1, 2], quantum chemistry [3], and quantum simulations [4]. However, during the present-day noisy intermediate-scale quantum (NISQ) era [5], current quantum computers are not precise enough due to the inherent instability of operations caused by quantum noises. As a critical step in between quantum hardware and quantum algorithms, quantum error correction [6–15] aims to correct the errors and decrease the logical error rate to low-enough values that are needed for practical applications. In recent years, there have been numerous experimental advancements in utilizing a variety of quantum error-correcting codes [16–26], including surface code [22–25], color code [27], and subsystem code [17]. However, due to the limited qubit resources on quantum chips (with typically one hundred qubits) and the limited precision of two-qubit gates (with typically 99.9% fidelity), most experimentally realized quantum codes, e.g. a surface code, only achieved a small distance (up to distance 7) and high logical error rate (down to  $10^{-3}$ ).

Repetition code is the only code that has been realized and demonstrated experimentally to have both large distances and extremely low error rates [26, 28–30]. This is essential for evaluating hardware performance [26, 30], detecting hardware defects like leakage [28, 29], and providing foundations for

building scalable fault-tolerant quantum computers. The low error rates also pave the way for detecting new physical mechanisms that significantly impact quantum hardware and facilitate the further suppression of quantum errors. For example, cosmic ray impacts [31] causing an error floor of  $10^{-6}$  have been identified on Google’s Sycamore device by using repetition code with distance 25 [24], and Google’s recent experiment with distance 29 has uncovered previously unobserved error mechanisms at an error floor on the order of  $10^{-10}$  [25].

To achieve an extremely low error rate, the decoding algorithm is essential, as it should be as accurate as possible to avoid introducing additional algorithmic errors. In recent experiments considering syndrome error (which using the circuit level noise model), decoding was carried out using the minimum-weight perfect matching (MWPM) algorithm and its variants [32–36]. MWPM algorithm is efficient as it has a polynomial time complexity in the number of qubits. However, as a minimum weight decoder, MWPM algorithms are inherently suboptimal compared to optimal maximum likelihood decoders (MLD). This suboptimality risks introducing additional logical errors during the decoding process.

In this work, we propose an exact maximum likelihood decoding algorithm with polynomial complexity for the repetition code with circuit-level noise. Our algorithm is based on the exact solution to the partition function of spin glasses on planar graphs, so we term our method PLANAR. Our algorithm is an optimal decoder for repetition code under circuit level noise, it eliminates the error introduced in the decoding stage with a correct error model. Using our algorithm, we first computed exact thresholds for repetition codes with circuit-level noise under depolarizing and superconduct-

\* These authors contributed equally

† [xuhk@baqis.ac.cn](mailto:xuhk@baqis.ac.cn)

‡ [feng\\_pan@sutd.edu.sg](mailto:feng_pan@sutd.edu.sg)

§ [panzhang@itp.ac.cn](mailto:panzhang@itp.ac.cn)

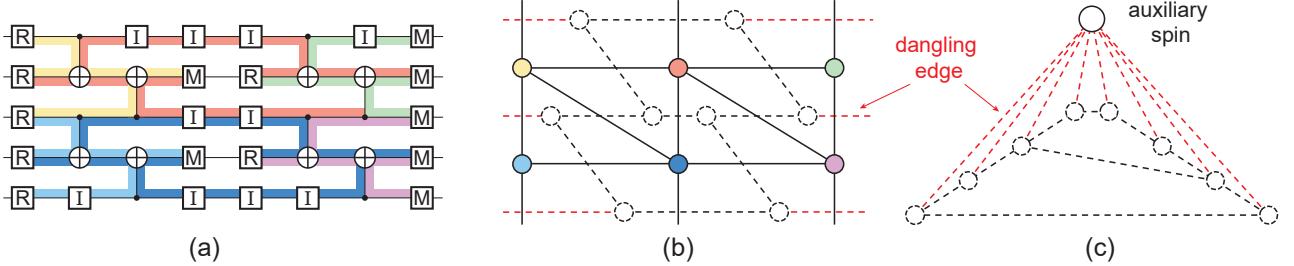


FIG. 1. (a) Syndrome measurement circuit and the detecting region of a distance-3 bit-flip repetition code with 2 rounds of measurements. We consider errors caused by the adjacent quantum operations including reset (R), measurement (M), idling (I), and two-qubit gates. Different detecting regions are labeled with different colors. (b) The error graph corresponds to (a). Solid vertices and lines stand for detectors and error mechanisms respectively. Each vertex corresponds to the detecting region with the same color in (a). Dashed vertices and their connections denote the dual graph. The red dashed lines indicate dangling edges that connect to only one detector. (c) The spin glass model used in our exact maximum likelihood decoding algorithm **PLANAR**, is mapped from the error graph in (b). An auxiliary spin is added and connected to all dangling edges of the dashed graph in (b).

ing SI1000 noise models [37], which has not been obtained to our best knowledge. Then, we applied our algorithm to Google’s repetition code experiments [25, 38], and show that compared to MWPM used by Google, **PLANAR** achieves an average improvement with 25% in logical error rate, and a higher suppression factor with  $\Lambda = 8.28$ . Finally, to evaluate the performance of our algorithm on experimental data with an unknown error model, we implemented a repetition code on superconducting systems using a 72-qubit quantum chip that lacks reset gates. By employing an error model derived through correlation analysis [26], our algorithm demonstrated up to an 8.76% improvement in logical error rates compared to MWPM.

*Repetition code with circuit level noise.* — We introduce the repetition code with measurement noise using an illustrative example with distance  $d = 3$  and  $r = 2$  rounds of repeated measurements, as shown in Fig. 1. The code comprises 3 data qubits and 2 measurement qubits that implement pairwise Z stabilizers. Ancilla qubits are measured and reset for two rounds, and the syndrome measurement results are used to infer whether a logical bit-flip error occurred. All qubits will be measured at the end to obtain the final logical state. In this circuit, errors can originate from various sources, including the operation of controlled-not gates, qubit idling (I), reset operations (R), and measurements (M). A standard approach to describing the errors and their effects on logical errors is the detector error model (DEM) [37, 39]. In DEM, a detector is essentially a linear exclusive-or (XOR) combination of multiple measurements and reset results. It focuses on detecting discrepancies in consecutive measurements across different circuit layers, offering a more compact representation than the full syndrome measurement results. As an illustrative example, the syndrome measurement circuits for a bit-flip repetition code and the corresponding detector error model corresponding to the distance-3 repetition code are presented in Fig. 1(a), where different detecting regions [37] are highlighted in different colors. Each detecting region corresponds to a detector associated with a parity check consecutively on 7 overall qubit measurements and reset. Within each

region, a specific error that activates the corresponding detector is termed error mechanism. Then we generate an error graph (in Fig. 1(b)) from the DEM, by mapping  $m = 6$  detectors and  $n = 15$  error mechanisms (that are shared by two detecting regions) into vertices and edges respectively. An error mechanism defined on an edge can be treated as the combination of an odd number of Pauli errors that anti-commute with the same set of detecting regions. It will trigger the corresponding detectors defined on the nodes. In this example, there are  $2^{15}$  possible error mechanisms grouped into two logical sectors, depending on the commutation relation of the error mechanism and the logical operator.

*Exact decoding.*— Given the syndrome measurement  $\gamma$ , which corresponds to events on detectors in the DEM, decoding is to determine whether a logical operator has been applied. In general, there are two kinds of decoding approaches: minimum weight and maximum likelihood. Minimum weight decoding, exemplified by minimum weight perfect matching, seeks the most likely error consistent with the syndrome measurements. In contrast, the maximum likelihood decoding identifies the most likely logical sector by computing  $P(l)$  for each logical sector  $l$ .  $P(l)$  stands for the total probability of all errors that are consistent with the syndrome measures and have the same commutation relations to the logical operator. Maximum likelihood decoding is theoretically optimal, but it is challenging to solve because it belongs to the #P class in computational complexity. There are no polynomial-time algorithms that can exactly solve this problem in general. In this work, we establish that, for the repetition code with circuit-level noise, exact algorithms exist for solving maximum likelihood decoding with polynomial complexity.

To demonstrate this, we use the example in Fig. 1. The maximum likelihood decoding computes the probability of all error configurations defined on the edges of Fig. 1(b) at a specific logical sector  $l$  and satisfies the syndrome measurements  $\gamma$  on the detectors defined on the nodes. To achieve this, we

can sum all Pauli operators  $s$  that do not affect the syndrome.

$$P(l) = \sum_s P[E(s, l, \gamma)], \quad (1)$$

where  $P[E(s, l, \gamma)]$  is the probability of generating a Pauli error  $E(s, l, \gamma) \in \{I, X\}^n$  (as a function of  $X$  stabilizer  $s$ , logical operator  $l$ , and syndrome  $\gamma$ ) in the error model. If we consider detectors as  $Z$  stabilizers and the error mechanisms it can detect as  $X$ -type errors, operators  $s$  are actually the  $X$  stabilizers [32] which are represented by dashed nodes in Fig. 1(b). The nodes and the edges (dashed) that connect them constitute the dual graph of the original detector error graph. Given a Pauli error  $E$ , computing  $P(l)$  in Eq.(1) can be mapped to the partition function computation of the Ising spin glass defined in Fig. 1(c). In this figure, each node defines a spin variable  $s_i = \pm 1$  denoting whether the corresponding  $X$  stabilizer is applied. On the edge  $(ij)$  connecting spins  $s_i$  and  $s_j$  there is a probability  $p_{ij}$  that an error mechanism  $E_{ij}$  shared by  $s_i$  and  $s_j$  is generated. This defines a spin-glass coupling [40]

$$J_{ij} = \frac{1}{2} \langle E_{ij}, Z_{ij} \rangle \log \frac{1 - p_{ij}}{p_{ij}}, \quad (2)$$

where  $E_{ij}$  represents the component of error  $E$  acting on the edge  $(ij)$ . The term  $\langle E_{ij}, Z_{ij} \rangle$  is the scalar commutator [40], which equals 1 if  $E_{ij}$  commutes to the Pauli  $Z$  operator acting on the edge  $(ij)$ , and  $-1$  otherwise.

The edges that are only covered by one spin correspond to a dangling edge and map to an external field acting on the spin. The effect of external fields can be fully captured by introducing an auxiliary spin. Finally and remarkably, this results in that the topology of the corresponding spin glass model with the auxiliary variable is a planar graph, as shown in Fig. 1(c). Thus, the exact partition function [41] can be computed using a polynomial algorithm such as the Kac-Ward formulism [42, 43].

*Results.*— We evaluate our approach using three examples. In the first example, we use synthetic data generated by depolarizing and SI1000 error models and compute the exact thresholds; in the second example, we revisit the decoding of Google’s recent repetition-code quantum memory experiments [25] using their experimental data and error model; in the last example, we conduct the quantum memory experiments using a 72-qubit quantum chip and evaluate the performance of our decoding approach with an error model derived from correlation analysis.

First, we study decoding from synthetic syndromes, where errors were drawn randomly from an error model and syndromes are computed (using Stim [39]) for errors and used for decoding. In the case when the error model used in decoding matches the actual error model in generating the data, as a typical inference problem, the decoding is on the Nishimori line [44, 45] and our exact decoder gives optimal decoding results and the exact threshold. In this work, we study two circuit-level error models, the standard depolarizing model and the SI1000 model, which was proposed to model the errors in superconducting devices [37]. The results are depicted in Fig. 2(a) and (b). We can see that, for both er-

ror models with different physical error rates, our planar decoder always gives lower logical error rates than the minimum weight perfect matching. Our data also gives the exact threshold 0.067(7) for the depolarizing noise and 0.020(5) for the SI1000 noise. It is clear that minimum weight perfect matching gives lower thresholds. The inset of Fig. 2(a) shows that the time complexity of PLANAR is  $O(N^{0.82})$ , with  $N$  denoting the number of spins and is linear in the number of detectors. Next, we apply our decoding algorithm to experimental data of quantum memory experiments. We first apply to Google’s repetition code experimental data on the WILLOW quantum chip [25] with 72 qubits. We randomly chose part of the data from Google’s  $d = 29$  repetition code experiment and used the subsample strategy [46] to obtain the data of smaller code as well as the error model. We calculated the logical error rate per round  $\epsilon_L = [1 - (1 - 2p_L)^{1/r}] / 2$  for various distances and  $r = 1000$  rounds using this data. The comparison of our results with those obtained by Google using MWPM is shown in Fig. 2(c). We can see that the improvements of our algorithm to MWPM reach 25% with  $d = 17$ . This reflects that in the original analysis of the error sources [25], 25% of error should be contributed by the decoding algorithm rather than the unknown sources. We also notice that the improvement increases with code distance because the logical error decays exponentially with distance  $d$ , with rate defined as  $\epsilon_L(d) \approx p_L/r \propto (1/\Lambda)^{(d+1)/2}$ . Here  $\Lambda$  is an important character of an algorithm and indicates how fast the logical error reduces with  $d$  increasing. Here using fitting, we obtained  $\Lambda_{\text{Planar}} \approx 8.28(4)$  which is significantly improved from  $\Lambda_{\text{MWPM}} \approx 8.11(5)$ . Note that, we used part of the dataset provided by Google, so we did not get exactly the same  $\Lambda$  as in [25].

