

Fault tolerance against amplitude-damping noise using Bacon-Shor codes

Long D. H. My^{1,2,*}, Akshaya Jayashankar^{3,*}, Prabha Mandayam^{4,5} and Hui Khoon Ng^{1,2,†}

¹*Yale-NUS College, Singapore*

²*Centre for Quantum Technologies, National University of Singapore, Singapore*

³*Centre for Quantum Engineering, Research and Education, TCG CREST, Sector V, Salt Lake, Kolkata 700091, India*

⁴*Department of Physics, Indian Institute of Technology Madras, Chennai 600036, India*

⁵*Center for Quantum Information, Computing and Communication, Indian Institute of Technology Madras, Chennai 600036, India*



(Received 25 April 2025; accepted 20 August 2025; published 23 September 2025)

Designing efficient fault-tolerance schemes is crucial for building useful quantum computers. Most standard schemes assume no knowledge of the underlying device noise and rely on general-purpose quantum error-correcting (QEC) codes capable of handling arbitrary errors. Biased-noise alternatives focus on only correcting a subset of some generic error basis (e.g., Pauli error basis) and lower resource needs by channeling the redundancy to dealing only with that subset. Yet, the most resource-efficient codes are expected to be those that directly target the specific noise process that afflicts the quantum device, rather than using a generic error-basis description. However, the question of whether such noise-adapted QEC protocols are amenable to fault-tolerant implementations remains largely unexplored. Here, we design a fault-tolerance scheme based on the Bacon-Shor codes that can protect against amplitude-damping noise in the device. We construct a universal set of logical gadgets tolerant to multiple damping errors and estimate the fault-tolerance threshold of our scheme. Our work thus establishes the possibility of achieving fault tolerance against amplitude-damping noise using noise-adapted quantum codes, while highlighting some of the unique challenges that arise in this context.

DOI: [10.1103/PhysRevResearch.7.033275](https://doi.org/10.1103/PhysRevResearch.7.033275)

I. INTRODUCTION

Quantum fault tolerance (FT) [1–5] provides the basic underpinnings of today’s global efforts toward realizing robust and scalable quantum computers. The framework of FT builds upon the theory of quantum error correction (QEC) [6–9] and offers a pathway to reliable quantum computation in the presence of noisy memory and gates. In a typical FT scheme, one chooses a QEC code to encode the computational—or “logical”—information into the physical qubits. Encoded operations are performed on the logical qubits, interleaved with periodic error correction to prevent catastrophic accumulation and spread of errors. FT theory offers a suite of techniques to construct noise-resilient logical quantum circuits from noisy physical gates and imperfect error-correction operations, provided the physical noise is below a certain threshold level.

The traditional approach to fault tolerance assumes no knowledge—apart from mild assumptions expected to be broadly applicable—of the underlying noise in the device and relies on general-purpose codes capable of correcting arbitrary errors. This requirement for robustness against arbitrary noise aligns well with the general philosophy of fault tolerance: We

ward against imprecision in our knowledge of the noise, by not relying on any such information. Nevertheless, this desire to accommodate arbitrary errors comes at the cost of requiring code choices that can be more resource demanding. For example, the smallest qubit code capable of correcting arbitrary errors on any one physical qubit demands five physical qubits [10]; instead, three physical qubits suffice if we assume that there is only bit-flip or only phase-flip noise, or four physical qubits if we assume only amplitude-damping noise [11].

Potential resource savings by incorporating properties of the noise are expected to be important for early fault-tolerance experiments when qubit numbers and gate fidelities are limited. Many early-stage error-correction demonstrations focus on correcting only say bit-flip or phase-flip noise (see, e.g., among many others, Refs. [12–14]), because they have only very few qubits or can only do a few gates, too few for general-purpose codes. While some of the recent experiments have taken impressive steps toward fault-tolerant operation [15,16], for scaling to large-scale computers, *noise-adapted* optimizations remain important. General-purpose codes are most efficient when the noise genuinely has no particular asymmetry or special properties, so that one is not “wasting” the redundancy introduced to protect against noise in all possible ways. Yet, real device noise is never truly symmetric, and one could seek noise-adapted fault-tolerance designs to first reduce those noise asymmetries by correcting the dominant noise types, before concatenating—at a higher effective level—with a general-purpose code to protect against residual, now more symmetric, noise.

Previous noise-adapted schemes have largely approached the problem from the following view of asymmetry in the

*These authors contributed equally to this work.

†Contact author: huikhooon.ng@nus.edu.sg

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

noise: We begin with an error basis—one that describes all possible errors—and declare that only a subset of the errors in that basis is dominant, and focus on correcting only those errors. This is the case for the biased Pauli noise models, where one begins with the Pauli error basis and chooses to focus on only X or Z errors. This motivated the many previous discussions on repetition codes [17–19] as well as rectangular surface codes for biased Pauli noise (see, for example, Refs. [20,21]). Similarly, in bosonic cat codes situations, where errors are decomposed as boson loss or phase errors, one chooses to focus on either loss or phase, depending on the dominant (sometimes engineered) noise process [22–26].

Starting from an error basis and then choosing to drop certain types of errors has a strong advantage when designing FT schemes. One could start from standard general-purpose FT schemes and adopt those designs in the noise-adapted situation, by simply pruning out parts of the FT circuits that deal with the dropped error types. This is particularly obvious for Calderbank-Shor-Steane-type codes with biased Pauli noise—one drops all the circuit components that implement the X -type checks (which deal with Z errors) when correcting only X errors and not Z errors. This, however, begs the following question: Rather than pruning existing FT designs, can we start from only the actual noise description—as a quantum process deduced from device characterization—and construct a fully fault-tolerant computational scheme?

We answered this question in part in our earlier work [27], where we constructed a universal set of fault-tolerant logical operations as well as a fault-tolerant error-correction gadget starting from the assumption that the noise is described by the amplitude-damping channel. Amplitude-damping noise cannot be phrased as biased noise in terms of the Pauli error basis and is the prototypical example that many studies turn to when hoping to examine beyond-Pauli situations. There are thus well-studied codes available that focus on correcting amplitude-damping noise [11,28–30]. Fault-tolerance studies of such codes were lacking, and our earlier work in Ref. [27] offered a first full construction of fault-tolerant gadgets for error correction and universal computation. There, we also pointed out the differences in circuit designs when discussing fault tolerance for such channel-adapted situations, compared to standard-Pauli or biased-Pauli scenarios. For example, that transversal constructions are always fault tolerant in Pauli settings no longer holds true in channel-adapted situations.

In our present work, we extend and complete our answer for amplitude-damping noise by showing that we can scale up our fault-tolerant designs to larger codes. Our earlier work [27] examined only the case of the smallest code, which corrects a single amplitude-damping error in any of the physical qubits within a single gadget. Fault-tolerant schemes come also with a prescription of how to scale up the protection, by using more physical qubits and operations, to correct more errors and hence compute more accurately. While we expect, in real-use cases, to eventually move to general-purpose codes to continue scaling up, the asymmetry in the noise may be strong enough to require first increasing protection against the amplitude-damping noise. We address this aspect here and develop a general FT scheme for amplitude-damping noise based on the two-dimensional Bacon-Shor code [31]. Specifically, we consider a Bacon-Shor code defined on a

square $(t + 1) \times (t + 1)$ lattice of physical qubits with a fixed gauge. By fixing the gauge, the Bacon-Shor code becomes a subspace code rather than its original subsystem formulation. In this form, the fixed-gauge Bacon-Shor code can correct amplitude-damping errors affecting up to t physical qubits [32]. The smallest instance, of the 2×2 lattice ($t = 1$), corresponds to the four-qubit code used as the basis of our earlier work [27]. Here, we explain how to adapt our earlier designs for $t > 1$ and construct a fault-tolerant error-correction gadget as well as a full set of universal logical operations. We demonstrate the scaling performance by estimating the logical infidelity of quantum memory and showing that the infidelity falls—below specific threshold levels—as the code size grows.

The rest of the paper is organized as follows: We start, in Sec. II, by reviewing the amplitude-damping noise as well as the $(t + 1) \times (t + 1)$ Bacon-Shor code. In Sec. III, we describe our construction of a fault-tolerant error-correction gadget for the Bacon-Shor code for general t . In Sec. IV, we focus on our construction of the logical controlled- Z (CZ) gadget, leaving the details of other logical gadgets to the Appendix (Appendices E and F) as they generalize easily from our earlier work. Finally, we estimate the logical fidelity and pseudotreshold for the quantum memory task using our scheme in Sec V and conclude in Sec. VI.

II. PRELIMINARIES

We begin by reviewing the assumed noise model, the Bacon-Shor codes, and the ideal (no-noise) error-correction procedure, in preparation for the description of our fault-tolerant construction.

A. Noise model

We assume amplitude damping as the dominant noise affecting every physical qubit in our quantum computer. Physical processes like spontaneous decay give rise to amplitude-damping noise, and this is the leading source of noise in some quantum computing platforms (see, e.g., Ref. [33]). Amplitude-damping noise is a single-qubit channel described by the completely positive and trace-preserving map,

$$\begin{aligned}\mathcal{E}_{\text{AD}}(\cdot) &= E_0(\cdot)E_0^\dagger + E_1(\cdot)E_1^\dagger, \\ E_0 &\equiv \frac{1}{2}[(1 + \sqrt{1-p})I + (1 - \sqrt{1-p})Z] \\ &= |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|, \\ E_1 &\equiv \frac{1}{2}\sqrt{p}(X + iY) = \sqrt{p}|0\rangle\langle 1| \equiv \sqrt{p}E.\end{aligned}\quad (1)$$

Here, I is the single-qubit identity, X , Y , and Z are the Pauli operators, and $\{|0\rangle, |1\rangle\}$ is the Z eigenbasis. $p \in [0, 1]$ is the damping parameter, assumed to be small for the weak noise regime needed for quantum computing. Below, we write $E \equiv |0\rangle\langle 1|$ for the damping operator.

Similar to the FT scheme of Ref. [27], we use the following unencoded (physical) operations to build the FT encoded gadgets:

$$\{\mathcal{P}_{|+\rangle}, \mathcal{P}_{|0\rangle}\} \cup \{\mathcal{M}_X, \mathcal{M}_Z, \text{CNOT}, \text{CZ}, X, Z, S, T\}. \quad (2)$$

Here, $\mathcal{P}_{|+\rangle}$ and $\mathcal{P}_{|0\rangle}$ refer to the preparation of the eigenstates of X and Z , respectively, and $\mathcal{M}_X$ and $\mathcal{M}_Z$ are measurements in the X and Z bases, respectively. CNOT $\equiv |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X$ is the two-qubit controlled-NOT gate, CZ $\equiv |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes Z$ is the two-qubit controlled-Z gate, $S \equiv |0\rangle\langle 0| + i|1\rangle\langle 1|$ is the phase gate, and $T \equiv |0\rangle\langle 0| + e^{-i\pi/4}|1\rangle\langle 1|$ is the $\pi/8$ gate.

We assume that errors in the state preparations, gates, measurements, and idling qubits are all due to $\mathcal{E}_{AD}$. Specifically, a noisy physical gate $\mathcal{G}$ is modeled as the ideal gate followed by the noise $\mathcal{E}_{AD}$ acting independently on each qubit participating in the gate. A noisy measurement is modeled as an ideal measurement preceded by the noise $\mathcal{E}_{AD}$, while a noisy preparation is an ideal preparation followed by $\mathcal{E}_{AD}$. We emphasize that the noise acts on each physical qubit individually and is assumed to be time and gate independent. More generally, one can regard the parameter p as an upper bound on the level of amplitude damping over time and gate variations.

In our discussion below, it is useful to organize the effect of $\mathcal{E}_{AD}$ in powers of p . To that end, we write

$$\mathcal{E}_{AD}(\cdot) = \frac{1}{4}(1 + \sqrt{1-p})^2 \mathcal{I}(\cdot) + p\mathcal{F}_a(\cdot) + \frac{p}{2}\mathcal{F}_z(\cdot) + \left[\frac{p^2}{16} + O(p^3)\right]Z(\cdot)Z, \quad (3)$$

where $\mathcal{I}(\cdot) \equiv (\cdot)$ is the identity channel, $\mathcal{F}_a(\cdot) \equiv E(\cdot)E^\dagger$ is the *damping error*, and $\mathcal{F}_z(\cdot) \equiv (1/2)[(\cdot)Z + Z(\cdot)]$ is referred to as the *off-diagonal error*. The latter is sometimes referred to as the backaction error in the context of superconducting qubits [34]. Written in this manner, $\mathcal{E}_{AD}$ can be thought of as leading to no error (or order 1) when the $\mathcal{I}$ part occurs, a first-order (in p) error when the $\mathcal{F}_a$ or $\mathcal{F}_z$ part occurs and a second-order error when a Z error occurs.

B. Bacon-Shor codes

Bacon-Shor codes are a class of stabilizer codes originally proposed in the context of *subsystem* quantum error correction [31,35]. Subsequently, it was shown that the class of Bacon-Shor codes achieves a higher fault-tolerance threshold than subspace stabilizer codes while lending itself to simpler fault-tolerant implementations that require only nearest-neighbor two-qubit measurements [36]. The $d \times d$ Bacon-Shor code is a distance- d CSS code encoding one logical qubit in d^2 physical qubits. For small lattice sizes, the Bacon-Shor code was shown to yield favorable pseudothresholds. Recently, the $[[4, 1, 2]]$ Bacon-Shor code was shown to achieve a pseudothreshold between 3% and 6% in the presence of depolarizing noise [37]. However, increasing the size of the code does not improve the logical rate indefinitely; rather, it reverses beyond a point. The code thus does not have a quantum accuracy threshold (i.e., the physical noise level below which increasing the code size reduces the logical error without limits), a problem that can be addressed to some extent by lattice-surgery techniques [38], by adjusting the measurement schedule [39], or by concatenation [36]. Concatenation, in particular, was shown to lead to an FT threshold of $\sim 10^{-4}$ against stochastic adversarial noise [36]. We will see this absence of a true FT threshold also in our construction, but this

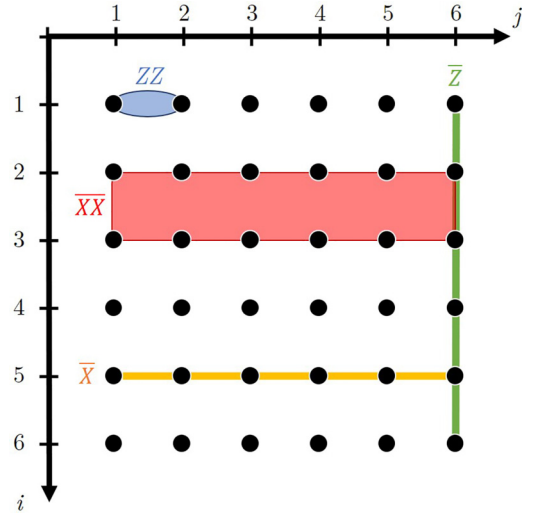


FIG. 1. A 6×6 Bacon-Shor code in the Z gauge. A Z -type stabilizer (marked ZZ in blue), an X -type stabilizer (marked XX in red), the logical X operator (marked $\bar{X}$ in orange), and the logical Z operator (marked $\bar{Z}$ in green) are shown.

reversal point occurs at a large code distance, likely beyond the regime we expect to use such noise-adapted codes (see Sec. V).

It was first noted in Ref. [40] that the $(t+1) \times (t+1)$ Bacon-Shor code can correct t amplitude-damping errors. Note that a code is said to correct t amplitude-damping errors if it is able to reduce the infidelity of the encoded state, to order $O(p^{t+1})$ after the error correction. Subsequently, explicit error-correction schemes were constructed for the Bacon-Shor code in Ref. [32].

