Distributed Quantum Error Correction for Chip-Level Catastrophic Errors

Qian Xu,[1,*] Alireza Seif,[1,*] Haoxiong Yan,[1] Nam Mannucci,[1] Bernard Ousmane Sane,[2] Rodney Van Meter,[3] Andrew N. Cleland,[1,4] and Liang Jiang[1,†]

[1]*Pritzker School of Molecular Engineering, The University of Chicago, Chicago, Illinois 60637, USA*
[2]*Graduate School of Media and Governance, Keio University, 5322 Endo, Fujisawa 252-0882, Japan*
[3]*Faculty of Environment and Information Studies, Keio University, 5322 Endo, Fujisawa 252-0882, Japan*
[4]*Center for Molecular Engineering and Material Science Division, Argonne National Laboratory, Lemont, Illinois 60439, USA*

(Received 11 April 2022; accepted 6 October 2022; published 6 December 2022)

Quantum error correction holds the key to scaling up quantum computers. Cosmic ray events severely impact the operation of a quantum computer by causing chip-level catastrophic errors, essentially erasing the information encoded in a chip. Here, we present a distributed error correction scheme to combat the devastating effect of such events by introducing an additional layer of quantum erasure error correcting code across separate chips. We show that our scheme is fault tolerant against chip-level catastrophic errors and discuss its experimental implementation using superconducting qubits with microwave links. Our analysis shows that in state-of-the-art experiments, it is possible to suppress the rate of these errors from 1 per 10 s to less than 1 per month.

DOI: 10.1103/PhysRevLett.129.240502

Introduction.—Extreme sensitivity to external noise is one of the main obstacles in building and operating large-scale quantum devices. Quantum error correction (QEC) solves this issue by encoding quantum information in a larger space so that the errors can be detected and corrected (see, e.g., Ref. [1] (Chap 10) and Ref. [2]). Existing QEC schemes mostly focus on local and uncorrelated error (or errors with finite-range correlations), see, e.g., [3,4]. Long-range correlations, however, e.g., due to coupling to a bosonic bath [5–7], can negatively impact the performance of QEC [8,9].

Recently, it has been shown that a cosmic ray event (CRE) can cause catastrophic correlated errors in superconducting qubits [10–13]. Upon impact of high-energy rays, phonons are created and spread in the substrate. These phonons then create quasiparticles in the superconducting material, which subsequently induces qubit decay [12]. Even though these events are rare, their effect is devastating as they cause fast correlated relaxation (T_1 error) in all the qubits in a chip that essentially erases the encoded quantum information [12], which is especially detrimental to long computational tasks that could take several hours [14]. Moreover, the adverse effect of CREs is not limited to superconducting qubits. Semiconductor spin qubits [15] and qubits based on Majorana fermions [16,17] also suffer from the charge noise and quasiparticle poisoning that are resulted from CRE, respectively. One system-specific approach to reducing the impact of CREs is through changing the design of the device, for example, by introducing phonon and quasiparticle traps [18–20] and enhancing phonon relaxation in the device [17].

In this Letter, we take a different approach and use a distributed error correcting scheme to detect and correct correlated errors by CREs. Distributed hardware architectures, connecting smaller nodes into a tightly coupled system using an interconnect network, have been proposed to achieve scalability for a single computation [21–29]. Here, we repurpose these architectures to improve fault tolerance. Our approach is system independent and works as long as a quantum network can be built to share entanglement between separate chips. In a network of chips, a CRE erases information from one chip, but as we show this event and the specific impacted chip can be detected [see Fig. 1(a)]. Since the location of the error is now known, we can use erasure QEC to correct the errors and recover the information [30–34]. We present a low-overhead erasure QEC scheme that is fault tolerant against the CREs and discuss its implementation using

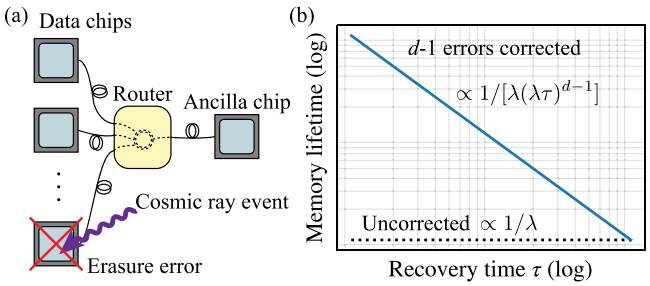

FIG. 1. (a) Information is encoded in an error correcting code that is distributed across multiple data chips. The CRE-induced erasure errors on the data chips are corrected using ancilla-assisted syndrome measurements. (b) A code that corrects $d-1$ erasure errors, suppresses the rate of the CRE-induced catastrophic events.

