Realization of an Error-Correcting Surface Code with Superconducting Qubits

Youwei Zhao,1,2,3,* Yangsen Ye,1,2,3,* He-Liang Huang,1,2,3,* Yiming Zhang,1,2,3 Dachao Wu,1,2,3 Huijie Guan,1,2,3 Qingling Zhu,1,2,3 Zuolin Wei,1,2,3 Tan He,1,2,3 Sirui Cao,1,2,3 Fusheng Chen,1,2,3 Tung-Hsun Chung,1,2,3 Hui Deng,1,2,3 Daojin Fan,1,2,3 Ming Gong,1,2,3 Cheng Guo,1,2,3 Shaojun Guo,1,2,3 Lianchen Han,1,2,3 Na Li,1,2,3 Shaowei Li,1,2,3 Yuan Li,1,2,3 Futian Liang,1,2,3 Jin Lin,1,2,3 Haoran Qian,1,2,3 Hao Rong,1,2,3 Hong Su,1,2,3 Lihu Su,1,2,3 Shiyou Wang,1,2,3 Yulin Wu,1,2,3 Yu Xu,1,2,3 Chong Ying,1,2,3 Jiale Yu,1,2,3 Chen Zha,1,2,3 Kaili Zhang,1,2,3 Yong-Heng Huo,1,2,3 Chao-Yang Lu,1,2,3 Cheng-Zhi Peng,1,2,3 Xiaobo Zhu,1,2,3 and Jian-Wei Pan1,2,3

1Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China
2Shanghai Branch, CAS Center for Excellence in Quantum Information and Quantum Physics, University of Science and Technology of China, Shanghai 201315, China
3Shanghai Research Center for Quantum Sciences, Shanghai 201315, China

(Received 2 January 2022; accepted 8 April 2022; published 11 July 2022)

Quantum error correction is a critical technique for transitioning from noisy intermediate-scale quantum devices to fully fledged quantum computers. The surface code, which has a high threshold error rate, is the leading quantum error correction code for two-dimensional grid architecture. So far, the repeated error correction capability of the surface code has not been realized experimentally. Here, we experimentally implement an error-correcting surface code, the distance-three surface code which consists of 17 qubits, on the Zuchongzhi 2.1 superconducting quantum processor. By executing several consecutive error correction cycles, the logical error can be significantly reduced after applying corrections, achieving the repeated error correction of surface code for the first time. This experiment represents a fully functional instance of an error-correcting surface code, providing a key step on the path towards scalable fault-tolerant quantum computing.

DOI: 10.1103/PhysRevLett.129.030501

Introduction.—Progress towards scalable fault-tolerant quantum computation relies on exploiting quantum error correction (QEC) to protect quantum systems against inevitable noises [1–7]. Notable experimental implementations on a variety of QEC architectures include surface code [8–11], REPETITION code [9], BOSONIC code [12–14], SHOR code [15,16], COLOR code [17,18], [5,1,3] code [19], etc. [20,21]. The surface code is so far the leading candidate for fault-tolerant quantum computation, featuring a high threshold error rate and requiring only nearest neighbor interactions on a two-dimensional square lattice [22–24]. This property makes it is perfectly compatible with the devices fabricated using planar photolithography, such as superconducting and quantum dot systems. In 2014, Barends et al. first realized high fidelity quantum gates at the fault-tolerant threshold for the surface code in a superconducting quantum processor [25]. Surface code experiments have since been progressively developed in terms of scale and utility. The entangling operations between two four-qubit surface codes using lattice surgery has been demonstrated in a ion-trap quantum processor [10]. Most recently, the surface code with distance-two has been realized with seven qubits to demonstrate the repeated error detection [8,9,11]. Until now, all the surface code experiments have only the ability to detect errors, and repeated error correction of surface code has not been implemented in any experiment. However, as the system size and circuit depth grow, the fidelity decreases exponentially, making it impractical to rely solely on error detection and dropping error events.