Here, we follow the proposal of Refs. [32,40], which employs the Bacon-Shor code with the choice of the Z gauge (see Fig. 1). The two-dimensional $n \times n$ Bacon-Shor code in the Z gauge is a subspace stabilizer code that encodes one qubit into n^2 qubits arranged in an $n \times n$ square lattice. The stabilizer group of the code is generated by the operators $Z_{i,j}Z_{i,j+1}$ [$1 \leq i \leq n$, $1 \leq j \leq n-1$, with i (j) as the row (column) index], acting on neighboring qubits in the same row, and $X_{i,\star}X_{i+1,\star}$ ($1 \leq i \leq n-1$), acting on all the qubits in two neighboring rows ($\star$ denotes all values of that index slot). This code can be thought of as the concatenation of a length- n repetition code (within each row) that protects against X errors, with a length- n repetition code protecting against Z errors (across the rows). The code space $\mathcal{C}$ is thus the span of the two states,

$$|\pm\rangle_L \equiv \left[\frac{1}{\sqrt{2}}(|0\rangle_{\text{row}} \pm |1\rangle_{\text{row}})\right]^{\otimes n}, \quad (4)$$

where $|0\rangle_{\text{row}} \equiv |0\rangle^{\otimes n}$ and $|1\rangle_{\text{row}} \equiv |1\rangle^{\otimes n}$ are the length- n repetition code codewords for the qubits in a single row; the outer tensor product in Eq. (4) goes over the rows. The logical operators are identified as $\bar{X} \equiv X_{i,\star}$ and $\bar{Z} \equiv Z_{\star,j}$, denoting, respectively, X operators acting on all the qubits in any one row of the lattice and Z operators acting on all the qubits in any one column of the lattice. This gives us also a useful shorthand: We write the $X_{i,\star}X_{i+1,\star}$ stabilizers simply as $\bar{X}\bar{X}$,

putting in an i index only when referring to a particular pair of rows.

The fact that the $n \times n$ fixed-gauge Bacon-Shor code can correct amplitude-damping errors on $t \equiv n - 1$ qubits can be seen by checking that the code and the amplitude-damping error set satisfy the Knill-Laflamme error-correction conditions [8] up to order $O(p')$ [40,41]. In fact, the well-known four-qubit code [11] that corrects a single amplitude-damping error is a 2×2 fixed-gauge Bacon-Shor code. We note in contrast that the $n \times n$ subsystem Bacon-Shor code can correct *arbitrary* errors on only $\sim n/2$ qubits. The choice of gauge is to facilitate the detection of amplitude-damping errors. In particular, the damping operator E_1 is detected by ZZ stabilizers, while the Z error arising from E_0 is detected by XX stabilizers. This procedure is explained in detail in the following section and in Appendix A.

C. Ideal error correction (EC) procedure

An EC procedure to correct amplitude-damping errors with Bacon-Shor codes was constructed in Ref. [32], employing only Pauli measurements and gates. This EC procedure is not fault tolerant and functions correctly only when there are no faults in the EC circuit itself. We thus refer to this as the “ideal” EC procedure. We note that this ideal EC procedure may be suboptimal (see, e.g., Ref. [29] for a numerically optimized recovery) in that it reduces the infidelity of the output state to the requisite $O(p'^{+1})$ but it need not give the minimal infidelity among all recovery maps. Its simple structure, however, provides the starting point for our fault-tolerant version below. We thus briefly review the components of the ideal EC scheme here; more details are given in Appendix A.

The ideal EC procedure involves three steps to deal with the $\mathcal{F}_a$, $\mathcal{F}_z$, and Z errors [see Eq. (3)] that can occur from amplitude-damping noise:

(1) *Damping syndrome extraction*: We measure all $Z_{i,j}Z_{i,j+1}$ stabilizers to deduce which rows of qubits suffered damping errors $\mathcal{F}_a$. This is followed by individual Z measurements on all qubits in the damped rows to further locate the errors. We then apply an X to each damped qubit to remove the X part of the damping error $E = X(I - Z)$, leaving the $(I - Z)$ part for the next step.

(2) *XX measurements*: Next, we measure all XX stabilizers, to simultaneously perform damping-error recovery and deal with Z -type errors (i.e., Z , $\mathcal{F}_z$, or $I \pm Z$ errors). On damped rows (those which had $\mathcal{F}_a$ errors), as discovered by the earlier step, the XX measurement projects the state back to the correct code state, up to Z errors that can be deduced from the measurement outcomes. On all rows, the XX measurement has the effect of either annihilating the $\mathcal{F}_z$ errors or detecting Z errors.

(3) *Recovery for Z errors*: The Z errors can be deduced and remedied in the recovery step by combining the syndrome information from all the ZZ and XX measurements of the earlier two steps.

The ideal EC procedure is clearly not fault tolerant. For example, a single $\mathcal{F}_a$ in the data qubits right before the XX measurement is converted by that measurement into an X error. This X error is uncorrectable for the $(t + 1) \times (t + 1)$ Bacon-Shor code tailored for amplitude-damping noise as it can be

confused with t damping errors. In the following section, we describe a construction of the EC gadget that is fault tolerant.

III. FAULT-TOLERANT EC GADGET

Our goal here is to construct an EC gadget that satisfies standard FT requirements (see Appendix B for a statement of these requirements). A schematic overview of the whole fault-tolerant EC gadget is given in Fig. 2. We first note that the ideal EC procedure described above is composed of smaller subcircuits $H_{i,i+1}$ [see Fig. 2(a)] since the stabilizer measurements involve at most two neighboring rows. Our first step toward constructing a fault-tolerant EC gadget is thus to make the subcircuits fault tolerant against the first-order damping error $\mathcal{F}_a$, while containing the spread of Z errors. We describe this construction in Sec. III A. Subsequently, in Sec. III B, we discuss how to obtain and decode the syndrome for Z errors in a fault-tolerant manner using these subcircuits, thus completing our construction of a fault-tolerant EC gadget as a combination of individual subcircuits fault tolerant against $\mathcal{F}_a$, $\mathcal{F}_z$, and Z errors.

A. Fault-tolerant subcircuits

A subcircuit is fault tolerant against $\mathcal{F}_a$ if it satisfies the following condition: For an incoming state to the subcircuit with s $\mathcal{F}_a$ errors, if r $\mathcal{F}_a$ errors occur during the subcircuit such that $s + r \leq t$, the output of the subcircuit differs from the ideal output by at most r $\mathcal{F}_a$ errors. This property guarantees that the whole EC circuit constructed from many subcircuits together is also fault tolerant against $\mathcal{F}_a$.

To achieve this, we exploit a similar technique used in Ref. [27] to make the four-qubit code fault tolerant against a single damping error: We temporarily extend each row to a $(2t + 1)$ repetition code, with the use of t ancillary qubits, which protects against t X errors without confusion with t damping errors. We switch back to having only $t + 1$ qubits in each row at the end of the subcircuit. Each subcircuit $H_{i,i+1}$ is made up of several components as shown in Fig. 2(b). We describe each component separately below.

1. Coupling

In the first step, we extend each row from $t + 1$ to $2t + 1$ qubits by coupling t out of the $t + 1$ data qubits to a set of t ancillary qubits initialized to the $|0\rangle$ state. Every data qubit in the set is coupled to a single ancillary qubit via a CNOT, with the data qubit as the control and the ancilla as the target. Each row thus temporarily becomes a $(2t + 1)$ repetition code capable of protecting against t X errors. We refer to these ancillary qubits coupled to the data qubits as the *coupling* ancillas, to distinguish them from ancillary qubits used elsewhere. The $(2t + 1)$ data and coupling ancillas are referred to as *extended* data qubits.

2. Damping syndrome extraction

Next, we have a damping syndrome extraction step to detect $\mathcal{F}_a$ errors occurring before and after the coupling step. Note that, in the coupling step, the CNOTs can propagate damping errors from the data qubits to the coupling ancillas, turning s damping errors in the incoming state to $2s$ damping

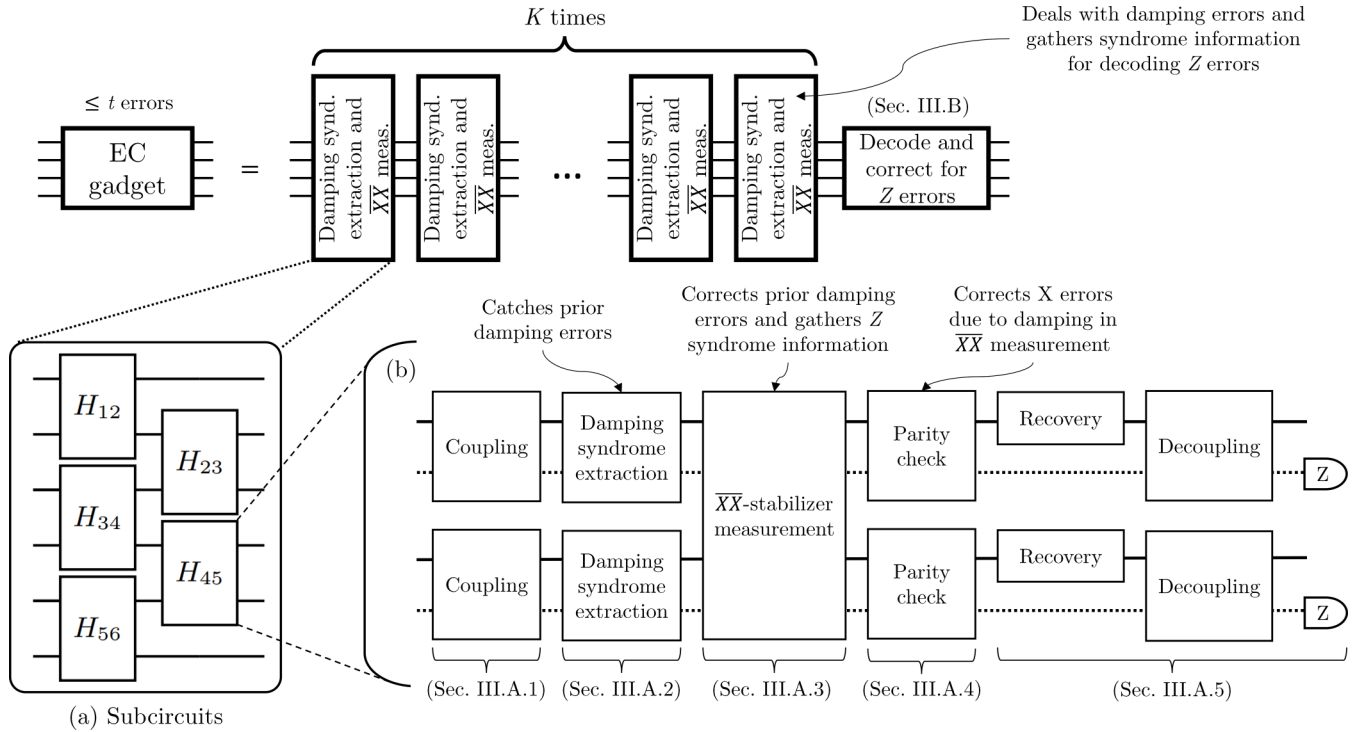


FIG. 2. Schematic overview of the fault-tolerant EC gadget, with references to the sections in the main text where the individual components are explained. Each solid horizontal line represents a row of $t + 1$ data qubits; each dotted line represents a row of t ancillary qubits. (a) The error correction gadget composed of smaller subcircuits, and (b) components of a subcircuit.

errors in the set of extended data qubits. If these errors are not taken care of before the $\bar{X}\bar{X}$ measurement, they get converted into $2s$ X errors that, together with further faults along the way, may exceed the capacity of the $(2t + 1)$ repetition code. A damping detection step is thus necessary immediately after the coupling step to catch and fix these errors, in addition to any incoming damping errors.

The damping detection is performed for each row separately. We start by measuring $Z_{i,j}Z_{i,j+1}$ ($1 \leq j \leq t$) stabilizers between two neighboring data qubits in the same row, using the parity measurement circuit of Fig. 3(a), employing an additional ancilla initialized to $|0\rangle$. At this stage, we only need to measure the parity of the data qubits, and not the coupling ancillas. Measuring the parity of the data qubits detects any of the $s \leq t$ damping errors occurring at the input or after the CNOT gates. The damping errors arising independently on the coupling ancillas (at most t) can be identified and corrected in the subsequent units once the errors on the data qubits have been fixed. Parity checks on the data qubits only thus suffice here.

One round of damping syndrome extraction applied to one row of data qubits involves t parity measurements between adjacent pairs of qubits, giving a length- t syndrome sequence

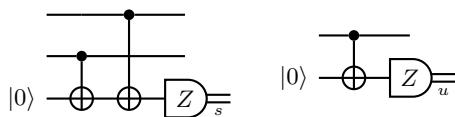


FIG. 3. Circuits for (a) a parity measurement between two qubits (left circuit) and (b) the nondestructive measurement of a qubit in the Z basis (right circuit).

of ± 1 . A nontrivial sequence indicates that damping errors have occurred in the data qubits and/or during the measurement. The state of the extended data qubits in that row is then reduced to a Z eigenstate (i.e., a string of 0s and 1s). At this point, we measure both the data qubits and the coupling ancillas in the Z basis using the circuit of Fig. 3(b); this measurement tells us which qubits have been damped to $|0\rangle$. Finally, we apply X to the damped qubits to remove the X part of the damping error $E = X(I - Z)$. This ensures that we do not carry forward more X errors than can be handled by the repetition code, as explained below.

If a trivial all $+1$ syndrome sequence is obtained after the first round of damping syndrome extraction, we cannot immediately conclude that there is no damping error—additional measurement faults could have given a falsely trivial syndrome. Instead, we have to repeat the parity measurements and keep doing so if we continue to get a trivial syndrome, until we have done t rounds. At this point, we can conclude that the row is undamped: Since at least one fault is needed for a nontrivial sequence, either the row is genuinely undamped or more than t faults have occurred so far, and this is beyond the scope of an FT syndrome measurement (see property P3 in Appendix B).

If a nontrivial sequence is obtained at any point during the repeated rounds of damping syndrome extraction, we stop the parity measurements and nondestructively measure all the extended data qubits in the faulty row in the Z basis to detect which qubits have been damped.

We note that, for our noise model, if a qubit has been damped to $|0\rangle$, a Z measurement—ideal or faulty—always gives the $+1$ outcome, and we never miss a damping event.

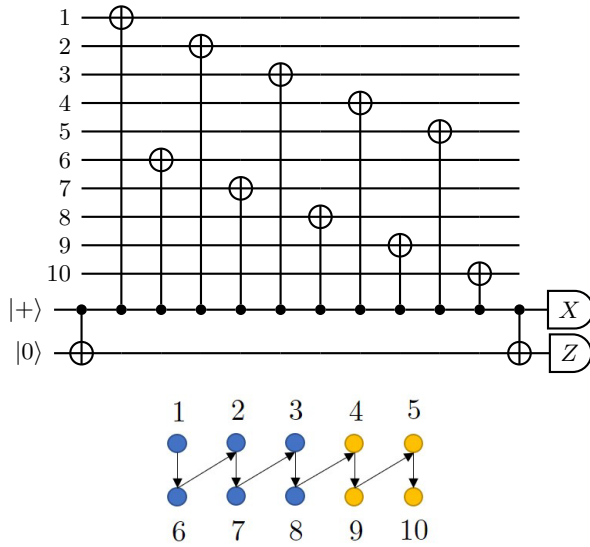


FIG. 4. The $\overline{XX}$ -stabilizer measurement between a pair of successive rows, shown for the 3×3 Bacon-Shor code. The top panel shows the circuit involving the $2 \times 5 = 10$ extended data qubits that make up two successive rows of a 3×3 Bacon-Shor code, with one ancillary qubit and one flag qubit (bottom qubits). The bottom panel shows the order in which the extended data qubits are coupled via the CNOTs to the ancillary qubit to assure fault-tolerance properties (see Sec. III A 3). The blue circles are the data qubits; the yellow ones are the coupling ancillas; the arrows indicate the order in which the ancillary qubit is coupled to the extended data qubits in the two rows.

However, a faulty Z measurement could give $+1$ outcome for a qubit that started in the state $|1\rangle$, and we might wrongly apply an X to an undamped qubit, effectively turning it into a damped qubit. We could have one of two possibilities happen here: (1) At least one parity measurement has a nontrivial outcome and the entire row of extended data qubits is measured; in this case, the number of errors in the outgoing state (equivalently, the number of errors in the state entering the next step) is limited by the number of faulty Z measurements. (2) No parity measurement has a nontrivial outcome; in this case, the number of errors in the outgoing state is limited by the number of errors on the coupling ancillas during the parity measurements and on the data qubits during the last layer of the parity measurements. In both cases, we note that the number of errors in the outgoing state is no more than the number of errors occurring inside the syndrome extraction step.