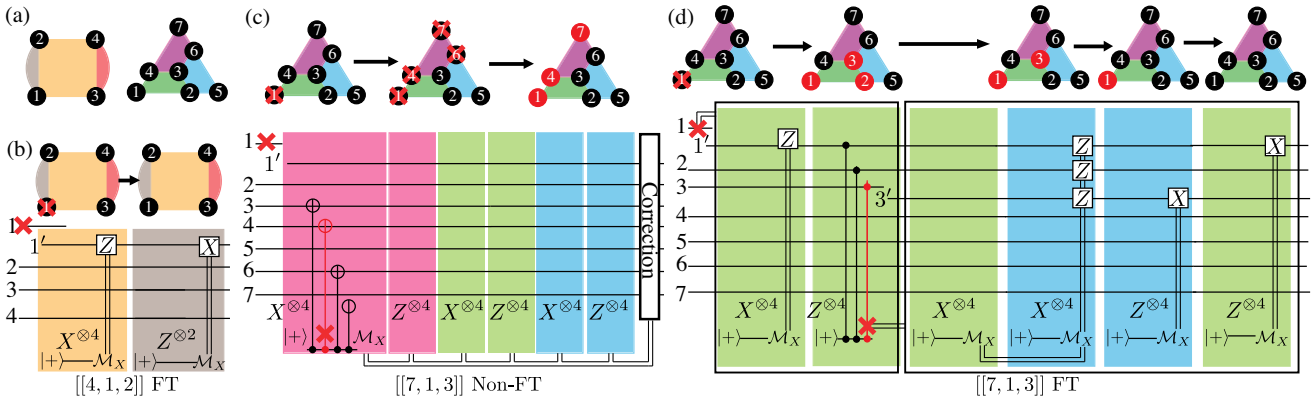

FIG. 2. (a) The illustration of the $[[4,1,2]]$ (left) and the $[[7,1,3]]$ (right) code. The colored plaquettes represent stabilizer generators that have supports on the surrounding vertices (data qubits). (b) The FT circuit for the $[[4,1,2]]$ code correcting one data erasure (red cross). Colored boxes represent an ancilla-assisted measurement of a stabilizer associated with a plaquette of the same color. The ancilla is initialized in $|+\rangle$, a sequence of CX/CZ gates between the ancilla and the data qubits are applied (not shown), and the ancilla is measured in the Pauli X basis. (c) The non-FT circuit for the $[[7,1,3]]$ code that is *nonadaptive*. An initial erasure error on a data qubit triggers the circuit, which measures all the six stabilizers in a fixed sequence (pink-green-blue) and applies the correction at the end. We explicitly show the CX gates in the first box ($X_3X_4X_6X_7$ stabilizer measurement) to illustrate an erasure error that propagates to multiple data errors and causes a logical failure. The top shows the evolution of the errors for the example trajectory. The red circles indicate qubits with potential Pauli errors (converted from the erasure errors). (d). The FT circuit for the $[[7,1,3]]$ corrects the errors *adaptively*. Suppose another erasure happens during the CZ gate (shown in red) between the ancilla and the third qubit while measuring $Z_1Z_2Z_3Z_4$. Upon detection of this error we stop the stabilizer measurement, discard and replace the ancilla and the third qubits and update the erasure-flag error set $\mathcal{E}$ to $\mathcal{E} = \{I, X_1\} \times \{I, P_3\} \times \{I, Z_1Z_2\}$, where P_3 indicates an arbitrary Pauli error on the third qubit. The correlated Z_1Z_2 error results from discarding the ancilla that is entangled with the first and the second qubits. We then measure the stabilizers $X_1X_2X_3X_4$, $X_2X_3X_5X_6$, $Z_2Z_3Z_5Z_6$, and $Z_1Z_2Z_3Z_4$ to correct possible errors within $\mathcal{E}$.

superconducting chips connected with microwave links (see, e.g., Refs. [35–38]), and provide logical-error estimates in state-of-the-art experimental systems. Our analysis indicates that under reasonable assumptions, we can suppress the damage from these catastrophic events to higher order and reduce the CRE-induced logical error rate from 1 every 10 s in Ref. [12] to less than 1 per month.