The ability of repeated error correction is essential for realizing large-scale quantum algorithms [26–29], but it is significantly more difficult than just error detection. On the one hand, more redundant qubits are required for error correction than for error detection. However, scaling the number of qubits while maintaining high-fidelity quantum operations remains a key challenge for quantum computing. On the other hand, one needs to know the exact number of qubits that are corrupted and more importantly, their location in the quantum state and the type of error.

In this Letter, we present the first implementation of repeated error detection and correction of the surface code. Specifically, by encoding the logical state using a distance-three surface code on the Zuchongzhi 2.1 superconducting quantum system [30,31], we show that the logical error can be reduced by up to about 26% after applying corrections in postprocessing. We also test the error detection performance of this code, and observe that the lifetime of the logical qubit is longer than those of any constituent physical qubits, when we postselect the instances that no
error is detected by both data qubits measurements and the stabilizer measurements in any cycle. Our studies, for the first time, demonstrate the feasibility of repeated quantum error correction using surface code, guiding future efforts to realize more powerful large-scale quantum error correction.

**Encoding the logical qubit with a distance-three surface code.**—The surface code, as suggested by Refs. [22–24] can be implemented on a two-dimensional array of physical qubits with only nearest-neighbor coupling, making its realization readily available on the Zuchongzhi 2.1 superconducting qubit platform. Here, we have chosen 17 out of the 66 qubits from the Zuchongzhi 2.1 system and created a distance-three surface code.

Its structure are depicted in Fig. 1(a). This 17-qubit surface code consists of 9 data qubits and 8 measurement qubits. The data qubits, which store the computation quantum state, are represented by gray dots with label $D_j, j = 1, 2, \ldots, 9$. There are two types of measurement qubits. The green dots whose label starts with $Z$ describe $Z$-ancilla qubits which measure the $Z$ parity of its adjacent data qubits, while a red dot, whose label has an $X$ in it, depicts an $X$-ancilla qubit that checks the sign of the product of Pauli $X$ operators acting on its neighboring data qubits. Taking measurement over all ancilla qubits projects the code space into the subspace spanned by the eigenstates of these parity check operators. The corresponding outcomes completely describe the state of the system. And any error occurred amid a process will manifest itself as a change in the measurement of these ancilla qubits. This enables us to keep track of the evolution of the system with only ancilla qubits without the risk of corrupting it.

Between each data qubit and measurement qubit, there resides a coupler denoted by a square rectangle. The introduction of couplers allows us to dynamically turn on and off interactions between nearest neighbor qubits [32]. On one hand, this allows us to implement entangling operations on adjacent pair of qubits while being protected from unwanted crosstalk. This makes it possible to pack all two-qubit gates in one surface code cycle into four layers. On the other, the architecture also alleviates the frequency crowding problems, making it easier to scale up the quantum system.

Figure 1(b) describes gate sequences in one surface code cycle. The ordering of the qubits here is in accordance with the spatial distribution as illustrated in Fig. 1(a). The gates are applied from left to right in 9 time steps. All operations encapsulated in the same colored box are applied simultaneously. The color of each two-qubit gate also matches that of the edge in Fig. 1(a) to facilitate understanding of the order of gates in a physical context.

The line with two dots on the ends represents a controlled-phase gate, i.e., CZ gate. The circuit is modified from the one used in Ref. [33] to facilitate gate implementations on our hardware. One of the changes is to replace all Hadamard gates by $R_Y(\pm \pi/2)$ rotations, aside from reshuffling of gate ordering. The replacement is based on the fact that a Hadamard gate can be decomposed as $H = ZR_Y(-\pi/2)$ or $H = R_Y(\pi/2)Z$ and $(I \otimes Z)\text{CZ}(I \otimes Z) = CZ$.

As a last step of the surface code cycle, all states of the ancilla qubits are measured in the $Z$ basis, where extra $Y$ gates, which are equivalent to Hadamard gates, are applied on the $X$-ancilla qubit to produce the needed $X$ measurement. In the meanwhile, a dynamical decoupling operation (DD) [9] is applied to the data target to mitigate the dephasing problem of the data qubits.