3. $\overline{XX}$ measurement

We now move to the measurement of the $\overline{XX}$ stabilizer. As in the ideal EC procedure, this step serves to recover code states from damping errors and deal with Z -type errors. A single ancillary qubit is initialized in $|+\rangle$ and then coupled, via CNOTs, to each of the extended data qubits in two successive rows; see Fig. 4 for the example of the 3×3 Bacon-Shor code. The ancillary qubit is then measured in the X basis. In the absence of faults in the $\overline{XX}$ measurement, with errors arising only from incoming errors prior to this $\overline{XX}$ step, an outcome of ± 1 indicates the presence or absence of Z errors in the outgoing state. To make this measurement fault tolerant,

we add a flag qubit [42] initialized to $|0\rangle$, as seen in Fig. 4. This qubit, along with a pair of CNOTs and a Z measurement, indicates whether a damping error has occurred in the ancillary qubit during the $\overline{XX}$ measurement. In the absence of faults, the measurement of the flag qubit always yields the outcome $+1$.

In the presence of faults in the $\overline{XX}$ measurement, however, the state can be contaminated by additional types of errors. We consider only $\mathcal{F}_a$ errors during the measurement procedure. $\mathcal{F}_Z$ and Z errors on the extended data qubits, from faulty CNOTs in the measurement circuit or in the waiting locations after the CNOTs, do not spread to the ancillary qubit and can be treated as incoming errors for the next measurement. The only errors that can cause problems are $\mathcal{F}_Z$ or Z errors on the ancillary and flag qubits, which can flip the $\overline{XX}$ -measurement outcome. We explain how to deal with those errors in Sec. III B; here, we focus only on the $\mathcal{F}_a$ errors.

There are various error scenarios that may occur during the $\overline{XX}$ measurement:

(1) A damping error $\mathcal{F}_a$ on one of the extended data qubits, after the damping syndrome extraction step but before the CNOT on that qubit, is projected to an $X(Y)$ error in the outgoing state when the measurement outcome is $+1$ (-1).

(2) A damping error $\mathcal{F}_a$ on one of the extended data qubits, due to the CNOT in the measurement circuit or from the idling time after the CNOT, causes a damping error in the outgoing state; these are in fact benign for the current measurement and are detected in the subsequent parity checks.

(3) A damping error $\mathcal{F}_a$ anywhere on the ancillary qubit propagates an X error to the flag qubit and correlated X errors to the extended data qubits in the outgoing state. Without the flag qubit, this can lead to logical errors when decoding, as it can be confused with the scenario that the extended data qubits are truly damped.

(4) A damping error $\mathcal{F}_a$ on the flag qubit can propagate a Z -type error to the ancillary qubit and/or flip the flag qubit, but it does not introduce any error to the extended data qubits. Two damping errors on both the flag and the ancillary qubits can lead to a random outcome in the measurement of the ancillary qubit, while flipping the outcome of the flag qubit to $+1$. This can cause a Z error when decoding, but this does not pose a problem since it is an $O(p^2)$ event that leads to an $O(p^2)$ error.

In what follows, we explain how to perform the $\overline{XX}$ measurement in a fault-tolerant manner in the presence of the error scenarios above. For that, we need the concept of an *error sequence* e for each row of extended data qubits. e specifies the error on each qubit in the outgoing state after the $\overline{XX}$ measurement: We write X if the qubit is affected by an X -type error (i.e., an $\mathcal{F}_a$, X , or Y error); we write I if there are no errors or only Z -type errors. For example, the error sequences on a pair of successive rows may look like $e_1 \equiv XIIXX$ for row 1 and $e_2 \equiv IXIXI$ for row 2. For each pair of successive rows with error sequences e_i and e_{i+1} , we let $\text{DIFF}(e_i, e_{i+1})$ be the number of columns in which e_i and e_{i+1} differ. For the example error sequences e_1 and e_2 above, $\text{DIFF}(e_1, e_2) = 3$ as the two sequences differ in columns 1, 2, and 5.

As we explain below, as long as $\text{DIFF} \leq t$ for every pair of successive rows, we can ensure fault-tolerant operation of the $\overline{XX}$ measurement. That $\text{DIFF} \leq t$ can be assured by

a clever ordering of the CNOTs between the extended data qubits and the ancillary qubit, as shown in Fig. 4. This statement is summarized in Lemma 1 below; the proof is in Appendix C.

Lemma 1. Let e_i and e_{i+1} be the X -type error sequences for a pair of successive rows, caused by $\ell \leq t$ errors in the $\overline{XX}$ -stabilizer measurement. It is possible to ensure $\text{DIFF}(e_i, e_{i+1}) \leq \ell$ by fixing the order in which the ancillary qubit is coupled to the extended data qubits, as given in Fig. 4.

Two points are worth noting about Lemma 1. First, the errors occurring inside the measurement can be any combination of the error scenarios mentioned above. Second, Lemma 1 is only applicable if the errors on the two rows in question originate from the $\overline{XX}$ measurement in that subcircuit and have not been propagated from other rows. Due to the way we execute subcircuits in an EC gadget [see Fig. 2(a) and Sec. III B below], this condition is satisfied.

After performing the $\overline{XX}$ measurement, we treat each row as a $(2t + 1)$ repetition code and use the parity measurements of Fig. 3(a) to obtain the syndrome sequence for each row. The syndrome sequence for a given row can, however, correspond to two complementary error sequences e and $\bar{e}$, where $\bar{e}$ is just e with the swap $I \leftrightarrow X$ for each qubit. For example, both $e = XXXI$ and $\bar{e} = IXXX$ correspond to the same syndrome sequence $(+1, -1, +1, +1)$. Lemma 1 then serves as our guide to deciding between the two possibilities: Based on the syndrome sequences for each pair of successive rows, we choose the error sequences e_1 and e_2 associated with the two rows such that $\text{DIFF}(e_1, e_2) \leq t$.

Assuming for the moment that the parity measurements following the $\overline{XX}$ measurement are perfect (imperfect parity measurements are handled in the next section), let us examine how that $\overline{XX}$ measurement behaves under the error scenarios mentioned above; see Appendix C for further details.

(1) When the ancilla is faulty, it propagates X errors to the data qubits such that the two error sequences satisfy $\text{DIFF}(e_1, e_2) = 1$, due to the way the cnots are applied (see Fig. 4).

(2) Any $s \leq t$ incoming damping errors in the extended data qubits propagate as $s \leq t$ X errors. The valid error sequence for the two rows, from the parity-check units, is therefore the one that satisfies $\text{DIFF}(e_1, e_2) \leq t$.

(3) Similarly, any $s \leq t$ outgoing damping errors after the $\overline{XX}$ measurement are identified by choosing the error sequence that satisfies $\text{DIFF}(e_1, e_2) \leq t$.

Note that the last two statements above are straightforward to see since both the rows that have $s \leq t$ damping or X errors are correctable by applying X operators on $s \leq t$ qubits on both the rows, such that they naturally satisfy $\text{DIFF}(e_1, e_2) \leq t$. Finally, the case where one could have a combination of the first and one of the other two error scenarios is treated in Appendix C. We have thus demonstrated that t damping errors in the $\overline{XX}$ measurement will not lead to uncorrectable X errors on more than t damping errors in the outgoing state.

4. Parity checks

We also need to deal with potential faults in the parity measurements [Fig. 3(a)] following the $\overline{XX}$ measurement. Faults

in the parity measurements can cause two problems. First, they can introduce more errors to the extended data qubits. However, regardless of how those errors are introduced, as long as the total number of errors in the $\overline{XX}$ measurement and the parity-check steps is no more than t , the condition of Lemma 1 is still satisfied and the error sequences at the end of the parity-check step still have the property stated in the lemma.

Second, if new faults occur on the ancilla qubits that are used in the parity measurements, they can give incorrect syndromes and consequently erroneous inferences of error sequences. Again, we address this simply by repeating the parity measurements until we get the same syndrome sequence $t + 1$ times. In the presence of no more than t faults, there must be at least one round without any faults among the $t + 1$ rounds in which we obtain the same syndrome sequence. By using the syndrome sequence that repeats $t + 1$ times, we can confidently infer the correct error sequence. In the worst-case scenario, we may need to repeat up to $t(t + 1) + 1$ rounds. This takes into account the possibility of fresh errors occurring as we repeat, but we are limited to no more than t errors in total, within fault-tolerance considerations. This simple repetition scheme, however, is not optimal; in Appendix D, we discuss various methods to reduce the number of repetitions required, thus improving the efficiency of the syndrome measurement process.

5. Recovery and decoupling

In the recovery step, we apply local X operations to the extended set of data qubits, based on the error sequences inferred from the syndromes obtained in the parity-check step (see Sec. III A 3 for the decoding algorithm). Finally, in the last step of the subcircuit, we decouple the coupling ancilla qubits from the original data qubits, using the same circuit as in the coupling step. The coupling ancilla qubits are then measured in the Z basis and the measurement outcomes are recorded for the purpose of determining damped rows later.

6. Fault tolerance of a subcircuit

We now come back to address the fault tolerance of the subcircuit against $\mathcal{F}_a$ errors: Fault tolerance of the subcircuit is ensured by the fault tolerance of its individual components. Assume that there are s incoming $\mathcal{F}_a$ errors to the subcircuit (both rows), and then $r_1, \dots, r_5$, and r_6 $\mathcal{F}_a$ errors in the coupling, damping syndrome extraction, $\overline{XX}$ measurement, parity check, recovery, and decoupling steps, respectively. Also, assume that $s + \sum_{i=1}^6 r_i \leq t$. We can leverage the fault-tolerance property at every step to show that there are no more than $r_4 + r_5 + r_6$ $\mathcal{F}_a$ errors in the outgoing state of the subcircuit, thereby showing that the subcircuit is fault tolerant against $\mathcal{F}_a$ errors. Specifically,

(1) Since $s + r_1 + r_2 \leq t$, the syndrome extraction step (Sec. III A 2) ensures that its outgoing state has no more than $r_1 + r_2$ $\mathcal{F}_a$ errors.

(2) Since $r_1 + r_2 + r_3 \leq t$, the $\overline{XX}$ -stabilizer measurement step (Sec. III A 3) ensures that the X -type error sequences of its outgoing state have the property stated in Lemma 1.

(3) This, together with the fact that $r_3 + r_4 \leq t$, ensures that correlated errors from the $\overline{XX}$ -stabilizer measurement

step are detected and its outgoing state differs from the detected state by no more than $r_4 \mathcal{F}_a$ errors (Sec. III A 4).

(4) In the recovery step, gates are applied transversally (Sec. III A 5). This ensures that there are no more than $(r_4 + r_5) \mathcal{F}_a$ errors at the end of this step.

(5) Finally, the decoupling step also applies transversal gates with respect to the data qubits (Sec. III A 5). This means there are no more than $(r_4 + r_5 + r_6) \mathcal{F}_a$ errors in the outgoing state of the subcircuit.

Furthermore, for the discussion on decoding Z-type errors (Z and $\mathcal{F}_z$) in the next section, it is important to note that a subcircuit does not spread Z-type errors badly. First, a Z-type error on any data qubit trivially passes through all steps except for the $\overline{XX}$ measurement, where it propagates to the measurement ancilla. However, the role of this ancilla is limited to the detection of the error and, moreover, the propagated error does not propagate back to the data qubits. Second, a Z-type error on any coupling ancilla behaves in the same manner as on a data qubit, except in the decoupling step, where it can propagate back to the data qubits. However, it can only propagate to at most one data qubit. This is also true for Z-type errors on measurement ancillas of parity measurements. Overall, the total number of Z-type errors in the outgoing state of the subcircuit is thus limited by the number of Z-type errors inside the subcircuit.

We note that our construction guarantees that the output of the EC gadget has no more than $(r_4 + r_5 + r_6)$ errors, more stringent than required by the fault-tolerance condition where the output can have up to $\sum_{i=1}^6 r_i$ errors [see property (P1) of Appendix B], equal to the number of errors that originate in the EC gadget itself. The construction can thus be optimized to relax some of the error propagation and correction requirements, but we opted for the current modular structure where components are made individually fault tolerant, for a simpler argument of fault tolerance of the full EC gadget.

B. Decoding Z errors

As noted earlier, each $\overline{XX}$ measurement between a pair of rows in a single subcircuit outputs only one syndrome bit. To have sufficient information for decoding Z errors, all the $\overline{XX}$ stabilizers between every pair of consecutive rows need to be measured. One round of Z-syndrome measurement consists of t subcircuits performed in two steps, thereby giving a t -bit syndrome sequence. In the first step, the subcircuits are performed between rows 1 and 2, 3 and 4, and so on, as shown in Fig. 2(a). In the second step, they are performed between rows 2 and 3, 4 and 5, and so on. Note that if there is an odd number of rows in the lattice, the last row stays idle in the first step and the first row stays idle in the second step. If the number of rows is even, the first and the last rows stay idle in the second step.

Due to faults in the $\overline{XX}$ measurements, the t -bit syndrome obtained after the first round of measurements is unreliable. We therefore perform many rounds of the $\overline{XX}$ measurements, using the same repetition scheme as in the parity-check step of a subcircuit. Specifically, we repeat until a syndrome sequence appears $t + 1$ times and use it to decode the Z error.

In addition to the $\overline{XX}$ syndrome, we also need information about the damped/undamped rows from the damping syndrome extraction step to complete the decoding, as explained in Appendix A. In the following, we explain how to get this information from the various syndrome bits obtained in the subcircuits.

Consider a subcircuit and a row involved in that subcircuit.

(1) *Damping syndrome extraction*: An outcome of -1 in any of the ZZ-stabilizer measurements or the individual Z measurements done on the row clearly indicates a damping error in the incoming state to the subcircuit or a damping fault in the syndrome extraction unit. In either case, we label the row as damped.

(2) *Parity checks*: An outcome of -1 in any of the parity measurements done on the row indicates a damping fault somewhere in the damping syndrome extraction, the $\overline{XX}$ measurement, or the parity-check step. If the flag qubit in the $\overline{XX}$ measurement yields an outcome of $+1$, then the row is genuinely damped or there are at least two damping errors on both the ancillary and flag qubits in the $\overline{XX}$ measurement. Regardless, we label the row as damped in this case. On the other hand, if the flag qubit yields an outcome of -1 , then it is possible that errors detected by the parity measurements are merely due to a damping error on the ancillary qubit in the $\overline{XX}$ measurement. In this case, we label the row as *potentially damped* row.

Without the flag qubit in the $\overline{XX}$ measurement, a potentially damped row might be incorrectly identified as a truly damped row. For instance, consider a 7×7 Bacon-Shor code capable of correcting up to six damping errors. The extraction of Z syndrome can be ambiguous for the following two scenarios: (a) four damping errors $\mathcal{F}_a$ each affecting one of the rows 1, 2, 3, and 4, along with an additional Z error on row 5, and (b) two damping errors on the ancilla qubits measuring $\overline{X}_1\overline{X}_2$ and $\overline{X}_3\overline{X}_4$, together with two Z errors on rows 6 and 7. The errors on the ancilla qubits can be arranged such that X errors appear on the first four rows. Both scenarios can thus produce an identical syndrome of $(+1, +1, +1, +1, -1, +1)$ for Z errors. For scenario (a), we need to apply Z's to the first five rows, whereas for scenario (b), we need to apply Z's to the last two rows; the flag qubit allows us to distinguish between the two.

Note that there are cases where a coupling ancilla is damped and the subcircuit under consideration detects and corrects the X part of the damping error. The $(I - Z)$ part, however, is transferred to the data qubits after the decoupling step and enters the next subcircuit. None of the parity measurements in the succeeding subcircuit will detect this error. However, we must be aware that a Z error, if projected out by the $\overline{XX}$ measurement in the next subcircuit, comes from a first-order damping error. To take this into account, we also declare a row to be damped/potentially damped if the preceding subcircuit has labeled it as damped/potentially damped.

