PHYSICS

Comparing Shor and Steane error correction using the Bacon-Shor code

Shilin Huang[1,2], Kenneth R. Brown[1,2,3,4], Marko Cetina[1,2,3]*

Quantum states decohere through interaction with the environment. Quantum error correction can preserve coherence through active feedback wherein quantum information is encoded into a logical state with a high degree of symmetry. Perturbations are detected by measuring the symmetries of the state and corrected by applying gates based on these measurements. To measure the symmetries without perturbing the data, ancillary quantum states are required. Shor error correction uses a separate quantum state for the measurement of each symmetry. Steane error correction maps the perturbations onto a logical ancilla qubit, which is then measured to check several symmetries simultaneously. We experimentally compare Shor and Steane correction of bit flip errors using the Bacon-Shor code implemented in a chain of 23 trapped atomic ions. We find that the Steane method produces fewer errors after a single round of error correction and less disturbance to the data qubits without error correction.

INTRODUCTION

Quantum computers promise to solve computationally difficult cryptographic and chemistry problems using billions or trillions of quantum gates (*1*, *2*). Current quantum devices are limited to hundreds of quantum gates, and many believe that the path forward is quantum error correction (EC), which enables us to make reliable qubits and gates out of unreliable components.

In the past few years, multiple experiments have demonstrated key elements of quantum EC. Highlights include improved state preparation with lower measurement error (*3*, *4*), extended coherence times (*5*), transversal logical gates (*3*, *6–8*), repeated rounds of syndrome extraction (*6*, *9–12*), magic-state injection (*8*), improved performance with code distance (*11*), and encoded circuits with multiple logical qubits (*13*). While no experiment has demonstrated an improvement in all aspects of an encoded qubit relative to its physical counterpart, these early works tested our assumptions about quantum EC protocols and the underlying noise models. These tests help us chart the path toward useful and scalable quantum EC.

Quantum EC encodes logical qubits into multiple physical qubits. The encoded code space has a number of symmetries whose actions are changed by the detectable errors. Stabilizer codes have binary symmetries corresponding to the +1 and −1 eigenvalues of the stabilizer operators, which form a commuting group of n-qubit Pauli operators (*14*). We measure the eigenvalues of a subset of the stabilizers that generates the whole code space. The error syndrome is a bit string where the 1's indicate that the corresponding stabilizer generator has an eigenvalue that is opposite of that of the code space.

A critical part of a quantum EC procedure is the fault-tolerant measurement of the error syndrome. We use the term Shor EC for all approaches that measure the stabilizer generators of the error syndrome one at a time. Traditionally, this requires the preparation of an entangled cat state as an encoded ancilla qubit for each generator (*15*). For some codes, a single qubit is an effective ancilla qubit (*16–18*), and for any code, a flag qubit system can be used instead of

entangled ancilla qubits (*19*). Knill and Steane present alternative methods where logical qubits in the same code are used to determine the error syndrome in one or two steps (*20*, *21*). Many other methods are possible (*22*), especially when considering the interplay between noisy syndrome bits and repeated measurements (*23*).

In this work, we experimentally compare Shor and Steane methods applied to the [[9,1,3]] subsystem Bacon-Shor code (*24*). The stabilizer group of this code is generated by four weight-6 operators: $S_1 = Z_1Z_2Z_3Z_4Z_5Z_6$, $S_2 = Z_4Z_5Z_6Z_7Z_8Z_9$, $S_3 = X_1X_2X_4X_5X_7X_8$, and $S_4 = X_2X_3X_5X_6X_8X_9$. The values of 12 weight-2 symmetry operators Z_iZ_{i+3} ($i \in \{1,2,3,4,5,6\}$) and X_iX_{i+1} ($i \in \{1,2,4,5,7,8\}$), referred to as gauge operators, can freely fluctuate without destroying the logical information. The logical $|0_L\rangle$ ($|+_L\rangle$) state consists of three rows (columns) of physical $\frac{1}{\sqrt{2}}(|+++\rangle + |---\rangle)$ [$\frac{1}{\sqrt{2}}(|000\rangle + |111\rangle)$] Greenberger-Horne-Zeilinger (GHZ) states for a particular gauge and can be prepared fault-tolerantly with only unitary operations (*18*). We prepare the $|0_L\rangle$ logical state using the circuit shown in Fig. 1A. In this circuit, any single- or two-qubit error of weight-1 is equivalent to a single error together with the action of a gauge operator. The logical operator $Z_L = Z_1Z_2Z_3$ can be fault tolerantly measured by transversal Z readouts followed by a majority vote among the row parities $Z_{3i+1}Z_{3i+2}Z_{3i+3}$ ($i = 0,1,2$).

The canonical method to determine the error syndrome of Bacon-Shor code is to measure weight-2 gauges (*24–26*). The values of four stabilizer generators can be obtained by products of gauge measurement outcomes. An alternative approach is to measure the four stabilizer generators directly with single qubit ancilla (*27*). This method depends on a specific ordering of the circuit that follows the underlying gauges (*28*). The circuit for measuring the Z stabilizers S_1 and S_2 is presented in Fig. 1B. From a broad perspective, both of these methods fall under the category of Shor EC. Measuring the full-syndrome requires 24 to 48 two qubit gates between the data qubits and syndrome qubits in an adaptive approach (*18*).

In Steane EC (*20*), a transversal controlled-not (CNOT) gate is applied between data and a fault tolerantly prepared $|+_L\rangle$ ($|0_L\rangle$) ancilla resource state used for correcting X (Z) errors, as illustrated in Fig. 1C. When correcting X (Z) errors, the data (ancilla) block is the control of CNOT, while the ancilla (data) block is the target. After the transversal CNOT gate, all qubits are measured in Z (X) basis.