**System calibration.**—Figure 2(a) displays the integrated histogram for a single-qubit gate, CZ gate, and readout error after performing calibration. Each single-qubit gate can be implemented in 25 ns with an average error of 0.092%. The duration of each CZ gates is set to be 32 ns, with the average error being 0.878%. The measurement operation takes 1.5 $\mu$s. With this setting, average readout error is 5.389%. Since the readout read line width $\kappa/2\pi$ is small in our experimental setup, to reduce the effect of photon residue in the resonator before gate recommence, we conservatively insert an idle operation of $3.5 \mu$s for a measurement operation.

Not only do we need to know how well we are doing with basic quantum operations, it is also necessary to assure ourselves that the complex state preparation and circuit
operation are well understood. For this purpose, we carried out a random circuit sampling task tailored for the surface code experiment, and compute the linear cross-entropy benchmarking fidelity, which is defined as

\[ F_{\text{XEB}} = 2^n(P(x_i)) - 1, \]  

(1)

to characterize the performance in a circuit level, where \( n \) is the number of qubits. \( P(x_i) \) represents the probability of bitstring \( x_i \) computed for the ideal quantum circuit.

The random circuit is made up of alternating layers of two-qubit gate and single-qubit gate, but it is slightly different from the previous works [30,31,34]. To properly extract the effect of gate sequence to the fidelity, the two-qubit gates are chosen from the set \{A, B, C, D\} in the same order as the surface code cycle. The single-qubit gate is set in the same fashion as how the Zuchongzhi 2.1 is benchmarked [30,31]. To be explicit, a random gate is selected from the pool of gates \{R_x(\pi/2), R_y(\pi/2) and \( R_{X+Y}(\pi/2) \}\) and applied to each of the 17 qubits, with the only requirement such that subsequent single-qubit gate cannot be the same. In total, we stack 21 single-qubit gate layers and 20 two-qubit gate layers in the circuit for the sampling task.

We have sampled 9 instances of random circuits with different random seeds [see Fig. 2(b)], and the average fidelity achieved is 0.030 ± 0.002. We also calculate a prediction of this value, which is 0.033, by taking the product of the Pauli fidelity of each single-qubit, two-qubit and measurement operation. As can be seen from the result, the fidelity of each gate is capable of predicting the circuit performance notwithstanding the inevitability of crosstalk and tailing in wave fronts.

**Experimental results.**—With enough information about fidelity in both gate and circuit levels, we are ready to carry out surface code experiments. For illustration purposes, we will consider the performance of the code in preserving the logical zero state \((0_L)\) and logical minus state \((-L)\) in the main text and leave the discussions of logical one state \((1_L)\) and logical plus state \((+L)\) surface code in the Supplemental Material [35], which includes Refs. [36–41]. To initialize a \((0_L)\) \((-L)\) state, the data qubits are prepared in \( |0\rangle^\otimes (|\rangle^\otimes) \) states with all ancilla qubits set to \( 0 \) states. Then one surface code cycle, as shown in Fig. 1(b), is applied which casts the Hilbert space into an eigenstate of all syndrome measurement operators. Depending on the data qubit being in a \( 0 \) or \(-\) state, this corresponds to a logical \((0_L)\) or \((-L)\) state, respectively. The state we obtain at this stage is described as round 1 in Fig. 3. The detection event fraction for round 1 equals the syndrome measurement values, which for perfect system without errors, should be 0.