(3) *Decoupling*: All the coupling ancillas are measured in the Z basis after being decoupled from the data qubits. If any of them yields a -1 outcome, it indicates a damping fault after the parity-check step. The fault may be detected by the damping syndrome extraction in the next subcircuit

if the fault is on a data qubit, or undetected by the next subcircuit if the fault is on a coupling ancilla [because only the $(I - Z)$ part remains, as explained above]. So, a row in the current subcircuit is also labeled as damped if any of the coupling ancilla measurements in the previous subcircuit has outcome -1 .

In any EC unit, a single row of data qubits is a part of many subcircuits. We say that the row is damped/potentially damped if any of the subcircuits in the EC unit flags it as damped/potentially damped. Note that a potentially damped row can later be marked as a damped row if a condition for damped row is matched. However, once the status of a row is marked as damped, it will not be changed to potentially damped or undamped (for the current EC unit).

Once the XX syndromes and the information about damped/potentially damped/undamped rows are collected, we decode and correct for Z errors using the following procedure: Recall from Sec. III A 3 that there are a pair of (complementary) error sequences e and $\bar{e}$ compatible with a given XX syndrome. Let n_d , n_p , and n_u denote the number of damped, potentially damped, and undamped rows, respectively. Let $z_d(e)$, $z_p(e)$, and $z_u(e)$ denote the number of Z errors in the damped, potentially damped, and undamped rows for the sequence e . Then, the number of Z errors in the damped, potentially damped, and undamped rows for the sequence $\bar{e}$ are $z_d(\bar{e}) = n_d - z_d(e)$, $z_p(\bar{e}) = n_p - z_p(e)$, and $z_u(\bar{e}) = n_u - z_u(e)$.

Next, we determine the more probable error sequence by summing up the number of Z errors with suitable weights. We assign a weight $w(e; n_d, n_p, n_u) = n_d + n_p + z_p(e) + 2z_u(e)$ to the error sequence e with respect to the configuration of n_d damped rows, n_p potentially damped rows, and n_u undamped rows. This definition is motivated by the relative orders of the different errors, in terms of the noise strength p . Specifically, a damped row requires at least a first-order damping event to occur; thus, n_d damped rows imply an event of at least $O(p^{n_d})$. Similarly, n_p potentially damped rows arise only when an event of order at least $O(p^{n_p})$ takes place. Finally, a Z error found in a potentially damped row can result from a first-order event (e.g., a true damping error), and thus, $z_p(e)$ is added to the weight. On the other hand, a Z error appearing in an undamped row necessitates at least a second-order event (e.g., a genuine Z error) that leads to the inclusion of $2z_u(e)$ in the weight.

Now, between e and $\bar{e}$, we choose the sequence with the lower weight as the output of the decoding. This is equivalent to comparing $z_p(e) + 2z_u(e)$ and $z_p(\bar{e}) + 2z_u(\bar{e})$. If there is no potentially damped row (i.e., $z_p = 0$), it reduces to comparing the number of Z errors in the undamped rows, as what is done for the ideal EC (see Appendix A).

This completes our description of the fault-tolerant EC gadget. Further optimizations of the EC gadget are possible. In particular, we note that there are multiple places in the gadget where we need to repeat the measurements until we obtain $t + 1$ identical repeats of the syndrome sequences. A simple counting gave a worst-case $t(t + 1) + 1$ repetitions of the requisite measurement. A more optimal approach gives a reduced requirement of no more than $(1/4)(t + 2)^2 + 1$ repetitions, fewer than $t(t + 1) + 1$ for $t > 1$; see Appendix D for further details.

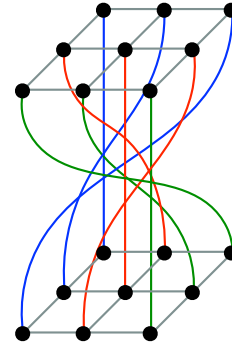


FIG. 5. Transversal implementation of the encoded CZ gate for 3×3 Bacon-Shor code. The physical CZ gates are between the pairs of qubits, one from the top code block, the other from the bottom code block, connected by a colored (red/blue/green) line. The colors mark the qubits in the same row/column for the top/bottom code blocks.

C. Fault tolerance of the EC gadget

We now pull together all the arguments that explain why the EC gadget satisfies the fault-tolerance property (P1) of Appendix B. As mentioned at the end of Sec. III A, a subcircuit is fault tolerant against $\mathcal{F}_a$ errors. Since an EC unit consists of multiple subcircuits, it is also fault tolerant against $\mathcal{F}_a$ errors.

In addition, a subcircuit contains the spread of Z errors. Hence, at the end of an EC unit, the outgoing state has no more Z errors than the number of errors entering and occurring inside the unit. To be fault tolerant against Z errors, however, we need to be sure that decoding the Z errors does not lead to a logical error. This amounts to having reliable information of the Z -type syndrome and the damped rows. The Z -type syndrome is reliable due to the repeat-until-success scheme we adopt (see Appendix D), and all damped rows are flagged according to the protocol described in the last section.

IV. LOGICAL GADGETS

To complete our discussion of fault-tolerant computation with Bacon-Shor codes, we describe how to construct a universal set of logical gadgets fault tolerant against amplitude-damping noise. Our construction of the single-qubit logical gadgets, as well as the preparation and measurement gadgets, follows the prescription of our earlier work [27] for the four-qubit code (the smallest Bacon-Shor code); see Appendixes E and F for a self-contained description. We hence confine our discussion here to the logical CZ gadget, as it presents some interesting features. We note that one can also construct a controlled-controlled- Z (CCZ) gadget, adding to the variety of multi-logical-qubit gates; see Appendix G.

The logical CZ gate for the Bacon-Shor codes can be realized transversally by performing physical CZ gates between qubits in each row of one code block and those in each column of the other code block. A specific implementation—there are many that work—is to have the physical CZ gate between qubit (i, j) of one code block and qubit (j, i) of another code block, as illustrated in Fig. 5 for the 3×3 code. That this works as the logical CZ gate can be understood by noting that

the logical states $|0\rangle_L$ and $|1\rangle_L$ can be written as

$$\begin{Bmatrix} |0\rangle_L \\ |1\rangle_L \end{Bmatrix} = \left(\frac{1}{\sqrt{2}} \right)^{n+1} \sum_{x \in \{0,1\}^n} [1 \pm (-1)^{|x|}] |x\rangle, \quad (5)$$

where x is an n -bit string, $|x|$ counts the number of 1s in x (the number of 1_{row} s in x), and $|0(1)_r\rangle \equiv |0(1)\rangle_{\text{row}} \equiv |0(1)\rangle^{\otimes n}$ from Eq. (4). The individual bits of x thus go over the different rows of qubits of the Bacon-Shor lattice, specifying if each row of qubits is in the $|0\rangle_{\text{row}}$ or the $|1\rangle_{\text{row}}$ state. Observe that $|0\rangle_L$ is a sum of terms, each of which has an even number of rows in $|1\rangle_{\text{row}}$; i.e., $|x|$ is even; $|1\rangle_L$ is composed only of terms with odd $|x|$. Since the logical Z is equivalent to applying Z to any column, if the state is $|0(1)\rangle_L$, an even (odd) number of columns would have had a logical Z applied to them, amounting to an overall identity (logical Z) operation, thus accomplishing the logical CZ gate.

It is not apparent that this construction is fault tolerant. As emphasized in Ref. [27], transversality in such a noise-adapted situation does not automatically mean fault tolerance, as the transversal gates can still convert a correctable error into an uncorrectable one. Now, an $\mathcal{F}_z$ or a Z error passes through the logical CZ without change and the transversality does assure fault tolerance. An $\mathcal{F}_a$ error, however, in the incoming state of, say, the top code block will propagate a Z error to the bottom block under a transversal CZ gate. An EC gadget following the bottom block may not, by itself, correctly decode the Z errors as our EC construction will incorrectly treat these first-order propagated Z errors as the second-order Z errors when trying to decode the detected errors. However, this problem can be resolved by incorporating the information about the damped rows in the first block. Specifically, when decoding Z errors for an output block of a logical CZ, the set of rows regarded as damped (used in the maximum-likelihood decoding of the Z -error syndromes described in Sec. III B) should include also those rows with qubits connected to damped qubits in the other block (as indicated by the damping syndrome extraction in the EC gadget of the other block). With this joint decoding of the two code blocks, we achieve fault tolerance of the logical CZ gadget.

V. LOGICAL INFIDELITY AND PSEUDOTHRESHOLD FOR QUANTUM MEMORY

To demonstrate the fault tolerance of our EC-gadget construction, here we estimate the logical infidelity for different code sizes $L = t + 1$ and compute the pseudothreshold levels for the basic task of quantum memory.

Following the standard fault-tolerance analysis of Ref. [43], the pseudothreshold can be estimated by considering an “extended memory gadget” comprising an idling (i.e., memory) step sandwiched between two EC units. Specifically, the pseudothreshold is that critical value of the noise strength, $p = p_{\text{th}}$, below which the fault-tolerant quantum memory outperforms the unencoded (i.e., not error-corrected) version. In particular, we compare the infidelity with and without error correction. For amplitude-damping noise with strength p , the fidelity of the unencoded qubit averaged over all input states falls as $1 - \alpha p$ for small p , where $\alpha = 1/3$. We write $1 - F(p)$ for

the infidelity between the encoded output and the input states, with $F(p)$ a function of p . Then, the pseudothreshold p_{th} is determined by the condition

$$1 - F(p_{\text{th}}) = \alpha p_{\text{th}}. \quad (6)$$

Often, the pseudothreshold for a fault-tolerant quantum computing scheme is computed by exact counting or full computer simulation, for a close estimation. As we are after only a simple gauge of the efficacy of our scheme, we opt instead for an approximate bound by counting the number of ways $t + 1$ errors can occur in the extended memory gadget. The $(t + 1)$ -error term gives the dominant contributor to the infidelity. Specifically, for a given noise strength p , we estimate

$$1 - F(p) < \left[\binom{2N + (t + 1)^2}{t + 1} - \binom{N}{t + 1} \right] p^{t+1}. \quad (7)$$

Here, N is the total number of circuit locations in our EC gadget and the second term in the brackets excludes the ways of having $t + 1$ errors in the first EC gadget in the extended memory gadget; those are counted already in the previous extended gadget.

Note that we have assumed that p is small enough such that larger-than- t -error terms give a contribution to the infidelity that is smaller than the overcounting we have here from counting all $(t + 1)$ -error locations, including those benign to the infidelity (usually quite a significant fraction). We need also then an estimate of N . We can do a careful counting using our EC-gadget construction (see Appendix H for the details), and we find $N = O(t^6)$, a rather rapid scaling with t due in part to the necessity to repeat the parity measurements for fault tolerance. The dominant part in a subcircuit is repeating the parity measurements, which take, in the worst case, $O(t^2)$ rounds with $O(t)$ location for each round. This gives $O(t^3)$ locations for a subcircuit. For the whole EC, there are $O(t)$ subcircuits. In the worst case, $\overline{XX}$ measurements are repeated $O(t^2)$ rounds for reliable decoding Z errors. This gives the $O(t^6)$ number of locations in total.

Figure 6 plots the upper bound on the infidelity for different values of the damping parameter p and for different Bacon-Shor code sizes n (i.e., $n \times n$ lattice, for $n = t + 1$). We show also the infidelity without error correction (dotted line), averaged over all pure states. We can then read off lower bounds on the pseudothresholds, i.e., the values of the damping parameter p below which, for a given $n \times n$ code, error correction gives improved infidelity values over the unencoded situation. These pseudothreshold values are given in Table I. We note that, for each damping strength, there is an optimal lattice size that gives minimal infidelity. This is similar to what was observed previously for Bacon-Shor codes for correcting Pauli errors [44]. This suggests that there is no nonvanishing accuracy threshold value for the damping strength below which increasing the code size always lowers the infidelity.

We note that our infidelity estimates, and hence the pseudothreshold values, here are likely far too pessimistic, due to our simple counting approach. For example, in our earlier work on the $n = 2$ case [27], a more careful estimate gave a pseudothreshold of 1.5×10^{-4} , compared to the corresponding $\sim 10^{-6}$ value in Table I. In particular, we count

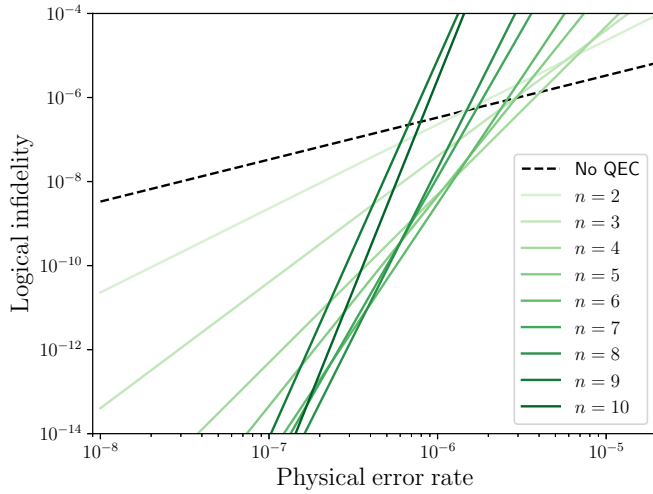


FIG. 6. Upper bounds on the infidelity of the output state after the extended memory gadget, plotted as a function of the amplitude-damping parameter p for various lattice sizes $n = t + 1$. The dotted line is the average infidelity for the unencoded state, while the green lines of varying shades are for the different lattice sizes $n = 2$ to $n = 10$, from light to dark green.

in Eq. (7) *all* ways of having $t + 1$ errors, not just those that do eventually lead to an uncorrectable situation. As is commonly observed in such analyses, a significant fraction of $> t$ -error situations are in fact benign; i.e., they give only correctable errors; this is certainly the case in our earlier $n = 2$ construction [27].

For our current general- n construction, we note that not all fault locations from the leading EC in the extended memory gadget will combine with fault locations in the trailing EC to form malignant sets of faults that do cause uncorrectable errors. Rather, only faults at locations near the end of the leading EC can combine with faults in the trailing EC to cause logical errors; most combinations of faults in the first EC can in fact be corrected. This is true even for subcircuits while considering only $\mathcal{F}_a$ errors—most of the combinations of $\mathcal{F}_a$ errors in a subcircuit can be eliminated by the subcircuit itself; only a small portion of faults occurring at the end of a subcircuit can propagate errors to combine with faults from subsequent subcircuits. In addition, during idle time, a qubit cannot be damped more than once, so we should exclude sets

TABLE I. Lower bounds on the pseudothreshold for lattice sizes up to $n = 10$.

n	p_{th}
2	1.46×10^{-6}
3	2.88×10^{-6}
4	4.08×10^{-6}
5	2.93×10^{-6}
6	2.57×10^{-6}
7	1.72×10^{-6}
8	1.49×10^{-6}
9	6.73×10^{-7}
10	7.94×10^{-7}

with multiple dampings on the same qubit. All these benign situations add to our upper bound on the infidelity in Eq. (7), but a more careful counting can exclude them to give a tighter upper infidelity bound and improved pseudothresholds.

Nevertheless, our goal here is simply to demonstrate how even smaller-sized Bacon-Shor codes are useful when adapted to a non-Pauli noise setting and can yield nontrivial pseudothresholds.

VI. CONCLUSION

We have shown how to construct an error-correction gadget as well as a set of universal logical operations based on the Bacon-Shor family of codes, all fault tolerant against amplitude-damping noise on individual physical components. This completes the work started in Ref. [27], giving a fully fault-tolerant approach for quantum computation capable of correcting amplitude-damping errors on multiple qubits. Our work demonstrates the possibility of fault-tolerant construction starting only from the noise description, without having to adapt the scheme from a more noise-agnostic protocol.

Due to the adaptation to a specific noise channel, rather than having a full error basis, guaranteeing fault tolerance here is not just about controlling the spread of errors, but also controlling the *types* of errors that emerge from the gates—they have to remain correctable by our noise-adapted code. As noted already in our earlier work, this is the basic reason why transversal constructions of operations, a cornerstone of general-purpose FT schemes, do not automatically assure fault tolerance in noise-adapted situations. Transversality guarantees a low spread of errors, but the types of errors that emerge are entirely determined by the gates used in the gadget construction and how the errors propagate through them.