[1]Duke Quantum Center, Duke University, Durham, NC 27701, USA. [2]Department of Electrical and Computer Engineering, Duke University, Durham, NC 27708, USA. [3]Department of Physics, Duke University, Durham, NC 27708, USA. [4]Department of Chemistry, Duke University, Durham, NC 27708, USA.
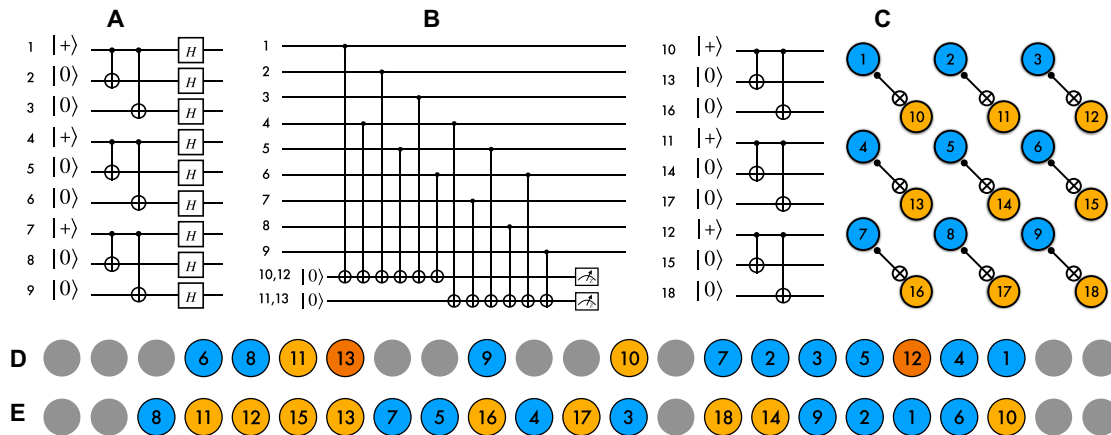*Corresponding author. Email: marko.cetina@duke.edu

Fig. 1. Implementation of Shor- and Steane-style error checks of the Bacon-Shor code using 23 trapped ions. (**A**) $|0\rangle$ state of the Bacon-Shor code is encoded into nine physical qubits by preparing three copies of a three-qubit GHZ-entangled state. (**B**) Shor-style checks are performed by mapping syndromes $S_1 = Z_1 Z_2 Z_3 Z_4 Z_5 Z_6$ and $S_2 = Z_4 Z_5 Z_6 Z_7 Z_8 Z_9$ to qubits 10 and 11 serving as ancillae. Separate ancilla qubits (12 and 13) are used to repeat the measurement. (**C**) Steane-style syndrome extraction involves preparing a $|+_L\rangle$ logical state using the circuit on the left applied to qubits 10 to 18 and then performing a transversal CNOT between the data qubit and the ancilla qubit. (**D**) Ion positions in the 23-qubit ion used for Shor-style syndrome extraction. (**E**) Ion positions in the 23 ion chain used for Steane-style syndrome extraction.

The error syndrome is obtained by parities of Z (X) stabilizer elements. We then assume that the syndrome is perfect and apply correction immediately based on the syndrome. Steane EC requires only 18 two qubit gates for the Bacon-Shor code to touch the data or teleport the data, respectively. For any code, when gate errors dominate, it is expected that Steane or Knill EC will outperform Shor EC. The cost is the preparation of the complicated ancilla qubits. For the case of the distance three Bacon-Shor code, the logical state can be directly prepared fault tolerantly using unitary preparation (*3*). This is a key advantage for small circuits and greatly benefits a Steane EC circuit.

RESULTS

We use an ion trap quantum computer to test the hypothesis that Steane EC induces less noise on the logical qubits than Shor EC. To perform this test, we prepare the logical $|0\rangle$ state of the Bacon-Shor code ($|0_L\rangle$) and check the value of its logical Z operator using an error-corrected measurement after a single EC cycle. We also limit our interrogation to measuring the X error syndrome. A full EC cycle would require measuring both the X and the Z error syndromes.

The experimental system, based on (*3*), consists of a near-equispaced linear chain of ^{171}Yb$^+$ ions storing ground-state hyperfine qubits. The ions are held in a Sandia HOA-2.1.1 trap housed in a room-temperature ultrahigh vacuum chamber and are manipulated using a 355-nm Raman process driven by single global laser beam together with up to 32 individually addressed beams. The state of each qubit is read out using an array of photomultiplier tubes coupled to multimode fibers.

To allow a direct comparison between fault-tolerant Shor and Steane EC, we extended our experimental system from 15 to 23 ions. The increase in the number of ions leads to a sharp increase in their axial motion and the attendant errors in their individual addressing (*29*). To counter these errors, we confined the ions to decrease the chain spacing from 4.4 to 3.8 μm and tailored our gates to suppress coupling to the chain's axial motion, as described in Materials and Methods. The mapping from qubits to ions for both Shor and Steane

EC experiments are shown in Fig. 1 (D and E, respectively). The CNOT and Hadamard gates are decomposed in a standard way into native Mølmer-Sørensen gates and single-qubit rotations (*30*). Details of the physical circuit model can be found in the Supplementary Materials.

To demonstrate our ability to control errors in our larger system, we perform a logical state preparation and measurement (SPAM) test. After preparing the $|0_L\rangle$ state, we measure all qubits in the Z basis and decode the logical Z-measurement outcome via majority voting. The logical SPAM error rate, defined as the probability of measuring $Z_L = -1$, is observed to be $0.20^{+0.27}_{-0.14}$ %, which is comparable with the previous result of 0.23(0.13)% in (*3*).

Shor EC

Our baseline experiment is a Shor-style syndrome extraction of the two Z stabilizers following the previous work (*3*). The Shor EC procedure for X errors involves measuring two weight-6 Z stabilizer generators S_1 and S_2 separately.

A fault-tolerant Shor EC cycle requires measuring the checks in multiple rounds. An adaptive protocol for general distance three quantum error-correction code (*27*, *31*) is to measure the stabilizer generators once and repeat the measurement if a nonzero error syndrome is detected. (*31*).

We use our 23 ion system to perform up to two rounds of Z-syndrome extraction by mapping Z stabilizers to four ancilla qubits as shown in Fig. 1B. We denote the Z syndrome obtained in round $r = 1,2$ by a two-bit string $s^{(r)} = s_1^{(r)} s_2^{(r)}$, where $s_i^{(r)} = 1$ (0) if the measurement of S_i in round r outputs -1 (1). We decode the experimental data using three different decoders:

1) Single-shot decoder: We perform only a single extraction round and immediately apply correction based on $s^{(1)}$. The correction operators corresponding to $s^{(1)} = 00, 01, 11$, and 10 are I, X_1, X_4, and X_7, respectively. This decoder is not fault-tolerant because a single X error between the two CNOT gates on qubits 4, 5, or 6 will always result in syndrome $s^{(1)} = 01$. After correction, an additional X error will be introduced on qubit 7; the original error is amplified and becomes uncorrectable.

2) Adaptive decoder I: If $s^{(1)} = 00$, then we do not perform the next extraction round and no correction is applied; otherwise, the next extraction round is triggered. As one error already occurred in round 1, we assume that $s^{(2)}$ is trustworthy and apply correction based on $s^{(2)}$ only. The correction operators corresponding to $s^{(2)} = 00, 01, 11,$ and 10 are $I, X_1, X_2,$ and X_3, respectively.