Next, to evaluate the generality and robustness of our decoder under different experimental conditions, we conduct bit-flip repetition-code experiments using a quantum chip with 72 tunable coupled Transmon qubits with code distances ranging from 3 to 11, and up to 11 rounds of repeated syndrome measurements. In the experiments, we deliberately removed the reset gates and other leakage suppression methods to assess the decoder’s effectiveness under a model that deviates from the standard Pauli error model. The error model employed in decoding is derived solely through the correlation analysis method [26]. Like MWPM, we only consider the correlations to align with the error graph shown in Fig. 1(b) and fail to account for multi-body, long-range correlations that are caused purely by non-Pauli errors. The results are shown in Fig. 2(d). We can see that at all distances and rounds, PLANAR outperforms MWPM, and the improvement reaches a maximum value of 8.76% with  $d = 9$ , because PLANAR captures better the Pauli residual noise. At  $d = 11$ , the relative improvement decreases abnormally. This reflects that with the system becoming larger, the errors that deviate from simple independent Pauli errors become dominant. In the process of decoding using PLANAR, we also discovered a qubit in the chip with significantly poor performance. We refer to Supporting Materials for details.

*Discussion.*— We have established that with a correct error model, the repetition code with circuit-level noise can be

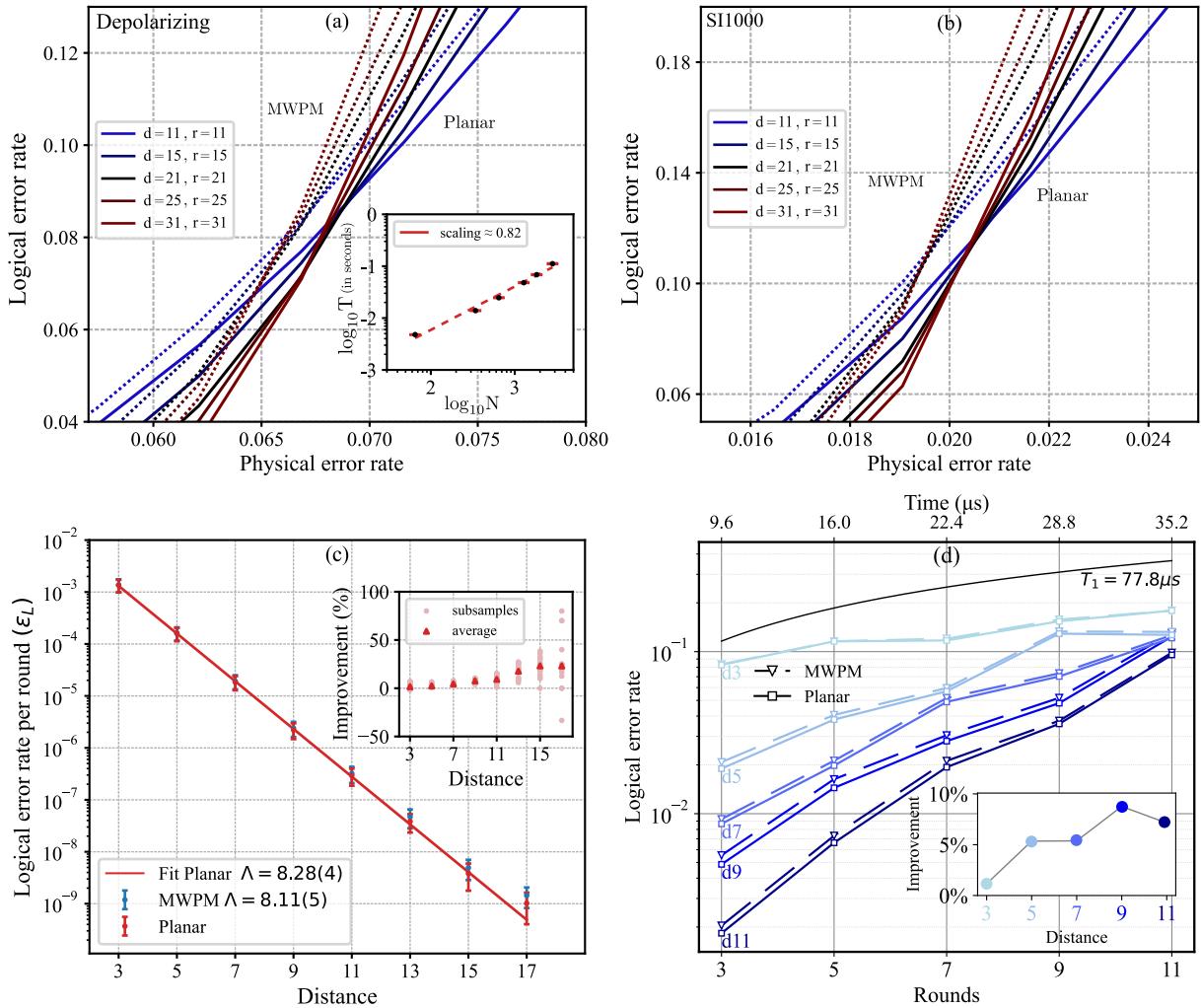


FIG. 2. Performance of our algorithm in decoding repetition codes under circuit-level noise using various syndrome data. (a) Synthetic syndrome data generated using the depolarizing noise model, which is also employed for decoding; (b) Synthetic syndrome data generated by the SI1000 noise model, which is also employed for decoding. In both (a) and (b), we compared the logical error rate obtained by PLANAR (solid lines) and MWPM (dotted lines). The inset of (a) displays the time complexity of PLANAR. Each data point represents an average over  $10^6$  syndrome shots. Notice that the threshold obtained by PLANAR is exact for both depolarizing noise and SI1000 noise. (c) Syndrome data and error model are derived from Google's quantum memory experiments [25]. We compute the logical error rate per round as a function of code distance. The red solid line represents the linear fitting using data ranging from  $d = 3$  to  $11$ , as defined in Ref. [25]. We randomly selected  $10^6$  samples from Google's datasets of each base  $X(Z)$  for  $d = [3, 5, 7, 9, 11, 13]$  and  $2 \times 10^6$  samples of each base  $X(Z)$  for  $d = [15, 17]$ . The inset displays the improvement of PLANAR over MWPM. Each data point is averaged over  $n_{\text{sub}}$  subsamples with X-basis and Z-basis. The light red points represent different subsamples. (d) Experimental data are generated on a 72-tunable-coupled-qubits chip, and the error model is derived through the correlation analysis method [26]. The black solid line represents the guideline for the physical error rate under  $T_1 = 77.8\mu\text{s}$ . The increasingly deeper shades of blue lines represent distances gradually increasing from 3 to 11. For each distance, solid lines and dashed lines indicate PLANAR and MWPM respectively. The inset displays the average performance improvement of the PLANAR method over MWPM. Each data point is averaged over 500,000 samples.

optimally decoded using an exact maximum likelihood decoder. We proposed a polynomial-time decoder – PLANAR decoder, utilizing the exact solution of planar spin glasses. Using extensive numerical and physical experiments, we show that under synthetic data with a correct error model, PLANAR decoder provides the exact threshold for the error model; under experimental data of Google's quantum memory experiments and more noisy data from a 72-qubit experiment without the reset gates, PLANAR decoder offers lower logical er-

ror rates and larger suppression factor  $\Lambda$  than the canonical minimum-weight perfect matching based method, it also provides useful information for analyzing hardware errors. The PLANAR decoder is also very fast, with computational complexity  $O(N^{0.82})$ . Where  $N$  is the number of spins in the spin glass model corresponding to the decoding problem, and is linear in number of detectors. Moreover, our numerical results with different kinds of data, circuits, hardware devices, and errors demonstrated the generality and robustness of our

decoder for the practical decoding problem of repetition code experiments.

We note that the PLANAR method is not limited to decoding repetition codes with circuit-level noise. If the decoding graph of a quantum code exhibits a planar structure under particular error models, the PLANAR method can be effectively applied. This encompasses both surface and rotated surface codes with independent code capacity noise. Consequently, our approach provides an efficient method for understanding exact thresholds and properties for codes with a planar structure. Moreover, the PLANAR method can be extended to codes beyond the planar structure. For instance, it can be utilized for decoding codes with the decoding graph possessing a finite genus  $g$ , with a computational complexity  $O(4^g)$ , yet remains polynomial with the number of spins. We leave this for future study.

## ACKNOWLEDGMENTS

This work is supported by Project 12047503, 12325501, 12247104, 12104056, and 92365206 of the National Natural Science Foundation of China, and Innovation Program for Quantum Science and Technology (No. 2021ZD0301802). F.P. acknowledges the support from the National Research Foundation, Singapore, A\*STAR under its CQT Bridging Grant and grants NRF2020-NRF-ISF004-3528. A Python package that is used to complete the numerical experiments of this manuscript can be found in [47].

- 
- [1] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM review* **41**, 303 (1999).
- [2] C. Gidney and M. Ekerå, How to factor 2048 bit rsa integers in 8 hours using 20 million noisy qubits, *Quantum* **5**, 433 (2021).
- [3] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Simulated quantum computation of molecular energies, *Science* **309**, 1704 (2005).
- [4] W. van Dam, H. Liu, G. H. Low, A. Paetznick, A. Paz, M. Silva, A. Sundaram, K. Svore, and M. Troyer, End-to-end quantum simulation of a chemical system, *arXiv preprint arXiv:2409.05835* (2024).
- [5] J. Preskill, Quantum computing in the nisq era and beyond, *Quantum* **2**, 79 (2018).
- [6] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, *Physical Review A* **86**, 032324 (2012).
- [7] D. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, Surface code quantum computing by lattice surgery, *New Journal of Physics* **14**, 123011 (2012).
- [8] E. T. Campbell, A theory of single-shot error correction for adversarial noise, *Quantum Science and Technology* **4**, 025006 (2019).
- [9] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, The XZZX surface code, *Nature communications* **12**, 2172 (2021).
- [10] S. Bravyi, A. W. Cross, J. M. Gambetta, D. Maslov, P. Rall, and T. J. Yoder, High-threshold and low-overhead fault-tolerant quantum memory, *Nature* **627**, 778 (2024).
- [11] M. B. Hastings and J. Haah, Dynamically generated logical qubits, *Quantum* **5**, 564 (2021).
- [12] Q. Xu, J. P. Bonilla Ataides, C. A. Pattison, N. Raveendran, D. Bluvstein, J. Wurtz, B. Vasić, M. D. Lukin, L. Jiang, and H. Zhou, Constant-overhead fault-tolerant quantum computation with reconfigurable atom arrays, *Nature Physics* **20**, 1084–1090 (2024).
- [13] A. J. Landahl, J. T. Anderson, and P. R. Rice, Fault-tolerant quantum computing with color codes, *arXiv preprint arXiv:1108.5738* (2011).
- [14] A. G. Fowler, 2-d color code quantum computation, *Physical Review A* **83**, 042310 (2011).
- [15] H. Bombín, Topological subsystem codes, *Physical Review A* **81**, 032301 (2010).
- [16] D. Honciuc Menendez, A. Ray, and M. Vasmer, Implementing fault-tolerant non-clifford gates using the [[8, 3, 2]] color code, *Physical Review A* **109**, 062438 (2024).
- [17] B. W. Reichardt, D. Aasen, R. Chao, A. Chernoguzov, W. van Dam, J. P. Gaebler, D. Gresh, D. Lucchetti, M. Mills, S. A. Moses, *et al.*, Demonstration of quantum computation and error correction with a tesseract code, *arXiv preprint arXiv:2409.04628* (2024).
- [18] Y. Hong, E. Durso-Sabina, D. Hayes, and A. Lucas, Entangling four logical qubits beyond break-even in a nonlocal code, *Physical Review Letters* **133**, 180601 (2024).
- [19] M. Da Silva, C. Ryan-Anderson, J. Bello-Rivas, A. Chernoguzov, J. Dreiling, C. Foltz, J. Gaebler, T. Gatterman, D. Hayes, N. Hewitt, *et al.*, Demonstration of logical qubits and repeated error correction with better-than-physical error rates, *arXiv preprint arXiv:2404.02280* (2024).
- [20] K. Mayer, C. Ryan-Anderson, N. Brown, E. Durso-Sabina, C. H. Baldwin, D. Hayes, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, *et al.*, Benchmarking logical three-qubit quantum fourier transform encoded in the steane code on a trapped-ion quantum computer, *arXiv preprint arXiv:2404.08616* (2024).
- [21] C. Ryan-Anderson, J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, *et al.*, Realization of real-time fault-tolerant quantum error correction, *Physical Review X* **11**, 041058 (2021).
- [22] Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, *et al.*, Realization of an error-correcting surface code with superconducting qubits, *Physical Review Letters* **129**, 030501 (2022).
- [23] D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, *et al.*, Logical quantum processor based on reconfigurable atom arrays, *Nature* **626**, 58 (2024).
- [24] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, *Nature* **614**, 676 (2023).
- [25] G. Q. AI and Collaborators, Quantum error correction below the surface code threshold, *Nature* (2024).
- [26] Google Quantum AI, Exponential suppression of bit or phase errors with cyclic error correction, *Nature* **595**, 383 (2021).
- [27] C. Ryan-Anderson, J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, *et al.*, Realization of real-time fault-tolerant quantum error correction, *Physical Review X* **11**, 041058 (2021).
- [28] M. McEwen, D. Kafri, Z. Chen, J. Atalaya, K. Satzinger,