Then, we repeat the surface code cycle up to 12 times. At the end of each stack, the states of all eight syndrome qubits, both the \( X \) and \( Z \) ancilla, are measured. For the last clock cycle, we also measure the state of the data qubits in the same basis as the syndrome measurement. This will give us an extra piece of information about the stabilizer. In total, for the circuit with \( n \) cycles, we record \( n \) sets of measurement about the ancilla qubits and one set of outcome for data qubit. Here we only show the error detection in the \( Z \) \( X \)-type ancilla qubits for logical \((0_L)\) \((-L)\) state in the main text, as the reading from the \( X \) \( Z \)-type ancilla qubits, which are only sensitive to the phase (bit) flip errors and thus in principle have no influence on the logical \( Z \) \( X \)-measurement of the logical \((0_L)\) \((-L)\) state, would not so informative. Full data can be found in the Supplemental Material [35].

A direct information one can extract from these data is error occurrence. To do so, the state of a measure qubit takes an XOR operator with that of its previous run, giving rise to the value of a stabilizer relating to this cycle. This operation serves the same purpose as a reset gate acting on the ancilla qubits, which is missing in our circuit. The stabilizer on the first run simply adopts the value of the reading from the first measure. Aside from these stabilizer values, an extra record is created by comparing the
calculated ancilla qubit state from the parity of the data qubits to the values measured in the last run. Whenever a change occurs in subsequent stabilizer values, a detection event of errors is fired.

Making use of the 13 rounds of stabilizer values, we can execute detection event analysis. The result is shown in Fig. 3. Figure 3(a) describes the fraction of a detection event (DEF) for the logical $|0_L\rangle$ state. Each line corresponds to the DEF for one $Z$-ancilla qubit. Here, the value of the first and last round is apparently lower than the others. This is a result of skipping idle operation which applies to all measurements except the last one. Since the first set of data is copied directly from the measurement without comparing with a physical data, the influence of an idle operation is also reduced. The curves in the middle are flat across different rounds, with slight lifting in the trend. It relates to leakage into a state out of the computational basis. Since the effect is minor, it indicates a low leakage error rate in our system. Overall, roughly 33% of measurements signaled a detection event. We attribute this high rate to our long measurement duration. Figure 3(b) describes the same quantity for the logical $|-L\rangle$ state. The experiment is done similarly except that we place a $R_y(-\pi/2)$ gate in the beginning and $R_y(-\pi/2)$ gate in the end of the circuit to each data qubit to produce the $|-L\rangle$ state after the first cycle, and measure all the data qubits in the $X$ basis after the final cycle. We see a similar result as we did for the logical $|0_L\rangle$ state, as expected.

The power of a surface code lies beyond it capability to detect errors. The pattern of the occurrence of these anomalies shed light on the errors happening on the data qubits. A powerful tool to visualize the correlation among each detection events is the correlation matrix. The correlation matrix describes the likelihood of observing a detection event at ancilla qubit $D_j$ at round $R_i$ given an error occurs at site $D_i$ at round $R_j$ which can be written as

$$p_{ij} = \frac{(p_{DiR_i} p_{DiR_j}) - (p_{DiR_i} p_{DiR_j})}{\sqrt{((p_{DiR_i} - (p_{DiR_i}))^2)((p_{DiR_j} - (p_{DiR_j}))^2)}} - \delta_{ij}, \quad (2)$$

with $\delta_{ij}$ being the Kronecker delta with $\delta_{ij} = 1$ if $i = j$ and 0 otherwise.

The measurement we have made and the resulting syndromes are converted into a correlation matrix as shown in Fig. 4(a) for logical $|0_L\rangle$ state and Fig. 4(b) for logical $|-L\rangle$ state. The main diagonal blocks describe simultaneous correlation of detections among various ancilla qubits. The patterns it presents are consistent with the topology of the surface code and can be explained with potential crosstalks between neighboring gates. The blocks away from the main diagonal by a distance of one or two correspond to detection events one or two QEC cycles apart. The peaks in detection correlation in these blocks are due to error accumulation along the cycles since no reset operations are taken between two surface code cycles. As to the patterns further away, they may be explained by leakage to state out of the computational basis. Here, we have truncated the correlation matrix at $p = 0$. The negative correlations are ignored which comes from us ignoring higher order correlations.