3) Adaptive decoder II: The adaptive decoder I can be improved by noticing that the second extraction round is unnecessary when $s^{(1)}$ is equal to 10 or 11. The weight-1 faults compatible with $s^{(1)} = 10$ are a single-qubit X error on qubits 1 to 3 before the two-qubit gates, or a single measurement error on $s_1^{(1)}$. We apply the correction X_1 immediately to correct the data-qubit error (up to a gauge) or convert the measurement error into a data-qubit error without amplification of errors. The only weight-1 faults compatible with $s^{(1)} = 11$ are single-qubit X errors on qubits 4, 5, or 6 before all two-qubit gates, which are corrected by X_2. If $s^{(1)} = 01$, then another extraction round is performed and the correction is applied based on $s^{(2)}$.

Our current experimental setup does not support mid-circuit measurements and therefore cannot execute adaptive circuits. To estimate the logical error rate of adaptive decoders, we post-select data from two experiments, denoted by E_1 and E_2, in which either one or two extraction rounds are deterministically performed. Let $\mu[s^{(1)}]$ be the probability that the first-round syndrome $s^{(1)}$ is measured, and let $\lambda_r[s^{(1)}]$ ($r = 1, 2$) be the conditional logical error rate in experiment E_r with first-round syndrome $s^{(1)}$ when $s^{(r)}$ is used for correction. The total logical error rates for single-shot decoder, adaptive decoder I, and adaptive decoder II are calculated as

$$p_L^{(SS)} = \mu(00)\lambda_1(00) + \mu(10)\lambda_1(10) + \mu(11)\lambda_1(11) + \mu(01)\lambda_1(01),$$
$$p_L^{(A1)} = \mu(00)\lambda_1(00) + \mu(10)\lambda_2(10) + \mu(11)\lambda_2(11) + \mu(01)\lambda_2(01), \quad (1)$$
$$p_L^{(A2)} = \mu(00)\lambda_1(00) + \mu(10)\lambda_1(10) + \mu(11)\lambda_1(11) + \mu(01)\lambda_2(01)$$

respectively. We note that the values $\mu[s^{(1)}]$ for E_1 and E_2 agree with negligible difference. In our experiment, we find that $p_L^{(SS)} = 9.7(1.2)\%$ (first row in Fig. 2), $p_L^{(A1)} = 10.0(1.5)\%$ (second row in Fig. 2), and $p_L^{(A2)} = 7.8(1.6)\%$ (third row in Fig. 2). We see that adaptive decoder II already outperforms single-shot decoder in the experiment. The experimental values of λ_r, which are included in the Supplementary

Materials, show that $\lambda_1[s^{(1)}] < \lambda_2[s^{(1)}]$ for $s^{(1)} = 00, 10, 11$, while $\lambda_1(01) > \lambda_2(01)$. This verifies that repetitive syndrome extraction helps when $s^{(1)} = 01$, for which single-shot decoder fails to maintain fault tolerance.

We also consider the use of one-round Shor-style circuit for post-selection of $|0_L\rangle$, whose logical error and rejection rates are $\lambda_1(00)$ and $1 - \mu(00)$. From Table 1, we see that Shor-style post-selection introduces much more logical X errors than direct preparation of $|0_L\rangle$ ($1.30^{+0.68}_{-0.49}\%$ versus $0.20^{+0.27}_{-0.14}\%$), with the extra cost of nonvanishing rejection rate.

Steane EC

Our Steane EC experiment involves the prepartion of the data qubits in the $|0_L\rangle$ state and the syndrome qubits in the $|+_L\rangle$ state. A transversal CNOT between the data qubits and the syndrome qubits allows X errors in the preparation of the data qubits to be mapped to the syndrome qubits (1c). For arbitrary data states, the logical ancilla state should be prepared in $|+_L\rangle$, which is not affected by the logical CNOT for any ideal input logical state. For the $|0_L\rangle$ data state, we can also use another $|0_L\rangle$ state as ancilla. We test both ancilla states and find a logical error on the data qubit of 4.93(0.49)% for a $|+_L\rangle$ ancilla (row 6 in Fig. 2) and a 4.75(0.5)% a $|0_L\rangle$ (row 7 in Fig. 2) after we correct based on the ancilla qubit measurements. We find that the Steane method yields lower logical data error compared to the Shor syndrome readout with P value of at least 3.7%.

Another way to measure the negative impact of the syndrome measurement circuit on the code is to decode the final logical measurement without the use of any syndrome information from the intermediate EC cycle. The corresponding logical error rate is termed as disturbance. For both Shor and Steane EC, we consider the disturbance of a single syndrome extraction round. We find that the resulting disturbance of the data qubits (Fig. 2) is at least 70(12)% less for Steane EC than for Shor EC .

When we use a Steane-style circuit with post-selection, our state will be in the wrong logical state with probability $O(p^3)$ for a probabilistic error model. Our state post-selected on errors measured when the ancilla qubit is prepared in the $|+_L\rangle$ state is $0.66^{+0.27}_{-0.21}\%$ (row 10 in Fig. 2), which outperforms the Shor-based post-selection (row 5 in Fig. 2). If we use an ancilla in the $|0_L\rangle$ state, then we can also
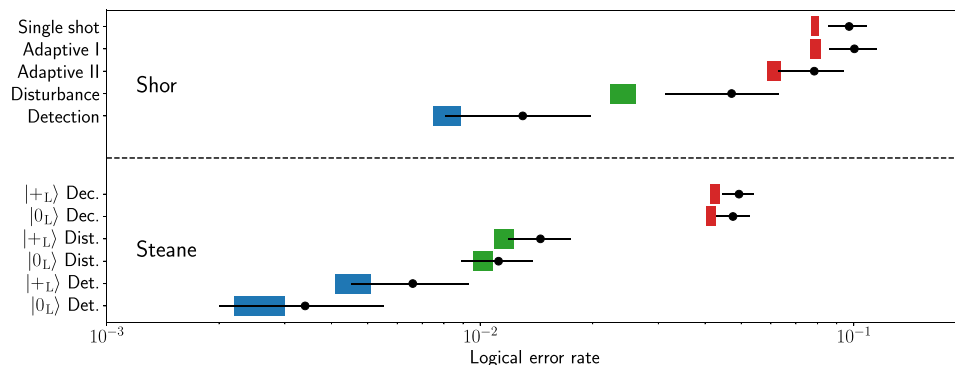
Fig. 2. Comparison of Shor-style and Steane-style EC protocols with the ancilla in the $|0_L\rangle$ and $|+_L\rangle$ logical states. Decoder (Dec., red) corrects the logical qubit with syndrome information before the final correction based on logical qubit measurement. Disturbance (Dist., green) only corrects based on the logical qubit measurement. Detection (Det., blue) only keeps the data when no errors are detected on the ancilla qubits. The Shor disturbance measurement reflects only a single round of syndrome measurement. The black dots are the experimental data, and the bars are numerical simulations. The errors correspond to 95% confidence intervals. The simulations are based on error characterization experiments and are not adjusted for the logical qubit experiments.

post-select on logical errors. This further improves the logical fidelity output to $0.34^{+0.21}_{-0.14}$ % (last row in Fig. 2), which is an improvement over not checking the error syndrome at the cost of a 34% rejection rate.