- C. Quintana, P. V. Klimov, D. Sank, C. Gidney, A. Fowler, *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, *Nature communications* **12**, 1761 (2021).
- [29] K. C. Miao, M. McEwen, J. Atalaya, D. Kafri, L. P. Pryadko, A. Bengtsson, A. Opremcak, K. J. Satzinger, Z. Chen, P. V. Klimov, *et al.*, Overcoming leakage in quantum error correction, *Nature Physics* **19**, 1780 (2023).
- [30] J. M. Günther, F. Tacchino, J. R. Wootton, I. Tavernelli, and P. K. Barkoutsos, Improving readout in quantum simulations with repetition codes, *Quantum Science and Technology* **7**, 015009 (2021).
- [31] X.-G. Li, J.-H. Wang, Y.-Y. Jiang, G.-M. Xue, X.-X. Cai, J. Zhou, M. Gong, Z.-F. Liu, S.-Y. Zheng, D.-K. Ma, *et al.*, Direct evidence for cosmic-ray-induced correlated errors in superconducting qubit array, *arXiv preprint arXiv:2402.04245* (2024).
- [32] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, Topological quantum memory, *Journal of Mathematical Physics* **43**, 4452 (2002).
- [33] O. Higgott and C. Gidney, Sparse blossom: correcting a million errors per core second with minimum-weight matching, *arXiv preprint arXiv:2303.15933* (2023).
- [34] Y. Wu and L. Zhong, Fusion blossom: Fast mwpm decoders for qec, in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Vol. 1 (IEEE, 2023) pp. 928–938.
- [35] A. G. Fowler, Optimal complexity correction of correlated errors in the surface code, *arXiv preprint arXiv:1310.0863* (2013).
- [36] O. Higgott, T. C. Bohdanowicz, A. Kubica, S. T. Flammia, and E. T. Campbell, Improved decoding of circuit noise and fragile boundaries of tailored surface codes, *Physical Review X* **13**, 031007 (2023).
- [37] M. McEwen, D. Bacon, and C. Gidney, Relaxing hardware requirements for surface code circuits using time-dynamics, *Quantum* **7**, 1172 (2023).
- [38] <https://zenodo.org/records/13273331>.
- [39] C. Gidney, Stim: A fast stabilizer circuit simulator, *Quantum* **5**, 497 (2021), 2103.02202 [quant-ph].
- [40] C. T. Chubb and S. T. Flammia, Statistical mechanical models for quantum codes with correlated noise, *Annales de l'Institut Henri Poincaré D* **8**, 269–321 (2021).
- [41] L. Onsager, Crystal statistics. i. a two-dimensional model with an order-disorder transition, *Phys. Rev.* **65**, 117 (1944).
- [42] M. Kac and J. C. Ward, A combinatorial solution of the two-dimensional ising model, *Physical Review* **88**, 1332 (1952).
- [43] J. K. Johnson, D. Oyen, M. Chertkov, and P. Netrapalli, Learning planar ising models, *The Journal of Machine Learning Research* **17**, 7539 (2016).
- [44] Y. Iba, The nishimori line and bayesian statistics, *Journal of Physics A: Mathematical and General* **32**, 3875 (1999).
- [45] A. Decelle, F. Krzakala, C. Moore, and L. Zdeborová, Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications, *Physical Review E* **84**, 066106 (2011).
- [46] J. Kelly, R. Barends, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, *et al.*, State preservation by repetitive error detection in a superconducting quantum circuit, *Nature* **519**, 66 (2015).
- [47] <https://github.com/CHY-i/planar>.
- [48] H. Bombin, R. S. Andrist, M. Ohzeki, H. G. Katzgraber, and M. A. Martin-Delgado, Strong resilience of topological codes to depolarization, *Physical Review X* **2**, 10.1103/PhysRevX.2.021004 (2012).
- [49] O. Higgott, Pymatching: A python package for decoding quantum codes with minimum-weight perfect matching, *ACM Transactions on Quantum Computing* **3**, 1 (2022).
- [50] P. Panteleev and G. Kalachev, Degenerate Quantum LDPC Codes With Good Finite Length Performance, *Quantum* **5**, 585 (2021).
- [51] J. Roffe, D. R. White, S. Burton, and E. Campbell, Decoding across the quantum low-density parity-check code landscape, *Physical Review Research* **2**, 10.1103/PhysRevResearch.2.043423 (2020).
- [52] G. Duclos-Cianci and D. Poulin, A renormalization group decoding algorithm for topological quantum codes, in *2010 IEEE Information Theory Workshop* (2010) pp. 1–5.
- [53] S. Bravyi, M. Suchara, and A. Vargo, Efficient algorithms for maximum likelihood decoding in the surface code, *Physical Review A* **90**, 032326 (2014).
- [54] H. Bombin and M. A. Martin-Delgado, Optimal resources for topological two-dimensional stabilizer codes: Comparative study, *Physical Review A* **76**, 10.1103/PhysRevA.76.012305 (2007).
- [55] D. Bacon, S. T. Flammia, A. W. Harrow, and J. Shi, Sparse quantum codes from quantum circuits, *IEEE Transactions on Information Theory* **63**, 2464 (2017), arXiv:1411.3334 [quant-ph].
- [56] X. Li, H. Xu, J. Wang, L.-Z. Tang, D.-W. Zhang, C. Yang, T. Su, C. Wang, Z. Mi, W. Sun, X. Liang, M. Chen, C. Li, Y. Zhang, K. Linghu, J. Han, W. Liu, Y. Feng, P. Liu, G. Xue, J. Zhang, Y. Jin, S.-L. Zhu, H. Yu, S. P. Zhao, and Q.-K. Xue, Mapping the topology-localization phase diagram with quasiperiodic disorder using a programmable superconducting simulator, *Phys. Rev. Res.* **6**, L042038 (2024).
- [57] S. Zhao, D. Tang, X. Xiao, R. Wang, Q. Sun, Z. Chen, X. Cai, Z. Li, H. Yu, and W.-H. Fang, Quantum Computation of Conical Intersections on a Programmable Superconducting Quantum Processor, *The Journal of Physical Chemistry Letters* **15**, 7244 (2024).
- [58] H. Xu, W. Liu, Z. Li, J. Han, J. Zhang, K. Linghu, Y. Li, M. Chen, Z. Yang, J. Wang, T. Ma, G. Xue, Y. Jin, and H. Yu, Realization of adiabatic and diabatic cz gates in superconducting qubits coupled with a tunable coupler, *Chinese Physics B* **30**, 044212 (2021).

## Supplementary information for “Exact Decoding of Repetition Code with Circuit Level Noise”

### Appendix A: Planar decoder For Code Capacity Noise

#### 1. Statistical Mapping

Under the Nishimori condition [40, 48]

$$J(\sigma) = \frac{1}{4} \sum_{\rho} \langle \sigma, \rho \rangle \cdot \log p(\rho), \quad \sigma, \rho \in \{I, x, y, z\}, \quad (\text{S1})$$

decoding of a QECC with  $n$  data qubits can be mapped to free energy problem of a random bonds Ising model (RBIM) [32]. Where the  $\langle , \rangle$  represents the scalar commutator which equals to 1 if two inputs are commute else -1. In such mapping, stabilizer generators will serve as spins while and data qubits will be (hyper)edges. The Hamiltonian will be formulated as

$$H_E(s) = - \sum_{\sigma} \sum_{i=1}^n \langle \sigma, E_i \rangle \cdot J(\sigma)_i \prod_{\{k | \langle \sigma, (g_k)_i \rangle = 1\}} s_k \quad (\text{S2})$$

where  $E$  is a given Pauli error, the subscript  $i$  represents the  $i$ -th qubit.  $(g_k)_i$  is the  $i$ -th component of  $k$ -th stabilizer. It can easily check the probability of an error  $E$  is:

$$\begin{aligned} P(E) &= e^{-H_E(s=1)} \\ &= \exp \left\{ \sum_{i=1}^n \sum_{\sigma} \langle \sigma, E_i \rangle \cdot J(\sigma)_i \right\} \\ &= \exp \left\{ \sum_{i=1}^n \sum_{\sigma} \langle \sigma, E_i \rangle \cdot \frac{1}{4} \sum_{\rho} \langle \sigma, \rho \rangle \cdot \log p(\rho) \right\} \\ &= \exp \left\{ \sum_{i=1}^n \frac{1}{4} \sum_{\sigma} \sum_{\rho} \langle \sigma, E_i \cdot \rho \rangle \cdot \log p(\rho) \right\} \\ &= \exp \left\{ \sum_{i=1}^n \sum_{\rho} \delta_{\rho, E_i} \cdot \log p(\rho) \right\} \\ &= \prod_{i=1}^n p(E_i) \end{aligned}$$

In the third line, we used the definition Eq. S1. And we used the two properties  $\langle A, B \rangle \cdot \langle A, C \rangle = \langle A, BC \rangle$  and  $\frac{1}{4} \sum_{\sigma} \langle \sigma, \rho \rangle = \delta_{\rho, I}$  in the next two equations. Additionally, the non-trivial configurations of spins means there are stabilizers acting on the error  $E$  and causing a new error  $E'$  which belongs to the same equivalence class as  $E$ . Under this mapping, summing all elements which equivalence to  $E$  is equal to sum all spin configurations. Therefore the coset probability of a quantum code equals to the partition function of the corresponding Ising model.

$$P(L) = \mathcal{Z}_L$$

Calculation of the partition function is a #P hard problem in general. However, there exist polynomial algorithms for special case such as 2D Ising model [41]. Another difficulty is

the  $Y$  errors will cause multi-interaction in Hamiltonian. This problem makes the most of quantum codes are not planar, for which the planar Ising model solution is not work. In order to solve this problem, it is straightforward to decode X-type and Z-type errors independently. Here we adopt a XZ-separated version of MLD. We consider the X-type and Z-type errors independently i.e. the Y error is the combination of X and Z. The whole code is divided into two parts. A n-qubits Pauli error will also cause two types of syndromes ( $s_x, s_z$ ). It needs to sum all Z-stabilizers for decoding  $s_x$  and sum all X-stabilizers for decoding  $s_z$ .

$$\begin{aligned} L^x &= \arg \max_{L_x} \sum_{S_x \in \mathcal{S}} P(L_x, S_x | s_z), \\ L^z &= \arg \max_{L_z} \sum_{S_z \in \mathcal{S}} P(L_z, S_z | s_x), \\ L &= L^x \cdot L^z. \end{aligned} \quad (\text{S3})$$

In fact the Eq. S3 is an approximation of Eq. S1 for general error model. And its accuracy is depend on the correlation between X and Z. Empirically, most of minimum weight algorithms are finding the minimum weight error of X and Z separately, like Minimum Weight Perfect Matching (MWPM) [33, 49] and belief propagation augmented by ordered statistics decoding (BPOSD) algorithm [50, 51]. Therefore, XZ-separation version of MLD will be better than MWPM theoretically. An approximately scheme that has been used to compute coset probabilities, utilize the XZ-separated MLD is renormalization group (RG) [52]. In this work we proposed that the exact coset probabilities can be calculated in polynomial time for some special codes. At first this method considers the XZ-separated MLD. Then it converts the code into a planar Ising model and calculates the partition function using the strict algorithm of planar Ising model.

There exists discussion of a special case code can be solved rigorously, the surface code [53]. However our work gives more general results. If the X or Z subgraph of a correction code corresponds to a planar graph, with XZ-uncorrelated noise, their MLD procedures can be transformed into partition function calculations of planar RBIM. And it can be solved in polynomial time in their system sizes. It is worth noting that it is only necessary that either the X or Z subgraph be planar, since X and Z are dual to each other, and when one is planar, the other is naturally planar. Surface codes, for example, have four-body interactions that vanish under uncorrelated noise as Fig. S1 show. Each edge connects to two spins and the graph is obviously planar. And there exist polynomial time algorithms for the calculation of exact free energy of the planar Ising model [42]. Therefore, it can decode Surface code with uncorrelated noise exactly in polynomial time. Further more it can also use the uncorrelated approximation to decoding more complex error model such as depolarizing model.

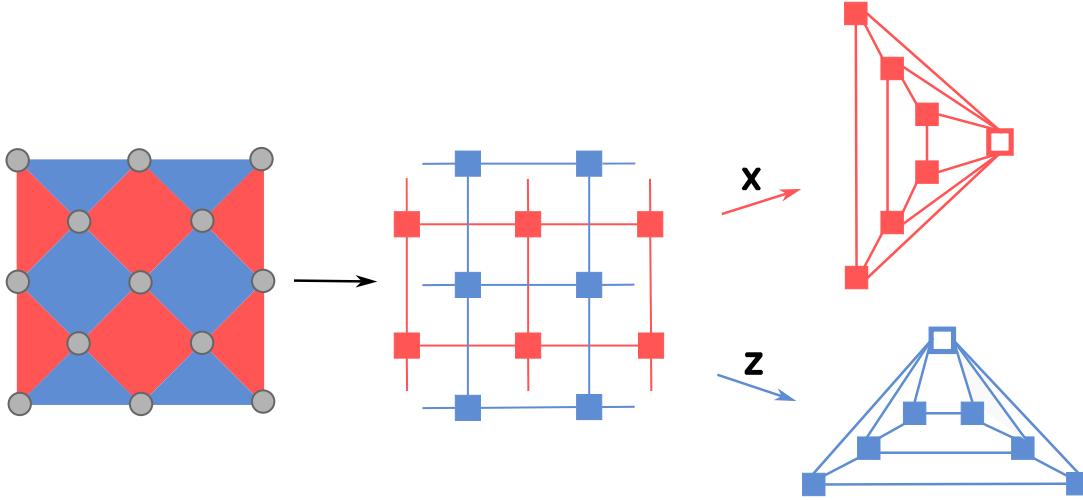


FIG. S1. Illustration of separating the distance 3 surface code. In the left graph, the red(blue) blocks are X(Z)-stabilizers and the gray nodes denote qubits which are (hyper)edges in RBIM. The red(blue)-subgraph at middle represents X(Z)-Ising model. The squares denote X(Z)-spins respectively. In the right graph, we add an additional spin for X(Z)-Ising model to construct a planar Ising model without external fields.