Setup.—We consider two levels of encoding on n chips. The first level uses an error correcting code (e.g., a surface code [39]) to protect the information in each chip. In the second level, we concatenate this code with a $[[n,1,d]]$ QEC code capable of correcting $d-1$ erasure errors [30], which is distributed across n separate chips. The operations in the first level should be protected by the first-level code. Therefore, operations in that level are followed by syndrome checks at every step. Upon a CRE impact on a specific chip, the qubits in a large region around the impact area experience a considerable reduction in their lifetime [12]. Consequently, most qubits in the impacted chip decay during an error correction cycle. The simultaneous decay of a large number qubits in a chip causes a significant increase in the number of error syndromes of the first-level encoding. The observation of a sharp jump in the number of error syndromes in a chip reveals the location of the erasure error in the second level [13,40], which subsequently triggers error correction in the second level. We expect that by correcting $d-1$ errors we would be able to suppress the rate of catastrophic events to $\propto \lambda(\lambda\tau)^{d-1}$, where λ is the CRE rate in a chip and τ is the time that it takes for the second-level error correction cycle [see Fig. 1(b)].

For example, we can use the $[[4,1,2]]$ code [30] to correct single erasure errors. As shown in Fig. 2(b), a single CRE event will trigger the QEC circuit to correct the erasure error and successfully restore the original encoding. However, if there is a second CRE erasure event during the erasure correction, the QEC circuit will fail to restore the encoded information, leading to a CRE-induced logical error rate proportional to $\lambda^2\tau$. Note that the QEC for $[[4,1,2]]$ is relatively simple because we only care about correcting single CRE errors and do not worry about CRE errors during the QEC operations. In order to use larger-distance codes, e.g., the $[[7,1,3]]$ code [41], to suppress the CRE errors to higher orders it is crucial to design the QEC circuit *fault tolerantly* so that possible CRE events during the QEC should not damage the encoded information.

Fault tolerant error correction for erasure errors.—We assume that by using sufficiently large surface codes in the first level, Pauli error rates due to the failure of the surface QEC are much lower than the rate of the CREs. As such, we only consider the errors induced by the CREs. For simplicity, we assume that a CRE-induced erasure error could propagate through a two-qubit gate and completely erase both involved qubits [42]. Upon detecting erasure errors on a chip, we replace the erased chip with a reserve chip. The data qubits on the new chip are randomly initialized. Hence, their erasure errors are converted to

detected Pauli errors randomly drawn from $\{I, X, Y, Z\}$ after the chip replacement. We propose a novel fault-tolerant QEC (FTQEC) scheme, called the erasure-flag scheme, that satisfies the fault-tolerant criteria [8,43,44] (see Supplemental Material [45]). The scheme *adaptively* performs nondestructive stabilizer measurements using one ancilla qubit on an ancilla chip [Fig. 1(a)]. A single erasure error that occurs on the ancilla could propagate into multiple data erasures on different data chips. However, we can detect the time and location of such bad errors. So similar to the flag FTQEC for Pauli errors [44,51,52], the access to this information enables us to design protocols that use minimal resources to tolerate the bad errors. Here, this information comes directly from the first-level QEC without requiring additional resources, e.g., flag qubits, in the second level.

The erasure-flag FTQEC protocol using a distance-d code is implemented as follows. (i) Upon detecting erasure errors on the data qubits, replace the erased data qubits (chips), initialize the erasure-flag error set $\mathcal{E}$ that contains the detected data errors, set $s = 0$ that counts the number of bad erasure errors that happen during the protocol and apply the following erasure-QEC. (ii) Measure a set of stabilizers of minimal size that can be used to correct the current $\mathcal{E}$. (A) If there are s_{new} bad erasures detected in the middle of a stabilizer measurement with $s + s_{\text{new}} \leq d - 1$, stop the measurement immediately, update s by adding s_{new}, replace the erased qubits (chips), update $\mathcal{E}$, and restart (ii). (B) Otherwise, apply a correction in $\mathcal{E}$ based on the measured syndromes.