Once an error has been detected, we can post-process the data and reduce the effect it brings to our system. One direct method is to drop erroneous data all in all. This simple treatment can greatly improve the fidelity of the logical state stored in our system by sacrificing efficiency of a dataset. Here fidelity refers to the fraction of logical states that is intact after a certain clock cycles. The results are shown in Figs. 5(a)–5(b). There are five lines in each graph. The purple line is obtained as described before. That is, the fidelity is calculated using data where no error is detected. Here the green line describes the situation where we only drop data if the final measurement of the data qubits implies an error. The green line corresponds to the case where detection is based on measurement on ancilla qubits only. The red lines are obtained with all data without any postprocessing. We have also included a gray dotted line describing the calculated fidelity based on the maximum physical relaxation time $T_1$ of participating physical qubits. By comparing these lines, we can see that both the measurement on data qubits and that on the ancilla qubit helps detecting errors and neither of them is capable of detecting all instances of error by itself. The inset describes the retained rate as a function of QEC round. Because of the exponential drop of the retained rate, we did not exercise analysis for rounds $> 5$ with the number of shots for each experiment set to 480 000.

We derived the logical error rate by fitting the curves with the expression used in Ref. [42]:

$$\mathcal{F}_L(k) = \frac{1}{2}[1 + (1 - 2\epsilon_L)^{k-k_0}]. \quad (3)$$

Here $k_0$ and $\epsilon_L$ relate to a shift in the round index and logical error rates which are to be fit. The fitted logical error rates are shown by each line. We also calculate the coherent
Calculated from the duration of single-qubit gate τ of logical qubit gate τ_{\text{photon depletion}}. Logical error rates modified from Ref. [42] by considering decoherence from physical qubits. Logical error rates are extracted from the measurements of either data qubits or ancilla qubits. The green (hexagon) line corresponds to the dropping scheme based on data qubits only while the navy (octagon) line is only affected by the measurement of ancilla qubits. The red (triangle) line describes the result with no postselection. The dotted line depicts the prediction based on data qubits only. To mitigate the distortion in the quantum computing system ζ, we introduce an idle operation in between for the same duration of τ, and measurement with waiting time for the number will be reduced with a longer code distance and shorter measurement time.

Conclusion and outlook.—Our experiment expands the surface code experiment’s capabilities beyond error detection to include error correction. We achieved high-fidelity operations of single-qubit gates, two-qubit gates, and readout simultaneously on a system, ensuring that we could observe the effectiveness of error correction, that is, the fidelity of the logical state is improved after error correction. This 17-qubit distance-three surface code, capable of correcting both bit- and phase-flip errors, already enables us to create a single fault-tolerant surface code memory. Future work will concentrate on realizing larger-scale surface codes to achieve the important goal of suppressing the logical error rate as the code distance increases. This necessitates further improvements to the quantum computing system’s performance, such as the number and quality of qubits, the fidelity of quantum gate operations, and rapid feedback of digital electronics.

The authors thank the USTC Center for Micro- and Nanoscale Research and Fabrication for supporting the sample fabrication. The authors also thank QuantumCTek Co., Ltd., for supporting the fabrication and the maintenance of room-temperature electronics. This research was supported by the National Key R&D Program of China, Grant No. 2017YFA0304300, the Chinese Academy of Sciences, Anhui Initiative in Quantum Information Technologies, Technology Committee of Shanghai Municipality, National Science Foundation of China (Grants No. 11905217, No. 11774326), Natural Science Foundation of Shanghai (Grant No. 19ZR1462700), and Key-Area Research and Development Program of Guangdong Province (Grant No. 2020B030300001).
H.-L. H. acknowledges support from National Natural Science Foundation of China (Grants No. 11905294), Youth Talent Lifting Project (Grant No. 2020-JCJQ-QT-030), China Postdoctoral Science Foundation, and the Open Research Fund from State Key Laboratory of High Performance Computing of China (Grant No. 201901-01).