## 2. Kac-Ward solution

We adopt the Kac-Ward determinant formula to compute the partition function of planar Ising model [42, 43]. The main conclusion of this formula is that the partition function of a zero-field Ising model defined on a planar graph  $G(E, V)$  is proportional to the determinant of a matrix.

$$Z = 2^n \left( \prod_{(i,j) \in E} \cosh \theta_{ij} \right) \det(I - W)^{\frac{1}{2}}, \quad (\text{S4})$$

where the  $\theta$  satisfies the definition  $\tanh \theta_{ij} \triangleq J_{ij}$ , and the  $W = AD$ ,  $D$  is a  $2|E| \times 2|E|$  diagonal matrix  $D_{ij,ij} = J_{ij}$ . The definition of  $A$  is under follow:

$$A_{ij,kl} = \begin{cases} \exp\left(\frac{i}{2}\phi_{ijl}\right), & j = k \text{ and } i \neq l \\ 0, & \text{otherwise} \end{cases}. \quad (\text{S5})$$

The  $\phi_{ijl}$  is clock-wise angle difference between directed edges  $(ij)$  and  $(jl)$ . The determinant can be computed by Gaussian elimination with the  $O(n^3)$  complexity. And the complexity of this calculation will reduce to  $O(n^{\frac{3}{2}})$ , if the sparse structure of the matrix is used.

## 3. Numerical Results

The first code family we considered is the Surface code. As previous discussion that the Surface code is equivalent to a Ising model with (three)four bodies interactions which is due to the existence of Y error on a single qubit. Under the uncorrelated noise,

$$\begin{aligned} p(I) &= (1 - p_x)(1 - p_z), \\ p(x) &= p_x(1 - p_z), \\ p(z) &= (1 - p_x)p_z, \\ p(y) &= p_xp_z \end{aligned} \quad (\text{S6})$$

The correlations between X and Z types of stabilizers disappear i.e.  $J(Y) = 0$  in Eq. S2. The Hamiltonian can be reduce to following term

$$\begin{aligned} H_E(s) = & - \sum_{i=1}^n (-1)^{\delta_{E_i,z}} J(x)_i \prod_{\{k|(g_k)_i=z\}} s_k \\ & - \sum_{i=1}^n (-1)^{\delta_{E_i,x}} J(z)_i \prod_{\{k|(g_k)_i=x\}} s_k \end{aligned} \quad (\text{S7})$$

From above equation, one can find the X and Z subsystem are independent. The partition function of the whole system is product of partition functions of two subsystem  $\mathcal{Z} = \mathcal{Z}_x \cdot \mathcal{Z}_z$ . And the  $L_x$  and  $L_z$  are independent consequently  $P(L_x, L_z) = P(L_x) \cdot P(L_z)$ . At this point, The graph of original ising model can be separated to X and Z subgraphs. Each subgraph is a  $d \times (d - 1)$  lattice. And spins at top and bottom(left and right) interact with the external fields. These external fields arise because qubits on the surface code boundary are only subject to a single X or Z check. This is shown as Fig. S1.(middle).

The presence of external fields causes the strict solution of the Ising model fail. However this problem can be handled by introducing an additional spin. Each boundary spins interact with this new spin and the strength of coupling is equal to  $J(\sigma)_\partial$ , the  $\partial$  denotes a boundary qubit of surface code as Fig. S1. (right). Obviously the new graph is also planar. The partition function of the new Ising model  $Z_a = 2Z$ . This is because the introduction of the new spin gives the system  $z_2$  symmetry. From the perspective of group theory, this additional spin represents a new stabilizer that can be obtained as a product of the original stabilizers. Therefore, it is still an element of the original stabilizer group. According to the rearrangement theorem, when summing over all generators, considering this element is equivalent to summing over all group elements twice.

Error correction will be achieved by calculating four times the partition functions. The syndrome and logical inputs are

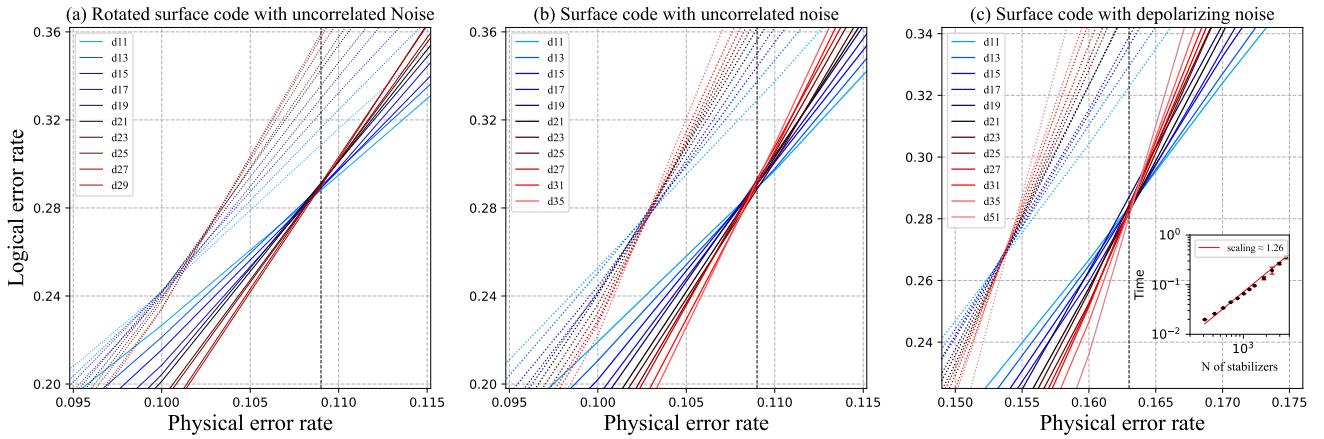


FIG. S2. Simulation on (Rotated) Surface codes with code capacity noise. We compared logical error rate calculated by Planar (solid lines) and MWPM (dashed lines). Each data point has  $10^5$  shots.

changed by changing  $E$  in Eq. S7. And they will not change the structure of graph. This error correction process is strictly MLD for uncorrelated noise, because we calculate the free energy exactly. At the same time it is a good approximation for correlated noise. We test *Planar* on uncorrelated error model. It will give the exact result of MLD at uncorrelated case. The threshold is about  $p_{uc} \approx 0.109$ , which can be read out from the Fig. S2(middle) approximately. This result is consistent with the analysis from a statistical physics perspective [40]. Numerical decoding are also simulated on depolarizing error model. The prior probability  $p_x = p_z = 2\epsilon/3$ , where  $\epsilon$  is physical error rate of depolarizing model. This modification thinks of the Y error as being composed of independent X and Z simply and is always used in MWPM [49] and BPOSD [51]. One can find the point where the curves intersect is between 0.162 and 0.164. This result is consistent with the theoretical value  $p = \frac{3}{2}p_{uc} \approx 0.164$  [48]. And the time scaling of our algorithm is about 1.26.

Another code family, we test the *Planar* is the rotated surface code (RSC) [54]. It is just cutting and rotation of the Surface code lattice and has almost the same threshold as the Surface code. However it has higher code rate, and used in a wide variety of experiments [24]. Same as the analysis of Surface code. The X(Z)-subgraph of RSC is also planar. And all open edges need to connect to a new spin. The difference is there exist boundary spins interacting with two external fields. A simple way to deal with these two dangling edges is to merge them together. And the weight of new edge is sum of weights of two original edges. And the numerical results are shown in Fig. S2.

## Appendix B: Circuit Level Noise and Detector Error Model

In order to better configure the parameters of the decoder, we need to closely align the error model with real QEC experiments. This requires us to consider in detail all possible error locations and error types that can occur when compiling a quantum circuit for a QEC code. We assume that the

error locations accompanies all effective quantum operations (including idling), such as quantum gates and measurements. Due to the nature of stabilizer codes, we consider all possible errors to be Pauli errors. These errors and their corresponding probabilities form an error model. The most commonly used and well-known model is depolarizing error model, assumes that Pauli  $X$ ,  $Y$ , or  $Z$  errors occur at any possible error location with the same probability. Another model used in this paper is the newly developed SI1000 [37], which is an empirical model derived from the statistical behavior of superconducting quantum computing equipment. Both models use a single parameter  $p$  to configure the Pauli errors across all possible locations and types. The difference is that SI1000 assigns different weights to error rates at different locations to better simulate the performance of superconducting devices. The specific configuration of the error model used in this paper is detailed in [37] Table D.2.

After fixing the compiled quantum circuit, and considering the circuit-level noise error model, certain errors are equivalent to each other (also known as gauge equivalent [55]) which means they would produce the same syndrome, i.e., resulting in identical detectors and having the same effect on the encoded logical qubit. Therefore, we can bundle these equivalent errors together and group them, with each group referred to as an "error mechanism." Similar to the circuit-level noise error model, each error mechanism can also be assigned a probability value. This probability represents the likelihood of any odd number of errors occurring within the group corresponding to that error mechanism. This bundled model, composed of error mechanisms and detectors, is referred to as the detector error model (DEM). As shown in Fig. S3, each edge represents an error mechanism with a weight equal to its probability value. The nodes it connects are the corresponding detectors, each of which can be in one of two states: flipped or not flipped.

In the experimental process, the correspondence between error mechanisms and detectors remains the same as in the simulation results once the circuit is finalized. However, the probability value associated with each error mechanism is no

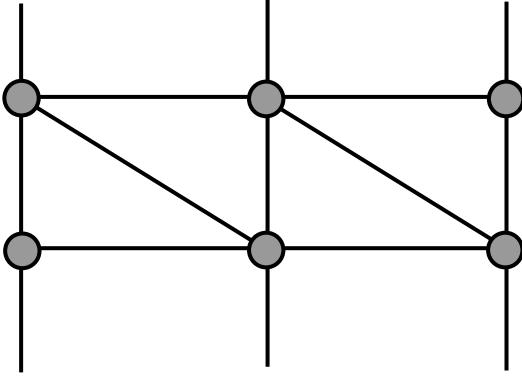


FIG. S3. DEM graph, each gray node represents a detector, while each edge corresponds to an error mechanism

longer derived from the combination of multiple corresponding Pauli errors. Instead, it is determined through correlation analysis [26], which directly uses multiple experimental samples to establish the relationships between detectors. This correlation is used in place of probability values to better approximate the error probability distribution in the actual circuit, thereby achieving more optimal decoding performance.

### Appendix C: Subsample Strategy

In this section, we provide a detailed explanation of the subsampling strategy. This approach involves deriving samples from a lower-distance code by extracting them from a higher-distance code. Figure S4 illustrates an example where  $d=3$  samples are obtained from a  $d=5$  code. For a  $d=5$  repetition code, there are 9 initial measurement results (default in the Z-basis), corresponding to 9 qubits, of which 5 are data qubits and 4 are ancilla qubits. During each round of syndrome measurement, the 4 ancilla qubits are measured to yield syndrome results. In the final round, all 5 data qubits are measured. From this setup, we can derive 3 subsamples corresponding to a  $d=3$  repetition code. Subsampling offers several practical advantages. First, it reduces the number of trials needed to acquire experimental data. Without subsampling, experiments would have to be conducted separately for each distance code. However, by employing subsampling, it suffices to perform experiments on the code with the highest distance and then extract data for lower-distance codes from the same dataset. Additionally, subsampling ensures that the experimental data for different distances are obtained under identical conditions, using the same equipment. This consistency enhances the reliability and comparability of the data across various distances.

Unlike MWPM, decoding subsampled data using *Planar* requires additional processing of the DEM of the higher-distance code. In Figure S5, we illustrate the Sub-DEM for the  $d=3$  subsample 1 repetition code derived from the DEM of a  $d=5$  repetition code. Compared to Figure S3, the red region in Figure S5 contains several additional beveled edges. For MWPM, these beveled edges can be directly handled and do not introduce any additional complexity. However, when

using *Planar*, it is necessary to combine the probabilities associated with these beveled edges with the probabilities of the transverse edges connected to the same node. The combined probability is calculated as  $p = p_1(1 - p_2) + p_2(1 - p_1)$ . After this calculation, the beveled edges are removed to streamline the decoding process.

## Appendix D: Experiment Details

### 1. Fabrication

The superconducting quantum processor used in this experiment consists of 72 qubits and 121 transmon-like couplers. As illustrated in Fig. S6, qubits are represented as circles, while the lines connecting the circles indicate couplers. The processor is fabricated using a flip-chip architecture [56], with the carrier chip measuring  $38 \text{ mm} \times 45 \text{ mm}$  and the top chip measuring  $18 \text{ mm} \times 20 \text{ mm}$ . The circuits on the top chip are fabricated on a c-plane sapphire wafer, while those on the carrier chip are fabricated on a silicon wafer.

The fabrication process begins with annealing the sapphire wafer in a chamber annealing furnace at  $1200^\circ\text{C}$  for 3 hours in an atmospheric environment. The silicon wafer is pre-treated for 2 minutes using an HF solution. A 200-nm-thick tantalum film is then deposited on both wafers via a DC magnetron sputtering system. The base structures, including transmission lines and flux lines on the carrier chip and resonators and qubit capacitors on the top chip, are patterned using photolithography with a laser direct-writing lithography (DWL) system. Following development, reactive ion etching (RIE) with  $\text{CF}_4$  gas is performed to remove unwanted metallic regions, finalizing the base structures for subsequent processing steps.