In our earlier work, we dealt with this issue by momentarily extending our four-qubit code to an eight-qubit one that could deal with additional varieties of errors beyond amplitude-damping ones, before switching back to the original four-qubit code. Here, we extend that approach to general Bacon-Shor codes, with the use of the coupling ancillas and the $(2t + 1)$ -qubit repetition code in the EC gadget. In addition, we present a curious solution to this problem in the cz gadget—we deal with the originally uncorrectable extra Z error, propagated to the second code block due to an amplitude-damping error in the first code block, by a joint decoding of the two code blocks to recognize this propagated Z error. Such a joint decoding to extend the fault-tolerance capabilities, requiring no joint quantum measurements but only pooling of the classical syndrome information, is likely useful in other contexts.

A different approach to achieving fault tolerance in such noise-adapted contexts would be to employ only gates that are “noise-structure-preserving,” so that incoming errors arising from the noise channel remain among the correctable error types after propagation through the gate. Such noise-structure-preserving gates have been identified for Pauli errors in the past [18,45]. For amplitude-damping noise, however, we have not been able to find a set of noise-structure-preserving physical gates with sufficient variety to construct a universal set of logical operations.

Of course, our gadget constructions here are far from optimal. The simplistic repetition of the syndrome checks, reminiscent of the repeated syndrome rounds of surface-code error correction, is but an easy-to-analyze route to ward against errors in the syndrome measurements; Appendix D offers some optimizations but certainly more can be done. Our component-by-component approach to assuring the fault tolerance of the EC gadget gave a clear path to showing the fault-tolerance properties, but led to fewer output errors than required by FT rules; a more holistic construction no doubt can lower resource demands and thereby improve the pseudothreshold numbers and truly realize the potential of noise adaptation.

Nevertheless, we have shown the possibility of a fully fault-tolerant quantum computation scheme adapted for amplitude-damping noise. We hope our work will motivate exploration into schemes adapted for other types of noise, including ones that are directly relevant to current quantum computing experiments.

ACKNOWLEDGMENTS

H.K.N. acknowledges partial support from the National Research Foundation, Singapore, and A*STAR under its CQT Bridging Grant. P.M. acknowledges funding support from the National Quantum Mission, an initiative of the Department of Science and Technology, Government of India.

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: CORRECTING AMPLITUDE-DAMPING NOISE WITH THE BACON-SHOR CODE

Here, we give further details of the ideal EC procedure of Ref. [32] to correct amplitude-damping noise using the Bacon-Shor code and illustrate its use with the 3×3 example.

1. Ideal EC procedure

The ideal EC procedure constructed in Ref. [32] involves three steps performed in sequence: a two-part syndrome extraction for damping errors, a Z-error syndrome extraction, and a recovery step. We explain each step in detail below.

(1) Damping-error syndrome extraction

(a) Measure the $Z_{i,j}Z_{i,j+1}$ stabilizers, for all i (row) and j (column). For each measurement, the two qubits involved are said to have even (odd) parity if the outcome is $+1$ (-1). In each row, if at least one parity outcome is odd, we declare that the row is damped; otherwise, the row is undamped.

(b) For each damped row, we measure all the qubits in the row in the Z basis to determine the positions of the damping errors. If the measurement outcome is $+1$, the qubit has a damping error; if the outcome is -1 , the qubit is not damped. This can be seen by considering the effect of a damping error E on an arbitrary encoded state: E has effect only if the qubit is originally in the $|1\rangle$ state; the damped row is disentangled from the rest of the rows, regardless of

TABLE II. List of possible syndromes and the projected errors after measuring an $\bar{X}\bar{X}$ stabilizer in the presence of damping errors. The two rows involved are denoted as 1 and 2. $\bar{I}_1$ ($\bar{I}_2$) denotes all identities in row 1 (row 2), $\bar{Z}_1$ ($\bar{Z}_2$) denotes a single Z operator on one of the qubits in row 1 (row 2). It is assumed that the X part of the damping error is corrected, leaving only the $(I - Z)$ part.

Damped rows	Initial error	Outcome	Final error
None	$\bar{I}_1\bar{I}_2$	$+1$	$\bar{I}_1\bar{I}_2$
None	$\bar{Z}_1\bar{I}_2$	-1	$\bar{Z}_1\bar{I}_2$
1	$(\bar{I}_1 - \bar{Z}_1)\bar{I}_2$	$+1$	$\bar{I}_1\bar{I}_2$
1	$(\bar{I}_1 - \bar{Z}_1)\bar{I}_2$	-1	$\bar{Z}_1\bar{I}_2$
1, 2	$(\bar{I}_1 - \bar{Z}_1)(\bar{I}_2 - \bar{Z}_2)$	$+1$	$\bar{I}_1\bar{I}_2$ or $\bar{Z}_1\bar{Z}_2$
1, 2	$(\bar{I}_1 - \bar{Z}_1)(\bar{I}_2 - \bar{Z}_2)$	-1	$\bar{Z}_1\bar{I}_2$ or $\bar{I}_1\bar{Z}_2$

the original encoded state, and the damped qubit ends up in the state $|0\rangle$.

We note that the above steps can identify the positions of up to $t = n - 1$ damping errors in each row [46]. At the end of this syndrome extraction step, we are ready to correct for the X part of the damping error $E = X(I - Z)$ by applying X to the damping locations, leaving $(I - Z)$ part for the subsequent detection. Alternatively, we can delay the application of these local X 's until the end of the EC procedure. This will not affect the detection of Z errors in the next part.

(2) *Z-error syndrome extraction*: Measure all the $\bar{X}_i\bar{X}_{i+1}$ stabilizers and obtain the syndrome bits. An $\bar{X}\bar{X}$ measurement on undamped rows has the effect of detecting Z errors. If one of the rows is damped, as detected in Step 1(a), the $\bar{X}\bar{X}$ measurement either removes that damping error and returns the correct code state (if outcome $+1$ is obtained), or it projects the qubits onto the correct code state up to a Z error on one of the qubits in the damped row (if outcome -1 is obtained). The measurement of $\bar{X}\bar{X}$ stabilizers can result in different postmeasurement errors depending on whether any of the rows involved were damped prior to the measurement. Table II gives the scenarios that can occur for the two rows before the measurement, along with the measurement outcomes and the subsequent errors.

Measuring the $\bar{X}\bar{X}$ stabilizers also has the effect of annihilating the off-diagonal $\mathcal{F}_z$ errors. To see this, note that $\mathcal{F}_z(\cdot) \equiv (1/2)[(\cdot)Z + Z(\cdot)]$ and

$$(1 + \bar{X}_i\bar{X}_{i+1})(Z\rho + \rho Z)(1 + \bar{X}_i\bar{X}_{i+1}) = 0, \quad (\text{A1})$$

where ρ is a state in the code space and the $\mathcal{F}_z$ error is on row i or $i + 1$. However, an even number of $\mathcal{F}_z$ errors may not be eliminated by an $\bar{X}\bar{X}$ measurement. For example, if there are two $\mathcal{F}_z$'s on a row, it has the effect of a single Z error, since

$$\mathcal{F}_z \otimes \mathcal{F}_z(\cdot) = \frac{1}{2}[(\cdot) + Z(\cdot)Z]. \quad (\text{A2})$$

The double $\mathcal{F}_z$ errors are not annihilated but contribute to the probability of measuring Z errors. We need to consider two cases: (a) even numbers of $\mathcal{F}_z$'s on both the rows and (b) odd numbers of $\mathcal{F}_z$'s on both the rows. In case (a), the postmeasurement errors are Z errors and would be corrected as if they originated from genuine Z errors. However, in case

TABLE III. List of possible syndromes and the projected errors after measuring an $\bar{X}\bar{X}$ stabilizer in the presence of $\mathcal{F}_z$ errors. The two rows involved are denoted as 1 and 2. The table shows various scenarios: even number of $\mathcal{F}_z$'s in both rows (cases 1 and 2), even number in one row and odd number in the other row (cases 3 and 4), and odd number in both rows (cases 5 and 6). $\bar{I}_1$ ($\bar{I}_2$) denotes no Z error in row 1 (2), while $\bar{Z}_1$ ($\bar{Z}_2$) denotes one Z error on one of the qubits in row 1 (2).

Case	Row 1	Row 2	Outcome	Final error
1	Even	Even	+1	$\bar{I}_1\bar{I}_2$ or $\bar{Z}_1\bar{Z}_2$
2	Even	Even	-1	$\bar{Z}_1\bar{I}_2$ or $\bar{I}_1\bar{Z}_2$
3	Odd	Even	± 1	Eliminated
4	Even	Odd	± 1	Eliminated
5	Odd	Odd	+1	$\bar{Z}_1\bar{Z}_2(\cdot) + (\cdot)\bar{Z}_1\bar{Z}_2$
6	Odd	Odd	-1	$\bar{Z}_1(\cdot)\bar{Z}_2 + \bar{Z}_2(\cdot)\bar{Z}_1$

(b), the postmeasurement errors still have the off-diagonal form (see Table III). Nevertheless, these errors are eliminated after considering all the $\bar{X}\bar{X}$ -stabilizer measurements. The reason is that there is at least one row with no $\mathcal{F}_z$ error, given that there are no more than $t = n - 1$ damping errors, and the measurement involving this row cannot result in errors of the off-diagonal form.

(3) *Recovery*: As mentioned earlier, the X part of the damping error can be corrected using the information about the damping locations obtained from the damping-error syndrome extraction step. Meanwhile, decoding Z errors involves both the syndromes obtained from the $\bar{X}\bar{X}$ measurements and the information about the damped rows obtained from the syndrome extraction unit. As can be seen from Table II, Z errors can be projected out from damping errors, but those are first-order errors and should be distinguished from genuine second-order Z errors.

To understand how we can decode the Z error, it is important to note that the measurement of $\bar{X}_i\bar{X}_{i+1}$ fixes the parity of the Z error for the i th and $(i + 1)$ th rows. If $\bar{X}_i\bar{X}_{i+1} = +1$, it means that both rows either have no error or both have Z errors. Conversely, if $\bar{X}_i\bar{X}_{i+1} = -1$, only one of them has a Z error. Starting from the first row, the $\bar{X}\bar{X}$ syndromes will determine the errors for all the other rows. Since there are two possibilities for the Z error in the first row—either no error or one Z error—there are two bit strings compatible with the obtained syndromes, denoted as f and $\neg f$. Since each damped row carries a weight of at least p , the determining factor for which sequence is more likely is actually the number of undamped rows having a Z error. Let f' be the sequence (between f and $\neg f$) with a smaller number of 1s among the undamped rows. We correct Z errors by applying a Z operation to the i th qubit if $f'_i = 1$. The specific qubit within a row to which the Z operation is applied does not matter because they differ only by a stabilizer operation.

2. Example: The 3×3 Bacon-Shor code

To help the reader better understand the ideal EC procedure described above, we explore the example of the 3×3 Bacon-Shor code, capable of protecting against $t = 2$ amplitude-damping errors. We label the qubits in the 3×3

TABLE IV. List of possible outcomes and the corresponding errors for the measurement of $\bar{X}_1\bar{X}_2$ in various scenarios. $\bar{Z}_i$ denotes a Z error on the i th row, regardless of the exact position of the error, because they are equivalent up to a stabilizer operator.

Initial error	Example	X syn	Projected error
No damped row	$\bar{I}_1\bar{I}_2$	+1	$\bar{I}_1\bar{I}_2$
Z error	$\bar{Z}_1\bar{I}_2$	-1	$\bar{Z}_1\bar{I}_2$
1 damped row	$(\bar{I}_1 - \bar{Z}_1)\bar{I}_2$	+1	$\bar{I}_1\bar{I}_2$
1 damped row	$(\bar{I}_1 - \bar{Z}_1)\bar{I}_2$	-1	$\bar{Z}_1\bar{I}_2$
2 damped rows	$(\bar{I}_1 - \bar{Z}_1)(\bar{I}_2 - \bar{Z}_2)$	+1	$\bar{I}_1\bar{I}_2$ or $\bar{Z}_1\bar{Z}_2$
2 damped rows	$(\bar{I}_1 - \bar{Z}_1)(\bar{I}_2 - \bar{Z}_2)$	-1	$\bar{Z}_1\bar{I}_2$ or $\bar{I}_1\bar{Z}_2$

lattice as q_{ij} , for $i, j = 1, 2, 3$, with i (j) as the row (column) index. The set of stabilizer generators are

$$S = \{Z_{11}Z_{12}, Z_{12}Z_{13}; Z_{21}Z_{22}, Z_{22}Z_{23}; Z_{31}Z_{32}, Z_{32}Z_{33}; \bar{X}_1\bar{X}_2, \bar{X}_2\bar{X}_3\}, \quad (\text{A3})$$

and we have the logical states written explicitly as

$$\begin{aligned} |0\rangle_L &= \frac{1}{2}(|0_r 0_r 0_r\rangle + |0_r 1_r 1_r\rangle + |1_r 0_r 1_r\rangle + |1_r 1_r 0_r\rangle), \\ |1\rangle_L &= \frac{1}{2}(|1_r 1_r 1_r\rangle + |1_r 0_r 0_r\rangle + |0_r 1_r 0_r\rangle + |0_r 0_r 1_r\rangle), \end{aligned} \quad (\text{A4})$$

where $|0\rangle_r \equiv |000\rangle$ and $|1\rangle_r \equiv |111\rangle$ denote the three-qubit states for each row and $|\cdot\rangle_r$ are for the three rows. The $|\pm\rangle_L$ states can be written as

$$|\pm\rangle_L = \frac{1}{2\sqrt{2}}(|0\rangle_r \pm |1\rangle_r)^{\otimes 3}. \quad (\text{A5})$$

We now describe the three steps of the ideal EC procedure for this 3×3 example.

(1) *Damping-error syndrome extraction*: We measure the six parity-check operators (two per row) $\{Z_{11}Z_{12}, Z_{12}Z_{13}; Z_{21}Z_{22}, Z_{22}Z_{23}; Z_{31}Z_{32}, Z_{32}Z_{33}\}$. For each row, if at least one of the measurement outcomes is -1 , we declare that the row is damped and then measure all three qubits in that row in the Z basis to determine the locations of damping errors. The row is otherwise undamped.

As an example, consider the case where there is one damping error on q_{11} versus the case with damping errors on both q_{12} and q_{13} . In both cases, $(Z_{11}Z_{12}, Z_{12}Z_{13}) = (+1, -1)$, but in the first case, $(Z_{11}, Z_{12}, Z_{13}) = (-1, +1, +1)$, while in the second case $(Z_{11}, Z_{12}, Z_{13}) = (+1, -1, -1)$. Measuring Z's on the individual qubits on the damped row thus locates the damping errors.

Once the damping locations have been identified, we are ready to correct for the X part of the damping error $E = X(I - Z)$ by applying X to the damping locations, leaving the $(I - Z)$ part for the subsequent steps. Alternatively, we can delay the application of these local X's until the end of the EC procedure. This will not affect the detection of Z errors in the next step.

(2) *Z-error syndrome extraction*: We next measure the two check operators $\{\bar{X}_1\bar{X}_2, \bar{X}_2\bar{X}_3\}$. The outcome $\bar{X}_1\bar{X}_2 = +1$ (-1) means that the parity, in the $\bar{X}$ basis, of the first and second rows is the same (different). Table IV shows the different possible syndrome outcomes for $\bar{X}_1\bar{X}_2$ and the corresponding errors. The same holds for $\bar{X}_2\bar{X}_3$.

In the case of a damping error $\mathcal{F}_a$, the measurements project the state to the code space or to error spaces with different Z errors. In the case of an off-diagonal $\mathcal{F}_z$, we can consider the following examples to highlight the discussion of the previous section:

(a) *Odd number of $\mathcal{F}_z$'s on one row:* $\mathcal{F}_z$ on q_{11} introduces the term $Z_{11}(\cdot) + (\cdot)Z_{11}$, eliminated by $\bar{X}_1\bar{X}_2$.