Simulation and perspective

We developed a simulation of the experiment based on unitary evolution of the full wave function where errors are added probabilistically as unwanted unitary transformations. The applied unitaries are derived from experimentally measured errors in SPAM, single qubit gates, two-qubit gates including cross-talk errors, and the effects of axial heating of the ion chain. The full details are in the Supplementary Materials.

The simulation is not fit to the logical error data and the parameters are instead determined from independent characterization and calibration experiments. This results in non-identical errors on different qubits in the chain. We see broad agreement between the simulated and the experimental data across the measured parameters.

The simulation allows us to determine the best steps forward for reducing the logical error. We find that the dominant source of error is the heating of the lowest axial mode at the rate of 180 phonons/ms, which causes noise on the Rabi frequencies on all ions (29). The single- and two-qubit gates in the EC circuit experience random under-rotation errors that grow with time. Our simulation method considers the time correlations of under-rotation errors within the circuit. While small single-qubit under-rotation errors are reduced

by SK1 composite pulse sequence (32), the XX gates are strongly affected. Another significant source of error is the dephasing during two-qubit gates. In our experiments, the effect of cross-talk on logical gate performance is negligible.

The axial heating can be reduced by cooling the trap (33), preparing the trap surface (34), or both (35), and its effects can be countered by sympathetic cooling (29). The reduction of axial temperature together with a decrease in the axial phase gradients of the addressing beams should also reduce the two-qubit dephasing (36). Dephasing can be further reduced by suppressing the relative intensity fluctuations of the red and blue sidebands in a (fiber-coupled) double-pass acousto-optic modulator (AOM) setup.

A limitation in our experiment is that detecting error syndromes does not help the decoding of the logical qubit without post-selection. In all cases, it would be better not to apply the syndrome extraction gadget. Using the simulator, we explore a scenario where the heating rate is reduced to 0, and Z (X) errors during two-qubit gates is reduced by a factor of 5 (4). We assume that the initial temperature of the lowest axial mode remains unchanged at 5.5 mK.

In this limit, we find that the logical error rate decreases to 0.81(0.06)% for Steane EC and 2.45(0.16)% for Shor EC (Table 2). Both methods see substantial improvements in errors, and Steane EC continues to outperform the Shor EC. The post-selected states would be prepared with a lower rejection rate. For Steane with $|0_L\rangle$ ancilla, we see a reduction of the rejection rates by a factor of 3 and of the error rates by a factor of 4 over the current simulated values (Table 1).

Table 1. Comparison of post-selection protocols for between Shor- and Steane-style error detection circuits. $|0_L\rangle$ We calculate the logical error rate (LER) based on the experimental (EXP) data, our simulation (SIM) of the experiment, and our predicted simulation for future experimental improvements (IMP). Post-selection leads to a rejection rate (RR), which is the fraction of trials that are not post-selected.

Postselection protocol		Direct $|0_L\rangle$ preparation	Shor	Steane ($|+_L\rangle$ ancilla)	Steane ($|0_L\rangle$ ancilla)
LER	EXP	$0.20^{+0.27}_{-0.14}$ %	$1.30^{+0.68}_{-0.49}$ %	$0.66^{+0.27}_{-0.21}$ %	$0.34^{+0.21}_{-0.14}$ %
	SIM	$0.22^{+0.03}_{-0.03}$ %	$0.82^{+0.07}_{-0.07}$ %	$0.46^{+0.05}_{-0.05}$ %	$0.26^{+0.04}_{-0.04}$ %
	IMP	$0.14^{+0.03}_{-0.02}$ %	$0.42^{+0.05}_{-0.04}$ %	$0.06^{+0.02}_{-0.02}$ %	$0.06^{+0.02}_{-0.01}$ %
RR	EXP		(35.32 ± 1.87)%	(27.10 ± 1.44)%	(33.81 ± 1.32)%
	SIM		(33.04 ± 0.29)%	(33.67 ± 0.36)%	(32.16 ± 0.35)%
	IMP		(20.71 ± 0.25)%	(15.98 ± 0.25)%	(12.32 ± 0.23)%

Table 2. Comparison between Shor and Steane EC protocols. We calculate the LER and disturbance (DSTB) based on the EXP data, our SIM of the experiment, and our predicted simulation for future experimental IMP. DSTB is the logical error rate when no ancilla readout information is used. Details of the protocols are in the Results section of the main text.

EC protocol		Shor			Steane	
		Single shot	Adaptive I	Adaptive II	$|+_L\rangle$	$|0_L\rangle$
LER	EXP	(9.72 ± 1.16)%	(10.04 ± 1.45)%	(7.84 ± 1.57)%	$4.93^{+0.51}_{-0.48}$ %	$4.75^{+0.51}_{-0.47}$ %
	SIM	(7.87 ± 0.17)%	(7.92 ± 0.26)%	(6.11 ± 0.26)%	(4.25 ± 0.13)%	(4.15 ± 0.12)%
	IMP	(5.68 ± 0.14)%	(3.11 ± 0.17)%	(2.45 ± 0.16)%	(0.81 ± 0.06)%	(1.08 ± 0.07)%
DSTB	EXP		(4.71 ± 1.59)%		$1.45^{+0.30}_{-0.26}$ %	$1.12^{+0.26}_{-0.23}$ %
	SIM		(2.41 ± 0.10)%		(1.16 ± 0.07)%	(1.02 ± 0.06)%
	IMP		(1.49 ± 0.08)%		(0.18 ± 0.03)%	(0.21 ± 0.03)%

DISCUSSION

Our work has demonstrated the promise of Steane-style quantum EC. The next level of investigation requires intermediate measurements to measure and reset the Steane ancilla. A larger system capable of pairing up multiple ion chains would allow us to perform Knill-style syndrome extraction via teleportation. Knill's method is both efficient and prevents the unwanted accumulation of leakage errors to states outside of the qubit space. Leakage errors present a challenge to Shor and Steane-style EC without additional leakage reduction gadgets (37). Further work should focus on examining codes with larger distance where the challenge of preparing logical ancilla state increases, but the disturbance to data introduced by Steane/Knill EC circuit also decreases. It would also be interesting to investigate intermediate error-correction circuits in between Shor and Steane methods (22).