For the top chip, after the base circuits are defined, the chip undergoes a piranha solution bath ( $2:1 \text{ H}_2\text{SO}_4:\text{H}_2\text{O}_2$ ) for 20 minutes, followed by an ultrasonic bath and a 10-minute rinse in deionized (DI) water. This piranha treatment is effective in improving the surface quality of the tantalum films. Next, Dolan-bridge  $\text{Al}/\text{AlO}_x/\text{Al}$  Josephson junctions are patterned using a double-angle shadow evaporation process in a 4-chamber e-beam evaporation system. After a lift-off process, the top chip is prepared for packaging.

For the carrier chip, aluminum air-bridges are fabricated on top of the base circuits. The bridge-shaped profile is created using a reflow method after the S1813 photoresist is patterned and developed with MF319 developer. Subsequently, a 500-nm-thick aluminum layer is deposited by evaporation, followed by lift-off steps. Indium bump structures are then patterned using a 2-layer resist photolithography process, and a 20- $\mu\text{m}$ -thick indium film is deposited via thermal evaporation. The lift-off of indium is achieved by immersing the chip in an  $80^\circ\text{C}$  NMP bath for over 8 hours.

Once both chips are ready, flip-chip bonding is performed using an SET Accura system. Accurate alignment of the two chips is ensured by a coaxial optical microscope, and a pressure of 6 kg is applied to achieve a uniform gap of  $20 \mu\text{m}$  between the chips. This alignment and gap uniformity are

d=5 dataset		d=3 subsample 1		subsample 2		subsample 3	
Initial	000000000	Initial	00000	Initial	00000	Initial	00000
1	0 0 0 0	1	0 0	1	0 0	1	0 0
2	1 0 0 1	2	0 1	2	0 0	2	1 0
3	0 1 0 0	3	0 0	3	1 0	3	0 1
4	1 0 0 0	4	0 0	4	0 0	4	1 0
5	0 0 0 1	5	0 1	5	0 0	5	0 0
Final	0 1 0 1 0	Final	0 1 0	Final	1 0 1	Final	0 1 0
		1					
		2					
		3					

FIG. S4. Example of deriving subsamples of d=3 code from d=5 dataset.

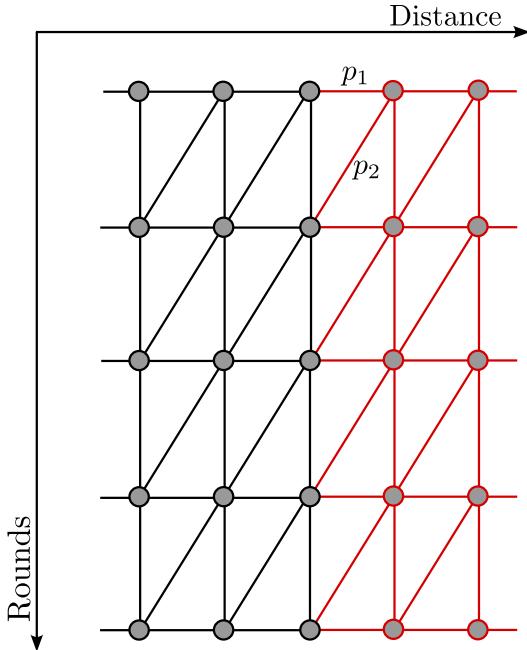


FIG. S5. Sub-DEM of d=3 in DEM of d=5 repetition code. The red nodes and lines corresponding to the DEM of subsample 1 in Fig.S4.

subsequently verified using a scanning electron microscope (SEM).

## 2. Measurement setup

Our experimental setup shown in Fig. S7 is identical to that described in Ref. [57]. The room-temperature electronics include systems for qubit readout, fast Z control, XY control, and Josephson parametric amplifier (JPA) control. For microwave generation and mid-circuit measurements, we utilize RF Direct Digital Frequency Synthesis (DDS) with a measurement repeat period of  $3.2\ \mu\text{s}$ .

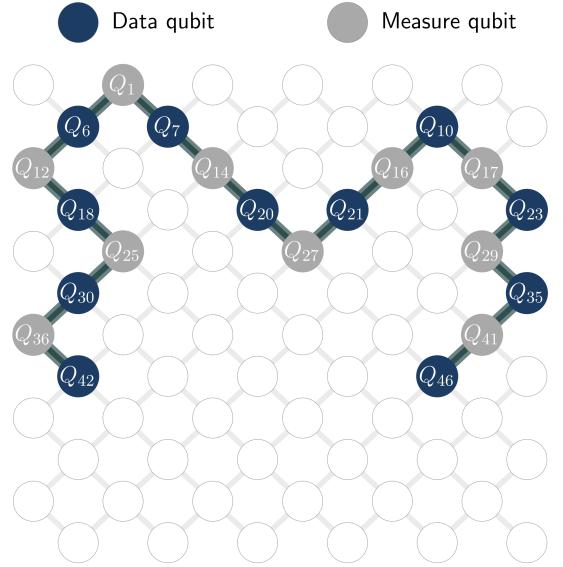


FIG. S6. Layout of active qubits in the 72-tunable-coupled-qubit chip. gray circles represent measurement qubits, while dark blue circles indicate data qubits used in the distance-11 repetition code. In the experiment, codes of different distances utilize overlapping sets of qubits. The labels inside the circles denote the unique identifier of the qubits.

For fast Z control, the equipment is equipped with hardware-implemented 6th-order infinite impulse response (IIR) filters on each channel, enabling real-time compensation for Z distortions. To further mitigate the effects of noise and harmonic distortions, a bandpass filter and an infrared (IR) filter are incorporated into the circuit design.

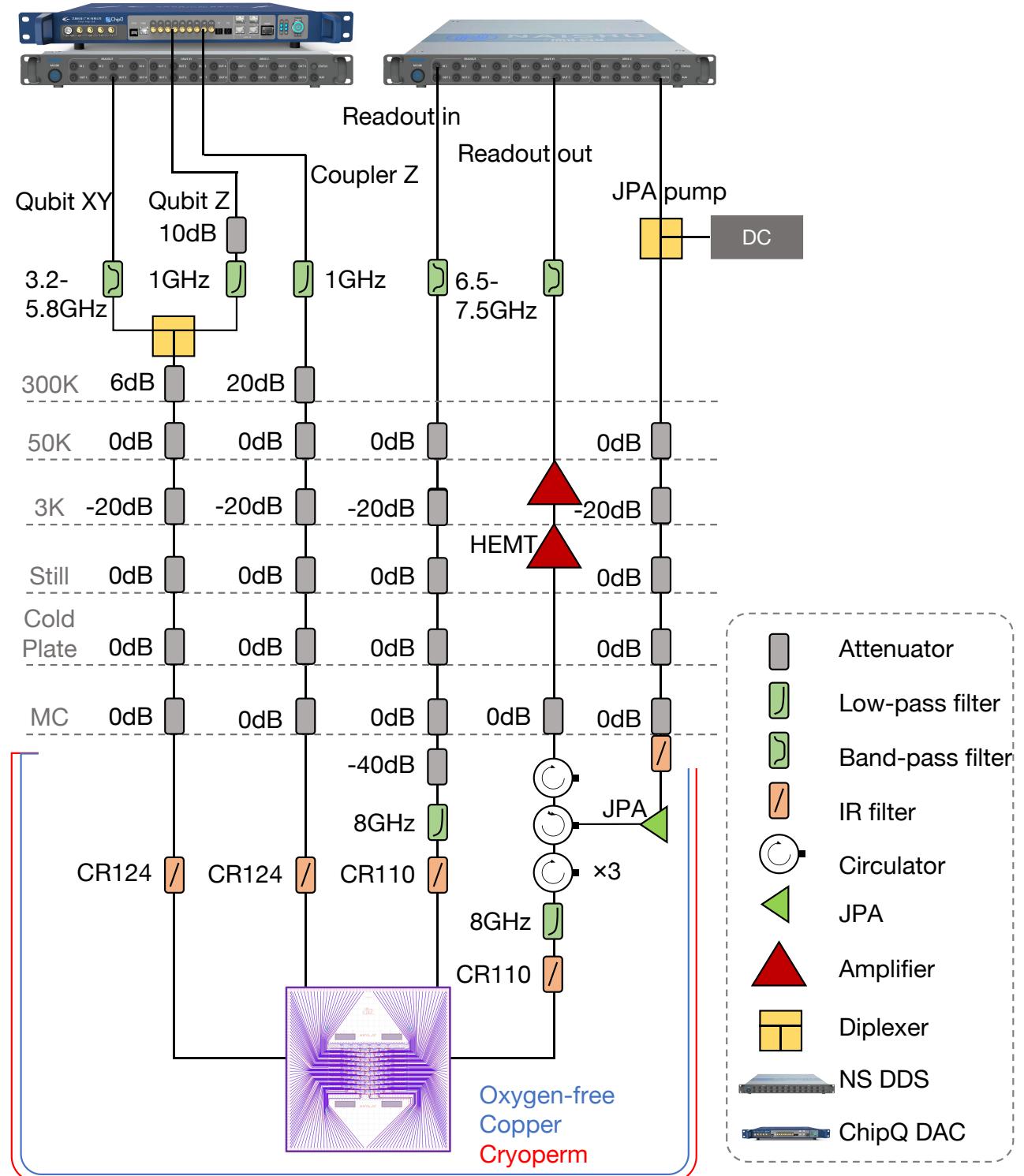
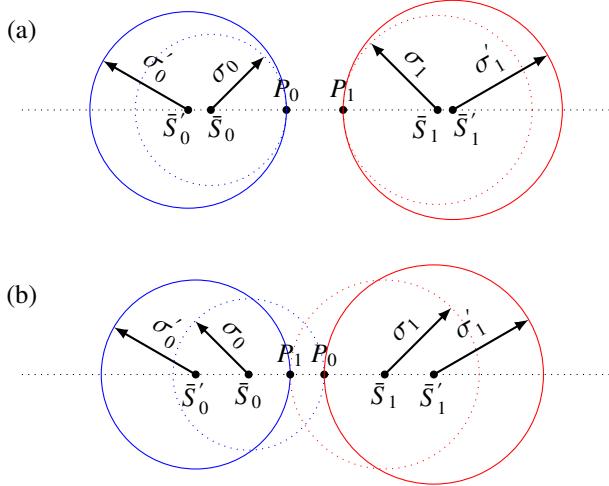


FIG. S7. A schematic of the measurement system, including cryogenic and room temperature wiring, various microwave devices and control system, amplifying circuit, measurement electronics, and electromagnetic shielding.

### 3. Optimization of readout

#### a. Method to state classification

Experimentally, the IQ signal of a qubit's readout resonator can be used to represent the qubit's state. To classify the states  $|0\rangle$  and  $|1\rangle$  of a qubit, the qubit should be initialized in  $|0\rangle$  and  $|1\rangle$  using the  $I$  gate and  $X$  gate, respectively. The IQ signals for these states are then measured multiple times and labeled as  $S_0$  and  $S_1$ . The data points from  $S_0$  and  $S_1$  can be plotted on the complex plane, where a boundary that optimally separates the two sets is determined. This boundary delineates the regions on the complex plane corresponding to the distributions of the two states.



**FIG. S9. State classification circles of qubits in our experiment.**  $\bar{S}_i$  and  $\sigma_i$  are expectation and standard deviation of  $S_i$  ( $i=0,1$ ).  $\bar{S}'_i$  and  $\sigma'_i$  are center and radius of state  $|i\rangle$  classification circle of qubit. Circle  $\bar{S}'_i$  is shown in **a** for  $|\bar{S}_0 - \bar{S}_1| \geq \sigma_0 + \sigma_1$ , while in **b** for  $|\bar{S}_0 - \bar{S}_1| < \sigma_0 + \sigma_1$ .

To improve readout fidelity and sampling efficiency during bitstring measurement, we classify the  $|0\rangle$  and  $|1\rangle$  states of a qubit using two regions defined by circles on the complex plane. The method for determining the parameters of these circles (shown as solid circles in Fig. S9) is as follows. Let  $\bar{S}_i$  and  $\sigma_i$  denote the expectation and standard deviation of  $S_i$  ( $i = 0, 1$ ), where  $\bar{S}_i$  is a complex value. As illustrated in Fig. S9, the dotted circle centered at  $\bar{S}_i$  has a radius  $\sigma_i$ , while the solid circle  $\bar{S}'_i$ , representing the classification boundary for state  $|i\rangle$ , has a larger radius  $\sigma'_i$  ( $> \sigma_i$ ). The experimental IQ signals and their respective expectations  $\bar{S}_i$  are shown in Fig. S8.