The fault tolerance of the protocol is guaranteed by the following two key ingredients. (a) Bad erasures can be immediately detected so that we can keep track of the erasure-flag error set resulting from the bad errors. (b) The erasure-flag error set is correctable (different errors either have different syndromes or differ by a stabilizer) if there are fewer than d faults. We show that the erasure-flag scheme can be applied to the four-qubit and seven-qubit codes using proper QEC circuits (Fig. 2). The FT circuit for the $[[4, 1, 2]]$ code [Fig. 2(b)] corrects a single data erasure at the input. A non-FT circuit for the $[[7, 1, 3]]$ code [Fig. 2(c)] is triggered by a data erasure error at the input and *nonadaptively* measures a full set of stabilizers in a fixed sequence. However, an extra erasure that occurs on the ancilla chip during a stabilizer measurement could propagate into multiple data errors and cause a logical failure. Therefore, the non-FT circuit fails to correct some consecutive double erasures. In contrast, the *adaptive* FT circuit [Fig. 2(d)], which keeps track of the possible error set and measures only a minimal set of stabilizers, can tolerate up to two consecutive erasures on arbitrary qubits. We note that the erasure-flag scheme is not guaranteed to work for any stabilizer code and it is mostly suitable for small codes or codes with certain structures, e.g., the topological surface codes with arbitrary distance (see Supplemental Material [45]).

Analysis of the erasure error rates.—Following Ref. [12], we model CREs on each chip by an independent and uncorrelated Poisson process $N(t)$, such that $P[N(t) = k] = (\lambda t)^k/k! \exp(-\lambda t)$, where λ is the rate of the events [53]. The value of λ depends on the geometry and other specifications of the chip [17], but for simplicity we use the reported value of $1/\lambda = 10$ s in Ref. [12]. Since the events in each chip are assumed to be independent, the introduction of additional chips increases the overall CRE rate linearly. Using the FT implementation of a QEC code that corrects $d - 1$ erasure errors in a cycle, a catastrophic event might occur if there are more than $d - 2$ additional events during the recovery time τ following the first event that triggers error correction. Such a catastrophic event leads to a logical failure at the second level of encoding. The rate of these catastrophic events is obtained by taking the product of the rate of the CREs that trigger error correction and the probability that more than $d - 2$ CREs happen in time τ following the first CRE. For a code over n chips, the former is $n\lambda$. However, since we need an ancilla chip for our QEC scheme, the latter factor should be calculated using the rate $(n + 1)\lambda$ (the reserve chips do not contribute to the logical error rates since they do not carry any logical information). Therefore, we find the rate of the catastrophic events, $\Gamma = n\lambda\{1 - \exp[-(n + 1)\lambda\tau] \sum_{k=0}^{d-2}[(n + 1)\lambda\tau]^k/k!\}$. For $n\lambda\tau \ll 1$, we have $\Gamma \approx n\lambda[(n + 1)\lambda\tau]^{d-1}/(d - 1)!$, which shows the desired error suppression in this regime. While we considered the worst-case scenario, not all weight-d (or higher) errors are catastrophic, and some are still correctable. Therefore, by considering the longest error correction and recovery time for τ [Figs. 2(b) and 2(d)], this analysis gives a lower bound on the memory lifetime, Γ^{-1}, limited by the $d - 1$ coincident CREs within τ (solid lines in Fig. 3).