The procedure outlined here is ideal for quantum architectures that enable low-error rate shuttling to connect distant qubits without multiple swap operations. For ions, a charge-coupled device–type architecture (38) that uses longer than two-ion chains as a baseline would work well. Ion temperature can be lowered after merging chains using sympathetic cooling and benefits from precooling one of the two merged chains. For neutral atoms, a change in encoding procedures from the Steane code (13) would enable testing these results in that system. In future quantum dot systems, which allow coherent transport of spins, a similar result will hold (39). Steane EC is also promising for a multilayer architecture where a logical ancilla qubit layer can be directly coupled to the data qubit layer.

MATERIALS AND METHODS

Experimental system

To perform two-qubit Mø lmer-Sø rensen gates, we use individually addressed Raman beams to induce state-dependent amplitude-modulated forces that couple the internal states of the ions to one set of the ions' radial modes (Fig. 3). For robustness to mode frequency drift, we choose the mean frequencies of these forces (lines 1 to 7 on Fig. 3) half-way between the frequencies of the neighboring radial modes (dashed lines on Fig. 3) and use a time-reversal symmetric modulation waveforms that null the mean motional displacement of each mode during the gate (40).

Predominant axial motion of the ions at the lowest axial mode frequency of 174 kHz causes the gate waveforms experienced by the ions to acquire modulation sidebands (lines 1a to 7a in Fig. 3). To suppress
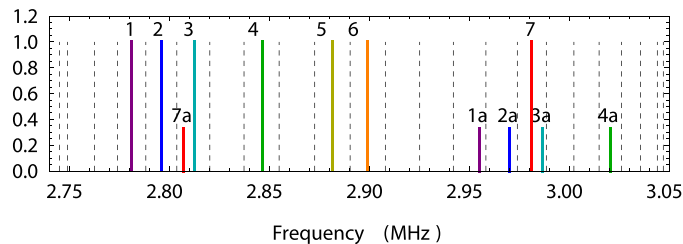

Fig. 3. Choice of carrier frequencies for amplitude-modulated XX gates in long ion chain. The frequencies of the radial modes of 23 ions (dashed lines) compared to the mean frequencies of the state-dependent amplitude-modulated forces (lines 1 to 7) that are used to perform entangling gates. Because of motion of the lowest axial mode of the ion chain, the gate waveforms as seen by the ions acquire sidebands at ±174 kHz (lines 1a to 7a).
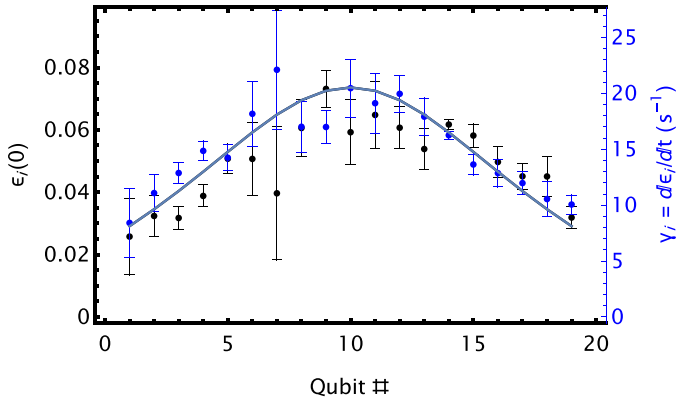
the resulting errors, we choose the XX gate waveforms for which these sidebands are separated from the radial modes by at least 3.3 kHz.

To mitigate the impact of the axial ion motion on the single-qubit gates, we first align the addressing beams to minimize the phase gradient of the Raman coupling along the chain axis.

The effects of Raman amplitude fluctuations are countered by using SK1 pulses (32). Each 13-µs-long pulse segment of the compound pulse is further Gaussian-shaped so as to null the first-order contribution of the ions' motion at the lowest axial frequency to the resulting unitary.

Physical error rates and error modeling

The logical error rate for each experiment is numerically estimated via Monte-Carlo sampling of state vector evolutions. The leading source of errors in the experiment is the motion of the ion chain in the lowest axial mode (mode 0) with frequency $\omega_0 = 2\pi \times 174$ kHz, which misaligns the ions relative to the individual-addressing beams. These beams are nearly Gaussian-shaped with waist $w = 450$ nm, corresponding to instantaneous carrier Rabi frequency of i-th ion $\Omega_i = \Omega_{i,\max}\exp\left(-x_i^2/w^2\right)$, where x_i is the instantaneous position of this ion.

Because, in our experiments, the energy E_0 of mode 0 corresponds to hundreds of motional quanta $\hbar\omega_0$, we describe this motion classically. In this limit, Ω_i varies with the classical phase ϕ of the ions' motion as $\Omega_i = \Omega_{i,\max} \exp\left(-2\epsilon_i\sin^2\phi\right)$, where $\epsilon_i = b_{i0}^2 \hbar E_0 / mw^2\omega_0^2$ is a unitless decay parameter, m is the ion mass, and b_{i0} is the participation of ion i in mode 0. Averaging over the motional phase ϕ gives the mean Rabi frequency $\overline{\Omega}_i = \Omega_{i,\max} f\left(\epsilon_i\right)$ of ion i, where $f(x) = e^{-x}I_0(x)$, with I_0 as the modified Bessel function of the first kind.

Uncertainty in the energy E_0 introduces rotation errors in both single- and two-qubit gates (29). A carrier rotation of ion i by angle θ becomes an imperfect rotation by angle $\theta f(\epsilon_i)$, while a two-qubit $XX(\theta)$ between ions i and j gate becomes $XX[\theta f(\epsilon_i)f(\epsilon_j)]$. Note that, in our model, ϵ_i and ϵ_j are perfectly correlated because they are both derived from the same random number n of motional quanta as $\epsilon_{i,j} = nu_{i,j}$, with $u_{i,j} = b_{(i,j)0}^2 \hbar / \left(mw^2\omega_0\right)$ fixed unitless numbers that we calculate.

We assume that at beginning of the circuit ($t = 0$) mode 0 is in thermal equilibrium. In this case, ϵ_i will be Boltzmann-distributed with mean $\overline{\epsilon}_i(0)$, which we determine by fitting carrier Rabi oscillations of all the qubits to the model from (29) (Fig. 4). We fit $\overline{\epsilon}_i$ to $\overline{n}u_i$ to obtain the mean initial phonon number $\overline{n}(0) = 660(20)$. At the beginning of the simulation, we sample n from from the Boltzmann distribution with mean $\overline{n}(0)$.