Note added.—Recently, we became aware of a similar work by Knirren et al. [46], which was carried out independently.

*These three authors contributed equally.

[2] D. A. Lidar and T. A. Brun, Quantum Error Correction
76, 076001 (2013).
(1996).
Sci. 63, 180501 (2020).
[8] C. K. Andersen, A. Remm, S. Lazar, S. Knirren, N. Lacroix,
G. J. Norris, M. Gabureac, C. Eichler, and A. Wallraff, Nat.
Phys. 16, 875 (2020).
[10] A. Erhard, H. P. Nautrup, M. Meth, L. Postler, R. Stricker,
M. Stadler, V. Negnevitsky, M. Ringbauer, P. Schindler,
Muthusubramanian, C. Zachariadis, F. Battistel, M.
Beekman, N. Haider, W. Vlothuizen et al., arXiv:
2102.13071 (2021).
H. Wang, Y. Song, C.-L. Zou et al., Nat. Phys. 15, 503
(2019).
[14] N. Ofel, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas,
B. Vlastakis, Y. Liu, L. Frunzio, S. Girvin, L. Jiang et al.
H.-Y. Tang, Q. Zhao, X.-L. Wang, K. Fujii, L. Li, N.-L. Liu,
K. Nemoto, W. J. Munro, C.-Y. Lu, A. Zeilinger, and J.-W.
(2021).
Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina et al.,
[17] D. Nigg, M. Mueller, E. A. Martinez, P. Schindler, M.
Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt,
J. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N.
[19] M. Gong, X. Yuan, S. Wang, Y. Wu, Y. Zhao, C. Zha, S. Li,
Z. Zhang, Q. Zhao, Y. Liu et al., Natl. Sci. Rev. 9, nwab011
(2022).
[20] H.-L. Huang, M. Narożniak, F. Liang, Y. Zhao, A. D.
Castellano, M. Gong, Y. Wu, S. Wang, J. Lin, Y. Xu, H.
Deng, H. Rong, J. P. Dowling, C. Z. Peng, T. Byrnes, X.
Zhu, and J. W. Pan, Phys. Rev. Lett. 126, 090502
(2021).
Yang, L. Li, N.-L. Liu, J. P. Dowling, T. Byrnes et al.,
Optica 6, 264 (2019).
190504 (2007).
Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell
[26] H.-L. Huang, Y. Du, M. Gong, Y. Zhao, Y. Wu, C. Wang,
S. Li, F. Liang, J. Lin, Y. Xu, R. Yang, T. Liu, M.-H. Hsieh, H.
Deng, H. Rong, C.-Z. Peng, C.-Y. Lu, Y.-A. Chen, D. Tao,
(2021).
Zhao, C. Liu, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan,
[29] H.-L. Huang, Q. Zhao, M. C. Liu, Z.-E. Su, X.-L. Wang,
L. Li, N.-L. Liu, B. C. Sanders, C.-Y. Lu, and J. W. Pan,
Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro et al.,
[33] R. Versluis, S. Poletto, N. Khammassi, B. Tarasinski, N.
Haider, D. J. Michalak, A. Bruno, K. Bertels, and L.
[34] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R.
Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell
supplemental/10.1103/PhysRevLett.129.030501 for the
system parameters of the quantum processor and experimental
details.
J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R.
[37] Y. Ye, S. Cao, Y. Wu, X. Chen, Q. Zhu, S. Li, F. Chen,
M. Gong, C. Zha, H.-L. Huang, Y. Zhao, S. Wang, S. Guo,
H. Qian, F. Liang, J. Lin, Y. Xu, C. Guo, L. Sun, N. Li, H.
(2021).
[38] Y. Sung, L. Ding, J. Braumüller, A. Vepsäläinen, B.
Kannan, M. Kjaergaard, A. Greene, G. O. Samach, C. McNally,
D. Kim, A. Melville, B. M. Niedzielski, M. E. Schwartz,
J. L. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver,