The point  $P_i$  is defined as the intersection of the dotted circle  $\bar{S}_i$  with the line segment  $\bar{S}_0\bar{S}_1$ . When the condition  $|\bar{S}_0 - \bar{S}_1| \geq \sigma_0 + \sigma_1$  is satisfied, the solid classification circle  $\bar{S}'_i$  is internally tangent to the dotted circle  $\bar{S}_i$  at  $P_i$ . Under this condition, the centers and radii of the classification circles,  $\bar{S}'_0$  and  $\bar{S}'_1$ , can be written as:

$$\bar{S}'_0 = \bar{S}_0 + (\sigma'_0 - \sigma_0) \frac{\bar{S}_1 - \bar{S}_0}{|\bar{S}_1 - \bar{S}_0|}, \quad (\text{S1a})$$

$$\bar{S}'_1 = \bar{S}_1 + (\sigma'_1 - \sigma_1) \frac{\bar{S}_0 - \bar{S}_1}{|\bar{S}_0 - \bar{S}_1|}. \quad (\text{S1b})$$

Otherwise, circle  $\bar{S}'_i$  is externally tangent with circle  $\bar{S}_j$  at  $P_j$  ( $i \neq j$ ), and  $\bar{S}'_0$  and  $\bar{S}'_1$  write

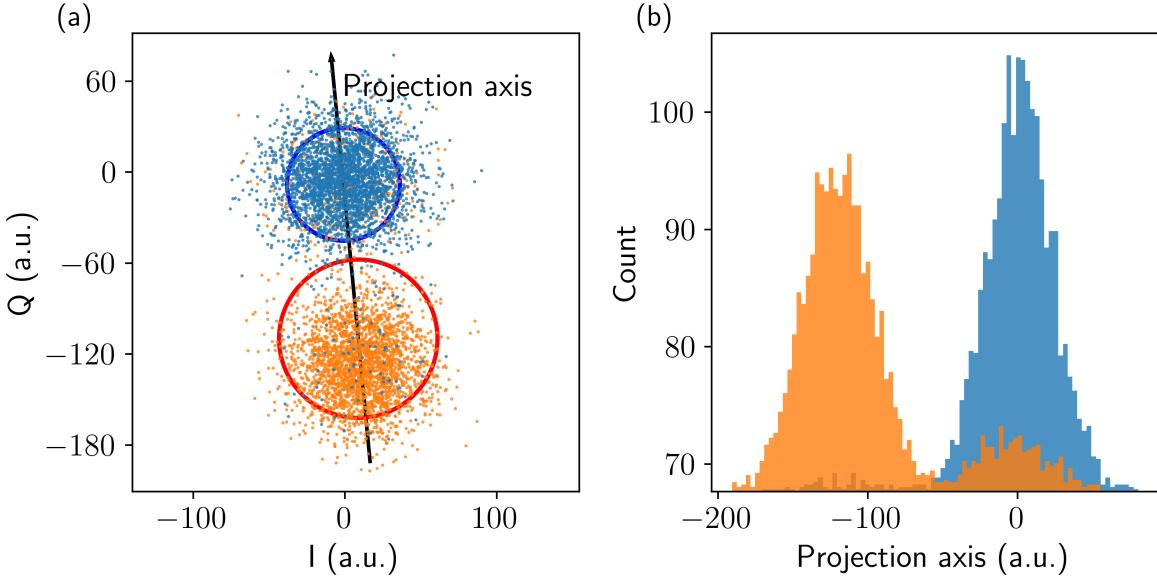
$$\bar{S}'_0 = \bar{S}_1 + (\sigma_1 + \sigma'_0) \frac{\bar{S}_1 - \bar{S}_0}{|\bar{S}_1 - \bar{S}_0|}, \quad (\text{S2a})$$

$$\bar{S}'_1 = \bar{S}_0 + (\sigma_0 + \sigma'_1) \frac{\bar{S}_0 - \bar{S}_1}{|\bar{S}_0 - \bar{S}_1|}. \quad (\text{S2b})$$

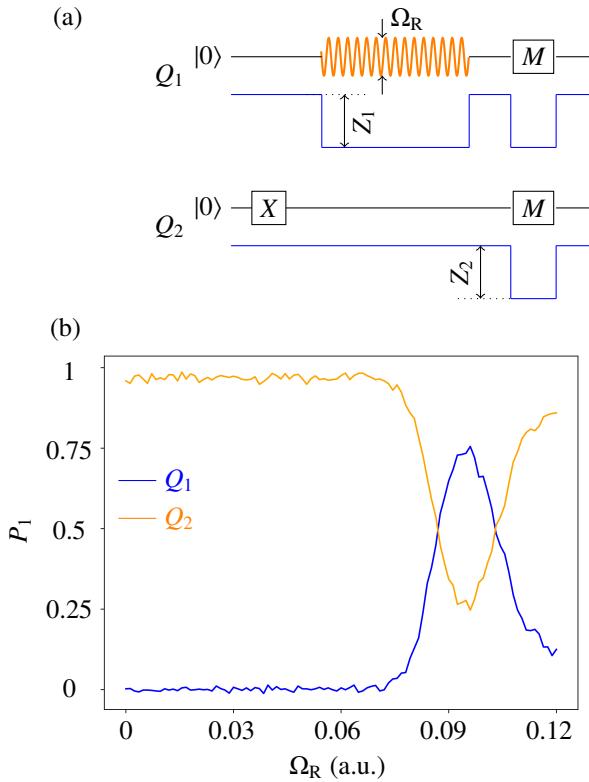
During a bitstring measurement, if the IQ signal of a qubit falls within the circle  $\bar{S}'_0$  or  $\bar{S}'_1$ , the qubit is classified as being in state  $|0\rangle$  or  $|1\rangle$ , respectively. Otherwise, the measurement is discarded. As a result of discarding some experimental data, the total number of samples is reduced. To increase the number of usable experimental samples, we set  $\sigma'_i$  significantly larger than  $\sigma_i$ .

#### b. Readout effects on qubits

It is well known that an AC-Stark shift occurs in qubits when a readout microwave is applied. In this scenario, the frequency of a qubit may shift closer to the frequencies of other qubits or two-level systems (TLS), potentially introducing unwanted interactions during qubit readout. To mitigate these interactions and improve readout fidelity, the amplitude of the readout microwave can be adjusted to reduce the AC-Stark effect, or a Z bias can be applied to the qubits to tune their frequencies apart from each other during measurements.



**FIG. S8. IQ signal of experiment.** The blue and yellow points are experimental IQ signal of state  $|0\rangle$  and  $|1\rangle$ , respectively. The dotted blue and red circles are experimentally corresponding to the circles  $\bar{S}_0$  and  $\bar{S}_1$ , respectively. The upper histograms represent the statistical distributions of the two states along the direction of  $\bar{S}_0$  and  $\bar{S}_1$ .



**FIG. S10. Measurement of Readout effects on qubits.** As shown in **a**, resonate drive (yellow pulse) with amplitude  $\Omega_R$  is applied on the readout resonator of  $Q_1$  before measurement. During driving and measurement, Z bias  $Z_1$  and  $Z_2$  are applied on  $Q_1$  and  $Q_2$ , respectively. Then the  $P_1$  of the two qubits are measured and shown in **b**.

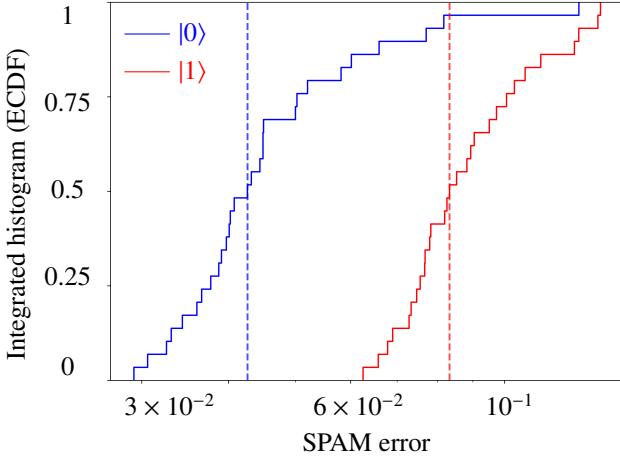
As shown in Fig. S10a, we utilize a pulse sequence to determine the amplitude ( $R_{\text{amp}}$ ) of the readout microwave and the Z biases ( $Z_1$ ,  $Z_2$ ) of the qubits. This pulse sequence can also be applied to specific qubits, such as  $Q_2$  in Fig. S10a, to identify the Z bias settings that yield optimal readout performance.

Since multiple readouts of ancilla qubits are involved in our experimental circuit, it is necessary to examine the effects of these readouts on the data qubits. To this end, we resonantly drive the readout resonator of  $Q_1$  with varying  $R_{\text{amp}}$  and measure the probabilities ( $P_1$ ) of  $Q_1$  and  $Q_2$  remaining in their respective states. The results, shown in Fig. S10b, indicate that tuning the readout amplitude of  $Q_1$  influences the state exchange between  $Q_1$  and  $Q_2$ . The appropriate  $R_{\text{amp}}$  for subsequent experiments should lie within a range where  $P_1$  of  $Q_2$  remains near zero, ensuring reliable qubit state readouts.

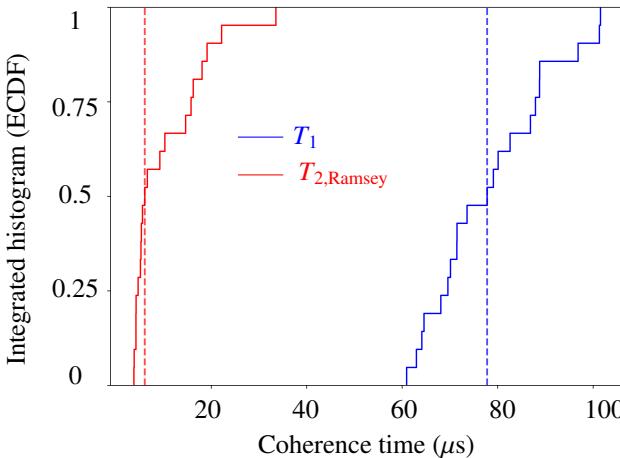
After completing these calibrations, we recheck the readout effects when only one qubit is excited prior to measurement. The IQ signal distributions of all qubits on the complex plane are then measured and analyzed to characterize these effects.

### c. Readout fidelity and decoherence time of qubits

Due to thermalization and the long resonator ringdown times ( $1/\kappa$ ), qubits cannot be effectively initialized without resetting their states, and the readout duration cannot be shortened, resulting in a readout time of  $2 \mu\text{s}$ . The readout errors (post-selected using the method described in subsection A) are shown in Fig. S11, while the coherence times are presented in Fig. S12.



**FIG. S11. Integrated histograms of state prepared and measurement(SPAM) error.** Integrated histograms for SPAM errors of qubits' states,  $|0\rangle$  (blue) and  $|1\rangle$  (red). Their median (dashed lines) are 0.043 and 0.083, respectively.



**FIG. S12. Integrated histograms of coherence time.** Integrated histograms for coherence time of qubits,  $T_1$  (blue) and  $T_{2,\text{Ramsey}}$  (red). Their median (dashed lines) are 77.8  $\mu\text{s}$  and 6.1  $\mu\text{s}$ .

#### 4. Calibration of quantum gates

In our experiment, single-qubit gates are implemented with the qubits and couplers maintained at their idle points, minimizing the impact of XY and ZZ couplings on the gate preparation. The controlled-Z (CZ) gate between two qubits (e.g.,  $Q_1$  and  $Q_2$ ) is realized by tuning the frequency of the coupler via Z bias, while the qubit frequencies remain fixed at their idle points [58].

The parameters of the CZ gate depend on the Z bias applied to the coupler, which follows an inverted hyperbolic cosine function. During the execution of the CZ gate, state leakage out of the computational subspace must be suppressed, and the conditional phase should accumulate to  $\pi$ . Additionally,

the distortion and delay in the coupler's Z bias are calibrated using the method described in Ref. [57].

##### a. Idle point of coupler

We position the two qubits ( $Q_1$ ,  $Q_2$ ) at their idle points, where their large detuning minimizes the influence of XY coupling on single-qubit gates. Additionally, the coupler ( $C$ ) is set to its idle point to minimize ZZ coupling between the qubits. As shown in Fig. S13a, an echo-type pulse sequence is used to measure the conditional phase of  $Q_1$  and  $Q_2$ . Tuning the coupler's frequency can effectively be modeled as applying a CPHASE gate ( $U_\phi$ ) on the two qubits, provided that state leakage and qubit exchange are negligible. The operation  $U_\phi$  can be expressed as:

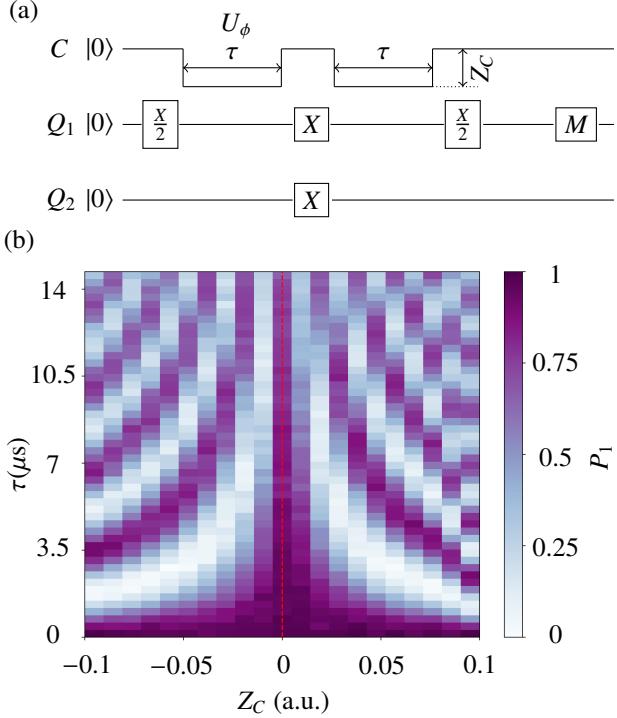
$$U_\phi = \text{diag}(1, e^{-i\phi_2}, e^{-i\phi_1}, e^{-i\phi_{12}}), \quad (\text{S3a})$$

$$\phi_{12} = \phi_1 + \phi_2 + \varphi, \quad (\text{S3b})$$

where  $\phi_1$  and  $\phi_2$  are the local Z rotation phases of  $Q_1$  and  $Q_2$ , respectively, and  $\varphi$  is conditional phase. Considering the effect of tuning the coupler's frequency on the two qubits as  $U_\phi$ , the exciting probability  $P_1$  of  $Q_1$  can be derived as:

$$P_1 = \frac{1}{2}(1 - \cos \varphi), \quad (\text{S4})$$

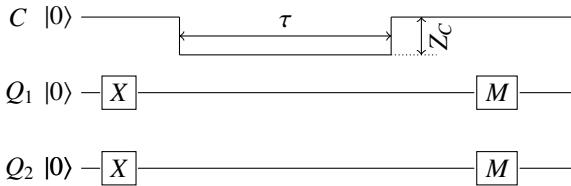
The ZZ coupling is given by  $\alpha_{zz} = \varphi/\tau$ , which corresponds to the oscillation frequency of  $P_1$ . By varying  $\tau$  and the coupler bias  $Z_c$ , we measure the probability  $P_1$  of  $Q_1$ , as shown in Fig. S13b. The idle bias of the coupler is determined where  $\alpha_{zz}$  reaches its minimum value, indicated by the red dashed line in Fig. S13b.