Without mid-circuit sympathetic cooling, the energy of the lowest axial mode increases with time, causing the gates applied near the end of the circuit to perform worse. We determine the rate of increase γ_i of ϵ_i by fitting Rabi oscillations at later times in the circuit (Fig. 4). A fit of γ_i to $\overline{n} \times u_i$ yields the axial heating rate $\dot{n} = 180$ (4) phonons/ms. This is consistent with a noisy electric field that is homogenous across the chain and corresponds to the heating rate of 8.3(2) phonons/ms of a single ion with axial frequency of $2\pi \times 174$ kHz.

We model heating of mode 0 as a biased random walk in n. Given $n(r\Delta t)$, we set

$$n[(r+1)\Delta t] := \begin{cases} n(r\Delta t)+1 & \text{withprob.} \quad (n+1)\times\overline{n}\Delta t \\ n(r\Delta t)-1 & \text{withprob.} \quad n\times\overline{n}\Delta t \\ n(r\Delta t) & \text{otherwise} \end{cases} \tag{2}$$

Fig. 4. Effect of axial heating on ion addressing. The decay parameter (29) ϵ_i of carrier Rabi oscillations as function of qubit number at the beginning of the circuit (black) and the rate of its increase with time γ_i (blue). Fits of the data to ϵ_i, $\gamma_i \sim b_{i0}^2$ (solid line) yield the initial mean phonon number in the lowest axial mode $\bar{n} = 660(20)$ and the heating rate of the lowest axial mode of the chain of $\dot{\bar{n}} = 180(4)$ phonons/ms.

To determine the imperfect rotation angle for a gate starting at time t, for each raw single-qubit rotation $\sigma_\phi(\theta)$ in the x-y plane on ion i, the actual rotation angle $\tilde{\theta}$ is set to be

$$\tilde{\theta} := \theta\left[1 + u_i\bar{n}(0)\right]f\left[u_i n(t)\right] \quad (3)$$

The extra coefficient $\left[1 + u_i\bar{n}(0)\right]$ compensates the average under-rotation at $t = 0$. Similarly, for each two-qubit $XX(\theta)$ rotation between ions i and j starting at time t, the actual rotation angle $\tilde{\theta}$ is set to be

$$\tilde{\theta} := \theta\left[1 + u_i\bar{n}(0)\right]\left[1 + u_j\bar{n}(0)\right]f\left[u_i n(t)\right]f\left[u_j n(t)\right] \quad (4)$$

We ignore the fluctuation of $n(t)$ within a single- or two-qubit gate duration.

Another contributing factor to logical error is dephasing during the entangling gates. Sources of this dephasing include axial phase gradients in the individual addressing beams combined with axial motion of the ions, fluctuations, in the ratio of the red and blue sidebands and fluctuations in coupling to other Zeeman levels. We model the dephasing during the XX gate between qubits i and j as a pair of equal single-qubit phase-flip Pauli channels

$$\mathcal{E}_Z(\rho) = \left[1 - p_Z^{(i,j)}\right]\rho + p_Z^{(i,j)}Z\rho Z \quad (5)$$

on qubits i and j following the gate. We determine the error probability $p_Z^{(i,j)}$ for each ion pair by performing K paired maximally entangling gates with opposite rotation angles ($+XX/-XX$) inside a Ramsey sequence on both ions. We fit the resulting Ramsey fringes to extract the contrast for both ions and set $p_Z^{(i,j)}$ to 1/2 of the average measured contrast loss per gate (Fig. 5). We assume $p_Z^{(i,j)}$ to be time independent.

To model dephasing on idling qubits due to optical phase fluctuations and unwanted ac shifts, for each idling circuit location with duration t, we apply a Pauli Z error with probability $\Gamma_{\text{deph}}t$, where Γ_{deph} is the measured dephasing rate. For each measurement, we flip the outcome from $|1\rangle$ ($|0\rangle$) to $|0\rangle$ ($|1\rangle$) with probability $p_{1\to0} = 4 \times 10^{-3}$ ($p_{0\to1} = 1.5 \times 10^{-3}$).

Acoustic and electrical cross-talk in the 32-channel AOM leads to cross-talk in individual addressing. We characterize this cross-talk by measuring the ratio of the carrier Rabi frequencies for a single addressed ion and the remaining qubits (Fig. 6). The effects of cross-talk on single-qubit gates are suppressed using SK1 pulses.

We estimate the contribution of two-qubit cross-talk errors on logical infidelity. When a two-qubit gate $X_iX_j(\theta)$ is applied between ions i and j with carrier Rabi frequency $\Omega(t)$, each other ion $k \neq i, j$ will also experience a spin-dependent force with carrier Rabi frequency $\Omega_k(t) = (\chi_{i,k} + \chi_{j,k})\Omega(t)$, where $\chi_{i,k}$ are the carrier cross-talk ratios. After the two-qubit drive is finished, two small two-qubit rotations $X_iX_k\left(\theta_k^{(1)}\right)$ and $X_jX_k\left(\theta_k^{(2)}\right)$ will be applied between ion pairs (i, k) and (j, k). Let $A_{i,j,k}$ be the rotation angle on ion pair (i, k) when the two-qubit AM pulse sequence on the pair (i, j) is applied on the pair (i, k) instead. Then, we have

$$\theta_k^{(1)} = \left(\chi_{i,k} + \chi_{j,k}\right)A_{i,j,k} \quad (6)$$

and

$$\theta_k^{(2)} = \left(\chi_{i,k} + \chi_{j,k}\right)A_{j,i,k} \quad (7)$$
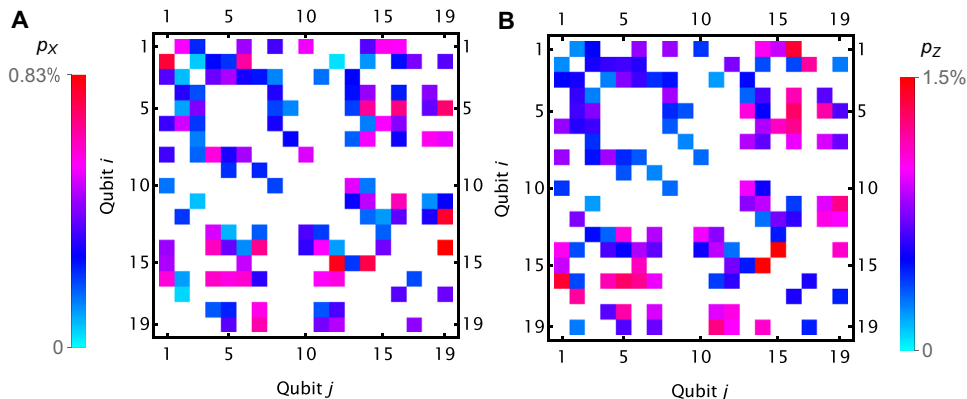