In contrast, for the non-FT implementation of the $[[7, 1, 3]]$ code, we obtain an upper bound on the memory lifetime (dashed line in Fig. 3). In this case, some double events cause a logical failure. For an upper bound, we only consider the case where the first erasure error occurs on the edge chips in Fig. 2(a). Following this event, depending on

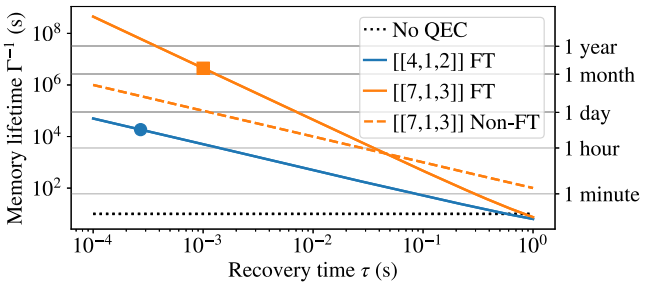

FIG. 3. The solid (dashed) lines show the lower (upper) bound of the lifetime with (without) the fault-tolerant implementation. The dotted line shows the expected lifetime λ^{-1} without error correction. The circle ($[[4, 1, 2]]$) and square ($[[7, 1, 3]]$) markers show estimates of the improved lifetime with error correction for maximal recovery time using experimentally feasible parameters.

the affected chip, there are one or two stabilizer measurements during which an ancilla erasure can lead to logical failure. Therefore, we consider CREs on these six edge chips with the rate 6λ as triggering events and find the probability of an additional event on an ancilla during one of the stabilizer measurements. Since different stabilizer measurements [colored boxes in Fig. 2(c)] have the same number of interchip gates, we assume that they each take $\tau/6$. Therefore, we find the upper bound of $1/\{6\lambda[1-\exp(-\lambda\tau/6)]\}$ for Γ^{-1} (Fig. 3). In Supplemental Material [45] we numerically verify that the logical error rates are suppressed to the desired order for both the FT and non-FT schemes. Furthermore, we show improved theoretical estimates for Γ^{-1} under a specific model of the hardware operations.

Experimental implementation.—Our proposed scheme can be implemented experimentally in superconducting devices by coupling multiple data chips to an ancilla chip through a router [54,55] [Fig. 1(a)]. The ancilla chip is used to collect the syndrome information by coupling a syndrome patch to the data patches (all encoded in a surface code) associated with different stabilizers. The detail of coupling of the ancilla chip with one of the data chips is illustrated in Fig. 4. To implement an entangling gate, e.g., CX, between the syndrome patch S and a data patch D, we introduce an ancilla patch A on the ancilla chip and apply the measurement-based gate [56] (see the inset of Fig. 4). The measurement of joint Pauli operators ZZ (XX) between the surface patches A and S (D) is implemented by lattice surgery [56], i.e., merging and then splitting the Z (X) boundaries of the two involved patches. In our distributed architecture, we need to nonlocally merge the boundaries of the A and D patches that sit on different chips by adding new plaquettes (dashed boxes in Fig. 4) that connect the boundaries. Each of the new plaquettes has two ancilla qubits (black dots in Fig. 4), each sitting on one chip and is locally coupled to two data qubits on the boundary of the surface patches. To measure a new plaquette stabilizer, we apply a nonlocal CX gate between the two ancilla qubits to create a Bell state $(1/\sqrt{2})(|00\rangle+|11\rangle)$, then apply two CX/CZ gates between the ancilla qubits and their coupled data qubits, then we apply another nonlocal CX gate between the ancillas and finally measure one of the ancillas. The nonlocal physical CX gate between the ancilla qubits can be implemented by teleportation-based gates that use preshared and purified bell pairs between two chips [57–60].

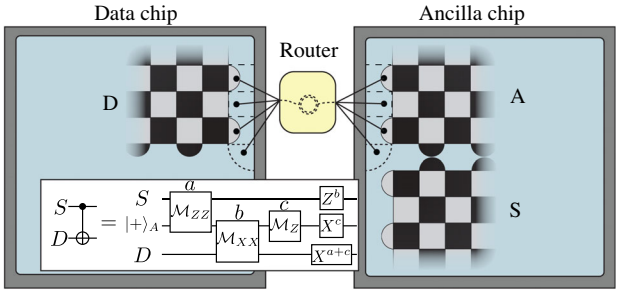

FIG. 4. Detail of the coupling of the ancilla chip in Fig. 1(a) to a data chip. We show three surface patches S (syndrome), A (ancilla), and D (data). Each patch is encoded in a rotated surface code, with data qubits on the vertices and X-type (Z-type) syndrome qubits on the black (white) plaquettes. To extract error syndromes, a CX gate between S and D patches is implemented by introducing the A patch and applying the measurement-based circuit shown in the inset.