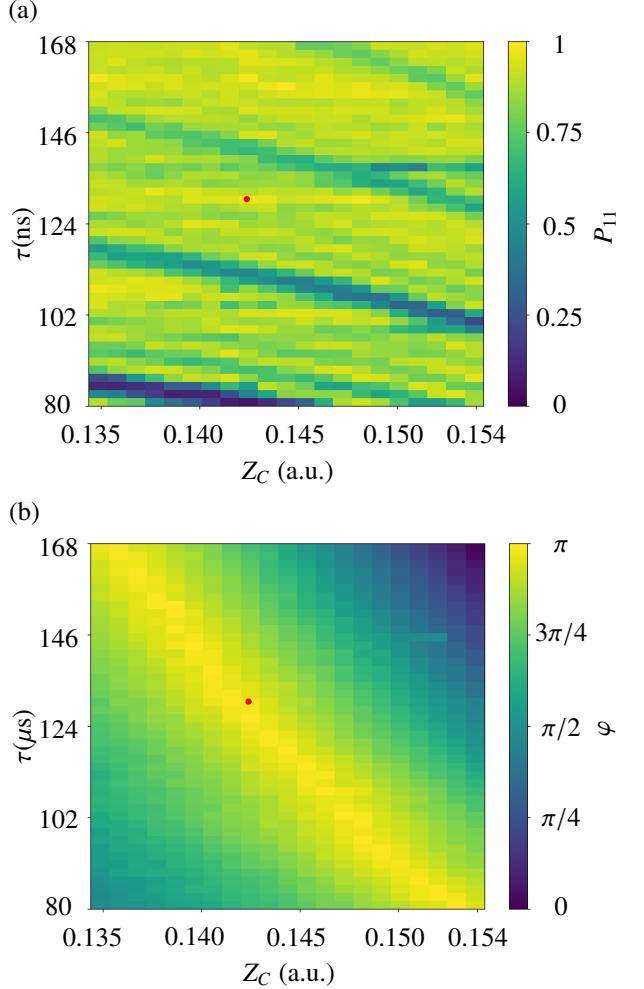
**FIG. S13. Measurement of conditional phase and ZZ coupling.** **a** shows an echo type sequence to measure conditional phase. Among which,  $\tau$  and  $Z_C$  are length and strength of Z bias of coupler ( $C$ ), respectively. **b** shows the probability ( $P_{11}$ ) of  $Q_1$  in state  $|1\rangle$  varying with  $\tau$  and  $Z_C$ .

#### b. Calibration of CZ gate

The parameters of the CZ gate include the Z bias of the coupler, the pulse length ( $\tau$ ), and the amplitude ( $Z_C$ ). To realize a CZ gate, these parameters must suppress state leakage to ensure that a  $U_\phi$  gate is applied to the two qubits while simultaneously accumulating a conditional phase of  $\pi$ . We implement the pulse sequence shown in Fig. S14 on the two qubits and the coupler, measuring the joint probability ( $P_{11}$ ) of the two qubits to experimentally characterize state leakage. Once leakage occurs, the value of  $P_{11}$  will decrease. Fig. S15a shows the variation of  $P_{11}$  with the experimental parameters  $Z_C$  and  $\tau$  of the CZ gate. The optimal parameters must satisfy the condition  $P_{11} \approx 1$  to successfully achieve a  $U_\phi$  gate.



**FIG. S14. Characterization of state leakage.** Using the pulse sequence can measure the joint probability  $P_{11}$  to characterize state leakage.  $\tau$  and  $Z_C$  are length and amplitude of Z bias of coupler ( $C$ ), respectively.



**FIG. S15. State leakage and conditional phase.** **a** and **b** show the experiment results of  $P_{11}$  and  $\varphi$  varying with  $\tau$  and  $Z_C$  shown in Fig. S13a and Fig. S14, respectively.

The circuit used for conditional phase calibration is similar to Fig. S13a, except this time we first bias the qubits and the coupler at their idle biases, and then apply a Z pulse with an amplitude of  $Z_C$  to the coupler at the position where the CZ gate is implemented. To obtain a more precise measurement of  $\varphi$ , we replace the second  $\frac{X}{2}$  pulse on  $Q_1$  with a tomography sequence and perform a tomography measurement. Conditional phase  $\varphi$  extract from tomography are shown in Fig. S15b. The parameters  $\tau$  and  $Z_C$  for the CZ gate must simultaneously satisfy  $P_{11} = 1$  and  $\varphi = \pi$ . These optimal values, labeled as  $\tau_{cz}$  and  $Z_{cz}$ , respectively, are marked as red points in Fig. S15.

Subsequently, we use a pulse train, as shown in Fig. S16a, to further optimize  $Z_{cz}$  by fine-tuning it and measuring the probability  $P_1$  of  $Q_1$ , with the results shown in Fig. S16b. Considering the effect of the coupler's Z bias on the two qubits as  $U_\phi$ , the probability  $P_1$  of  $Q_1$  can be expressed as:

$$P_1 = \frac{1}{2}(1 + \cos 2N\varphi). \quad (\text{S5})$$

As shown in Fig. S16b, for different  $N$ ,  $P_1$  has a maximum if  $\varphi = \pi$ .

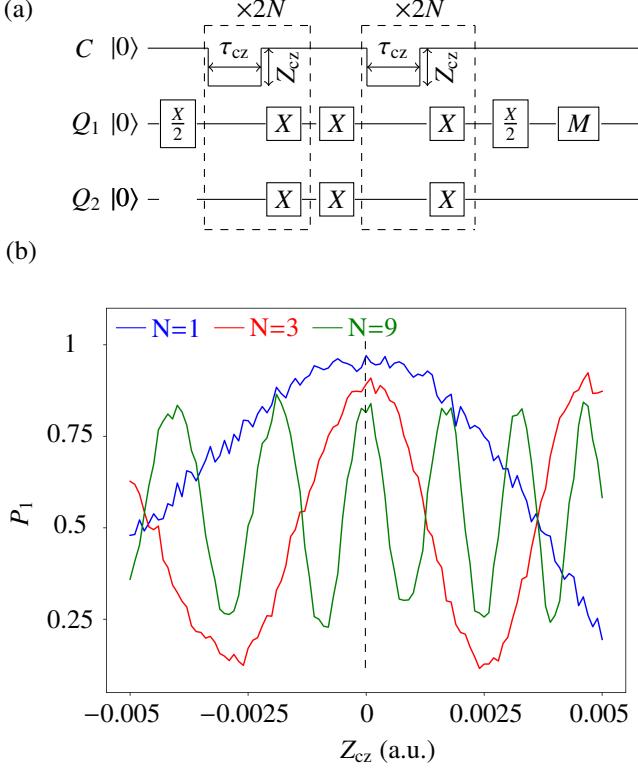


FIG. S16. **Optimization of coupler's Z bias of CZ gate.** **a** shows the pulse train to optimize coupler's Z bias of CZ gate. We repeat the pulse sequence in the dashed box  $2N$  times, measure the probability ( $P_1$ ) of  $Q_1$  and obtain the results shown in **b**.

To achieve a standard CZ gate, it is also necessary to measure and eliminate the local Z rotation phases  $\phi_1$  and  $\phi_2$ . For this purpose, a virtual Z gate is employed to cancel out the local Z rotation phases. As shown in Fig. S17a, we apply a virtual Z gate ( $Z_\phi$ ) on  $Q_i$  during the execution of the CZ gate. The probability  $P_1$  of  $Q_i$  then takes the form

$$P_1 = \frac{1}{2}(1 + \cos N(\phi_i + \phi)). \quad (\text{S6})$$

Here,  $P_1$  reaches its maximum value when  $Z_\phi$  cancels the single-qubit phase  $\phi_i$ , i.e.,  $\phi_i + \phi = 0$ , as indicated by the red point in Fig. S17b.

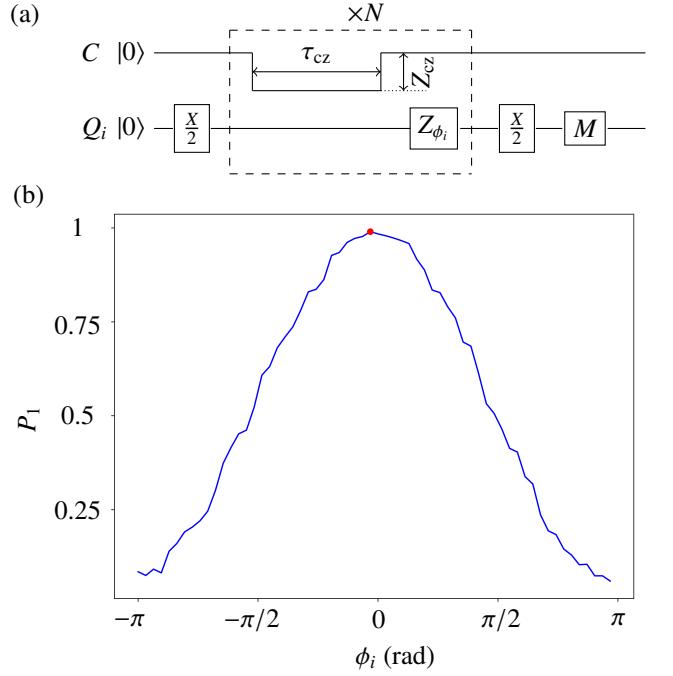


FIG. S17. **Elimination of single-qubit phase.** **a** shows a pulse sequence to eliminate single-qubit phase. The pulse sequence in dashed box is repeated  $N$  times and followed a probability ( $P_1$ ) measurement of  $Q_1$ , and the measurement results are shown in **(b)**.

### c. Leakage and fidelity of quantum gate

After performing gate calibrations, we use randomized benchmarking (RB) to evaluate the fidelity of single-qubit gates (SG) and CZ gates. As shown in Fig. S18, the median error rate for parallel single-qubit gates is below  $10^{-3}$ , while for parallel CZ gates it is below  $2 \times 10^{-2}$ . The isolated CZ gate (blue line) achieves an error rate of less than  $10^{-2}$ .

To estimate the average leakage rate per cycle, we measure the  $|2\rangle$  population of the qubit as a function of the number of cycles and fit the results to a rate equation, following the approach in Ref. [26]. The median leakage error rates are 0.0017 for parallel CZ+SG operations (blue) and 0.0002 for parallel SG operations (red).

It is noted that the fidelity of performing CZ gates in parallel is significantly lower than performing two-qubit gates independently. This is primarily caused by crosstalk, parasitic coupling, and the activation of unwanted couplings during the control process. Crosstalk can couple control signals from certain channels to unexpected channels, causing differences in the modulation signals experienced by the target qubits and the coupler when performing simultaneous CZ gates compared to those experienced during isolated calibration. Parasitic coupling and unwanted coupling lead to the spectator effect, where the states of surrounding qubits influence the evolution of the target qubit pair. These affect the gate fidelity. At the same time, these effects can also cause correlated errors in the system, as shown in Fig. S21b.

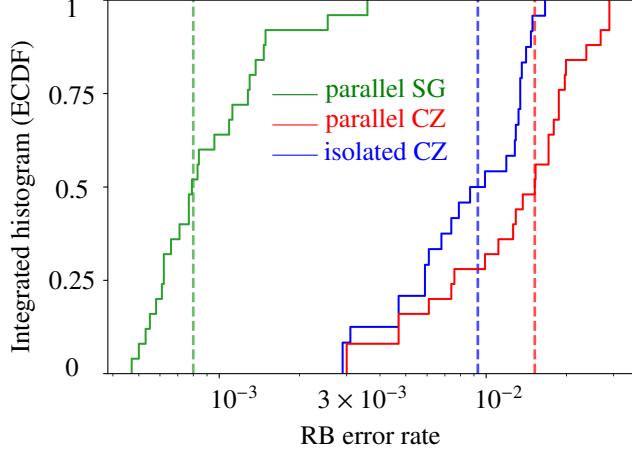


FIG. S18. **Integrated histograms of gate error.** Integrated histograms for gate errors of qubits, parallel SG (single-qubit gate), isolated CZ (CZ gate) and parallel CZ. Their median (dashed lines) are 0.0008 , 0.0095 and 0.016 respectively.

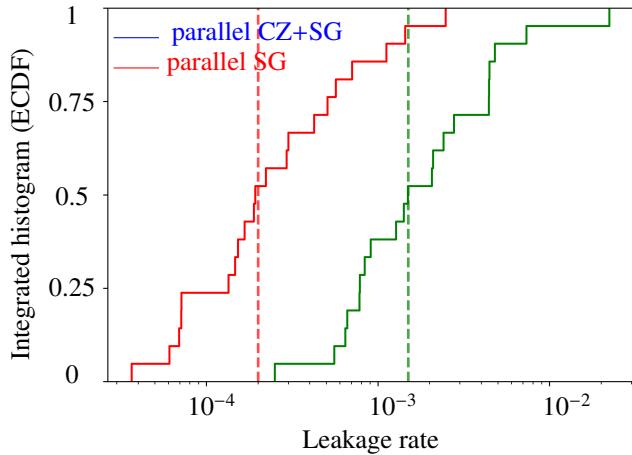


FIG. S19. **Integrated histograms of leakage error.** Integrated histograms for leakage errors of qubits, parallel CZ+SG(blue) and parallel SG(red). Their median (dashed lines) are 0.0017 and 0.0002, respectively.