Fig. 5. X-flip errors in XX gates. The measured probabilities p_X of an X error (**A**) and p_Z of a Z error (**B**), per qubit, per fully entangling XX gate between qubits i and j.
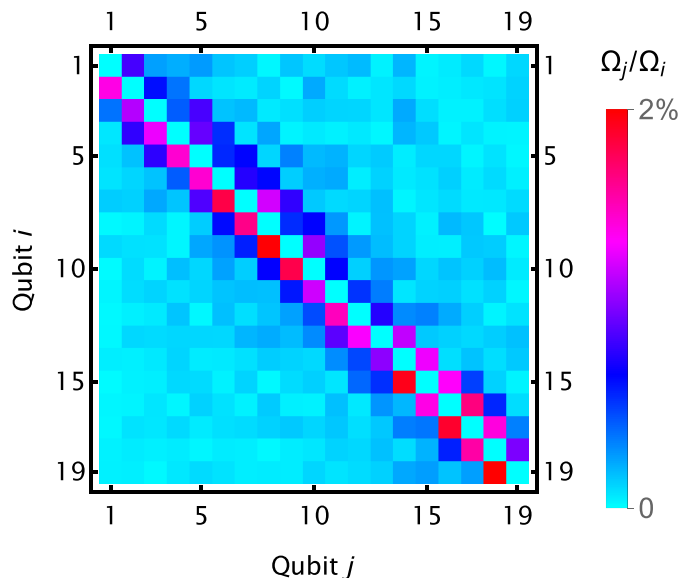
Fig. 6. Ion addressing cross-talk. The measured ratio of the Rabi frequency Ω_j of qubit j and the Rabi frequency Ω_i of the target resonantly driven qubit i.

The values $A_{i,j,k}$ can be analytically calculated from the pulse shapes, motional mode frequencies, and spin-motion coupling strengths.

Motional dephasing during two-qubit gates causes X errors (*41*) and contributes to the logical error rate. We determine the probability p_X of single-qubit Pauli X errors during the $X_iX_j(\theta)$ gate between ions i and j by preparing the target qubits in the $|0\rangle$ state and measuring the probability p of a single bit flip after K fully entangling gates. We set $p_X = p/K/2$ and show the results in Fig. 5B. See "Demonstration of fault-tolerant Steane quantum error correction" by the research groups led by M. Müller (Aachen & Jülich) and by T. Monz (Innsbruck) for related work (*42*).