We can estimate the length of the outer QEC cycle and the corresponding upper bound of the logical error rate based on realistic experimental parameters in the superconducting architecture. The most time-consuming physical operations are the two-qubit gates (∼100 ns [61,62]), measurements (∼200 ns [63]) and interchip state transfers (∼100 ns [37,38]). We assume that each surface patch is a 10×10 rotated surface code and each surface-level operation is followed by ten repeated rounds of surface QEC. For maximum parallelism for all the operations, we estimate that the maximum recovery time τ for the $[[4,1,2]]$ ($[[7,1,3]]$) code correcting 1 (2) erasure errors is approximately 270 μs (1000 μs). See Supplemental Material [45] for details. Based on these estimated recovery times, we obtain a lower bound of the memory lifetimes of approximately 5 hours using the four-qubit, and 51 days for seven-qubit codes (markers in Fig. 3).

Discussion.—In principle, our scheme can be extended to universal fault-tolerant computing. Multiple logical qubits in the second-level code can be encoded transversely across all the data chips such that different base qubits comprising each logical qubit sit on different chips and each chip contains exactly one base qubit from each logical qubit. Such an encoding is fault tolerant since each CRE only erases one base qubit in each logical qubit, which can be corrected by performing QEC on different logical qubits independently. The implementation of universal fault-tolerant gates can be adapted from existing protocols [64]. Furthermore, the resource overhead required for overcoming the CREs could be less than that required for the standard depolarizing noise. For example, we can prepare the magic states non-fault-tolerantly and verify them by performing erasure detection, without applying costly magic-state distillation [64,65]. Additionally, we can use Knill-type QEC [31] that utilizes teleportation to correct erasure errors. We explore and compare that approach to ours in Supplemental Material [45].

Last, we note that for now the second layer of QEC only suppresses the logical error rate due to the erasure errors. In the regime where the Pauli errors due to the failure of the surface codes are more detrimental, it is advantageous to minimize the total logical error rate by tailoring the outer codes to correct both the erasure and the Pauli errors. For instance, Ref. [34] shows that the surface code can handle a mixture of erasure and Pauli errors using a tailored decoder. We leave this topic to future work.

We thank Vignesh Raman for helpful discussions. We acknowledge support from the ARO (W911NF-18-1-0020, W911NF-18-1-0212), ARO MURI (W911NF-16-1-0349, W911NF-21-1-0325), AFOSR MURI (FA9550-19-1-0399, FA9550-21-1-0209), AFRL (FA8649-21-P-0781), DoE Q-NEXT, NSF (OMA-1936118, EEC-1941583, OMA-2137642), NTT Research, and the Packard Foundation (2020-71479). A. S. is supported by a Chicago Prize Postdoctoral Fellowship in Theoretical Quantum Science. B. O. S. and R. V. acknowledge the support from JST Moonshot R&D Grant (No. JPMJMS2061).