(b) *Even number of $\mathcal{F}_z$'s on one row:* Two $\mathcal{F}_z$'s on q_{11} and q_{12} introduce the term $Z_{11}Z_{12}(\cdot) + (\cdot)Z_{11}Z_{12} + Z_{11}(\cdot)Z_{12} + Z_{12}(\cdot)Z_{11} \propto \bar{I}_1(\cdot)\bar{I}_1 + \bar{Z}_1(\cdot)\bar{Z}_1$. This is an incoherent mixture of $\bar{I}_1$ and $\bar{Z}_1$ and gets detected as a Z error.

(c) *Even number of $\mathcal{F}_z$'s on two rows:* One $\mathcal{F}_z$ on q_{11} and another $\mathcal{F}_z$ on q_{21} introduce the term $Z_{11}Z_{21}(\cdot) + (\cdot)Z_{11}Z_{21} + Z_{11}(\cdot)Z_{21} + Z_{21}(\cdot)Z_{11}$. The measurement of $\bar{X}_1\bar{X}_2$ with the outcome $+1$ (-1) projects out the term $Z_{11}Z_{21}(\cdot) + (\cdot)Z_{11}Z_{21}$ ($Z_{11}(\cdot)Z_{21} + Z_{21}(\cdot)Z_{11}$). Now, there are off-diagonal terms remaining, but the measurement $\bar{X}_2\bar{X}_3$ eliminates these terms as it sees only one $\mathcal{F}_z$ on the second row.

(3) *Decoding and recovery:* As mentioned earlier, the X part of the damping error can be corrected using the information about damping locations from the damping-error syndrome extraction step. Meanwhile, decoding for Z error involves both the syndrome obtained from the measurement of $\{\bar{X}_1\bar{X}_2, \bar{X}_2\bar{X}_3\}$ and the information about damped rows obtained from the damping-error syndrome extraction unit.

To see why the latter is needed, we consider the following two scenarios, both being second-order events: Suppose in both scenarios, we obtain the Z -error syndrome $(\bar{X}_1\bar{X}_2, \bar{X}_2\bar{X}_3) = (+1, -1)$, which means the error is either $\bar{Z}_1\bar{Z}_2\bar{I}_3$ or $\bar{I}_1\bar{I}_2\bar{Z}_3$.

(a) *Scenario 1:* There is one damping error on the first row (any qubit) and another on the second. From Table IV, we deduce that $\bar{Z}_1\bar{Z}_2\bar{I}_3$ is the error and there is no way $\bar{I}_1\bar{I}_2\bar{Z}_3$ is projected out in this case.

(b) *Scenario 2:* There is one second-order Z error on the third row. In this case, the syndrome is always $(\bar{X}_1\bar{X}_2, \bar{X}_2\bar{X}_3) = (+1, -1)$ and the error is $\bar{I}_1\bar{I}_2\bar{Z}_3$.

This example highlights the fact that the Z -error syndrome alone is insufficient to distinguish all the correctable errors; the information about the damped rows must also be included. The general rule is that, between two errors compatible with the Z -error syndrome, we choose the one with the fewer number of Z errors on the undamped rows.

APPENDIX B: FAULT-TOLERANCE PROPERTIES

Here, we state the properties that the EC and logical gadgets must have for fault-tolerant operation against amplitude-damping noise.

An amplitude-damping channel affecting n qubits locally can be understood as n amplitude-damping errors affecting the data qubits, with n_1 first-order $\mathcal{F}_a$ or $\mathcal{F}_z$ errors, and n_2 second-order Z errors such that $n_1 + 2n_2 \leq n$. We use the term *fault* to refer to the scenario where noise afflicts a physical location within a gadget and *error* to refer to the resulting change on the quantum state.

For fault-tolerant operation against amplitude-damping noise, our gadgets must have the following properties (for more details, see Chapter 2 in Ref. [47]):

(P1) An EC gadget is fault tolerant against t errors if, for an input with s errors and an EC gadget with r faults such that $s + r \leq t$, the output state differs from a codeword by at most r errors and ideally decoding the output state gives the same codeword as ideally decoding the input state.

(P2) A preparation gadget is fault tolerant against t errors if, for $r \leq t$ faults in the gadget, the output differs from a codeword by at most r errors and ideally decoding the output state gives the same codeword as prepared by an ideal preparation.

(P3) A measurement gadget is fault tolerant against t errors if, for an input with s errors and a measurement gadget with r faults such that $s + r \leq t$, ideally decoding the outcomes gives the same result as ideally decoding and measuring the input state.

(P4) A k -qubit gate gadget is fault tolerant against t errors if, for input blocks with $s_1, s_2, \dots, s_k$ errors and a gadget with r faults such that $\sum_{i=1}^k s_i + r \leq t$, each output block has at most r errors, and jointly decoding the output blocks by combining their classical syndromes gives the same state as ideal decoding of the input blocks followed by the ideal k -qubit operation.

In the last property, jointly decoding the output blocks implies that the syndrome information from all the blocks can be classically combined to decode correlated errors due to error propagation within the logical gadget. This modification does not affect the fault-tolerance property of the whole circuit and is used when we construct the logical CZ gadget later.

For fault-tolerant simulation of an ideal circuit, we replace each operation in the ideal circuit by the corresponding logical gadget followed by an EC gadget in each output block (except the measurement gadget that has no output). To discuss the robustness of the simulation circuit, we consider *extended gadgets*. An *extended gadget* is defined as a logical gadget together with the leading and trailing EC gadgets. It can be shown that if there are no more than t faults inside every extended gadget of the simulation circuit, then the properties of the simulation gadgets stated above will ensure that the circuit simulates the ideal circuit correctly. This was proven rigorously in Ref. [47] for arbitrary local noise and applies to our situation of local amplitude-damping noise.

In particular, for the memory setting, where we want to store information for a certain time, the argument can be understood as follows: An extended memory gadget includes a waiting step sandwiched between two EC gadgets. Note that the extended gadgets are not separated, rather they overlap: The leading EC of one extended gadget is the trailing EC of the previous extended gadget. Assume there are no more than t faults inside any extended gadget. Consider a particular extended gadget with s errors in the incoming state, r_1, r_2, r_3 faults in the leading EC, the waiting step, and the trailing EC, respectively, such that $r_1 + r_2 + r_3 \leq t$. Assume that $s + r_1 \leq t$. Property (P1) then ensures that there are at most r_1 errors in the output of the leading EC, thus at most $r_1 + r_2 \leq t$ errors in the input to the trailing EC. Invoking property (P1) one more time, we see that the output of the trailing EC has at most r_3 errors. We also see that the incoming state to the following extended gadget, which is the input state to the current trailing EC, has at most $r_1 + r_2$ errors. If we now consider the extended gadget that follows, the total

number of errors in the incoming state and faults in the leading EC is $r_1 + r_2 + r_3 \leq t$. This justifies our assumption at the beginning that $s + r_1 \leq t$ since the previous extended gadget behaves in exactly the same way. By stringing together the chain of extended gadgets, we can conclude that the final output state has no more than t errors; thus, an ideal decoder attached at the end will decode it correctly.

APPENDIX C: PROOF OF LEMMA 1

Here, we provide a proof of Lemma 1 stated in the main text.

Lemma 1. Let e_i and e_{i+1} be the X -type error sequences for a pair of successive rows, caused by $\ell \leq t$ errors in the $\overline{XX}$ -stabilizer measurement. It is possible to ensure $\text{DIFF}(e_i, e_{i+1}) \leq \ell$ by fixing the order in which the ancillary qubit is coupled to the extended data qubits, as given in Fig. 4.

Proof. First, note that a damping error on the flag qubit does not propagate any error to the extended data qubits. It can only propagate Z -type error to the ancillary qubit and corrupt its measurement outcome.

Next, consider damping errors only in the ancillary qubit. As damping errors beyond the first one have no additional effect on the qubit, we need only consider a single $\mathcal{F}_a$ error. In general, an $\mathcal{F}_a$ error on the ancillary qubit can propagate X errors to multiple extended data qubits in a row, possibly causing more than t qubits in that row to be labeled with an X . However, we can ensure $\text{DIFF}(e_i, e_{i+1}) \leq 1$ for every pair of rows by a specific choice of the order in which the ancillary qubit is coupled to the extended data qubits. The zig-zag coupling pattern illustrated in Fig. 4 suffices. It is easy to see why this works using the example of the 3×3 code in Fig. 4. If the fault on the ancillary qubit occurs right after coupling with qubit 7, then the corresponding propagated error sequences are (taking $i = 1$) $e_1 = XXIII$ and $e_2 = XXIII$, and hence $\text{DIFF}(e_1, e_2) = 0$. On the other hand, if the fault occurs right after qubit 3, then $e_1 = XXXII$, $e_2 = XXIII$, and $\text{DIFF}(e_1, e_2) = 1$. In general, the zig-zag coupling pattern ensures that a fault in the ancillary qubit propagates X errors to two rows in a zig-zag manner, thus ensuring $\text{DIFF}(e_i, e_{i+1}) \leq 1$.

If, instead, there are faults only in the extended data qubits and no faults in the ancillary qubit, there are at most ℓ qubits marked as X and it is straightforward to see that $\text{DIFF}(e_i, e_{i+1}) \leq \ell$ for any pair of rows.

More generally, in addition to a fault in the ancillary qubit [which gives $\text{DIFF}_1(e_i, e_{i+1}) \leq 1$], we can allow for at most $\ell - 1$ faults distributed among the extended data qubits, leading to $\text{DIFF}_2(e_i, e_{i+1}) \leq \ell - 1$. Putting these together, the DIFF function for a pair of error sequences satisfies $\text{DIFF}(e_i, e_{i+1}) \leq \text{DIFF}_1(e_i, e_{i+1}) + \text{DIFF}_2(e_i, e_{i+1}) \leq \ell$. ■

APPENDIX D: OPTIMIZING SYNDROME EXTRACTION

In Sec. III, we constructed an EC gadget that is fault tolerant against $\leq t$ amplitude-damping faults. There, we did not consider questions of optimality and resource costs. Here, we carry out a more careful analysis that gives an improved (lower) number of repetitive syndrome measurements needed for fault-tolerant operation.

We recall that, in the EC gadget, there are three syndrome extraction steps that have to be repeated:

(1) *Damping-error syndromes:* We repeat the parity measurements on the data qubits of a row at most t times to verify that the trivial syndrome sequence is correctly diagnosed.

(2) *$\overline{XX}$ measurements:* In the parity-check step after the $\overline{XX}$ measurement, we repeat the parity measurements on the extended data qubits of a row until a syndrome sequence appears $t + 1$ times.

(3) *Decoding Z errors:* While decoding the Z errors, we repeat the Z -syndrome measurement until a syndrome sequence appears $t + 1$ times.

All three steps above can be cast into the problem of fault-tolerant syndrome measurement for the repetition code. In steps (1) and (2), the repetition code is defined on a row of qubits with ZZ stabilizers between two consecutive qubits; in step (3), the repetition code is defined across the rows of the lattice with $\overline{XX}$ stabilizers between two consecutive rows. Also, in step (1), we have to verify that the syndrome sequence is indeed the trivial syndrome (all $+1$ s), whereas, in steps (2) and (3), we have to verify an *a priori* unknown syndrome sequence.

For step (1), if the input state has damping errors but a trivial syndrome sequence is obtained, there must be at least one fault in that round of measurements. Repeating t times, while continuing to obtain a trivial syndrome sequence, means that the last round must be free of measurement faults, and damping errors in the input can be detected since there is at least one error in the input and there are $t - 1$ faults in the first $t - 1$ rounds; if this is not the case, we would have had $> t$ faults, beyond the requirements of fault tolerance.

For steps (2) and (3), we prove the following estimate on the required number of syndrome extraction rounds:

Lemma 2 (Fault-tolerant syndrome extraction). To ensure fault tolerance of the EC gadget, it is sufficient to repeat each of the steps (2) and (3) no more than $(1/4)(t + 2)^2 + 1$ times.

Proof. We want to estimate the minimum number of rounds of syndrome extraction required to make the EC gadget fault tolerant. For step (2) or (3), the true syndrome sequence is unknown and we choose a sequence based on majority voting. If a sequence appears $t + 1$ times, it is safe to decode based on that sequence as there must be at least one round without faults, or we violate the $\leq t$ faults precondition for fault tolerance. We imagine a worst-case scenario where a syndrome sequence appears t times before a fault happens and changes the syndrome to another one. This can repeat $t + 1$ times as we allow t faults and one no-fault case. Therefore, naively, if we repeat $t(t + 1) + 1$ times, we are assured that there will be one sequence appearing $t + 1$ times.

However, we notice that there are two types of faults. For the first type, a fault only introduces new to the data qubits without affecting the measurement outcomes in a round. They are not harmful because even if we decode using the obtained syndrome, the decoded error differs from the true error by no more than the number of faults in that round. For the second type, a fault changes the measurement outcomes in a round with or without introducing errors to the data qubits. In this case, we are at risk of introducing more errors if using the obtained syndrome to decode.

We consider a more specific scenario where there are r_1 type-1 faults and r_2 type-2 faults such that $r_1 + r_2 = r \leq t$. We imagine, in the worst-case scenario, a faulty syndrome sequence appears r_2 times due to r_2 faults of the second type. We have to repeat enough times so that at least one syndrome sequence without faults appears more than r_2 times. Following the earlier argument, it suffices to repeat $N_{\text{rep}} = r_2 + r_2(r_1 + 1) + 1 = -r_2^2 + (r + 2)r_2 + 1$ times since we are then assured that there is one sequence without type-2 faults appearing $r_2 + 1$ times. Maximizing N_{rep} with respect to r_2 gives $N_{\text{rep}}^{\text{max}} = (1/4)(r + 2)^2 + 1 \leq (1/4)(t + 2)^2 + 1$. Therefore, repeating $(1/4)(t + 2)^2 + 1$ times and choosing to decode the syndrome that appears most often is sufficient for fault tolerance. ■

We conclude by discussing two further approaches to reduce the number of syndrome measurement repetitions, thereby making the EC gadget less costly.

(1) *Redundant parity measurement*: We can detect faulty syndrome measurements by adding one more redundant parity measurement between the first and the last qubit [or between the first and the last row in case (3)]. One round of measurements, therefore, involves $t + 1$ parity measurements, giving a length- $(t + 1)$ syndrome sequence of ± 1 .

In the absence of faults in the ancilla measurements, i.e., none of the measurement outcomes are flipped, the sequence always has an even number of -1 s, regardless of the number of faults in the data qubits. An odd number of -1 s arises only if there are measurement faults. Therefore, we call a syndrome sequence with an even number of -1 s a *valid* syndrome sequence and one with an odd number of -1 s an *invalid* one. Obtaining an invalid sequence thus implies that at least one measurement fault has occurred during the current round of measurements. Also, an even number of measurement faults are needed to turn a valid syndrome sequence into another valid syndrome sequence. It means that every time we obtain a valid sequence, either it is the correct sequence without measurement faults or there are at least two measurement faults in the sequence. Effectively, we can reduce the number of repetitions by half.

We then have the following situation for the three steps above: For step (1): In the syndrome extraction, if the trivial sequence is obtained, we repeat the parity measurements up to $\lfloor t/2 \rfloor + 1$ times before concluding that there are no damping errors. Whenever we obtain an invalid sequence, we conclude that there is at least one fault and continue as if we obtained a nontrivial sequence. For steps (2) and (3), replacing r_2 by $r_2/2$ in the argument in the proof above, we arrive at $(1/8)(t + 2)^2 + 1$ for the maximum number of repetitions. So, for the parity-check step and the Z -syndrome measurement, we repeat $(1/8)(t + 2)^2 + 1$ times and choose the majority among the set of valid syndromes.

(2) *Keeping track of the number of faults*: Up till now, we have allowed for up to t faults in the syndrome measurements. However, if we keep track of the number of faults that have already occurred in the EC gadget, then we can further reduce the number of repetitions needed before we violate the fault-tolerance precondition of $\leq t$ faults. For example, if there are at least t_1 faults already, then the syndrome measurements

need to be fault tolerant against only $t - t_1$ faults, thereby reducing the number of measurement repetitions.