Supplementary Materials

This PDF file includes:
Supplementary Text
Figs. S1 to S3
Tables S1 to S3

REFERENCES AND NOTES

1. C. Gidney, M. Ekerå, How to factor 2048 bit RSA integers in 8 hours using 20 million noisy qubits. *Quantum* **5**, 433 (2021).
2. M. E. Beverland, P. Murali, M. Troyer, K. M. Svore, T. Hoefler, V. Kliuchnikov, G. H. Low, M. Soeken, A. Sundaram, A. Vaschillo, *Assessing requirements to scale to practical quantum advantage*. arXiv:2211.07629 [quant-ph] (2022).
3. L. Egan, D. M. Debroy, C. Noel, A. Risinger, D. Zhu, D. Biswas, M. Newman, M. Li, K. R. Brown, M. Cetina, C. Monroe, Fault-tolerant control of an error-corrected qubit. *Nature* **598**, 281–286 (2021).
4. N. H. Nguyen, M. Li, A. M. Green, C. Huerta Alderete, Y. Zhu, D. Zhu, K. R. Brown, N. M. Linke, Demonstration of shor encoding on a trapped-ion quantum computer. *Phys. Rev. Appl.* **16**, 024057 (2021).
5. V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, M. H. Devoret, Real-time quantum error correction beyond break-even. *Nature* **616**, 50–55 (2023).
6. C. Ryan-Anderson, J. G. Bohnet, K. Lee, D. Gresh, A. Hankin, J. P. Gaebler, D. Francois, A. Chernoguzov, D. Lucchetti, N. C. Brown, T. M. Gatterman, S. K. Halit, K. Gilmore, J. A. Gerber, B. Neyenhuis, D. Hayes, R. P. Stutz, Realization of real-time fault-tolerant quantum error correction. *Phys. Rev. X* **11**, 041058 (2021).
7. C. Ryan-Anderson, N. C. Brown, M. S. Allman, B. Arkin, G. Asa-Attuah, C. Baldwin, J. Berg, J. G. Bohnet, S. Braxton, N. Burdick, J. P. Campora, A. Chernoguzov, J. Esposito, B. Evans, D. Francois, J. P. Gaebler, T. M. Gatterman, J. Gerber, K. Gilmore, D. Gresh, A. Hall, A. Hankin, J. Hostetter, D. Lucchetti, K. Mayer, J. Myers, B. Neyenhuis, J. Santiago, J. Sedlacek, T. Skripka, A. Slattery, R. P. Stutz, J. Tait, R. Tobey, G. Vittorini, J. Walker, D. Hayes, Implementing fault-tolerant entangling gates on the five-qubit code and the color code. arXiv:2208.01863 [quant-ph] (2022).
8. L. Postler, S. Heuβen, I. Pogorelov, M. Rispler, T. Feldker, M. Meth, C. D. Marciniak, R. Stricker, M. Ringbauer, R. Blatt, P. Schindler, M. Müller, T. Monz, Demonstration of fault-tolerant universal quantum gate operations. *Nature* **605**, 675–680 (2022).
9. S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code. *Nature* **605**, 669–674 (2022).
10. N. Sundaresan, T. J. Yoder, Y. Kim, M. Li, E. H. Chen, G. Harper, T. Thorbeck, A. W. Cross, A. D. Córcoles, M. Takita, Demonstrating multi-round subsystem quantum error correction using matching and maximum likelihood decoders. *Nat. Commun.* **14**, 2852 (2023).
11. Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit. *Nature* **614**, 676–681 (2023).
12. Y. Zhao, Y. Ye, H.-L. Huang, Y. Zhang, D. Wu, H. Guan, Q. Zhu, Z. Wei, T. He, S. Cao, F. Chen, T.-H. Chung, H. Deng, D. Fan, M. Gong, C. Guo, S. Guo, L. Han, N. Li, S. Li, Y. Li, F. Liang, J. Lin, H. Qian, H. Rong, H. Su, L. Sun, S. Wang, Y. Wu, Y. Xu, C. Ying, J. Yu, C. Zha, K. Zhang, Y.-H. Huo, C.-Y. Lu, C.-Z. Peng, X. Zhu, J.-W. Pan, Realization of an error-correcting surface code with superconducting qubits. *Phys. Rev. Lett.* **129**, 030501 (2022).
13. D. Bluvstein, S. J. Evered, A. A. Geim, S. H. Li, H. Zhou, T. Manovitz, S. Ebadi, M. Cain, M. Kalinowski, D. Hangleiter, J. P. B. Ataides, N. Maskara, I. Cong, X. Gao, P. S. Rodriguez, T. Karolyshyn, G. Semeghini, M. J. Gullans, M. Greiner, V. Vuletic, M. D. Lukin, Logical quantum processor based on reconfigurable atom arrays. *Nature* **626**, 58–65 (2024).
14. D. Gottesman, Stabilizer codes and quantum error correction. arXiv:quant-ph/9705052 [quant-ph] (1997).
15. P. W. Shor, FOCS '96 (IEEE Computer Society Press, 1996), pp. 56–65.
16. A. G. Fowler, A. M. Stephens, P. Groszkowski, High threshold universal quantum computation on the surface code. *Phys. Rev. A* **80**, 052312 (2009).
17. Y. Tomita, K. M. Svore, Low-distance surface codes under realistic quantum noise. *Phys. Rev. A* **90**, 062320 (2014).
18. M. Li, D. Miller, K. R. Brown, Direct measurement of Bacon-Shor code stabilizers. *Phys. Rev. A* **98**, 050301 (2018).
19. R. Chao, B. W. Reichardt, Flag fault-tolerant error correction for any stabilizer code. *PRX Quantum* **1**, 010302 (2020).
20. A. M. Steane, Active stabilization, quantum computation, and quantum state synthesis. *Phys. Rev. Lett.* **78**, 2252–2255 (1997).
21. E. Knill, Quantum computing with realistically noisy devices. *Nature* **434**, 39–44 (2005).
22. S. Huang, K. R. Brown, Between shor and steane: A unifying construction for measuring error syndromes. *Phys. Rev. Lett.* **127**, 090505 (2021).
23. C.-Y. Lai, Y.-C. Zheng, T. A. Brun, Fault-tolerant preparation of stabilizer states for quantum calderbank-shor-steane codes by classical error-correcting codes. *Phys. Rev. A* **95**, 032339 (2017).
24. D. Bacon, Operator quantum error-correcting subsystems for self-correcting quantum memories. *Phys. Rev. A* **73**, 012340 (2006).
25. P. Aliferis, A. W. Cross, Subsystem fault tolerance with the bacon-shor code. *Phys. Rev. Lett.* **98**, 220502 (2007).
26. D. Poulin, Stabilizer formalism for operator quantum error correction. *Phys. Rev. Lett.* **95**, 230504 (2005).
27. M. Li, M. Gutiérrez, S. E. David, A. Hernandez, K. R. Brown, Fault tolerance with bare ancillary qubits for a [[7,1,3]] code. *Phys. Rev. A* **96**, 032341 (2017).
28. M. Li, D. Miller, M. Newman, Y. Wu, K. R. Brown, 2D compass codes. *Phys. Rev. X* **9**, 021041 (2019).
29. M. Cetina, L. Egan, C. Noel, M. Goldman, D. Biswas, A. Risinger, D. Zhu, C. Monroe, Control of transverse motion for quantum gates on individually addressed atomic qubits. *PRX Quantum* **3**, 010334 (2022).
30. D. Maslov, Basic circuit compilation techniques for an ion-trap quantum machine. *New J. Phys.* **19**, 023035 (2017).
31. R. Chao, B. W. Reichardt, Quantum error correction with only two extra qubits. *Phys. Rev. Lett.* **121**, 050502 (2018).
32. K. R. Brown, A. W. Harrow, I. L. Chuang, Arbitrarily accurate composite pulse sequences. *Phys. Rev. A* **70**, 052318 (2004).
33. J. Labaziewicz, Y. Ge, P. Antohi, D. Leibrandt, K. R. Brown, I. L. Chuang, Suppression of heating rates in cryogenic surface-electrode ion traps. *Phys. Rev. Lett.* **100**, 013001 (2008).
34. D. A. Hite, Y. Colombe, A. C. Wilson, K. R. Brown, U. Warring, R. Jördens, J. D. Jost, K. S. McKay, D. P. Pappas, D. Leibfried, D. J. Wineland, 100-Fold reduction of electric-field noise in an ion trap cleaned with in situ argon-ion-beam bombardment. *Phys. Rev. Lett.* **109**, 103001 (2012).

35. J. A. Sedlacek, J. Stuart, D. H. Slichter, C. D. Bruzewicz, R. McConnell, J. M. Sage, J. Chiaverini, Evidence for multiple mechanisms underlying surface electric-field noise in ion traps. *Phys. Rev. A* **98**, 63430 (2018).

36. R. T. Sutherland, Q. Yu, K. M. Beck, H. Häffner, One- and two-qubit gate infidelities due to motional errors in trapped ions and electrons. *Phys. Rev. A* **105**, 022437 (2022).

37. P. Aliferis, B. M. Terhal, Fault-tolerant quantum computation for local leakage faults. *Quantum Info. Comput.* **7**, 139–156 (2007).

38. D. Kielpinski, C. Monroe, D. J. Wineland, Architecture for a large-scale ion-trap quantum computer. *Nature* **417**, 709–711 (2002).

39. J. Taylor, H.-A. Engel, W. Dür, C. Marcus, P. Zoller, M. Lukin, Fault-tolerant architecture for quantum computation using electrically controlled semiconductor spins. *Nat. Phys.* **1**, 177–183 (2005).

40. M. Kang, Y. Wang, C. Fang, B. Zhang, O. Khosravani, J. Kim, K. R. Brown, Designing filter functions of frequency-modulated pulses for high-fidelity two-qubit gates in ion chains. *Phys. Rev. Appl.* **19**, 014014 (2023).

41. Y. Wang, S. Crain, C. Fang, B. Zhang, S. Huang, Q. Liang, P. H. Leung, K. R. Brown, J. Kim, High-fidelity two-qubit gates using a microelectromechanical-system-based beam steering system for individual qubit addressing. *Phys. Rev. Lett.* **125**, 150505 (2020).

42. L. Postler, F. Butt, I. Pogorelov, C. D. Marciniak, S. Heußen, R. Blatt, P. Schindler, M. Rispler, M. Müller, T. Monz, Demonstration of fault-tolerant steane quantum error correction. arXiv:2312.09745 [quant-ph] (2023).