*These authors contributed equally to this work.
†liang.jiang@uchicago.edu

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, England, 2010).
[2] D. A. Lidar and T. A. Brun, *Quantum Error Correction* (Cambridge University Press, Cambridge, England, 2013).
[3] J. P. Clemens, S. Siddiqui, and J. Gea-Banacloche, Phys. Rev. A **69**, 062313 (2004).
[4] N. H. Nickerson and B. J. Brown, Quantum **3**, 131 (2019).
[5] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **65**, 062101 (2002).
[6] D. V. Averin and R. Fazio, J. Exp. Theor. Phys. Lett. **78**, 664 (2003).
[7] R. Klesse and S. Frank, Phys. Rev. Lett. **95**, 230503 (2005).
[8] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Inf. Comput. **6**, 97 (2006).
[9] D. Aharonov, A. Kitaev, and J. Preskill, Phys. Rev. Lett. **96**, 050504 (2006).
[10] A. P. Vepsäläinen, A. H. Karamlou, J. L. Orrell, A. S. Dogra, B. Loer, F. Vasconcelos, D. K. Kim, A. J. Melville, B. M. Niedzielski, J. L. Yoder *et al.*, Nature (London) **584**, 551 (2020).
[11] L. Cardani, F. Valenti, N. Casali, G. Catelani, T. Charpentier, M. Clemenza, I. Colantoni, A. Cruciani, G. D'Imperio, L. Gironi *et al.*, Nat. Commun. **12**, 2733 (2021).
[12] M. McEwen, L. Faoro, K. Arya, A. Dunsworth, T. Huang, S. Kim, B. Burkett, A. Fowler, F. Arute, J. C. Bardin *et al.*, Nat. Phys. **18**, 107 (2022).
[13] R. Acharya, I. Aleiner, R. Allen, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, J. Atalaya, R. Babbush *et al.*, arXiv:2207.06431.
[14] C. Gidney and M. Ekerå, Quantum **5**, 433 (2021).
[15] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. L. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, Rev. Mod. Phys. **85**, 961 (2013).
[16] D. Rainis and D. Loss, Phys. Rev. B **85**, 174533 (2012).
[17] J. M. Martinis, npj Quantum Inf. **7**, 90 (2021).
[18] I. Nsanzineza and B. L. T. Plourde, Phys. Rev. Lett. **113**, 117002 (2014).
[19] U. Patel, I. V. Pechenezhskiy, B. L. T. Plourde, M. G. Vavilov, and R. McDermott, Phys. Rev. B **96**, 220501(R) (2017).
[20] F. Henriques, F. Valenti, T. Charpentier, M. Lagoin, C. Gouriou, M. Martínez, L. Cardani, M. Vignati, L. Grünhaupt, D. Gusenkova *et al.*, Appl. Phys. Lett. **115**, 212601 (2019).
[21] L. M. Duan, B. B. Blinov, D. L. Moehring, and C. Monroe, Quantum Inf. Comput. **4**, 165 (2004).
[22] L. Jiang, J. M. Taylor, A. S. Sørensen, and M. D. Lukin, Phys. Rev. A **76**, 062323 (2007).
[23] J. Kim and C. Kim, Quantum Inf. Comput. **9** (2009).
[24] Y. L. Lim, S. D. Barrett, A. Beige, P. Kok, and L. C. Kwek, Phys. Rev. A **73**, 012304 (2006).
[25] N. H. Nickerson, Y. Li, and S. C. Benjamin, Nat. Commun. **4**, 1756 (2013).
[26] D. K. L. Oi, S. J. Devitt, and L. C. L. Hollenberg, Phys. Rev. A **74**, 052313 (2006).
[27] R. Van Meter, T. D. Ladd, A. G. Fowler, and Y. Yamamoto, Int. J. Quantum. Inform. **08**, 295 (2010).
[28] R. Van Meter and S. Devitt, IEEE Comput. **49**, 31 (2016).
[29] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L. M. Duan, and J. Kim, Phys. Rev. A **89**, 022317 (2014).
[30] M. Grassl, T. Beth, and T. Pellizzari, Phys. Rev. A **56**, 33 (1997).
[31] E. Knill, Phys. Rev. A **71**, 042322 (2005).
[32] E. Knill, R. Laflamme, and G. Milburn, arXiv:quant-ph/0006120.
[33] M. Silva, arXiv:quant-ph/0405112.
[34] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, Nat. Commun. **13**, 4657 (2022).
[35] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Nature (London) **558**, 264 (2018).
[36] C. J. Axline, L. D. Burkhart, W. Pfaff, M. Zhang, K. Chou, P. Campagne-Ibarcq, P. Reinhold, L. Frunzio, S. M. Girvin, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Nat. Phys. **14**, 705 (2018).
[37] Y. Zhong, H.-S. Chang, A. Bienfait, É. Dumur, M.-H. Chou, C. R. Conner, J. Grebel, R. G. Povey, H. Yan, D. I. Schuster, and A. N. Cleland, Nature (London) **590**, 571 (2021).
[38] A. Gold, J. P. Paquette, A. Stockklauser, M. J. Reagor, M. S. Alam, A. Bestwick, N. Didier, A. Nersisyan, F. Oruc, A. Razavi, B. Scharmann *et al.*, npj Quantum Inf. **7**, 142 (2021).
[39] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).
[40] Z. Chen, K. J. Satzinger, J. Atalaya, A. N. Korotkov, A. Dunsworth, D. Sank, C. Quintana, M. McEwen, R. Barends, P. V. Klimov *et al.*, Nature (London) **595**, 600 (2021).
[41] A. Steane, Proc. R. Soc. A **452**, 2551 (1996).
[42] In practice, the propagation of erasure errors at the surface-code level through the gates depends on both the microscopic detail of the erasures and the implementation of the gates. Here, we consider the worst-case scenario, where a single erasure can completely erase all the qubits involved in a multiqubit gate.
[43] D. Gottesman, An introduction to quantum error correction and fault-tolerant quantum computation, arXiv:0904.2557.
[44] C. Chamberland and M. E. Beverland, Quantum **2**, 53 (2018).
[45] See Supplemental Material http://link.aps.org/supplemental/10.1103/PhysRevLett.129.240502 for numerical simulations of different erasure-QEC schemes, a comparison between the erasure-flag scheme and the Knill-type QEC