APPENDIX E: BASIC FAULT-TOLERANT COMPONENTS

We want to construct a universal set of logical gadgets that are fault tolerant. To do that, we first need some basic fault-tolerant operations, which we describe in this section. In the next section (Appendix F), we make use of these basic building-block operations to construct our logical gadgets. The basic operations we will construct here are the preparation and measurement gadgets.

1. Measurement gadgets

Next, we describe fault-tolerant gadgets that implement the destructive measurements of logical Z and X .

a. Logical- Z measurement

The logical- Z measurement can be realized by measuring each physical qubit in the Z basis. In the absence of any errors, the Z measurements on one row of qubits will give either all -1 s or all $+1$ s [see Eq. (5)], which we refer to as a (-1) -row or a $(+1)$ -row, respectively. The outcome of the logical- Z measurement is $+1$ if there is an even number of $(+1)$ -rows over the whole lattice [again, see Eq. (5)]; it is -1 if there is an odd number of them.

That this simple logical- Z measurement is fault tolerant can be seen as follows: We first note that we need only consider damping errors $\mathcal{F}_a$'s since $\mathcal{F}_Z$ and Z do not affect the Z measurement. A faulty physical Z measurement damps a qubit to the $|0\rangle$ state and gives the outcome $+1$. A damped row is thus a mixture of -1 s and $+1$ s. For at most t faults occurring during the measurements, a damped row has at most $t - 1$ s and at least one 1 since there are $t + 1$ qubits in a row. A damped row is therefore always reliably detected. Such a damped row should be regarded as $(+1)$ -row when deducing the logical- Z measurement outcome since a damping error has effect only when the qubit is originally in the state $|1\rangle$.

b. Logical- X measurement

The logical- X measurement can be similarly realized by measuring each physical qubit in the X basis. However, unlike the logical- Z measurement, this transversal implementation is not fault tolerant. This can be understood by noticing that, in the no-fault situation, the logical- X outcome is equal to the parity of the measurement outcomes of any one row of qubits. A single damping error in a row randomizes the parity of that row. As we allow for damping in more than half the rows, we cannot deduce the correct logical outcome even by looking at the individual parity values for all rows.

Instead, we use the circuit in Fig. 7 for a fault-tolerant logical- X measurement. For each row, an ancilla initialized in $|+\rangle$ is coupled to all qubits in the row via CNOT gates. Then, we disentangle one data qubit, say the first qubit, from the other qubits in the row. We then measure this qubit and the ancilla in the X basis. The other qubits in the row are measured in the Z basis.

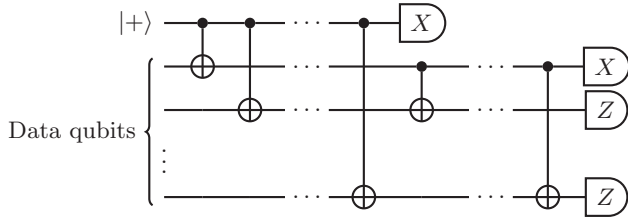


FIG. 7. Logical-X measurement on a single row. The logical outcome is given by the outcomes of the two X measurements. The row is *valid* only if both the X measurements give the same outcome and every Z measurement gives +1.

For input states $|\pm\rangle_L$, the only states of relevance to the logical-X measurement, the joint states of the row with the ancilla (denoted by a) before the final X and Z measurements are, in the absence of any faults, $|\pm\rangle_a |\pm\rangle |0\rangle^{\otimes t}$. The outcome of the ancilla or the first qubit is, therefore, the logical outcome. If there are faults, the measurement outcomes can deviate from the ideal case. We say that a row is invalid if at least one of the Z measurements has outcome -1 or if the outcome of the ancilla measurement does not match the outcome of the first data qubit. The logical outcome is then determined as the majority outcome among the valid rows.

The fault tolerance of the measurement circuit in Fig. 7 can be understood as follows: For each row, a single damping error or a single $\mathcal{F}_Z$ anywhere either makes the row invalid or does not affect the outcome of the measurements performed on the row, as explained below.

(1) A fault in the initialization or measurement of the ancilla may randomize the outcome of the ancilla measurement, but it has no effect on the data qubits. Therefore, if the outcome of the ancilla measurement matches that of the first data qubit, the latter will be correct.

(2) A fault in the measurement of the first data qubit also randomizes its outcome, but it has no effect on the ancilla's outcome. Therefore, the row's outcome is correct if the two outcomes match.

(3) A fault in the CNOTs or an error in the incoming state results in at least one of the Z measurements having outcome 1.

A second-order error, which can be a Z error, two $\mathcal{F}_Z$ errors, or two $\mathcal{F}_a$ errors, might lead to a valid row with an incorrect outcome. For s errors in the input and r faults in the measurement gadget such that $s + r \leq t$, more than half the rows will still be valid and correct, and hence, the logical outcome is always determined correctly from majority voting.

2. Preparation gadgets

Finally, we describe fault-tolerant preparation of the logical state $|+\rangle_L$.

a. Preparation of cat state

The $(t + 1)$ -qubit cat state is defined as

$$|\beta_0\rangle \equiv \frac{1}{\sqrt{2}}(|0\rangle^{\otimes(t+1)} + |1\rangle^{\otimes(t+1)}). \quad (\text{E1})$$

This state serves as the cornerstone for the preparation of other logical states, for example, the logical $|+\rangle_L$ and the

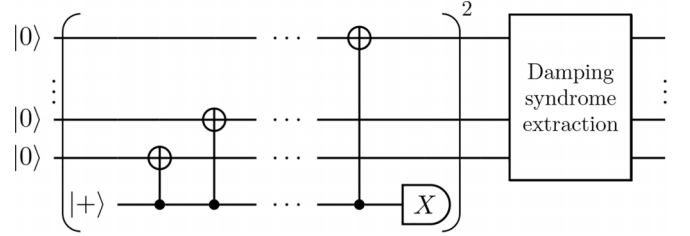


FIG. 8. Preparation and verification of the cat state. The $X_1 X_2 \dots X_{t+1}$ measurement is done twice followed by the damping syndrome extraction unit. The output state is accepted if all the measurements give +1 outcomes.

logical resource states (see Appendix F). To prepare this state, the idea is to measure the operator $X_1 X_2 \dots X_{t+1}$ on the input state $|0\rangle^{\otimes(t+1)}$ and postselect the outputs based on the measurement outcomes. We measure $X_1 X_2 \dots X_{t+1}$ using the circuit in Fig. 8, where an ancillary qubit initialized in the $|+\rangle$ state is coupled to the data qubits through CNOTs and finally measured in the X basis. We do the measurement twice (see below for the reason); i.e., the outgoing state after the first measurement becomes the input for the second measurement. Finally, we pass the outgoing state after the second measurement through the damping syndrome extraction unit, as described in Sec. III A for the EC gadget. The final output state is accepted only if all the measurements, namely, two X-type measurements and all the parity measurements, give +1 outcomes.

That this preparation procedure satisfies the fault-tolerance condition (P2) of Appendix B can be understood as follows: We consider up to t faults anywhere in the preparation gadget, which includes two $X_1 X_2 \dots X_{t+1}$ measurements and the damping syndrome extraction unit. We will check that the gadget either accepts the output state with the correct fault-tolerance property or rejects it.

(1) Z or $\mathcal{F}_Z$ errors pass through all the checks but they do not spread. Hence, t Z or $\mathcal{F}_Z$ errors remain as t Z or $\mathcal{F}_Z$ errors in the output state. The errors may change the outcome of the X measurements, but they have no effect on the output state.

(2) Damping errors $\mathcal{F}_a$ on the data qubits during the two measurements cause $\mathcal{F}_a$ and/or X errors in the outgoing state of the second measurement. Damping of the ancillary qubits in the middle of the measurements also propagates X errors to the data qubits. The damping syndrome extraction unit (even with faults inside it) ensures that these errors will be detected and the output state will be rejected. Moreover, s faults inside the damping syndrome extraction unit lead to at most s damping errors in the output state (see Sec. III A). Thus, if there are only faults inside the syndrome extraction, the output state is accepted with no more errors than the number of faults inside the unit.

(3) More worrying are damping errors after the preparation of the first ancillary qubit or right before the X measurement of the first ancillary qubit. The outgoing state of the first $X_1 X_2 \dots X_{t+1}$ measurement is $|0\rangle^{\otimes(t+1)}$ for the first case and $|1\rangle^{\otimes(t+1)}$ for the second. Note that both states pass the damping syndrome extraction unit. Moreover, $|0/1\rangle^{\otimes(t+1)}$ can be thought of as the cat state $|\beta_0\rangle$ with $(I \pm Z)$ error. Therefore, if we stop after the first measurement, there might

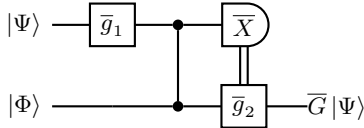


FIG. 9. Single-logical-qubit gadget implementing the gate $\bar{G} = \bar{H}, \bar{S},$ or $\bar{T}$. $|\Psi\rangle$ is the incoming state on which $\bar{G}$ is to be applied. For $\bar{G} = \bar{H}$, $|\Phi\rangle = |+\rangle_L$, $\bar{g}_1$ is the identity, and $\bar{g}_2 = \bar{X}$; for $\bar{G} = \bar{S}$, $|\Phi\rangle = |\Phi_S\rangle$, $\bar{g}_1 = \bar{H}$, and $\bar{g}_2 = \bar{Y} = \bar{X}\bar{Z}$; for $\bar{G} = \bar{T}$, $|\Phi\rangle = |\Phi_T\rangle$, $\bar{g}_1 = \bar{H}$, and $\bar{g}_2 = \bar{S}\bar{X}$. The gate $\bar{g}_2$ is applied, conditioned on the outcome of the $\bar{X}$ measurement. We note that the $\bar{X}$ and $\bar{Z}$ contained in the $\bar{g}_2$'s do not need to be applied physically; see the main text.

be a chance that one damping error causes a Z error in the output state and there is no way to tell whether this error comes from a first-order damping error. This violates the fault-tolerance property (P2) stated in Appendix B. The second $X_1 X_2 \cdots X_{t+1}$ measurement is added to remedy this. It projects the $|0\rangle_L^{\otimes(t+1)}$ state to the cat state (half of the time) and hence removes the Z error. The $|0\rangle_L^{\otimes(t+1)}$ state is still possible as an output state if both of the ancillary qubits are damped. However, this is a second-order event and the fault-tolerance condition is still fulfilled.

b. Preparation of $|+\rangle_L$

Observe that $|+\rangle_L = |\beta_0\rangle^{\otimes(t+1)}$. Thus, the preparation of $|+\rangle_L$ can be implemented by fault tolerantly preparing $t+1$ copies of the cat state.

APPENDIX F: A UNIVERSAL SET OF LOGICAL GADGETS

We are now ready to construct a universal set of logical gadgets, using the basic fault-tolerant components of the previous section. In particular, our universal set of gadgets comprises

- (1) preparation of the $|0\rangle_L$ and $|+\rangle_L$ states;
- (2) measurement of logical X and Z ;
- (3) two-qubit logical gate CZ ;
- (4) single-qubit logical gates H , S , and T .

Here, the logical H is the Hadamard gate, $\bar{H} \equiv |+\rangle_L\langle 0| + |-\rangle_L\langle 1|$; the logical S is the phase gate, $\bar{S} \equiv |0\rangle_L\langle 0| + i|1\rangle_L\langle 1|$; and the logical T (or $\pi/8$) is the gate $\bar{T} \equiv |0\rangle_L\langle 0| + e^{i\pi/4}|1\rangle_L\langle 1|$.

We note that our construction requires no logical X or Z gates. $\bar{X}$ and $\bar{Z}$ can be tracked virtually when the operations are all Clifford, with no need for physical gates. In the presence of the single non-Clifford logical T gate, we note that $\bar{T}\bar{Z}\bar{T}^\dagger = \bar{Z}$, while $\bar{T}\bar{X}\bar{T}^\dagger = \bar{S}\bar{X}$. $\bar{Z}$ thus simply commutes past $\bar{T}$, while for the $\bar{X}$, we apply the $\bar{S}$ and commute the remaining $\bar{X}$ past the operations.

Apart from the preparation of $|0\rangle_L$ and the single-logical-qubit gadgets, all the other gadgets are already constructed in the earlier sections of the paper. We focus our attention on the single-logical-qubit gates first. Our construction of all three single-logical-qubit gadgets makes use of the same teleportation structure shown in Fig. 9. Each requires a resource state as input: $|+\rangle_L$ for $\bar{H}$, $|\Phi_S\rangle \equiv (1/\sqrt{2})(|0\rangle_L + i|1\rangle_L)$ for $\bar{S}$, and $|\Phi_T\rangle \equiv (1/\sqrt{2})(|0\rangle_L + e^{i\pi/4}|1\rangle_L)$ for $\bar{T}$. The circuits also require the use of the logical- X measurement gadget,

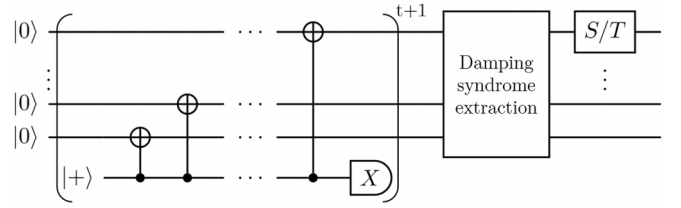


FIG. 10. Preparation of single-row states $|\beta_{S/T}\rangle$. The $X_1 X_2 \cdots X_{t+1}$ measurement is done $t+1$ times before applying a damping syndrome extraction and a single-qubit gate S/T . The output state is accepted if all the measurements have $+1$ outcomes.

as well as an additional logical gadget, conditioned on the measurement outcome, to complete the gate teleportation.

Note that we already have fault-tolerant gadgets for every operation in the gate-teleportation circuits in Fig. 9, as well as for the preparation of $|+\rangle_L$. We thus already have a fault-tolerant $\bar{H}$. This also gives us a fault-tolerant preparation of $|0\rangle_L$: We apply the fault-tolerant $\bar{H}$ to the fault-tolerantly prepared $|+\rangle_L$. What remains then is to describe the fault-tolerant preparation of $|\Phi_S\rangle$ and $|\Phi_T\rangle$. The fault tolerance of the teleportation gadgets then simply follows from the fact that their components are all separately fault tolerant.

The preparation of $|\Phi_S\rangle$ and $|\Phi_T\rangle$, written together as $|\beta_{S/T}\rangle$, starts by first preparing the single-row states $|\beta_{S/T}\rangle$, defined as

$$\begin{aligned} |\beta_S\rangle &= \frac{1}{\sqrt{2}}(|0\rangle^{\otimes(t+1)} + i|1\rangle^{\otimes(t+1)}), \\ |\beta_T\rangle &= \frac{1}{\sqrt{2}}(|0\rangle^{\otimes(t+1)} + e^{i\pi/4}|1\rangle^{\otimes(t+1)}). \end{aligned} \quad (F1)$$

The idea is to prepare a cat state first and apply a single-qubit gate to get the correct phase factor. We use the circuit shown in Fig. 10, similar to the one used to prepare the cat state. We start by doing the measurement of $X_1 X_2 \cdots X_{t+1}$ a total of $t+1$ times, by feeding the outgoing state of each measurement as the input of the next measurement. The outgoing state of the last measurement is passed through the damping syndrome extraction unit and finally, we apply an S or a T gate (depending on whether we are preparing $|\beta_S\rangle$ or $|\beta_T\rangle$) to one of the qubits. We accept the output state to be properly prepared only when all the measurement outcomes are $+1$.

The fault tolerance of this preparation step can be understood in a very similar way to the analysis given for the cat state. The only difference is the number of repetitions for $X_1 X_2 \cdots X_{t+1}$ measurement. We noted previously that a damping error on the ancilla qubit may cause the outgoing state of a measurement to be one of the trivial states $|0\rangle_L^{\otimes(t+1)}$. Applying an S/T gate to this state is meaningless since the relative phase is already lost. Repeating the measurement t times ensures that at least one of them has its ancilla qubit undamped. The accepted output (if no damping after the damping extraction) is thus a genuine cat state and the single-qubit gate S/T should add the correct relative phase to the state.