##### 5. $p_{ij}$ analysis for configuring decoding priors

To enhance the accuracy of the decoding process, it is crucial to provide the decoder with a set of prior probability distributions that closely reflect the actual noise conditions of the hardware. These probabilities correspond to the error mechanisms in the decoder error model (DEM) and form the weights of the edges in the decoding graph. This set of prior values is derived through a method known as “ $P_{ij}$  analysis.” This

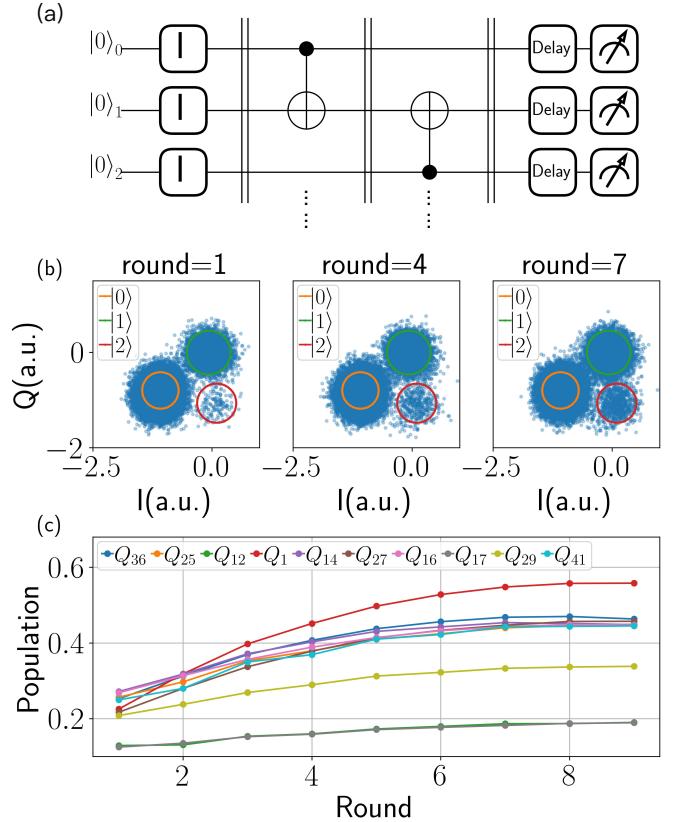


FIG. S20. **Leakage analysis.** (a) The circuit unit for characterizing leakage errors in the repetition code experiment, with repeated measurements conducted in multiple rounds. Even-numbered qubits represent data qubits, odd-numbered qubits represent measurement qubits, and ellipses indicate the extension of the circuit structure for additional qubits. (b) The in-phase and quadrature components of specific qubit’s readout signals during the execution of the circuit in (a). The orange, green, and red circles represent the positions of the readout signals corresponding to qubits in states 0, 1, and 2, respectively. As the number of measurement rounds increases, the proportion of state-2 increases significantly, indicating that more frequent measurements lead to greater leakage of quantum information out of the computational subspace. (c) The average detection events for different qubits as a function of the number of measurement rounds, distinguishing only between readout signals of states 0 and 1. Higher average values exceeding 0.5 indicate more severe state leakage errors. It is evident that Q1 experiences the most significant state leakage errors.

method aims to calculate the correlation between specific detector configurations by statistically analyzing the frequency of their simultaneous occurrences across numerous samples, using this correlation to characterize probability. Further details can be found in [26].

Fig. S21(b) presents the results of the  $P_{ij}$  correlation analysis for a distance-11 round-3 repetition code, with its corresponding decoding graph shown above. The x and y axes are divided by thicker gray lines to indicate different detectors, while the smaller divisions within these blocks, marked by thin gray lines, represent different rounds of the same detector. The heatmap is symmetric along the line  $y = x$ , as the

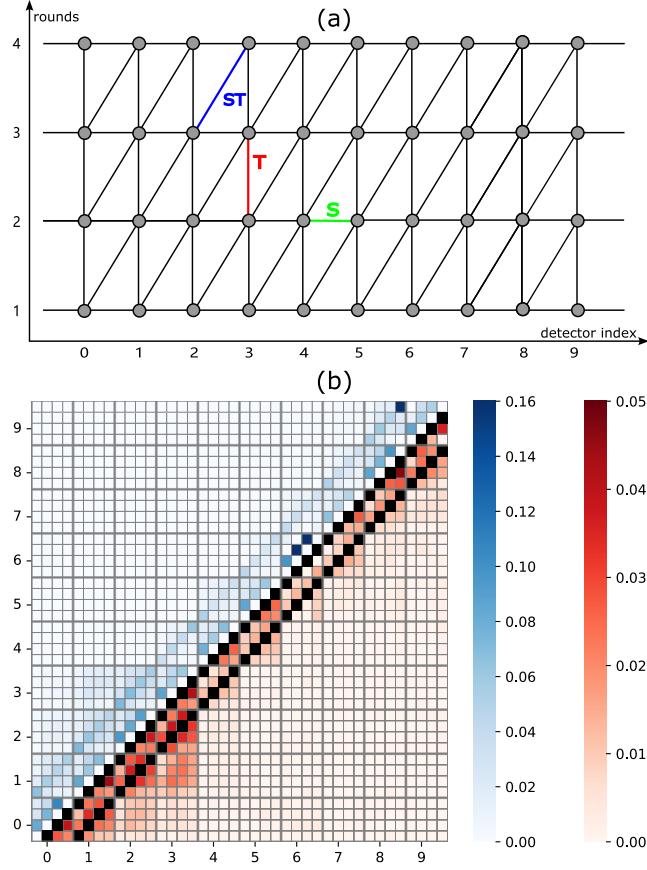


FIG. S21. (a) The Detector Error Model (DEM) for a distance-11 round-3 repetition code is depicted, featuring four rounds of detectors. The final round of detectors is derived from the final measurement of all data qubits. Different types of edges are represented in various colors. (b) The heatmap of the  $P_{ij}$  correlation analysis is shown. The large blocks along the x and y axes, divided by thicker gray lines, correspond to the detector indices in (a). Within each large block, smaller divisions marked by thin gray lines represent different rounds, corresponding to the y-axis in (a)

correlation between  $i$  and  $j$  is identical to that between  $j$  and  $i$ . To better visualize these correlations, we employ two color bars to distinguish between different value scales: the upper-left section of the heatmap uses a blue color bar for higher correlations, while the lower-right section uses a red bar for relatively lower correlations. The black blocks in the lower-right part are masked to subtract the highest values, which are displayed in the symmetric upper-left location. These masked blocks correspond to error mechanisms that connect adjacent rounds of the same detector and neighboring detectors within the same round (i.e., red T (timelike error) edge and green S (spacelike error) edge in Fig. S21(a)). Notably, the time-

like error probability for the 6-th detector shows an unnaturally high value as the rounds progress, indicating poor performance of the 6th ancilla qubit, which aligns with our evaluation shown in Fig. S20. Additionally, the pronounced red areas spilling over in the lower-left corner indices 0, 1, 2, and 3 of the detector can be attributed to the effects caused by the diffusion of leakage and crosstalk. At the end of the article, we have included a heatmap Fig. S22 of the  $P_{ij}$  correlation analysis for a distance-11 round-7 repetition code. In this heatmap, the presence of spacetime error (ST) edges is clearly visible.

However, this method relies on a strong assumption: it presumes that correlations exist only between selected specific detectors. This assumption holds in an ideal circuit-level error model, where it is assumed that only Pauli errors are present. Under this assumption, each Pauli error causes the simultaneous flipping of at most two adjacent detectors, meaning that each Pauli error that fits this assumption can be incorporated into a single edge (error mechanism), as illustrated in Fig. S3. Nevertheless, real quantum hardware includes a small number of error mechanisms beyond the error model, which can lead to simultaneous flipping of multiple neighboring detectors or long-range correlations between detectors. These additional error types not only interfere with the calculation of probabilities for error mechanisms within the model but also are not currently accounted for by the decoder, thereby reducing decoding performance.

Another issue arising from these errors beyond the standard circuit-level error model is that they can sometimes lead to deviated or even non-physical values when calculating the probabilities of error mechanisms at the boundary (i.e., edges linked to only one detector). This problem occurs because the method for calculating probabilities at the boundary involves subtracting the probabilities of all other edges connected to the same detector from the total probability of that detector flipping. The additional multi-body correlations caused by non-Pauli errors increase the sum of the probabilities to be subtracted, potentially resulting in deviated or negative probabilities for boundary edges.

Part of this issue can be attributed to the non-reset repetition code we use, which cannot effectively suppress leakage diffusion. When the number of distance and rounds are limited, the impact of beyond-Pauli errors on performance is not so significant, only some deviations occur. However, when the distance and rounds are equal to or greater than 7, negative probability issues may sometimes arise, necessitating correction. We simply adjust those possible deviation and the one to three non-physical boundary probability values to an empirical value (equal to a certain empirical constant), which still yields physical results after correction. However, when the number of distance and rounds is greater than or equal to 11, the corrected decoding results do not demonstrate a decreasing trend in the logical error rate with increasing distance.

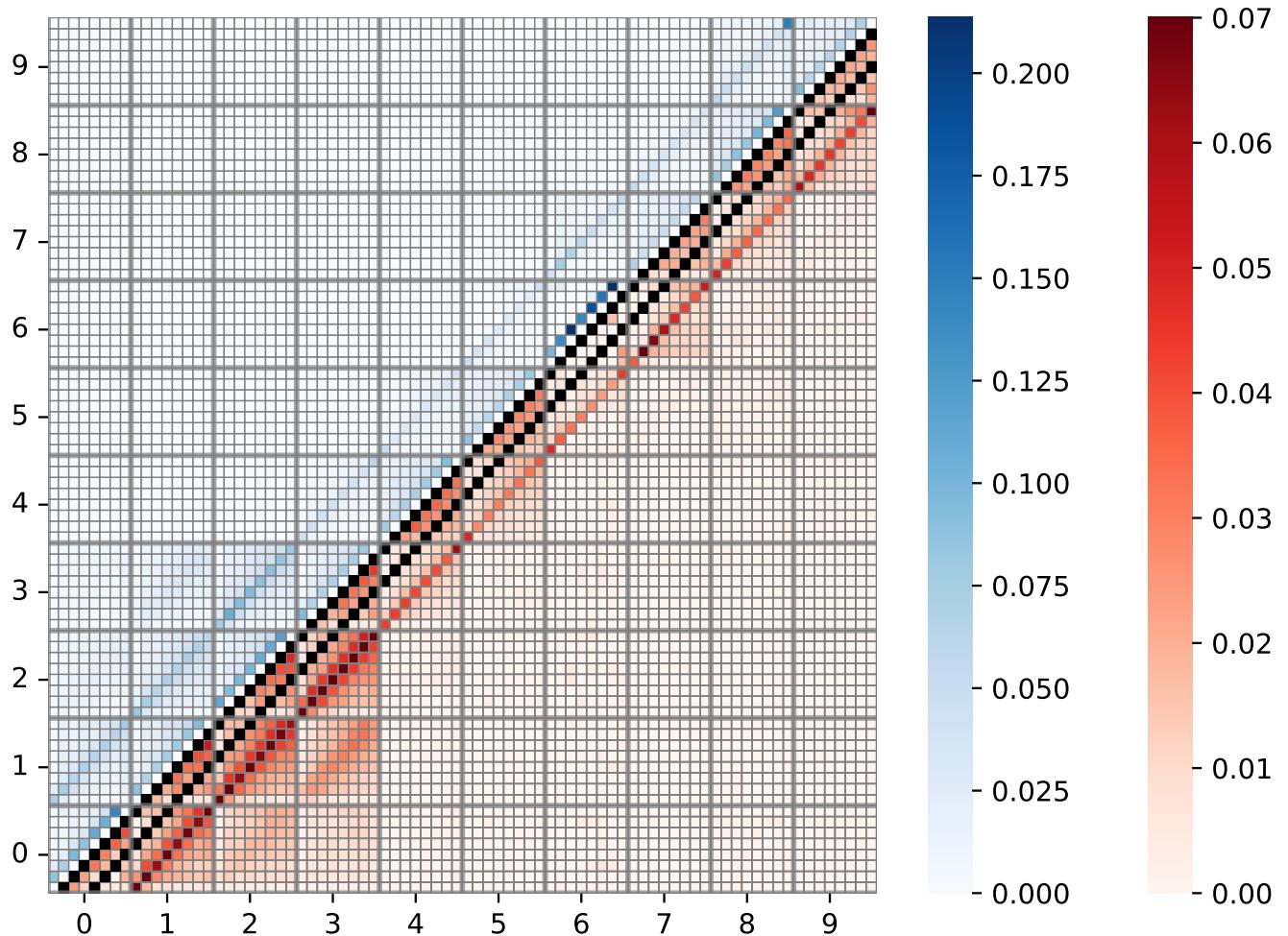


FIG. S22. The heatmap of the  $P_{ij}$  correlation analysis for distance 11, rounds 7. The presence of spacetime (ST) error edges are clearly visible