scheme, and a discussion on the application of the erasure-flag scheme to more codes and the system performance implications, which includes Refs. [46–50].

[46] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. (N.Y.) **43**, 4452 (2002).

[47] A. Y. Kitaev, Ann. Phys. (Amsterdam) **303**, 2 (2003).

[48] R. D. Van Meter III, Architecture of a quantum multicomputer optimized for Shor's factoring algorithm, Ph.D. thesis, Keio University, 2006.

[49] M. Ahsan, R. Van Meter, and J. Kim, ACM J. Emerging Technol. Comput. Syst. **12**, 39 (2015).

[50] M. Y. Siraichi, V. F. d. Santos, S. Collange, and F. M. Q. Pereira, in *Proceedings of the 2018 International Symposium on Code Generation and Optimization, CGO 2018* (ACM, New York, NY, 2018), pp. 113–125.

[51] R. Chao and B. W. Reichardt, Phys. Rev. Lett. **121**, 050502 (2018).

[52] R. Chao and B. W. Reichardt, npj Quantum Inf. **4**, 42 (2018).

[53] In principle, there might be spatial correlations between these events, but they can be treated independently as long as the correlation length is finite and the chips are separated beyond that length.

[54] H.-S. Chang, K. J. Satzinger, Y. P. Zhong, A. Bienfait, M.-H. Chou, C. R. Conner, É. Dumur, J. Grebel, G. A. Peairs, R. G. Povey, and A. N. Cleland, Appl. Phys. Lett. **117**, 244001 (2020).

[55] C. Zhou, P. Lu, M. Praquin, T.-C. Chien, R. Kaufman, X. Cao, M. Xia, R. Mong, W. Pfaff, D. Pekker, and M. Hatridge, arXiv:2109.06848.

[56] C. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, New J. Phys. **14**, 123011 (2012).

[57] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).

[58] Y. Wan, D. Kienzler, S. D. Erickson, K. H. Mayer, T. R. Tan, J. J. Wu, H. M. Vasconcelos, S. Glancy, E. Knill, D. J. Wineland, A. C. Wilson, and D. Leibfried, Science **364**, 875 (2019).

[59] K. S. Chou, J. Z. Blumoff, C. S. Wang, P. C. Reinhold, C. J. Axline, Y. Y. Gao, L. Frunzio, M. Devoret, L. Jiang, and R. Schoelkopf, Nature (London) **561**, 368 (2018).

[60] H. Yan, Y. Zhong, H.-S. Chang, A. Bienfait, M.-H. Chou, C. R. Conner, E. Dumur, J. Grebel, R. G. Povey, and A. N. Cleland, Phys. Rev. Lett. **128**, 080504 (2022).

[61] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Nature (London) **574**, 505 (2019).

[62] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan *et al.*, Phys. Rev. Lett. **127**, 180501 (2021).

[63] J. Heinsoo, C. K. Andersen, A. Remm, S. Krinner, T. Walter, Y. Salathé, S. Gasparinetti, J.-C. Besse, A. Potočnik, A. Wallraff, and C. Eichler, Phys. Rev. Appl. **10**, 034040 (2018).

[64] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).

[65] E. Knill, arXiv:quant-ph/0402171.