Now, we need to transform $|\beta_{S/T}\rangle$ into the logical states $|\Phi_{S/T}\rangle$ on the full lattice. To do this, we first prepare the other t rows in a state denoted as $|0\rangle_L^{t \times (t+1)}$, that is, a logical $|0\rangle_L$ state of a $t \times (t+1)$ Bacon-Shor code. This can be done as

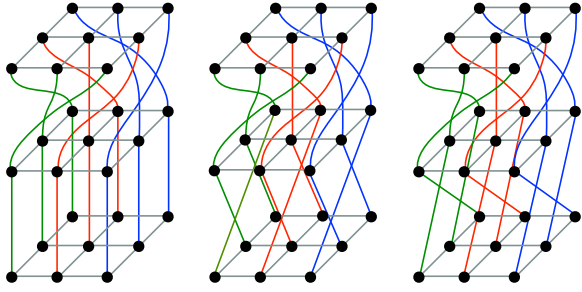


FIG. 11. Implementation of the encoded CCZ gate for the 3×3 $[(t+1) \times (t+1)]$ Bacon-Shor code in three $(t+1)$ steps, going from left to right. The qubits in different code blocks that participate in the same physical CCZ gate are linked by the same color.

follows: We prepare $|0\rangle_L$ fault tolerantly as described earlier and measure one row in the Z basis. If the outcomes are all -1 s, we obtain the desired state $|0\rangle_L^{t \times (t+1)}$; otherwise, if the outcomes are all 1 s or a mixture of 0 s and 1 s, we have, instead, the state $|1\rangle_L^{t \times (t+1)}$. In this case, we can simply keep track of the Pauli error in the software. Next, we measure the $\bar{X}\bar{X}$ stabilizer between the row prepared in the $|\beta_{S/T}\rangle$ state and any row of the $|0\rangle_L^{t \times (t+1)}$ state. This should be done using a subcircuit discussed earlier in Sec. III.

An accepted $|\beta_{S/T}\rangle$ state may still have damping errors caused by faults inside the damping extraction or the S/T gate at the end of the circuit in Fig. 10. In this case, the relative phase in the $|\beta_{S/T}\rangle$ state would be annihilated before it is transferred to the full lattice. The damping syndrome extraction in the subcircuit is therefore needed to detect this error. We thus accept the output state from the subcircuit only if its damping syndrome extraction unit detects no damping.

APPENDIX G: THE CCZ GADGET

The previous section gives a universal set of logical operations sufficient for arbitrary computation with the Bacon-Shor code in the presence of amplitude-damping noise. Nevertheless, we point out that one can expand the set of fault-tolerant operations to include a three-qubit logical gadget—the CCZ gate—on a general $(t+1) \times (t+1)$ Bacon-Shor code that is tolerant to $\leq t$ errors arising from the amplitude-damping noise. The CCZ gate also provides an alternate pathway to universality, since the CCZ together with the Hadamard gate implemented by teleportation forms a universal set of logical gates, without requiring the S and T gadgets described earlier.

Logical CCZ is implemented by applying a set of physical CCZs transversally for $t+1$ times. In each step, the rows of a control lattice are connected to the columns of the other control lattice and, in turn, the rows of the second control lattice are connected to the rows of the target lattice. In each step, the connections to the rows of the target lattice are permuted in a cyclic manner. Such a piecewise FT implementation has also been demonstrated in Ref. [48], for the standard Pauli-noise setting. However, our construction is instead specifically tailored to the case of amplitude-damping noise and the gauge-fixed Bacon-Shor code.

Figure 11 illustrates an example of the logical CCZ for the 3×3 Bacon-Shor code. Here, we intersperse the CCZ

gadget with EC units between two executions of the transversal CCZs. The syndrome bits allow us to decide how to correct the final qubits for a fault-tolerant CCZ. For instance, note that a damping error on a qubit in the first lattice could propagate a CZ operation on the qubits in the other two blocks,

$$\text{CCZ}(\mathcal{F}_a \otimes I \otimes I) \text{CCZ} = \mathcal{F}_a \otimes \text{CZ}. \quad (\text{G1})$$

(G1) In such a case, we use the syndrome bit information to decide how to apply CZ to undo the propagated errors. Other such checks can be done to ensure the fault tolerance of the gadget, but we leave the details for the reader to verify.

APPENDIX H: PSEUDOTHRESHOLD CALCULATION

Finally, we prove the result stated in the main text that the EC gadget has $O(t^6)$ fault locations; this was used in Sec. V to estimate the pseudothreshold. We first count the number of locations in a single EC subcircuit [see Fig. 2(b)], assuming the efficient syndrome extraction strategy described in Appendix D:

(1) *Coupling*: There are $2t+1$ locations per row, and hence $2(2t+1)$ locations in total.

(2) *Damping syndrome extraction*: For a single row, each round of parity measurements is done in three time steps, the first two steps involving the coupling of $t+1$ data qubits to $t+1$ ancilla and the last step is the measurement of the ancilla. During that time, the additional ancillary qubits are all idle. Therefore, there are $3(3t+2)$ locations per round. At most $\lfloor t/2 \rfloor + 1$ rounds are performed, and hence, there are $6(3t+2)(\lfloor t/2 \rfloor + 1)$ possible locations for both rows in a given subcircuit.

(3) *Recovery for X errors*: We can delay this step until the end of the subcircuit.

(4) *$\bar{X}\bar{X}$ -stabilizer measurement*: There are $2(2t+1)+2$ qubits involved (including extended data, ancillary, and flag qubits) and the measurement is done in $2(2t+1)+2$ time steps including ancilla-flag decoupling and measurements at the end. Together with the four locations in the preparation of ancillary and flag qubits, there are $16(t+1)^2+4$ locations in total.

(1) *Parity checks*: In this step, parity measurements are done on the full set of extended data qubits. Hence, each round has $3 \times 2(2t+1)$ locations. If we repeat the parity measurements until a valid syndrome sequence appears $(\lfloor t/2 \rfloor + 1)$ times, in the worst case we need $(t+1)(\lfloor t/2 \rfloor + 1) + 1$ rounds. A more optimized scheme needs only $\lfloor (t+2)^2/8 \rfloor + 1$ rounds. In total, there are at most $12(2t+1)(\lfloor (t+2)^2/8 \rfloor + 1)$ locations for both rows.

(2) *Decoupling and recovery*: We decouple and measure all the extra qubits in the Z basis. The recovery for X errors takes up one more time step. In total, there are $6(2t+1)$ locations.

Taking all the above steps together, in total there are at most $N_{\text{sub}} = 16t^2 + 72t + 40 + 12(2t+1)(\lfloor (t+2)^2/8 \rfloor + 1) + 6(3t+2)(\lfloor t/2 \rfloor + 1)$ locations in a subcircuit.

For odd t , there are an even number of rows in the lattice and we can perform $(t+1)/2$ subcircuits in parallel. Recall from Sec. III that one round of Z -syndrome extraction needs $t+1$ subcircuits and this is repeated until a valid sequence appears a certain number of times. Since a second-order

event is required to convert a valid Z -syndrome sequence into another valid sequence, in the worst case, similarly to the parity-measurement step in a subcircuit, we need to repeat it $\lfloor (t/2 + 2)^2/4 \rfloor + 1$ times. Therefore, in total, there are at most $N_{\text{odd}} = (\lfloor (t/2 + 2)^2/4 \rfloor + 1)(t + 1)N_{\text{sub}}$ locations in an EC gadget for odd t .

For even t , there are an odd number of rows in the lattice. Hence, we can perform $t/2$ subcircuits in parallel with one row left idle. The total number of locations in one round

is therefore $(t/2)N_{\text{sub}} + N_{\text{idle}}$, where $N_{\text{idle}} = (t + 1)(3(t + 1)\lfloor t/2 \rfloor + 4t + 9)$ is the number of locations in the idling row. As in the odd t case, we also need to repeat the syndrome extraction $\lfloor (t/2 + 2)^2/4 \rfloor + 1$ times in the worst case. However, in this case we can schedule the subcircuits so that after $t + 1$ parallel steps, we measure the Z syndrome $t/2$ times. Therefore, in total, there are roughly at most $N_{\text{even}} = \lfloor t/8 + 2 \rfloor (t + 1)[(t/2)N_{\text{sub}} + N_{\text{idle}}]$ locations in an EC gadget for even t .

-
- [1] E. Knill, R. Laflamme, and W. Zurek, Threshold accuracy for quantum computation, [arXiv:quant-ph/9610011](#).
 - [2] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, Limitations of noisy reversible computation, [arXiv:quant-ph/9611028](#).
 - [3] P. W. Shor, Fault-tolerant quantum computation, in *Proceedings of 37th Conference on Foundations of Computer Science* (IEEE, USA, 1996), p. 56.
 - [4] J. Preskill, Fault-tolerant quantum computation, in *Introduction to Quantum Computation and Information*, edited by H.-K. Lo (World Scientific, Singapore, 1998), pp. 213–269.
 - [5] D. Gottesman, Theory of fault-tolerant quantum computation, *Phys. Rev. A* **57**, 127 (1998).
 - [6] P. W. Shor, Scheme for reducing decoherence in quantum computer memory, *Phys. Rev. A* **52**, R2493 (1995).
 - [7] A. R. Calderbank and P. W. Shor, Good quantum error-correcting codes exist, *Phys. Rev. A* **54**, 1098 (1996).
 - [8] E. Knill and R. Laflamme, Theory of quantum error-correcting codes, *Phys. Rev. A* **55**, 900 (1997).
 - [9] D. Gottesman, Class of quantum error-correcting codes saturating the quantum Hamming bound, *Phys. Rev. A* **54**, 1862 (1996).
 - [10] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Perfect quantum error correcting code, *Phys. Rev. Lett.* **77**, 198 (1996).
 - [11] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Approximate quantum error correction can lead to better codes, *Phys. Rev. A* **56**, 2567 (1997).
 - [12] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, Experimental repetitive quantum error correction, *Science* **332**, 1059 (2011).
 - [13] Z. Chen *et al.*, Exponential suppression of bit or phase errors with cyclic error correction, *Nature (London)* **595**, 383 (2021).
 - [14] W. P. Livingston, M. S. Blok, E. Flurin, J. Dressel, A. N. Jordan, and I. Siddiqi, Experimental demonstration of continuous quantum error correction, *Nat. Commun.* **13**, 2307 (2022).
 - [15] D. Bluvstein *et al.*, Logical quantum processor based on reconfigurable atom arrays, *Nature (London)* **626**, 58 (2024).
 - [16] R. Acharya *et al.*, Quantum error correction below the surface code threshold, *Nature (London)* **638**, 920 (2025).
 - [17] I. Goulay and J. F. Snowdon, Concatenated coding in the presence of dephasing, *Phys. Rev. A* **62**, 022308 (2000).
 - [18] P. Aliferis and J. Preskill, Fault-tolerant quantum computation against biased noise, *Phys. Rev. A* **78**, 052331 (2008).
 - [19] P. Aliferis, F. Brito, D. P. DiVincenzo, J. Preskill, M. Steffen, and B. M. Terhal, Fault-tolerant computing with biased-noise superconducting qubits: A case study, *New J. Phys.* **11**, 013061 (2009).
 - [20] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh error threshold for surface codes with biased noise, *Phys. Rev. Lett.* **120**, 050505 (2018).
 - [21] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Tailoring surface codes for highly biased noise, *Phys. Rev. X* **9**, 041031 (2019).
 - [22] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang *et al.*, Extending the lifetime of a quantum bit with error correction in superconducting circuits, *Nature (London)* **536**, 441 (2016).
 - [23] C. Chamberland, K. Noh, P. Arrangoiz-Arriola, E. T. Campbell, C. T. Hann, J. Iverson, H. Putterman, T. C. Bohdanowicz, S. T. Flammia, A. Keller *et al.*, Building a fault-tolerant quantum computer using concatenated cat codes, *PRX Quantum* **3**, 010329 (2022).
 - [24] J. Guillaud and M. Mirrahimi, Repetition cat qubits for fault-tolerant quantum computation, *Phys. Rev. X* **9**, 041053 (2019).
 - [25] A. S. Darmawan, B. J. Brown, A. L. Grimsmo, D. K. Tuckett, and S. Puri, Practical quantum error correction with the xzzx code and Kerr-cat qubits, *PRX Quantum* **2**, 030345 (2021).
 - [26] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, Stabilization and operation of a Kerr-cat qubit, *Nature (London)* **584**, 205 (2020).
 - [27] A. Jayashankar, M. D. H. Long, H. K. Ng, and P. Mandayam, Achieving fault tolerance against amplitude-damping noise, *Phys. Rev. Res.* **4**, 023034 (2022).
 - [28] R. Lang and P. W. Shor, Nonadditive quantum error correcting codes adapted to the amplitude damping channel, [arXiv:0712.2586](#).
 - [29] A. S. Fletcher, P. W. Shor, and M. Z. Win, Channel-adapted quantum error correction for the amplitude damping channel, *IEEE Trans. Inf. Theory* **54**, 5705 (2008).
 - [30] A. Jayashankar, A. M. Babu, H. K. Ng, and P. Mandayam, Finding good quantum codes using the Cartan form, *Phys. Rev. A* **101**, 042307 (2020).
 - [31] D. Bacon, Operator quantum error-correcting subsystems for self-correcting quantum memories, *Phys. Rev. A* **73**, 012340 (2006).
 - [32] Á. Piedrafita and J. M. Renes, Reliable channel-adapted error correction: Bacon-Shor code recovery from amplitude damping, *Phys. Rev. Lett.* **119**, 250501 (2017).
 - [33] L. Chirrolli and G. Burkard, Decoherence in solid-state qubits, *Adv. Phys.* **57**, 225 (2008).
 - [34] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of quantum error-correcting codes for a bosonic mode, *Phys. Rev. X* **6**, 031006 (2016).

- [35] D. W. Kribs, R. Laflamme, D. Poulin, and M. Lesosky, Operator quantum error correction, [arXiv:quant-ph/0504189](#).
- [36] P. Aliferis and A. W. Cross, Subsystem fault tolerance with the Bacon-Shor code, *Phys. Rev. Lett.* **98**, 220502 (2007).
- [37] A. Shlosberg, A. M. Polloreno, and G. Smith, Towards demonstrating fault tolerance in small circuits using Bacon-Shor codes, [arXiv:2108.02079](#).
- [38] C. Gidney and D. Bacon, Less Bacon more threshold, [arXiv:2305.12046](#).
- [39] M. S. Alam, J. Chen, and T. R. Scruby, Bacon-Shor board games, [arXiv:2504.02749](#).
- [40] R. Duan, M. Grassl, Z. Ji, and B. Zeng, Multi-error-correcting amplitude damping codes, in *2010 IEEE International Symposium on Information Theory* (IEEE, USA, 2010), pp. 2672–2676.
- [41] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, Caltech, 1997.
- [42] R. Chao and B. W. Reichardt, Quantum error correction with only two extra qubits, *Phys. Rev. Lett.* **121**, 050502 (2018).
- [43] P. Aliferis, D. Gottesman, and J. Preskill, Quantum accuracy threshold for concatenated distance-3 codes, *Quantum Inf. Comput.* **6**, 97 (2006).
- [44] J. Napp and J. Preskill, Optimal Bacon-Shor codes, *Quantum Inf. Comput.* **13**, 490 (2013).
- [45] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais *et al.*, Bias-preserving gates with stabilized cat qubits, *Sci. Adv.* **6**, eaay5901 (2020).
- [46] Note that the code can actually tolerate up to $n - 1$ damping errors in each row, for any $n - 1$ rows, i.e., maximally $(n - 1)^2$ damping errors. However, it can only tolerate $n - 1$ damping errors in *arbitrary* locations, as there are cases where n damping errors cause a logical error, e.g., n errors in one row or one column.
- [47] P. Aliferis, Level reduction and the quantum threshold theorem, [arXiv:quant-ph/0703230](#).
- [48] T. J. Yoder, Universal fault-tolerant quantum computation with Bacon-Shor codes, [arXiv:1705.01686](#).